

# The PACE Trial Invalidates the Use of Cognitive Behavioral and Graded Exercise Therapy in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Review

Mark Vink\*

Family Physician, Soerabaja Research Center, Amsterdam, The Netherlands

\*Corresponding author: Mark Vink, Family Physician, Soerabaja Research Center, Amsterdam, The Netherlands, E-mail: [markvink.md@outlook.com](mailto:markvink.md@outlook.com)

Received date: 12 Jan 2016; Accepted date: 24 Mar 2016; Published date: 30 Mar 2016.

Citation: Vink M (2016) The PACE Trial Invalidates the Use of Cognitive Behavioral and Graded Exercise Therapy in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Review. *J Neurol Neurobiol* 2(3): doi <http://dx.doi.org/10.16966/2379-7150.124>

Copyright: © 2016 Vink M. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## Abstract

The main findings reported in the PACE trial were that cognitive behavioral therapy (CBT) and graded exercise therapy (GET) were moderately effective treatments for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS), and fear avoidance beliefs constituted the strongest mediator of both therapies. These findings have been challenged by patients and, more recently, a number of top scientists, after public health expert Tuller, highlighted methodological problems in the trial. As a doctor who has been bedridden with severe ME for a long period, I analyzed the PACE trial and its follow-up articles from the perspectives of a doctor and a patient. During the PACE trial the eligibility criteria, both subjective primary outcomes, and most of the recovery criteria were altered, creating an overlap of the eligibility and recovery criteria; consequently, 13% of patients were considered "recovered," with respect to 1 or 2 primary outcomes, as soon as they entered the trial. In addition, 46% of patients reported an increase in ME/CFS symptoms, 31% reported musculoskeletal and 19% reported neurological adverse events. Therefore the proportion negatively affected by CBT and GET would be between 46% and 96%, most likely estimated at 74%, as shown in a large survey recently conducted by the ME Association. Medication with such high rates of adverse events would be withdrawn with immediate effect. There was no difference in long-term outcomes between adaptive pacing therapy, CBT, GET and specialist medical care, and none of them were effective, invalidating the biopsychosocial model and use of CBT and GET for ME/CFS. The discovery that an increase in exercise tolerance did not lead to an increase in fitness means that an underlying physical problem prevented this; validates that ME/CFS is a physical disease and that none of the treatments studied addressed this issue.

**Keywords:** Chronic Fatigue Syndrome; CBT; Cognitive Behavioral Therapy; GET; Graded Exercise Therapy; Myalgic Encephalomyelitis; PACE Trial; Pacing

## Abbreviations

6MWT: 6 Minute Walk Test; APT: Adaptive Pacing Therapy; ATS: American Thoracic Society; BMJ: British Medical Journal; CBT: Cognitive Behavioral Therapy; CDC: Centers for Disease Control and Prevention; CFS: Chronic Fatigue Syndrome; CGI: Clinical Global Impression; CMO: Chief Medical Officer, the UK government's principal medical adviser; GET: Graded Exercise Therapy; IOM Institute of Medicine; ME: Myalgic Encephalomyelitis; ME/CFS: Myalgic Encephalomyelitis/Chronic Fatigue Syndrome; SD: Standard Deviation; SF-36: Short Form-36; SMC: Specialist Medical Care; VO<sub>2</sub>max: Maximal Oxygen Uptake; WHO: World Health Organization

## Introduction

Following an extensive review of the literature, the American Institute of Medicine (IOM) concluded that "myalgic encephalomyelitis (ME) and chronic fatigue syndrome (CFS) are serious, debilitating conditions that affect millions of people in the United States and around the world"; that it is a "medical - not a psychiatric or psychological - illness" without a "known cause or effective treatment" which "can cause significant impairment and disability" [1] rendering 25% of patients homebound or bedridden [2] yet "the term chronic fatigue syndrome can result in trivialization and stigmatization" of this "complex, multisystem, and often devastating disorder" [1].

Most doctors are unaware of the seriousness of ME or that it has been classified as a neurological disease by the World Health Organization (WHO) since 1969 [3]; therefore, patients often receive "hostility from their health care provider" and are "subjected to treatment strategies that exacerbate their symptoms" (i.e., CBT and GET) [1].

This is in stark contrast to the conclusions of the PACE trial and follow up articles, which noted that "Chronic fatigue syndrome (CFS, also called Myalgic encephalomyelitis/encephalopathy or ME) is a debilitating condition with no known cause or cure" [4] but concluded that CBT and GET are safe and cured 22% of patients [5]. The trial's conclusions and methodology have been criticized by patients, who have been ignored thus far; however, a recent analysis by public health expert and investigative reporter Tuller highlighted methodological problems in the PACE trial [6-8]. Further issues concerning the methodology of reporting were highlighted by Coyne [9] and Laws [10].

The PACE trial, which cost 8 million dollars [6], was a multicenter trial, the largest CBT and GET trial for ME/CFS conducted thus far, involving 641 patients, established because systematic reviews found that CBT and GET were promising treatments for CFS/ME [11], but "the published trials" were "criticized for being too small, too selective, and for using different outcome measures" [4] but also because of "an absence of data for safety outcomes, and high dropout rates" [11]. The main aim of the

trial was “to provide high quality evidence...about the relative benefits...as well as adverse effects, of the most widely advocated treatments for CFS/ME” (i.e., CBT and GET) [4]. And the trial compared “pacing, defined as adaptive pacing therapy (APT), CBT and GET, when added to specialist medical care (SMC) with SMC alone” [11].

## Objectives

Healthcare, including psychological interventions, should be evidence based, and healthcare evaluation should incorporate the patient’s perspective [12-14]. As a doctor who has been bedridden with severe ME for a long period after GET caused a severe relapse from which I have not recovered, I am in a unique position to combine the patient and doctor perspectives, to elucidate the safety and effectiveness of CBT and GET in ME/CFS by reviewing and analyzing the outcomes of the PACE trial and follow-up articles, in particular the objective ones, to answer the following questions objectively:

- Are CBT and/or GET effective treatments for ME/CFS?
- Are these treatments safe?
- Is recovery from ME/CFS with these treatments possible?
- Were there any other important findings, and if so, what were they?

Full-text articles, including published supplementary material, were retrieved and analyzed, using PubMed to find them. Treatment manuals, patient newsletters, and the PACE trial protocol were also used and this review includes extensive use of direct quotes from the PACE trial and follow-up articles to avoid ambiguity as to whether the information contained therein was fabricated or misinterpreted.

## Background Information

On June 22, 2003, the PACE trial was registered with the International Standard Randomised Controlled Trial Number registry (ISRCTN, number 54285094), “a registry and curated database containing the basic set of data items deemed essential to describe a study at inception” [15], and placed in the “Mental and Behavioural Disorders” condition category [16], even though ME has been classified as a neurological disease by the WHO, which uses CFS as a synonym for ME, since 1969 [3]. Use of the WHO International Classification of Diseases is mandatory in England [17], where the trial took place, and as stated by Dr. L’Hours from the WHO, “according to the taxonomic principles... it is not permitted for the same condition to be classified to more than one rubric” [18].

The “scientific title” of the PACE trial is listed as “a randomised controlled trial of adaptive pacing, cognitive behaviour therapy, and graded exercise, as supplements to standardised specialist medical care... for patients with the chronic fatigue syndrome/myalgic encephalomyelitis or encephalopathy” [16]. “All research and therapy staff and participants are unblinded to treatment allocation of individual participants” [4] and “There is no masking” [16]. In addition, there was no placebo control group; therefore it was not a randomized controlled trial but an unblinded trial that relied on 2 subjective self-reported primary outcomes: “fatigue (measured by Chalder fatigue questionnaire score) and physical function (measured by short form-36 subscale score)” [11]. Yet as concluded by Wood et al. [19], “In trials with subjectively assessed outcomes lack of adequate allocation concealment or of blinding tend to produce over-optimistic estimates of the effect of interventions”.

The “PACE trial protocol: Final version 5.0,” was submitted to the ISRCTN on “01.02.2006”, was “updated from protocol 3.1, 11.02.2005 and incorporates the following” 2 substantial amendments, “4.1,” dated “05.08.2005” and “5.1,” dated “01.02.2006” [20].

The protocol’s publication history shows that “Version 1” was submitted on “30 Oct 2006”, and “Version 2” was resubmitted “24 Jan 2007”, the protocol was accepted on Jan 31, 2007 and published on March 8, 2007 [21]; however, the trial involved the recruitment of “641 participants” “between March 18, 2005, and Nov 28, 2008” [11]. Therefore, the trial began on March 18, 2005, and the final revised protocol was published 2 years later, on March 8, 2007. As Evans noted, “A fundamental principle in the design of randomized trials involves setting out in advance the endpoints that will be assessed in the trial, as failure to prespecify endpoints can introduce bias into a trial and creates opportunities for manipulation” [22]. And therefore the final version of a protocol should be published before the trial starts.

However, Evans [22] also noted that “sometimes new information may come to light that could merit changes to endpoints during the course of a trial” and “Such changes can allow incorporation of up-to-date knowledge into the trial design. However,” they “can also compromise the scientific integrity of a trial” [22]. But “Changes in long-term trials” should be considered “as medical knowledge evolves or when assumptions made in design of the trial appear questionable” but only if the decision “to modify an endpoint” is made “independent of the data obtained from the trial to date,” and only an “external advisory committee that has not reviewed data from the trial” should make those changes. It is “not appropriate” that “decision makers” (i.e., “study sponsors, investigators, and DMCs” who may have “impressions” “of the trial to date,” which “may influence decisions regarding changes in endpoints”) make endpoint “revisions during the trial” [22]. Yet the authors state that “The statistical analysis plan was finalised, including changes to the original protocol, and was approved by the trial steering committee... before outcome data were examined” [11]. This indicates that the changes were implemented by the authors and approved by the trial steering committee, and this procedure differs from that described by Evans above [22]. This statement is of particular importance, as many endpoint changes were made during the PACE trial.

## Eligibility Criteria

According to the PACE trial, “chronic fatigue syndrome (CFS) is a condition characterised by chronic disabling fatigue and other symptoms, which are not better explained by an alternative diagnosis. Myalgic encephalomyelitis/encephalopathy (ME) refers to a severe debilitating illness thought by some to be a separate illness, but by others to be synonymous with CFS. In keeping with the MRC Research Advisory Group report and the CMO’s working group report, we will refer to the illness using both terms: CFS/ME” [4]. Which means that the PACE trial regarded ME and CFS as the same illness.

The eligibility criteria used in the PACE trial included fulfillment of the Oxford Criteria, with “a bimodal score of 6 of 11 or more on the Chalder fatigue questionnaire and a score of 60 of 100 or less on the short form-36 physical function subscale. 11 months after the trial began, this requirement was changed from a score of 60 to a score of 65 to increase recruitment” [11].

According to the first participant newsletter (June 2006), when this change was made in February 2006, 75–80 of 641 patients had been recruited [23], and according to the second participant newsletter (March 2007), “Three new PACE centres have recently started recruiting participants to the trial” and “A seventh centre in Bristol will start recruiting in the spring” [24]. This indicates that recruitment problems were attributed to the fact that only 3 of the 7 PACE centers were open when the trial began in March 2005, that it took almost 2 years to open an additional 3 centers and longer than 2 years to open all 7 centers.

The consequence of changing the eligibility criteria was that healthier people, for whom it is easier to exercise and who experience fewer problems doing so, were selected for participation in a trial to assess the efficacy and safety of 4 treatments, including an exercise treatment (GET), which could have affected related outcomes. This change also created an overlap in the Short Form 36 (SF-36) physical function entry (a score of  $\leq 65$ ) and recovery (a score of  $\geq 60$ ) criteria. This also occurred with a change to the Chalder Fatigue Questionnaire recovery score [5], where by a score of 18 represented both recovery and eligibility, so that 13% of patients were already considered to have recovered on 1 or 2 of the primary outcomes upon entering the trial [6], the only raw data released by the trial, other requests were deemed “vexatious” and refused [6] even though the UK Medical Research Council policy concerning research data sharing states that “publicly-funded research data are a public good, produced in the public interest” which “should be openly available to the maximum extent possible” [25]. And the Research Data Management Policy of Queen Mary University of London, the home of the PACE trial, states that, “publicly funded research data should be made openly available in a timely manner” [26].

Wicherts et al. [27] explored “authors’ reluctance to share data” in psychological studies and found that the “willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results”. They concluded that it was “rather disconcerting that roughly 50% of published papers in psychology contain reporting errors and that the unwillingness to share data was most pronounced when the errors concerned statistical significance” [27].

### The Trial’s ME/CFS Criteria

The main characteristic of ME is abnormally delayed muscle recovery following trivial activities, as noted by infectious disease specialist Dr. Melvin Ramsay [28], who witnessed and documented the outbreak of an unknown disease in the Royal Free Hospital, London, in 1955 which was initially thought to be atypical poliomyelitis and later became known as ME [28]. However, the Oxford Criteria used by the PACE trial “require fatigue to be the main symptom, accompanied by significant disability, in the absence of an exclusionary medical or psychiatric diagnosis (psychosis, bipolar disorder, substance misuse, an organic brain disorder, or an eating disorder)” [11], even though (chronic) fatigue is *not* required to diagnose ME as noted by Dr Dowsett [29] in 1992, who documented that “One of the most striking features of ME is that the patient is not tired all the time”. And on top of that, as noted by Ramsay, a diagnosis of ME should not be made without the abnormally delayed muscle recovery [28].

The Oxford Criteria, which do not exclude patients with “any depressive disorder and any anxiety disorder, including phobias, obsessive-compulsive disorder, and post-traumatic stress disorder” [11], include patients whose fatigue could be caused by a psychiatric disorder which could inflate CBT and GET efficacy and safety outcomes as, contrary to patients with a psychiatric problem, ME/CFS patients suffer from delayed recovery and worsening of symptoms following exercise, as objective evidence provided by Paul et al. [30] and Black et al. [31] for example, showed. Failure to exclude psychiatric disorders resulted in the inclusion of participants in which 33% and 47% had active depression and comorbid psychiatric disorders, respectively [11] and could have included patients with exercise phobia, leading to an erroneous impression of ME/CFS.

In 2003, a group of researchers led by Reeves from the Centers for Disease Control and Prevention (CDC), which included one of the principal investigators from the PACE trial, concluded via consensus, that “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” [32]. This means that the Oxford Criteria should no longer be used.

Unfortunately the PACE trial ignored their own recommendation, which is not surprising, as 2 of the 3 principal investigators and a PACE trial research center leader were involved in creating these criteria in 1991 [33]. Consequently, at baseline, only 67% of participants met the “international criteria for chronic fatigue syndrome” [11], which are the 1994 Fukuda Criteria [34], in which the main characteristic of ME is optional, unlike the 2011 International Consensus Criteria, in which this is a requirement [2] and only 56% of participants fulfilled the “London Criteria for Myalgic Encephalomyelitis” [11], yet Goudsmit, one of the authors of the London Criteria [35], noted, “The version used in the PACE trial was not written by any of those who were involved (“sic!”) with the London Criteria” [36]. According to Shepherd [37], the medical advisor of the ME Association and one of the authors of the original version of the London Criteria, “White et al modified the Task Force version...and called it version 2” omitting the requirement that “it is vital that the M.E. study groups we use in research are as ‘pure’ as possible, the existence of a parallel disease would be grounds for disqualification” [35]. However, as stated by David [38], “British investigators have put forward an alternative, less strict, operational definition which is essentially chronic...fatigue in the absence of neurological signs...with...psychiatric symptoms...as common associated features”.

The differences in percentages when using stricter criteria highlight the heterogeneity of the population because of using the Oxford Criteria. If the original London Criteria [35] had been used the 47% of participants with comorbid psychiatric conditions [11] would have been excluded.

### Subjective Outcomes

#### The Chalder fatigue questionnaire

One of the two primary outcomes of the PACE trial was “fatigue (measured by Chalder fatigue questionnaire score)... up to 52 week” [11] yet “the original bimodal scoring” was changed “to Likert scoring to more sensitively test our hypotheses of effectiveness” [11].

Table 1 shows the Chalder Fatigue Questionnaire [20] showing my own situation, using the original binary or bimodal scoring on the left and Likert scoring on the right [11]. Prior to viewing my scores, it is important to reiterate that I have been bedridden with severe ME for a long period, because I no longer have the muscle power to sit, stand, or walk. I am dependent on others and I was seen by a consultant psychiatrist, who excluded psychiatric disorders.

Table 1 shows that I score 4 on the binary and 18 on the Likert scale. My binary score indicates that I would have been ineligible, and therefore not ill enough, to enter the ME/CFS PACE trial, despite having severe ME, but the Likert score indicates that I would have been eligible for entry. However, one cannot be ill enough and not ill enough simultaneously. A Likert score of  $\leq 18$  was also one of the recovery criteria [11]; therefore, without having received any treatment, I would also be classed as recovered on the Chalder Fatigue Questionnaire even though my medical situation has not changed and I remain bedridden and dependent on others.

This highlights one of several problems with this scale, showing that 2 entirely different outcomes are produced by simply changing the scoring system. This demonstrates that these scores cannot be compared, and changing the scoring system in the midst of a trial, makes this instrument unreliable.

Furthermore, I did not score many points, despite having severe ME, because too many items were depression related, which is irrelevant in ME/CFS. Yet it is not surprising, as the scale, which does not provide a comprehensive reflection of fatigue-related severity, symptomology, or functional disability in ME/CFS [39], was developed by mental health professionals [40]. This also means that improvements could simply be

	Bimodal or binary scoring				Likert scoring			
	0	0	1	1	0	1	2	3
	* Less than usual	* No more than usual	* More than usual	* Much more than usual	* Less than usual	* No more than usual	* More than usual	* Much more than usual
* Do you have problems with tiredness?		0				1		
* Do you need to rest more?				1				3
* Do you feel sleepy or drowsy?		0				1		
* Do you have problems starting things?		0				1		
* Do you lack energy?		0				1		
* Do you have less strength in your muscles?				1				3
* Do you feel weak?		0				1		
* Do you have difficulty concentrating?		0				1		
* Do you make slips of the tongue when speaking?				1				3
* Do you find it more difficult to find the correct word?			1				2	
* How is your memory?		0				1		
<b>Total</b>	<b>4</b>				<b>18</b>			

**Table 1:** Chalder Fatigue Questionnaire which I filled in myself assessing my own situation

The 11 questions in the Chalder Fatigue Questionnaire, as used in the PACE trial.

Bimodal scoring allocates 0, 0, 1 and 1 and Likert scoring allocates 0, 1, 2, and 3 to each of these answers [20].

improvements in comorbid psychiatric disorders, present in 47% of trial participants [11], emphasizing the issues involved in failing to exclude people with psychiatric disorders. The National Institute of Health therefore concluded that “The Oxford Criteria...are flawed and include people with other conditions, confounding the ability to interpret the science” [41] and that “continuing to use the Oxford definition may impair progress and cause harm” and “we recommend that the Oxford definition be retired” [41,42].

### Objective outcomes

The authors concluded that “Both self-report and objective measures were used, and both were found to mediate treatment effects, lending credence to the results” [43] and “Our conclusions are supported by secondary outcomes, as both CBT and GET provided greater improvements than did APT and SMC” [11]. In other words, the authors concluded that both subjective and objective measures showed that CBT and GET were effective treatments. However, following successful treatment, all 4 treatment groups’ mean scores for the 2 self-report primary outcomes remained below the trial’s entry criteria [11,44], which meant that people remained sufficiently ill to re-enter the trial. And lost employment and Welfare benefit claims, 2 of the objective outcomes [45], showed that at 52-week follow up, the proportion of patients in receipt of income benefits had increased from 10% to 13% and from 14% to 20% in the CBT and GET group respectively [45]. The proportion of participants in receipt of illness/disability benefits increased from 32% to 38% and from 31% to 36% in the CBT and GET group respectively [45] and there was a 100%

increase in the proportion of participants in receipt of income protection or private pensions in both these groups [45] at 52 week follow up.

However, the authors also concluded that 22% of participants were considered to have recovered in both treatments groups [5]. Treatment concluded at 36 weeks, and follow up was performed at 52 weeks; therefore, the participants, particularly the 22% who were considered recovered, had 16 weeks to find employment so that the number of participants in receipt of benefits should have decreased dramatically, but instead it went up.

### The self-paced step test

The “submaximal self-paced step test”, “strongly and reliably predicts the maximal aerobic capacity VO<sub>2</sub>max, is sensitive to change” [46], and it was used as an objective measure of fitness in the PACE trial. The results of the self-paced step test [11] show the levels of fitness in response to the 4 treatments and it highlights the finding that CBT and GET produced the least and the pacing group the greatest improvement even though patients in the GET group had been exercising “five times a week”, in the form of walking, or cycling or swimming [43]. The overall conclusion was “that fitness... did not appear to mediate treatment effects” [43]. In other words, fitness did not improve, and none of the treatments were effective.

### The Six-Minute Walk Test (6MWT)

The 6MWT provides a good representation of the patient’s ability to perform submaximal activities of daily living and may be useful in serial evaluation of patient status and/or response to therapeutic interventions

[47]. The results of the 6MWT showed improvements with all 4 treatments. The GET group improved most and walked 379 m within 6 min, with an adjusted difference of 35 m beyond that of the SMC group after 52 weeks. The CBT group walked 354 m, which was an improvement of 4 m beyond that of the APT group but 1.5 m less than that of the SMC group [11]. None of these improvements reached statistical significance, for which an increase of at least 86 m was required [48].

The average age of PACE trial participants was 38 years [11], and normal distances for healthy men and women aged 38 years are 659 and 600 m, respectively [49]. Patients with severe MS manage 389 m [50]; those with Class 3 heart failure (with Class 4 considered the worst), who experience difficulty in completing everyday activities and have a 1-year mortality rate of 10–15% [51], manage 402 m [52]; and patients with end-stage lung or heart disease walk 335 m or less [53]. Yet ME/CFS patients who had completed a full course of “successful” CBT provided by qualified and experienced therapists only managed 19 m more, and those for whom treatment was most effective (i.e., GET) only managed 379 m and were outperformed by community dwelling 80–89 year olds with chronic health problems, who walked 417 m [54]. It is worth noting that people in the GET group engaged in “30 min of physical exercise five times a week,” and this was “most commonly walking” [43], so that walking only one fifth of it during the 6MWT should be easy and they should easily reach the age related normal levels at the 52-week follow-up, in the absence of an underlying physical problem, which is the assumption of the biopsychosocial model.

Patients who are able to walk 400 m or less, are placed on the waiting list for a lung transplant [55], and within a year of transplant, their 6MWT results return to normal [56]. However, following “effective” treatment, ME/CFS patients would remain on the transplant list.

Yet according to the secondary mediation analysis, “Exercise tolerance as measured by the number of metres walked in a fixed time was a strong mediator of GET alone” [43]. However in the supplementary material for the secondary mediation analysis, the significance of the 6MWT was deemphasized [57], because the researchers noted an “increase in exercise tolerance (walking distance) without an increase in exercise capacity (fitness)” [43] which means that the patients had optimized their walking distance without increasing their fitness, despite engaging in “physical exercise five times a week” [43], because the therapies used did not address the underlying metabolic problem preventing improvements in fitness. However, rather than providing an explanation to this effect, the following reasons (in full with cited reference numbers) were provided for failing to use 6MWT outcomes.

“Six-Minute Walk Test Participants were asked to walk as far as possible in six minutes, and the distance walked in metres was recorded. This is a measure of exercise tolerance (14). Due to concerns about patients with CFS coping with physical exertion, no encouragement was given to participants as they performed the test, in contrast to the way this test is usually applied (15,16). Rather than provide encouragement, we told participants, “You should walk continuously if possible, but can slow down or stop if you need to”.

Furthermore we had as little as 10 metres of walking corridor space available in centres rather than the 30 to 50 metres of space used in other studies (15-17); this meant that participants had to stop and turn around more frequently. Due to the modifications and the associated measurement error we considered this test as an internally referenced measure of behaviour change or exercise tolerance, not a measure of physical fitness” [57]. It is reasonable to expect, in a trial costing 8 million dollars [6], that the 3 principal investigators, who are mental health, and not 6MWT professionals, should have obtained guidelines for performing the 6MWT (one of the trial’s secondary outcomes).

Reference number 15 in the above statement refers to the American Thoracic Society (ATS) 6MWT guidelines published in March 2002 [58], the PACE trial began 3 years later in March 2005 [11] and the full text is available free of charge on the ATS website [58].

There are a number of inaccuracies in the PACE trial statement above, as the following quotes from the ATS guidelines show. “The self-paced 6MWT assesses the submaximal level of functional capacity. Most patients do not achieve maximal exercise capacity during the 6MWT; instead, they choose their own intensity of exercise and are allowed to stop and rest during the test” [58]. This indicates that the 6MWT is a measure of exercise tolerance rather than physical fitness and patients can stop and rest whenever they need to. The guidelines stipulate that “the walking course must be 30 m in length”. However, “some have used 20- or 50-m corridors” and there is “no significant effect of the length of straight courses ranging from 50 to 164 ft” (i.e., 15–50 meters) [58].

Standardized “encouragement” during the walk includes “Keep up the good work. You have only 2 minutes left” and “You are doing well. You have only 1 minute to go”. However, the guidelines also stipulate “Do not use other words of encouragement (or body language to speed up)” [58]. Therefore, it is standard practice to inform patients without encouraging them to walk further than they can.

The PACE trial was based on the biopsychosocial model, whereby there is nothing physically wrong in ME; however, this model is contradicted in the above statement, as the researchers expressed “concerns about patients with CFS coping with physical exertion” [57], and this refers to walking for only 6 minutes, which constitutes an indirect acknowledgement of the invalidity of the biopsychosocial model.

## The Actometer

The Actometer, an objective and reliable measure of activity [59], was meant to be used, at the beginning and the end of the trial, to assess improvement objectively; the last was deemed “too great a burden” [60] for patients, even though it weighs only 26 g [59], patients had consented to use it, had completed “moderately effective” treatment [11], and 22% of those in the CBT and GET groups had recovered [5]; therefore, it should have been easier and less of a burden.

## Are CBT and GET Safe?

The PACE trial protocol noted, “There is a discrepancy between patient organisation reports of the safety of CBT and GET and the published evidence of minimal risk from RCTs” [4]. The PACE trial concluded that “all four treatments tested are safe” [11], which is not surprising given that the protocol stated, “A risk assessment has been undertaken, and the therapies are of low risk to participants” [4]. The CBT and GET treatment manuals informed patients that these treatments were safe and effective [61,62], even though the trial was established to determine if they were or not, and used patient-rated subjective primary outcomes. This raises the question as to how safety was defined and assessed.

According to the PACE trial, “safety was assessed primarily by recording all serious adverse events, including serious adverse reactions to trial treatments” [11] and “Adverse events were considered serious when they involved death, hospital admission, increased severe and persistent disability, self-harm, were life-threatening, or required an intervention to prevent one of these” [11]. However, in patient surveys most of these issues were not raised. Instead, they reported that CBT and GET could exacerbate ME/CFS symptoms and cause relapses [63].

According to the PACE trial protocol, a high proportion of participants’ symptoms worsened with these treatments because of “rigidly applied programmes that are not tailored to the patient’s disability” [4] and “PACE treatment manuals minimize this risk by being based on mutually agreed

and flexible programmes that vary according to the patient's response" [4]. If this were the case, GET would not have left me bedridden. However, the basis of GET is that patients should ignore their symptoms (because they are signs of deconditioning rather than illness) and adhere to an established program designed by a qualified therapist, which increases exercise incrementally without considering the patient's symptoms during the program. As soon as therapists tailor exercise programs to patients' symptoms GET is not GET anymore but has become a form of pacing.

In the PACE trial "Non-serious adverse events were common" [64]. Actually they were very common, as 93% of patients reported non-serious adverse events; 49% and 51% reported 4 and more than 4 such adverse events, respectively, and there were no differences between the 4 groups [64]. A simple way to explain ME is to state that the M stands for musculoskeletal and the E for neurological problems. Therefore, in the list of non-serious adverse events, only 3 groups were considered (i.e., those involving musculoskeletal, neurological, and ME-related non-serious adverse events).

In the PACE trial, 46% of participants reported an increase in ME/CFS symptoms; 31% and 19% reported musculoskeletal and neurological non-serious adverse events, respectively [64]. It's unclear how large the overlap was between these three groups and therefore the proportion affected by CBT and GET would have been between 46% and 96%. As noted by Kindlon [65], who pooled a number of surveys, 20% and 51% of patients reported that CBT and GET, respectively, exacerbated their symptoms; Kindlon [65] also found that this was 82% in patients with severe ME, and many were not severely affected or bedridden prior to GET. The real proportion is likely to be approximately 74%, as concluded in a recent ME Association survey involving more than 1,400 patients [63].

These figures differ entirely from the 2% and 1% of participants reporting serious adverse events in the CBT and other 3 groups, respectively, which the PACE trial used to declare all 4 treatments to be safe [11].

It is particularly interesting to note that these numbers are also very high in the adaptive pacing group, in which patients were told to remain within their limits, and the SMC group, which did not involve any exercise, so that very low rates of adverse events would be expected in both groups. The very high proportions of adverse events in these 2 groups suggest that many individuals experienced difficulty completing the step test and 6MWT as expected in patients with ME/CFS, who experience abnormal responses to exercise, as highlighted by 2-day exercise testing [66,67], abnormal gene expression and immunity following exercise [68,69], and a left shift in anaerobic threshold observed in a number of studies [70] which was recently explored by Vink [71], in a former Dutch national field hockey champion who is now bedridden with severe ME. Vink [71] concluded that an impaired oxidative phosphorylation is the reason why he can only perform trivial activities, and an impaired lactic acid excretion plays an important role in the abnormally delayed muscle recovery.

The Supplement to the secondary mediation analysis shows that 90% of patients in the APT group completed the primary outcome questionnaires at follow up, but only 71% performed the 6MWT, and only 62% completed the step test [57]; however, the trial reported that only "33 (5%) of 640 participants were lost to follow-up, but rates did not differ between groups" [11] and "by 52 weeks, only 33 (5%) were missing primary outcome data, with no significant difference between treatment groups" [5]. This suggests that 29% and 38% of the participants, respectively, withdrew from the study because the first step test and 6MWT exacerbated their symptoms, and as 93% reported adverse events, it suggests that these 29% and 38% constitute the tip of the iceberg.

## Recovery

The recovery article states that "We changed our original protocol's threshold score for being within a normal range on this measure from

a score of  $\geq 85$  to a lower score as that threshold would mean that approximately half the general working age population would fall outside the normal range. The mean (S.D.) scores for a demographically representative English adult population were 86.3 (22.5) for males and 81.8 (25.7) for females. We derived a mean (S.D.) score of 84 (24) for the whole sample, giving a normal range of 60 or above for physical function" [5].

Having recovered differs from being within the normal range, and the PACE trial did not examine the "demographically representative English adult population." Furthermore Bowling et al. [72] did not involve the "general working age population," as the participants ages ranged from 16 to 85, with 28.6% aged 65 years or older. The PACE trial cohort's mean age was 38 years [11] and therefore participants should be compared to healthy adults of the same age. According to the survey by Bowling et al. [72], which the PACE trial used, the mean physical functioning score for adults aged 35–44 years was 93.3, with a standard deviation of 13.4 [72] providing a normal score of  $\geq 79.9$ , rather than  $\geq 60$ , for physical functioning.

However, SF-36 physical functioning scores for healthy 38 year olds are not normally distributed but skewed to the right, with nearly everyone in the maximum range; and according to the BMJ's statistical resources for readers, standard deviations will then be grossly inflated, are not a good measure of variability anymore and are therefore inappropriate for use [73], as highlighted by the fact that, Bowling et al.'s [72] survey of 2,000 patients, 1,200 and 400 had median scores of 100 and 90, respectively. The SF-36 physical functioning score decreases with age and ill health, almost 600 patients (28.6%) were aged 65 years or older, and 22% and 16% had chronic and acute health problems, respectively; therefore, the SF-36 physical functioning recovery score should have been 100 rather than 60 or more.

According to the PACE recovery article, "Before we can determine the proportions recovered we need an operational definition of recovery itself. An ideal definition remains uncertain" [5]. However, this is inaccurate, as concluded by Kennedy [74], recovery "is the elimination of...symptoms and a return to premorbid levels of functioning".

The recovery article also stated, "it is important to note that recovery does not mean being free of all symptoms" [5]. However, if the symptoms of a disease remain, then it is an overly optimistic definition of recovery, including ill people. The following statement acknowledges this: "The prevalence of the case-level international (CDC) definition of CFS may have been inaccurate because we only examined for accompanying symptoms in the previous week, not the previous 6 months" [5], as most symptoms that have been present for less than a week are self-limiting and symptoms that have been there for 6 months or more are not. "We therefore use the term 'recovery' in this paper to mean recovery from the current episode of the illness" [5]. According to this, instead of being able to walk 5 or 6 yards to the toilet twice daily, I would be able to do so 3 times per day and otherwise remain bedridden highlighting the fact that recovery from the current episode of illness does not necessarily constitute recovery.

The protocol stated: "'Recovery' will be defined by meeting all four of the following criteria: (i) a Chalder Fatigue Questionnaire score of 3 or less, (ii) SF-36 physical Function score of 85 or above, (iii) a CGI score of 1, and (iv) the participant no longer meets Oxford Criteria for CFS, CDC criteria for CFS or the London criteria for ME" [4]. Yet even though the authors chose these themselves, during the trial they "changed three of the thresholds for measuring recovery from our original protocol" [5], reflected in the following 6 statements.

## The Chalder Fatigue Questionnaire (CFQ)

"We changed our original protocol's threshold score for being within a

normal range from a binary score of  $\leq 3$  out of 11” to “a population mean (S.D.) Likert score of 14.2 (4.6) out of a maximum score of 33” [5]. This indicates that both the score and the scoring system were changed.

### The SF-36 physical function subscale:

“We changed our original protocol’s threshold score for being within a normal range on this measure from a score of  $\geq 85$  to “60 or above” [5].

### Oxford criteria

“To satisfy the third criterion for severity of fatigue and disability, participants had to meet trial entry thresholds for fatigue (a binary score of  $\geq 6$  out of 11 on the CFQ) and abnormal levels of physical function (a score of  $\leq 65$  out of 100 on the SF-36 physical function subscale)” [5]. However, the original score required for the SF-36 was “a score of 60 of 100 or less” [11]. And the bimodal Chalder Fatigue Questionnaire score of  $\geq 6$  out of 11 was changed to a Likert scoring of 18 or more, therefore, the Oxford Criteria were also changed during the trial.

### The CDC CFS case definition

“For the purposes of this study, the four or more symptoms needed to be present within the previous week of the assessment date, rather than the previous 6 months” [5]. Most symptoms that are present for less than a week are acute and self-limiting; however, symptoms that have been present for 6 months are chronic and not self-limiting.

### The London criteria

“Specifically, these criteria included... no ‘primary’ depressive illness and no anxiety disorder present (which we interpreted as no co-morbid mood disorder of any kind)” [5]. However, these criteria also state that “the existence of a parallel disease would be grounds for disqualification” [35], which was omitted from the version used in the trial, without the knowledge of the original authors according to Goudsmit and Shepherd [36,37], 2 of the original authors. All psychiatric diseases should therefore have been excluded which did not occur.

### The self-rated CGI change score

“We considered scores of 1 (‘very much better’) or 2 (‘much better’) as evidence of the process of recovery, rather than our original protocol threshold of a score of 1 only, because we considered that participants rating their overall health as ‘much better’ represented the process of recovery” [5]. However, everyone who’s had a flu-like illness knows that, while one feels much better during convalescence, relative to the day upon which one fell ill, one is in the process of recovery but has not recovered yet which highlights the fact that the self-rated CGI change score is a global measure of change and clinical progress [75], rather than a measure of recovery, and is also “unreliable and too general to measure... treatment responses validly” [76].

Table 2 highlights the changes to the recovery criteria, made by the authors during an unblinded trial, rather than by an independent external trial committee with no access to the data to avoid any form of bias, as stipulated by Evans [22].

The recovery article also stated that “Although it seemed that slightly smaller proportions had recovered from the illness as a whole, when the criterion ‘not meeting the London criteria for ME’ was applied, we found that the differences were due to missing data rather than to change in recovery status” [5]. But how did they know this if the data were missing?

One of the conclusions of the recovery article was “that CBT and GET were both significantly more likely than APT and SMC to be associated with recovery at 52 weeks, even when using a conservative definition of recovery” [5]. A conservative definition of recovery ensures that the proportion of patients who have recovered is deliberately lower than the

actual proportion. However, the endpoint changes during the trial caused an overlap in entry and recovery criteria, whereby a score of  $\geq 60$  of 100 on the SF-36 physical functioning subscale represented recovery, yet the entry criteria required a score of  $\leq 65$ . Therefore, even if participants’ health deteriorated and their scores dropped from 65 to 60, they were classed as having recovered, even though a score of 60 out of 100 is normal for 75-84 year olds [72] but represents disability in 38 year olds. As found by Tuller [6], the consequence of this was that 13% of patients were already classed as recovered on 1 or 2 of the primary outcomes as soon as they entered the trial. Is that a conservative definition of recovery?

As mentioned previously and shown in Table 1, my bimodal Chalder Fatigue Score indicated that even though I am bedridden and dependent on others, I would not have been eligible for the trial. During the trial, the scoring system was changed, and without changing my answers, I suddenly scored the minimum (of 18) points required to enter the trial. Yet with the same Likert score of 18, and without having received any treatment or any change to my medical situation, I was also classed as recovered, despite remaining bedridden. Is that a conservative definition of recovery?

The authors also “changed some of the thresholds for measuring recovery from those of the original protocols; we made the changes before analysis and to more accurately reflect recovery” [5]. As shown in Table 2, extensive changes were made to the recovery criteria during the trial, which broadened the definition of recovery, rendering it less accurate and less conservative. “Finally”, the authors said, “we cannot be sure that recovery was sustained beyond the assessment at 52 weeks” [5]. As the definition of recovery included people who remained ill or disabled, they could not have sustained recovery beyond the 52-week assessment, because they had not recovered in the first place. The authors also noted that “Two studies of recovery in adults after CBT found similar proportions in recovery: 23% and 24%”; “but the definition for normal range used was the more liberal population mean  $\pm 2$  s.d. rather than the more conservative 1 s.d. that we used” [5]. With respect to SF-36 physical functioning scores, the standard deviation quoted in the PACE trial was 24. The minimum level of recovery defined in this manner was a score of 60 out of 100 in the PACE trial. The establishment of a second standard deviation would have led to a minimum score of 36 out of 100 to represent recovery which in a 38 year old indicates severe disability rather than recovery, putting “recovery” in CBT studies for ME/CFS into perspective. “Our finding that 22–56% of participants met various composite or single criteria for recovery or improvement a year after starting either CBT or GET is therefore consistent with previously published studies” [5]. Yet even though I am bedridden and dependent on others, I fulfill the single criterion for recovery using the Chalder Fatigue Questionnaire, the same score also classified me as sufficiently ill to enter the trial; however, if a patient fulfills only one of several criteria for recovery, the patient has not recovered, regardless of the definition of recovery.

The authors also concluded that “The proportions recovered in each treatment arm were similar in the subgroups meeting alternative definitions of CFS and ME, implying that these findings generalize to different definitions of CFS and ME” and “patients who have either CFS or ME...should therefore be offered either CBT or GET to provide the best chance of recovery with these treatments” [5]. Yet their definition of recovery labeled patients as recovered, regardless of whether they remained disabled or not.

### Fear Avoidance Beliefs

In January 2015, the secondary mediation analysis was published, and the authors stated that “Fear avoidance beliefs are characterised by fears that activity or exercise will make symptoms worse” [43] and concluded that “Our main finding was that fear avoidance beliefs were the strongest

	Changes made during the trial?	What changes were made?	Changes made by the authors themselves?
Chalder Fatigue Questionnaire	Yes	1. Binary score was changed to a Likert score	Yes
	Yes	2. Binary score of 3 out of 11 was changed to a Likert score of 18 or less out of a maximum of 33	Yes
SF-36 physical function subscale	Yes	"changed from a score of $\geq 85$ to "60 or above" [5].	Yes
Self-rated CGI change score	Yes	changed from "a score of 1 only" to "scores of 1 ('very much better') or 2 ('much better')" [5]	Yes
Oxford criteria for CFS	Yes	The SF-36 physical function score was changed from a score of 60 or less to 65 or less;	Yes
	Yes	The bimodal Fatigue score of $\geq 6$ out of 11 was changed to a Likert scoring of 18 or more	Yes
International CFS case definition	Yes	"the four or more symptoms needed to be present within the previous week of the assessment date, rather than the previous 6 months" [5]	Yes
ME case definition: the London criteria	No, changes were made before the trial	"because it is vital that the M.E. study groups we use in research are as 'pure' as possible, the existence of a parallel disease would be grounds for disqualification" [35], which was omitted from the version used in the trial; and the London criteria were changed without the knowledge of the original authors according to Goudsmit [36] and Shepherd [37], 2 of the original authors	Yes

**Table 2:** Changes made to the recovery criteria

Source: PACE trial recovery article [5]

mediator for both CBT and GET. Changes in both beliefs and behaviour mediated the effects of both CBT and GET, but more so for GET. The results support a treatment model in which both beliefs and behaviour play a part in perpetuating fatigue and disability in chronic fatigue syndrome" [43]. Yet in a rapid response in the BMJ two weeks later, the authors stated, "nor did we say that fear of exercise in CFS was "irrational"" and ME/CFS is "an illness where exercise increases symptoms" [77], thereby acknowledging that avoiding exercise is the appropriate course of action. And in 2005 when the PACE trial began, one of the principal investigators coauthored an article in which the conclusion was that "CFS patients without a comorbid psychiatric disorder do not have an exercise phobia" [78]. The same principal investigator also coauthored an article concluding that exercise causes immunological damage/abnormalities in ME/CFS [79]. Why a blood test to check these immunological parameters was not part of the "definitive randomised trial" [80] conducted to determine whether CBT and GET were effective and safe in ME/CFS is unclear.

The authors also stated, "the FINE trial found that fear avoidance, embarrassment avoidance, all-or-nothing and avoidance behaviour were cross-sectional mediators of the treatment effect" [43]. Yet the FINE trial itself reported that the treatment effect, in more severely affected patients, "is small and not statistically significant at one year follow-up" [81], which was the primary outcome of the trial. Therefore, these factors could not have been "mediators of the treatment effect," as there was no treatment effect. Yet the PACE trial authors concluded "This mediational analysis strengthens the validity of our theoretical model of CBT and supports the idea that a similar model is valid for GET by confirming the role of fearful beliefs and avoidance behaviour" [43], ignoring the study's evidence to the contrary. And when the authors concluded that "The increase in exercise tolerance (walking distance) without an increase in exercise capacity (fitness) might have been facilitated by the mediating effect of reduced fear avoidance beliefs" [43], they should have concluded that these patients had simply optimized walking distance without increasing fitness, despite engaging in "physical exercise five times a week" [43], because the therapies used did not address the underlying physical problem in ME/CFS.

## Removing the Naturally occurring Fluctuation of ME/CFS

"The main finding of this long-term follow-up study of the PACE trial participants is that the beneficial effects of the rehabilitative CBT and GET therapies on fatigue and physical functioning observed at the final 1 year outcome of the trial were maintained at long-term follow-up 2.5 years from randomisation" [44]. However, this was not the main finding which was that there was no difference in efficacy between treatments and none of them were effective [44].

A review by Whiting et al. [82] showed that "The relapsing nature of CFS suggests that follow-up should continue for at least an additional 6 to 12 months after the intervention period has ended, to confirm that any improvement observed was due to the intervention itself and not just to a naturally occurring fluctuation in the course of the illness". The first follow up was performed at 52 weeks, which was only 16 weeks after treatment completion at 36 weeks, and long-term follow up was performed at least 2 years subsequent to randomization [44], which was at least 1.25 years subsequent to treatment completion; therefore, the naturally occurring fluctuation of the illness was omitted from consideration.

The authors stated that "In interpreting the follow-up data it is important to note that many of the participants had received additional treatment for chronic fatigue syndrome since completing the trial" [44] and "roughly a quarter and a third of the participants originally allocated to APT and SMC respectively had received a therapeutically adequate amount (ten or more sessions) of CBT or GET after the trial final trial outcome" [44] which made it "possible that this additional treatment was important in improving the long-term outcome for these patients" [44]. However, the Supplementary appendix long-term follow-up, which "formed part of the original submission", was "peer reviewed" and "supplied by the authors" [83], shows that the majority of participants, i.e 76% and 83%, did not have any additional CBT respectively GET, after the trial had finished.

Furthermore, researchers cannot control the environment subsequent to completion of a trial; therefore, an effect cannot be attributed to the receipt of any form of additional post trial treatment. Another problem is the carryover effect, whereby the effect of treatment is carried over to



the second phase [84], which in this case was the phase following trial completion. As noted by Larun et al. [84] this is a greater problem “when the condition of interest is unstable,” and “both effects are very likely in CFS/ME”. Last but not least, the Supplementary appendix long-term follow-up [83] shows that in all 4 groups, patients 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. This suggests that additional CBT and GET, provided by qualified and experienced therapists from the trial subsequent to trial completion, could have been detrimental to patients’ health. And on top of that, in the CBT and SMC groups an adequate number of sessions of CBT and GET were more detrimental than an inadequate number of sessions [83].

## Discussion

The PACE trial was a large multicenter trial established as a decisive means of testing the safety and efficacy of CBT and GET in ME/CFS, based on the conclusions of the Medical Research Council’s Research Advisory Group and the CMO’s working group, that ME and CFS constitute the same illness [4], and the biopsychosocial model “which supposes that unhelpful interpretations of symptoms, fearful beliefs about engaging in activity, and excessive focus on symptoms are central in driving disability and symptom severity” [43] and that “both de-conditioning (loss of muscle strength and reduced exercise capacity) and avoidance of activity... maintain fatigue and disability” [43]. CBT and GET were designed “to identify and challenge unhelpful cognitions” [43] and change and reverse such cognitions by “gradually increasing physical activity to improve fitness and get the body used to activity again” [5], and by doing so cure ME/CFS. Yet the 2 subjective primary outcomes of the trial relied on patients’ interpretations of their symptoms, but these 2 approaches are at odds. If a model and treatment are based on patients’ incorrect interpretations of their symptoms, then primary outcomes should not rely on these. In addition to being unreliable, subjective data lack objectivity and are prone to outside influences, particularly in trials that aim to modify participants’ subjective beliefs [85]. There is a low correlation between objective and subjective activity measurements [59] not only in chronically ill but also in healthy people [86]. And the PACE trial authors themselves noted that “objective measures of physical activity have been found previously to correlate poorly with self-reported outcomes” in ME/CFS [5]. In other words, in a decisive trial, the primary outcomes should have been objective rather than subjective.

During the trial, extensive endpoint changes were made by the authors themselves as shown by the following examples: “we made the changes before analysis and to more accurately reflect recovery” [5] and “We changed our original protocol’s threshold score for being within a normal range from”... [5]. Yet during a trial, such alterations should only be made for compelling reasons and only by an independent trial steering committee without access to the data, rather than by the trial investigators, as stipulated by Evans [22]. According to Goldacre: “Switching your outcomes breaks the assumptions in your statistical tests. It allows the “noise” or “random error” in your data to exaggerate your results (or even yield an outright false positive, showing a treatment to be superior when in reality it’s not) leading to the wrong conclusions and “in medicine, that’s not a matter of academic sophistry - it causes avoidable suffering” [87]. And according to Ioannidis [88], “Flexibility increases the potential for transforming... “negative” results into “positive” results”, and the greater the flexibility of the outcomes, the less likely it is that the research findings are accurate [88]. This is highlighted by the fact that 13% of participants were considered to have recovered, with respect to 1 or 2 of the recovery criteria [6], as soon as they entered the trial, as the above-mentioned endpoint changes created an overlap between entry and recovery criteria.

At 52-week follow up the objective outcomes showed the following: the step test, a reliable objective measure of fitness [46], showed no significant improvements in any of the 4 treatment groups; the 6MWT results showed that ME/CFS patients would remain on the waiting list for a lung transplant [55] following treatment deemed effective by the PACE trial; employment rates did not differ significantly between the 4 treatment groups, the number of patients claiming state sick pay and disability benefits increased following CBT and GET [45], and the number of patients in receipt of income protection or private pensions had actually doubled in the CBT and GET groups [45] and the authors deemed it too much of a burden [60] for patients to wear a 26 g weighing Actometer [59] at the end of the trial even though the PACE trial concluded that CBT and GET were moderately effective, and that 22% of patients in the CBT and GET groups recovered following successful treatment [5]; therefore, it should have been less of a burden to wear the Actometer at the end of the trial than at the beginning and these figures should have decreased dramatically and not gone up.

With respect to the definition of recovery, the PACE trial authors stated, “it is important to note that recovery does not mean being free of all symptoms” [5]. In other words, “recovery” was synonymous with “no recovery”. The long-term follow-up study showed that patients in all 4 groups still fulfilled the entry criteria for fatigue and physical functioning, the 2 primary outcomes, following effective treatment [44]. So that, similar to the above-mentioned objective outcomes, both primary outcomes showed that patients remained sufficiently ill to re-enter the trial. At long-term follow up, no differences were found between the 4 treatments, which is known as a null effect, and none of the treatments were effective [44]. And as noted by Ioannidis “investigators working in any field are likely to resist accepting that the whole field in which they have spent their careers is a “null field”” [88]; which might explain the authors’ reluctance to release trial data.

Further, 46% of patients reported increases in their ME/CFS symptoms, 31% reported musculoskeletal, representing the M in ME, and 19% reported neurological adverse events, representing the E in ME [64]. So that the proportion of participants negatively affected by CBT and GET is between 46% and 96%, and most likely estimated at 74%, as recently reported by a large survey conducted by the British ME Association, which involved 1,428 patients [63] and medication with such high rates of adverse events should be withdrawn from the market with immediate effect. This is in stark contrast to the PACE trial authors’ response to a National Institutes of Health treatment review of ME/CFS [89], in which the PACE trial authors stated, “It is important not to overemphasise the harms associated with an effective treatment when there are so few others available” [90]. Yet we should not ignore or underplay the harm associated with CBT and GET for people with ME/CFS, and we should also not ignore that the PACE trials long-term follow up showed that none of the treatments were effective for ME/CFS [44]. This is in total contrast to the promising results of the Norwegian Rituximab trials, which suggest that ME/CFS is an autoimmune disease and showed that symptoms were alleviated in 66% of patients, two thirds of whom remained in complete remission at 36-month follow up [91].

The PACE trial’s null effect, confirms the experiences of patients, which was also highlighted by Falk Hvidberg et al. [92] who concluded that of patients with 21 diseases, including chronic renal failure, cancer, and stroke, those with ME/CFS demonstrated the lowest quality of life confirming the findings of the 1996 health status report [93], illustrating the fact that nothing has changed within the last 20 years with respect to ME/CFS patients’ health, providing further proof that CBT and GET, which most ME patients have tried because they desperately want to recover, are not effective.

The discovery that an increase in exercise tolerance did not lead to an increase in fitness [43], as observed in healthy people, shows that there is an underlying physical problem preventing this, which means that none of the 4 treatments in the trial addressed this issue and validates that ME/CFS is a physical disease. Therefore, the PACE trial, along with the FINE Trial, which found that these treatments were ineffective in more severely affected patients [81], disproved the biopsychosocial model as an explanation of ME/CFS, confirming the conclusion of the American Institute of Medicine, following an extensive review of the literature, that ME/CFS is a serious and debilitating physical, rather than a psychological or psychiatric, disease [1].

## Conclusion

The PACE trial was the largest trial of its kind. During the trial extensive endpoint changes were made, creating an overlap in entry and recovery criteria so that 13% of patients were already recovered on 1 or 2 of the primary outcomes upon entering the trial [6]; the PACE trial's definition of recovery labeled patients as recovered, regardless of whether they remained disabled or not. Its null effect invalidates the use of CBT and GET in ME/CFS and invalidates the biopsychosocial model of deconditioning and fear avoidance.

The PACE trial found that the proportion negatively affected by CBT and GET was between 46% and 96%, most likely estimated at 74%, as shown in a large survey recently conducted by the ME Association [63]. Medication with such high rates of adverse events would be withdrawn immediately. The PACE trial's discovery of an "increase in exercise tolerance (walking distance) without an increase in exercise capacity (fitness)" [43] means there's an underlying physical problem preventing this and that none of the treatments were addressing this issue but also that ME/CFS is a physical disease, confirming the conclusion of the IOM [1]. And therefore, from now on our focus should be on biomedical (instead of psychosocial) research to find effective treatment for this debilitating disease.

## Conflict of interest

No competing interests.

## Funding

No funding was involved in supporting this work.

## Acknowledgements

I would like to thank my parents for typing out my speech memos and for proof reading the manuscript and Elsevier for the Language Editing of the manuscript.

## References

- Committee on the Diagnostic Criteria for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome, Board on the Health of Select Populations, Institute of Medicine (2015) *Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness*, Source Washington (DC): National Academies Press (US).
- Carruthers BM, van de Sande MI, De Meirleir KL, Klimas NG, Broderick G, et al. (2011) Myalgic encephalomyelitis: International Consensus Criteria. *J Intern Med* 270: 327-338.
- WHO (2010) ICD-10-CM Diagnosis Codes-Diseases of the nervous system G00-G99, Other disorders of the nervous system G89-G99, Other disorders of brain G93-94, Diagnosis Code G93.3.
- White PD, Sharpe M, Chalder T, DeCesare JC, Walwyn R (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 medical care alone for patients with the chronic fatigue syndrome/myalgic encephalomyelitis or encephalopathy. *BMC Neurol* 7: 6.
- White PD, Goldsmith K, Johnson AL, Chalder T, Sharpe M (2013) Recovery from chronic fatigue syndrome after treatments given in the PACE trial. *Psychol Med* 43: 2227-2235.
- Tuller D (2015) TRIAL BY ERROR: The Troubling Case of the PACE Chronic Fatigue Syndrome Study.
- Tuller D (2015) TRIAL BY ERROR: The Troubling Case of the PACE Chronic Fatigue Syndrome Study (second installment).
- Tuller D (2015) TRIAL BY ERROR: The Troubling Case of the PACE Chronic Fatigue Syndrome Study (final installment).
- Coyne J (2015) Uninterpretable: Fatal flaws in PACE Chronic Fatigue Syndrome follow-up study.
- Laws K (2015) PACE - Thoughts about Holes.
- White PD, Goldsmith KA, Johnson AL, Potts L, Walwyn R, 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. *Lancet* 377: 823-836.
- Kielmann T, Huby G, Powell A, Sheikh A, Price D, et al. (2011) From awareness to involvement? A qualitative study of respiratory patients' awareness of health service change. *Health Expect* 14: 321-333.
- Thornton S (2014) Beyond rhetoric: we need a strategy for patient involvement in the health service. *BMJ* 348: g4072.
- Asadi-Lari M, Tamburini M, Gray D (2004) Patients' needs, satisfaction, and health related quality of life: Towards a comprehensive model Health and Quality of Life Outcomes 2: 32.
- About: What is the ISRCTN registry?
- ISRCTN registry, 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*.
- NHS Information Centre (2011) HRG4 Operational Information Standard - Post Implementation/Maintenance Review. Version No: 1.0.
- WHO (2004) ICD-11: ME CFS: Current codings in ICD-10 for Postviral fatigue syndrome; [Benign] myalgic encephalomyelitis and Chronic fatigue syndrome. L'Hours A, WHO HQ, Geneva.
- Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, et al. (2008) Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study *BMJ* 336: 601-605.
- PACE (2006) PACE trial protocol: Final version 5.0. ISRCTN54285094.
- White PD, Sharpe M, Chalder T, DeCesare JC, Walwyn R (2007) Pre-publication history 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.
- Evans S (2007) When and How Can Endpoints Be Changed after Initiation of a Randomized Clinical Trial? *PLoS Clin Trials* 2: e18.
- Participants' Newsletter Issue 1. June 2006.
- Participants' Newsletter Issue 2. March 2007.
- MRC policy on research data sharing, 2015.
- Queen Mary University of London (2015) Research Data Access and Management Policy.

27. Wicherts JM, Bakker M, Molenaar D (2011) Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results. *PLoS One* 6: e26828.
28. Ramsay MA (1988) *Myalgic Encephalomyelitis and Post Viral Fatigue States: the saga of Royal Free Disease*. 2nd edition, Gower Medical Publishing, London.
29. Dowsett EG (1992) Conversation piece. Interview by P. D. Welsby. *Postgrad Med J* 68: 63–65.
30. Paul L, Wood L, Behan WM, Maclaren WM (1999) Demonstration of delayed recovery from fatiguing exercise in chronic fatigue syndrome. *Eur J Neurol* 6: 63-69.
31. Black CD, McCully KK (2005) Time course of exercise induced alterations in daily activity in chronic fatigue syndrome. *Dyn Med* 4: 10.
32. Reeves WC, Lloyd A, Vernon SD, Klimas N, Jason LA, et al. (2003) Identification of ambiguities in the 1994 chronic fatigue research case definition and recommendations for resolution. *BMC Health Serv Res* 3: 25.
33. Sharpe MC, Archard LC, Banatvala JE, Borysiewicz LK, Clare AW, et al. A report—chronic fatigue syndrome: guidelines for research. *J R Soc Med* 84: 118-121.
34. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, et al. (1994) The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Ann Intern Med* 121: 953-959.
35. Dowsett EG, Goudsmit E, Macintyre A, Shepherd CB (1994) The London criteria In: Report from The National Task Force on Chronic Fatigue Syndrome (CFS), Post Viral Fatigue Syndrome (PVFS), Myalgic Encephalomyelitis (ME). Westcare 96-98.
36. Goudsmit E (2015) London Criteria for M.E. - for website discussion. Comment dated 31 March 2011.
37. Shepherd C (2015) London Criteria for M.E. - for website discussion, comment dated 27 February 2011.
38. David AS (1991) Post viral syndrome and psychiatry. *British Medical Bulletin* 47: 966-988.
39. Haywood KL, Staniszewska S, Chapman S (2011) Quality and acceptability of patient-reported outcome measures used in chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME): a systematic review. *Qual Life Res*.
40. Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, et al. (1993) Development of a fatigue scale. *J Psychosom Res* 37: 147-153.
41. Green CR, Cowan P, Elk R, O'Neil KM, Rasmussen AL (2014) DRAFT EXECUTIVE SUMMARY: NATIONAL INSTITUTES OF HEALTH-Pathways to Prevention Workshop: Advancing the Research on Myalgic Encephalomyelitis/ Chronic Fatigue Syndrome.
42. Green CR, Cowan P, Elk R, O'Neil KM, Rasmussen AL (2014) Final Report: NATIONAL INSTITUTES OF HEALTH-Pathways to Prevention Workshop: Advancing the Research on Myalgic Encephalomyelitis/ Chronic Fatigue Syndrome, EXECUTIVE SUMMARY.
43. Chalder T, Goldsmith KA, White PD, Sharpe M, Pickles AR (2015) Rehabilitative therapies for chronic fatigue syndrome: a secondary mediation analysis of the PACE trial. *Lancet Psychiatry* 2: 141–152.
44. Sharpe M, Goldsmith KA, Johnson AL, Chalder T, Walker J, et al. (2015) Rehabilitative treatments for chronic fatigue syndrome: long-term follow-up from the PACE trial. *Lancet Psychiatry* 2: 1067-1074.
45. McCrone P, Sharpe M, Chalder T, Knapp M, Johnson AL, 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.
46. Petrella RJ, Koval JJ, Cunningham DA, Paterson DH (2001) A self-paced step test to predict aerobic fitness in older adults in the primary care clinic. *J Am Geriatr Soc* 49: 632-638.
47. Pollentier B, Irons SL, Benedetto CM, Dibenedetto AM, Loton D, et al. (2010) Examination of the Six Minute Walk Test to Determine Functional Capacity in People with Chronic Heart Failure: A Systematic Review. *Cardiopulm Phys Ther J* 21: 13-21.
48. Wise RA, Brown CD (2005) Minimal clinically important differences in the six-minute walk test and the incremental shuttle walking test. *COPD* 2: 125-129.
49. Chetta A, Zanini A, Pisi G, Aiello M, Tzani P, et al. (2006) Reference values for the 6-min walk test in healthy subjects 20-50 years old. *Respir Med* 100: 1573-1578.
50. Goldman MD, Marrie RA, Cohen JA (2008) Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. *Multiple Sclerosis* 14: 383-390.
51. Reisfield GD, Wilson GR (2007) Prognostication In Heart Failure. *J Palliat Med* 10: 245-246.
52. Lipkin DP, Scriven AJ, Crake T, Poole-Wilson PA (1986) Six minute walking test for assessing exercise capacity in chronic heart failure. *Br Med J (Clin Res Ed)* 292: 653-655.
53. Faggiano P, D'Aloia A, Gualeni A, Lavatelli A, Giordano A (1997) Assessment of oxygen uptake during the 6-minute walking test in patients with heart failure: preliminary experience with a portable device. *Am Heart J* 134: 203-206.
54. Steffen TM, Hacker TA, Mollinger L (2002) Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. *Phys Ther* 82: 128-137.
55. Kadikar A, Maurer J, Kesten S (1997) The six-minute walk test: a guide to assessment for lung transplantation. *J Heart Lung Transplant* 16: 313-319.
56. Seoane L, Alex S, Pirtle C, Gupta M, Taylor DE, et al. (2010) Utility of the 6-minute walk test following lung transplantation. *Ochsner J* 10: 227-230.
57. Chalder T, Goldsmith KA, White PD, Sharpe M, Pickles AR (2015) Supplement to: Rehabilitative therapies for chronic fatigue syndrome: a secondary mediation analysis of the PACE trial. *Lancet Psychiatry* 2: 141-152.
58. ATS Statement (2002) Guidelines for the Six-Minute Walk Test. *Am J Respir Crit Care Med* 166: 111-117.
59. Scheeres K, Knoop H, Meer vd, Bleijenberg G (2009) Clinical assessment of the physical activity pattern of chronic fatigue syndrome patients: a validation of three methods. *Health Qual Life Outcomes* 7: 29.
60. White PD, Sharpe MC, Chalder T, DeCesare JC, Walwyn R, for the PACE trial management group, Response to comments on "Protocol for the PACE trial", 28 July 2008.
61. Burgess M, Chalder T (2004) PACE Manual for Participants Cognitive Behaviour Therapy for CFS/ME, MREC version 2 – November 2004.
62. Bavinton J, Dyer N, White PD (2004) On behalf of the PACE trial management group, Final Trial Version: (MREC Version 1), Graded Exercise Therapy, PACE Information for Participants, GET participants manual MREC Version 1; 16/11/04.
63. ME Association (2015) Our CBT, GET and Pacing Report calls for major changes to therapies offered for ME/CFS.
64. Dougall D, Johnson A, Goldsmith K, Sharpe M, Angus B, et al. (2014) Adverse events and deterioration reported by participants in the PACE trial of therapies for chronic fatigue syndrome. *J Psychosom Res* 77: 20-26.
65. Kindlon T (2011) Reporting of harms associated with graded exercise therapy and cognitive behavioral therapy in myalgic encephalomyelitis/ chronic fatigue syndrome.

66. Keller BA, Pryor JL, Giloteaux L (2014) Inability of myalgic encephalomyelitis/chronic fatigue syndrome patients to reproduce VO2peak indicates functional impairment. *J Transl Med* 12: 104.
67. Snell CR, Stevens SR, Davenport TE, Van Ness JM (2013) Discriminative validity of metabolic and workload measurements for identifying people with chronic fatigue syndrome. *Phys Ther* 93: 1484-1492.
68. Light AR, Bateman L, Jo D, Huguen RW, Vanhaisma TA, et al. (2012) Gene expression alterations at baseline and following moderate exercise in patients with Chronic Fatigue Syndrome and Fibromyalgia Syndrome. *J Intern Med* 271: 64-81.
69. Nijs J, Nees A, Paul L, DeKooning M, Ickmans K, et al. (2014) Altered immune response to exercise in patients with chronic fatigue syndrome/myalgic encephalomyelitis: a systematic literature review. *Exerc Immunol Rev* 20: 94-116.
70. Lane RJM, Woodrow D, Archard LC (1994) Lactate responses to exercise in chronic fatigue syndrome. *J Neurol Neurosurg Psychiatry* 57: 662-663.
71. Vink M (2015) The Aerobic Energy Production and the Lactic Acid Excretion are both Impeded in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *J Neurol Neurobiol* 1.
72. Bowling A, Bond M, Jenkinson C, Lamping DL (1999) Short Form 36 (SF-36) Health Survey Questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *J Public Health Med* 21: 255-270.
73. 2. Mean and standard deviation. *BMJ statistical resources for readers*.
74. Kennedy S (2002) Full remission: a return to normal functioning. *J Psychiatry Neurosci* 27: 233-234.
75. Busner J, Targum SD (2007) The Clinical Global Impressions Scale Applying a Research Tool in Clinical Practice. *Psychiatry (Edgmont)* 4: 28-37.
76. Forkmann T., Scherer A., Boecker M., Pawelzik M., Jostes R., et al. (2011) The clinical global impression scale and the influence of patient or staff perspective on outcome. *BMC Psychiatry* 11:83.
77. Chalder T, Goldsmith K, White P, Sharpe M, Pickles A, Mediators of response to behavioural treatments in CFS, 29 January 2015.
78. Gallagher AM, Coldrick AR, Hedge B, Weir WR, White PD (2005) Is the chronic fatigue syndrome an exercise phobia? A case control study. *J Psychosom Res* 58: 367-373.
79. White PD, Nye KE, Pinching AJ, Yap TM, Power N, et al. (2004) Immunological changes after both exercise and activity in chronic fatigue syndrome: A pilot study. *Journal of Chronic Fatigue Syndrome* 12: 51-66.
80. Walwyn R, Potts L, McCrone P, Johnson AL, DeCesare JC, et al. (2013) A randomised trial of adaptive pacing therapy, cognitive behaviour therapy, graded exercise, and specialist medical care for chronic fatigue syndrome (PACE): statistical analysis plan. *Trials* 14: 386.
81. Wearden AJ, Dowrick C, Chew-Graham C, Bentall RP, Morriss RK, et al. (2010) 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:c1777.
82. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, et al. (2001) Interventions for the Treatment and Management of Chronic Fatigue Syndrome - A Systematic Review. *JAMA* 286: 1360-1368.
83. Sharpe M, Goldsmith KA, Johnson AL, Chalder T, Walker J, White PD (2015) Supplement to: Rehabilitative treatments for chronic fatigue syndrome: long-term follow-up from the PACE trial. *Lancet Psychiatry*.
84. Larun L, Brurberg KG, Odgaard-Jensen J, Price JR (2015) Exercise therapy for chronic fatigue syndrome. *Cochrane Database Syst Rev* Art No: CD003200.
85. Hooper M (2011) Report: complaint to the relevant executive editor of the lancet about the pace trial articles published by the lancet.
86. van den Berg-Emons RJ, L'Ortye AA, Buffart LM, Nieuwenhuijsen C, Nooijen CF, et al. (2011) Validation of the Physical Activity Scale for individuals with physical disabilities. *Arch Phys Med Rehabil* 92: 923-928.
87. Belluz J (2015) How researchers dupe the public with a sneaky practice called "outcome switching".
88. Ioannidis JPA (2005) Why most published research findings are false. *PLoS Med* 2: e124.
89. Smith MEB, Haney E, McDonagh M, Pappas M, Daeges M, et al. (2015) Treatment of myalgic encephalomyelitis/chronic fatigue syndrome: A systematic review for a national Institutes of Health pathways to prevention workshop. *Ann Intern Med* 162: 841-850.
90. White PD, Clauw DJ, van der Meer JW, Moss-Morris R, Taylor RR (2015) Treatment of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *Ann Intern Med* 163: 885.
91. Fluge Ø, Risa K, Lunde S, Alme K, Rekeland IG, et al. (2015) B-Lymphocyte Depletion in Myalgic Encephalopathy/Chronic Fatigue Syndrome. An Open-Label Phase II Study with Rituximab Maintenance Treatment. *PLoS ONE* 10: e0129898.
92. Falk Hvidberg M, Brinth LS, Olesen AV, Petersen KD, Ehlers L (2015) The Health-Related Quality of Life for Patients with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS). *PLoS ONE* 10: e0132421.
93. Komaroff AL, Fagioli LR, Doolittle TH, Gandek B, Gleit MA, et al. (1996) Health status in patients with chronic fatigue syndrome and in general population and disease comparison groups. *Am J Med* 101: 281-290.