Review Article

JHP

Journal of Health Psychology
1–16

DOI: 10.1177/1359105319847261
journals.sagepub.com/home/hpq
SAGE

Assessment of the scientific rigour of randomized controlled trials on the effectiveness of cognitive behavioural therapy and graded exercise therapy for patients with myalgic encephalomyelitis/chronic fatigue syndrome: A systematic review

SA Ahmed[1], JC Mewes[1] and HJM Vrijhoef[1,2,3]

Abstract

Cognitive behavioural therapy and graded exercise therapy have been promoted as effective treatments for patients with myalgic encephalomyelitis/chronic fatigue syndrome. However, criticism on the scientific rigour of these studies has been raised. This review assessed the methodological quality of studies on the effectiveness of cognitive behavioural therapy and graded exercise therapy. The methodological quality of the 18 included studies was found to be relatively low, as bias was prominently found, affecting the main outcome measures of the studies (fatigue, physical functioning and functional impairment/status). Future research should focus on including more objective outcome measures in a well-defined patient population.

Keywords

chronic fatigue syndrome, cognitive behavioural therapy, graded exercise therapy, myalgic encephalomyelitis, systematic review

Introduction

Controversy surrounds myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) (Geraghty, 2017; Petrie and Weinman, 2017; Tuller, 2017). Subject of an ongoing fierce debate regarding the treatments for ME/CFS patients is the PACE trial (White et al., 2011) which led to recommending a programme of steadily increasing activity graded exercise therapy (GET) and cognitive behavioural therapy (CBT). The main points of discussion include the thresholds used for primary outcome measures, a broad set of inclusion criteria rather than more strict sampling criteria, contamination of the trial by promoting the success of GET and CBT to study participants during the trial and the favouring of subjective self-reported measurement instruments over objective tests of

[1]Panaxea, The Netherlands
[2]Maastricht University Medical Center, The Netherlands
[3]Vrije Universiteit Brussel, Belgium

Corresponding author:
SA Ahmed, Panaxea, Matrix II, unit 1.09, Science Park 400, 1098 XH Amsterdam, The Netherlands.
Email: anam.ahmed@panaxea.eu

physical function when defining recovery (Geraghty, 2017). After broadening the original Ramsay criteria for ME (Dowsett et al., 1990), the defining characteristics of ME (e.g. post-exertional malaise (PEM)) were no longer found mandatory (Box 1).

Box 1. Descriptions of the Oxford and Centers for Disease Control and Prevention (CDC) criteria.

Oxford criteria (Sharpe et al., 1991): unexplained physical and mental fatigue for at least 6 months, myalgia and sleep and mood disturbances; exclusion of other diseases.
CDC criteria (Fukuda et al., 1994): unexplained, persistent or relapsing fatigue for at least 6 months and the presence of at least four of the eight following symptoms for at least 6 months: impaired in short-term memory and/or concentration, sore throat, tender lymph nodes, muscle pain, joint pain, headaches, unrefreshing sleep and PEM (more than 24 hours).

After questioning for almost 30 years whether ME/CFS is a highly debilitating organic disease or a psychiatric one, reports of The National Academy of Medicine and The Health Council of the Netherlands, and revisions of the guidelines by the US-based Centers for Disease Control and Prevention (CDC) and The National Institute for Health and Care Excellence (NICE), seem to turn the tide. The reports propose new diagnostic criteria, stress the importance of their dissemination, acknowledge that ME/CFS is accompanied by substantial (physical) limitations and recommend that more research is required on pathogenesis, diagnosis and treatment (Institute of Medicine, 2015; The Health Council of the Netherlands, 2018). Revision of the guidelines for ME/CFS patients resulted in removing information about GET and CBT and stating that there is no approved treatment for ME/CFS (CDC, 2018). Moreover, the results of a review of the guidelines on the diagnosis and management of ME/CFS will be announced in 2020 (NICE, 2017).

In light of the debate around the scientific rigour of studies on therapies for ME/CFS, existing reviews assessing the effectiveness of these therapies (e.g. Larun et al., 2017; Smith et al., 2015) need to be updated, given the main methodological concerns raised more recently. Hence, our scientific literature review was guided by the following research questions:

1. What is the methodological quality of studies on the effectiveness of CBT and GET for ME/CFS patients?
2. (a) To what extent are CBT and GET significant effective treatments for ME/CFS patients?
   (b) Which outcome measures and cut-off scores were used to operationalize improvement and/or recovery?
3. What is known about harmful/unwanted effects of treatments for ME/CFS patients, and how proportionate are these to the intended effects?
4. Do study findings differentiate taking into consideration ME/CFS patients with or without PEM, the hallmark symptom of ME/CFS?

Methods

Search strategy

A systematic literature review was conducted searching the databases PubMed, Medline, Embase, PsycINFO, PEDro, the Cochrane Database of Clinical Trials and Web of Science for articles on the effectiveness of CBT and GET from January 1980 to April 2017. The search string (CFS OR chronic fatigue syndrome OR ME OR myalgic encephalomyelitis) AND (cognitive behavioural therapy OR graded exercise therapy) AND (randomized controlled trial OR controlled clinical trial OR random allocation OR double-blind method OR single-blind method OR clinical trial) was used. We also searched the reference lists of the eligible articles identified for the review. Two authors (S.A.A. and J.C.M.) screened all the references that were found in the literature search on title

and abstract. Full articles were retrieved, if the title and abstract were considered relevant. Two authors (S.A.A. and H.J.M.V.) independently assessed whether the articles fulfilled the inclusion criteria. Any disagreements were resolved by consensus. Included studies had to be randomized controlled trials (RCTs), written in English or Dutch, and the authors evaluated the effectiveness of CBT or GET (as reported by the authors of the article), included adult patients with ME/CFS (regardless of the criteria used), and measured at least one of the following outcome measures: fatigue, physical functioning, improvement/recovery, depression, pain, sleep quality, quality of life, employment status and harmful/adverse/side effects.

Data extraction and analysis

Data were extracted into a standardized form. Extracted data included inclusion and exclusion criteria, diagnostic criteria used, study size, drop-out, participants characteristics, methodological quality, use and publication of a study protocol, description of the interventions and control groups, outcome measures, measurement instruments, the outcome measures being subjective or objective, the cut-off scores for outcome measures, the definition of recovery, results and adverse/harmful effects. We defined length of follow-up as the follow-up period after the end of the intervention period. The extracted data were analysed descriptively.

The risk of bias tool of the Cochrane Collaboration was used to assess the methodological quality of the included studies (Higgins, 2011). This tool analyses whether the random sequence generation, the allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting and other biases have a high, low or unclear risk of bias. The category 'sample size power calculation' was added to this tool. In this tool, a plus sign represents a low risk of bias, a minus sign represents a high risk of bias and an empty box an unclear risk of bias. Discrepancies were resolved through consensus. Whenever selective reporting was encountered, additional literature searches were done, to see if the measured but non-reported data/outcomes could be found elsewhere.

Results

A total of 1535 publications were identified. After screening title and abstract, 1316 studies were excluded (Figure 1). Of the 54 articles identified for full-text screening, 18 articles met eligibility criteria, including 14 clinical trials.

Study- and patient characteristics

All studies were published between 1996 and 2016. Nine studies were conducted in the United Kingdom (Deale et al., 1997, 1998, 2001; Fulcher and White, 1997; O'Dowd et al., 2006; Sharpe et al., 1996, 2015; White et al., 2011, 2013), four in the Netherlands (Prins et al., 2001; Tummers et al., 2010; Vos-Vromans et al., 2016; Wiborg et al., 2015), two in the United States (Jason et al., 2007; Lopez et al., 2011) and one each in Australia (Wallman et al., 2004), New Zealand (Moss-Morris et al., 2005) and Spain (Nunez et al., 2011). Ten studies used the CDC criteria for inclusion of their patients in their study and five used the Oxford criteria (Fulcher and White, 1997; Sharpe et al., 1996, 2015; White et al., 2011, 2013). The inclusion criteria for the patient population were defined ambiguously in the trial of Deale, as the original study (Deale et al., 1997) mentioned that patients need to fulfil both the UK (Oxford) diagnostic criteria and the revised criteria of the CDC, and the follow-up studies (Deale et al., 1998, 2001) only the UK criteria.

Prins et al. (2001) stated that they used the CDC criteria; however, screening of the mandatory four out of the eight additional symptoms did not occur.

The mean age ranged from 31 to 45.9 years. In all studies, more females than males were included. The number of participants in the RCTs ranged from 49 to 641. Final measurement took place at the end of the intervention at 12 weeks (Lopez et al., 2011) and at 16 weeks (Tummers et al., 2010). The follow-up period

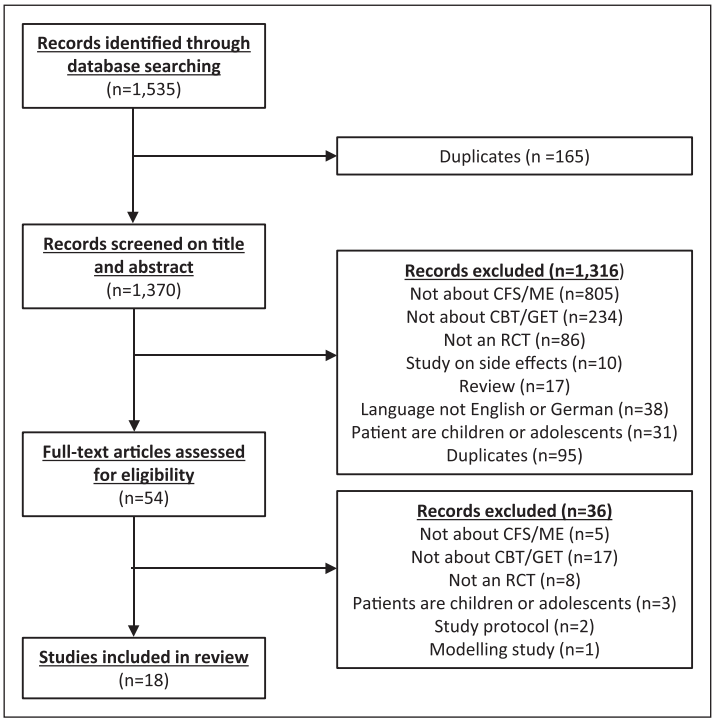

Figure 1. Flowchart of study selection.

after the end of the intervention was 4 weeks (Wallman et al., 2004), 12 weeks (Fulcher and White, 1997), 6 months (Deale et al., 1997, 1998; Moss-Morris et al., 2005; Wiborg et al., 2015), 12 months (Jason et al., 2007; Nunez et al., 2011; O'Dowd et al., 2006; Sharpe et al., 1996; Vos-Vromans et al., 2016) and 14 months (Prins et al., 2001). White et al. (2011, 2013) had a follow-up period of 12 months after randomization and Sharpe et al. (2015) a minimum of 24 months. Deale et al. (2001) had a follow-up period of 60 months. A total of 13 studies presented information on patients with some form of psychiatric diagnosis. (Deale et al., 1997, 1998, 2001; Fulcher et al., 1997; Jason et al., 2007; Moss-Morris et al., 2005; Nunez et al., 2011; O'Dowd et al., 2006; Sharpe et al., 1996, 2015; White et al., 2011, 2013).

Methodological quality

Figure 2 shows the risk of bias assessment. A low risk of bias is seen in the categories 'random sequence generation', 'selective reporting' and 'other bias'. The majority of the studies had an unclear risk of bias in the categories 'allocation concealment' and 'blinding of participants and personnel'. Studies scored a high risk of bias in the categories 'blinding of outcome assessment', 'incomplete outcome data' and 'sample size power calculation'. Moss-Morris et al. (2005), Nunez et al. (2011) and Wiborg et al. (2015) scored a low risk of bias in five categories, making them the least biased studies. Fulcher and White (1997), O'Dowd et al. (2006), Prins et al. (2001) and White et al. (2011) scored a high risk of bias in three or four categories, making them the most biased studies.

Existence and publication of the trial protocol was also analysed. Six studies mentioned its trial protocol (O'Dowd et al., 2006; Sharpe et al., 2015; Vos-Vromans et al., 2016; White et al., 2011, 2013; Wiborg et al., 2015), two of these not being publicly available (O'Dowd et al., 2006; Wiborg et al., 2015). Deviations

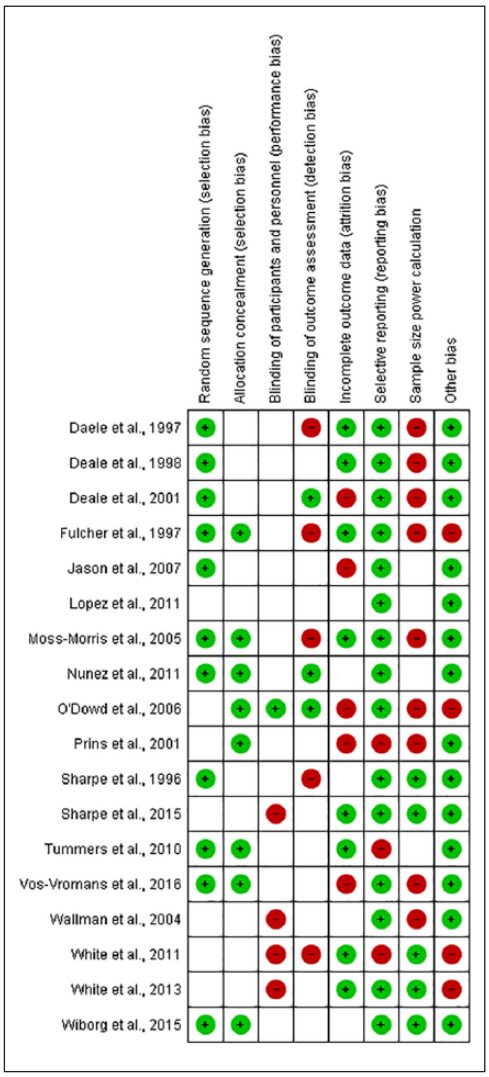

Figure 2. Summary of risk of bias of included studies.

from its protocol are mentioned by O'Dowd et al. (2006) and White et al. (2011, 2013).

Type of intervention and control groups

Of eight studies, CBT was either the main intervention (Deale et al., 1997, 1998, 2001; Wiborg et al., 2015) or the intervention compared to other interventions (Jason et al., 2007; O'Dowd et al., 2006; Prins et al., 2001; Vos-Vromans et al., 2016). In seven studies, CBT was part of the intervention (Lopez et al., 2011; Nunez et al., 2011; Sharpe et al., 1996, 2015; Tummers et al., 2010; White et al., 2011, 2013). GET was the main intervention in two studies (Fulcher and White, 1997; Wallman et al., 2004) and in five studies part of the intervention (Moss-Morris et al., 2005; Nunez et al., 2011; Sharpe et al., 2015; White et al., 2011, 2013). Seven studies made multiple comparisons of interventions (Jason et al., 2007; O'Dowd et al., 2006; Prins et al., 2001; Sharpe et al., 2015; White et al., 2011, 2013; Wiborg et al., 2015) (Table 1). As a control, five studies used standard medical care (Moss-Morris et al., 2005; Nunez et al., 2011; O'Dowd et al., 2006; Sharpe et al., 1996; Tummers et al., 2010), four studies relaxation groups (Deale et al., 1997, 1998, 2001; Jason et al., 2007), two studies flexibility and relaxation groups (Fulcher and White, 1997; Wallman et al., 2004), three studies specialist medical care (Sharpe et al., 2015; White et al., 2011, 2013), one study a psychoeducational seminar control (Lopez et al., 2011), one study a natural course control group (Prins et al., 2001) and one study a waiting list control condition (Wiborg et al., 2015). The duration of the interventions ranged from 2.5 to 8 months. The total number of sessions varied from 8 to 16 (Table 1).

Outcome measures, measuring instruments and effectiveness

Table 2 depicts the measuring instruments used per outcome measure, as reported by the authors of the papers. The main outcomes assessed in the studies were fatigue, physical functioning and functional impairment or status. For fatigue, the measurement instrument used most often was the Chalder Fatigue Questionnaire and the Checklist Individual Strength (CIS), for physical functioning the Short Form-36 (SF-36) and for functional impairment/status this was again the SF-36 and the Sickness Impact Profile (SIP). Table 2 not only shows the wide variety of outcome measures and instruments used, but

Table 1. Overview of interventions in the included studies.

Name of the first author (year of publication)	Comparison	Duration	Number of sessions/duration per session	Interval session	Supervised by	Individual (I) versus group (G)	Intervention components
Deale et al. (1997, 1998, 2001)	***CBT*** versus relaxation	4–6 months	13/not stated	Weekly or fortnightly	Not stated	I	Patient engagement and education, homework component, planned activity and rest, a sleep routine, graded activity increasement and cognitive restructuring of unhelpful beliefs and assumptions. Small, feasible timeframes of activity and rest were realized (thrice-daily 5 minutes walk vs weekly 45 minutes walk)
Fulcher and White (1997)	***Graded exercise treatment*** versus flexibility and relaxation	3 months	Not stated/not stated	Weekly	Exercise physiologist	Not stated	Home exercises (mainly walking, other: cycling, swimming) 5 days/week, 5–15 minutes at an initial intensity of 40 per cent of VO2 max, increased by 1–2 minutes with max. 30 minutes (maximum intensity of 60% of VO2 max)
Jason et al. (2007)	***CBT*** versus ***COG*** versus ***ACT*** versus relaxation	6 months	13/45 minutes	Fortnightly	CBT: not stated COG: clinical psychologist ACT: exercise physiologist	I	CBT: Based on Deale et al. (1997) COG: Patient engagement and education, stress reduction techniques for disease symptoms, restraints, and emotional distress, home relaxation exercises, techniques to uplift mood, cognitive restructuring of maladaptive thinking and cognitive strategy development of effective coping techniques ACT: Flexibility and strength assessment (shoulder flexibility test, sit and reach test and hand dynamometer test), patient engagement and education, exercise prescription (homework) and maintaining functional gains. Safe, achievable exercise targets were set. Exercise programme: frequency – three times a week, increase in exercise was gradually realized.
Lopez et al. (2011)	***CBSM*** versus control (PE)	4 months	12/2 hours	Weekly	Not stated	G	(1) relaxation part (20–30 minutes): relaxation techniques instructions and patient's view on usefulness techniques; (2) didactic and discussion part (90 minutes): cognitive restructuring of ongoing stressors, general stress management, coping skills, and interpersonal communication. Homework included.
Moss-Morris et al. (2005)	***GET*** + ***SMC*** versus SMC	3 months	Not stated/not stated	Weekly	Not stated	I	Individual exercise plan, 4–5 times/week and 10–15 minutes at an intensity of 40 per cent of VO2 max. First 6 weeks: increase in duration (3–5 minutes/week), after 6 weeks: increase in intensity (5 beats/minute/week). Final goal: 5 days/week, 30 minutes, at an intensity of 80 per cent of VO2 max.

Table 1. (Continued)

Name of the first author (year of publication)	Comparison	Duration	Number of sessions/duration per session	Interval session	Supervised by	Individual (I) versus group (G)	Intervention components
Nunez et al. (2011)	***Multidisciplinary treatment*** versus SMC	CBT: 2.5–3 months; GET: 3 months	9/CBT: 90 minutes; 9/GET: 60 minutes	CBT: Biweekly; GET: thrice-weekly	CBT: not stated; GET: physiotherapist	G	CBT, GET and conventional pharmacological symptomatic treatment. CBT: psychoeducational interventions, progressive muscle relaxation techniques, sleep routine, coping skills on pain-inducing attitudes, cognitive restructuring, information on symptoms, better attention and memory increasing techniques. GET: gradual exercise increases (5 minutes/session), flexibility exercise and relaxation therapy. Maximum exercise load: 40 minutes/day.
O'Dowd et al. (2006)	***CBT*** versus EAS versus SMC	4 months	8/2 hours	Fortnightly	Not stated	G	Clarification illness beliefs and its management, monitoring activity levels, structured (increasing) exercise programme including strength, balance and stretching exercises, sleep pattern modification, mood management advice and goals setting. Exercise was increased if certain level was achieved for several days.
Prins et al. (2001)	***CBT*** versus ***guided support group*** versus natural course	8 months	CBT:16/1 hour	Not stated	Two therapists	I	Patient engagement and education, cognitive restructuring on somatic attributions, improving sense of control over symptoms, facilitating behavioural change, structured activity programme, plan for work rehabilitation, relapse prevention and further improvement of self-control
			Guided support groups: 11/90 minutes	Not stated	Psycho-therapist	G	Experience sharing (one central theme per meeting)
Sharpe et al. (1996)	***CBT*** + ***medical care*** versus SMC	4 months	16/1 hour	Weekly	Cognitive therapist	I	Clarification illness explanation (psychological and social factors), increase in activity level and strategy development (reduce perfectionism and self-criticism and increase problem-solving approach).
Sharpe et al. (2015); White et al. (2011, 2013)	***SpMC*** + ***CBT*** versus ***SpMC*** + ***APT*** versus ***SpMC*** + ***GET*** versus specialist medical care	SpMC: 12 months; CBT: 23 weeks	SpMC: 3; CBT:14/not stated	SpMC: not stated; CBT: First 4 sessions: weekly; the rest: fortnightly	Therapist	I	SpMC: patient education, generic advice and symptomatic pharmacotherapy. CBT: addressing unhelpful cognitions, establishment of an activity, rest and a regular sleep pattern, gradual increases in physical and mental activity and increase in problem-solving approach.
		SpMC: 12 months; APT: 23 weeks (5.75 months)	SpMC: 3; APT: 14/not stated	SpMC: not stated; APT: First 4 sessions: weekly; the rest: fortnightly			SpMC: patient education, generic advice and symptomatic pharmacotherapy. APT: paced activity, linkage between activity and fatigue, awareness development on early warnings of exacerbation, reducing demands and stress and planned rest and relaxation.
		SpMC: 12 months; GET: 23 weeks (5.75 months)	SpMC: 3; GET: 14/not stated	SpMC: not stated; GET: First 4 sessions: weekly; the rest: fortnightly			SpMC: patient education, generic advice, symptomatic pharmacotherapy. GET: exercise/physical activity with an increase in duration. To avoid overexertion, target HR range was 30 minutes of light exercise five times/week. When achieved, intensity and exercise nature was increased. Main exercise was walking.

(Continued)

Table 1. (Continued)

Name of the first author (year of publication)	Comparison	Duration	Number of sessions/duration per session	Interval session	Supervised by	Individual (I) versus group (G)	Intervention components
Tummers et al. (2010)	***Stepped care (guided self-instruction*** + ***CBT, if desired)*** versus SMC	Guided self-instruction: 4 months; CBT: 6 months	Guided self-instruction: not stated; CBT: 14/ not stated	Guided self-instruction: fortnightly; CBT: not stated	Not stated	I	(1) minimal intervention: self-instruction booklet with CFS information and weekly assignments; (2) If desired, CBT: patient engagement and education, diminish somatic attributions, activity programme with increases, rehabilitation plan and development of relapse avoidance strategy.
Vos-Vromans et al. (2016)	***CBT*** versus **MRT**	CBT: 6 months	CBT: 16/45–60 minutes	CBT: first 6 weeks: weekly; next 20 weeks: biweekly	CBT: therapist	Not stated	Cognitive restructuring of negative beliefs, a sleep routine and planned graded activity increasement.
		MRT: Part 1: 2 weeks; 2 weeks no treatment; Part 2: 10 weeks	MRT: not stated/ not stated	MRT: Part 1: total contact time – 8.5 hours; Part 2: weekly and biweekly, total contact time – 33 hours	Part 1: not stated; Part 2: physical therapist an occupational therapist	Part 1: not stated; Part 2: individual	(1) observation: patient assessment (interview, physical examination, baseline assessment and goal setting); (2) treatment phase: CBT, and if relevant, body awareness therapy, gradual reactivation, pacing, mindfulness, gradual normalization of sleep/ wake rhythm and social reintegration.
Wallman et al. (2004)	***GET with pacing*** versus relaxation/ flexibility	3 months	Not stated/not stated	Every second day	Not stated	I	Graded exercise programme with aerobic activity focusing on the major large muscles of the body, initial exercise duration of 5–15 minutes every second day with the intensity based on mean HR value. Exercises were walking, cycling or swimming.
Wiborg et al. (2015)	***CBT (8 patients, 2 therapists) versus CBT (4 patients, 1 therapist)*** versus control	6 months	14/2 hours	Not stated	Not stated	Group	Personal goal setting, fixing sleep/wake cycles, reducing the focus on bodily symptoms, challenging of illness beliefs, gradual activity increase and accomplishment of personal goals.

CBT: cognitive behavioural therapy; COG: cognitive therapy treatment; ACT: anaerobic activity treatment; CBSM: cognitive behavioural stress management; PE: psychoeducational seminar control; GET: graded exercise therapy; SMC: standard medical care; EAS: education and support groups; SpMC: specialist medical care; APT: adaptive pacing therapy; MRT: multidisciplinary rehabilitation treatment; VO2 max: maximal oxygen uptake; CFS: chronic fatigue syndrome; HR: heart rate.
Intervention(s): bold, italic and underlined.

Table 2. Outcome measures and measuring instruments used.

Name of the first author (year of publication)	Fatigue	Physical functioning	Functional impairment/ status	Distress and mood	Symptoms assessment	Change in health	Quality of life	Physiological assessment	Perceived exertion	Psychological well-being/ disorder	Disabilities	Causal attributions	Improvement	Anxiety and depression	Daily functioning	Self-efficacy/sense of control	Satisfaction (life/treatment)	No label/others
Deale et al. (1997)	FPR (#), CFQ (*#)		SF-36 PF (*), WSAS (#)	GHQ, BDI									CGIS (*)					Ass. rat. (*), IA, LTGR (#)
Deale et al. (1998)												BQ						
Deale et al. (2001)	CFQ		SF-36 PF	GHQ									CGIS (*)					RF (*), COS (*), ***emp. status (*)***
Fulcher and White (1997)	CFS (*), VAS (*)				HADS, PSQI, SF-36 PF (*)	CGIS (*)		***CQM (*), TWT (*), TCLC (*)***	BS (*)									
Jason et al. (2007)	FSS (&)		SF-36 PF (*), ***emp. status, 6-MWT (&)***							BDI (*), BAI (*), BPI (*), SES (*), PSS (&), QOLS, CGIS (*)								
Lopez et al. (2011)				PSS (#), POMS (#)	CDC-SC (#)		QOLI (#)											
Moss-Morris et al. (2005)	CFQ (*)	SF-36 PF			IMQ (*)			***VO2 peak; Exp. gass; HR***	BS				CGIS (*)					
Nunez et al. (2011)	FIS		SHAQ P; W (&)				SF-36 T; PF; BP (&*)											CM (&)
O'Dowd et al. (2006)	CFQ (*), BS	***Shuttles walked (*)***, ***Walking speed (*)***	SF-36 PCS, MCS (*)	HADS						GHQ								***Recall, reaction time, HUI***
Prins et al. (2001)	CIS F (*)		SIP (*), KPS (*)				EuroQol VAS (^)			SCL-90 (^)						SES (*)		Hours worked
Sharpe et al. (1996)	Non specific FS	***6-MWT***				SRS (*)								HADS	KPS (*)			Work status, days in bed
Sharpe et al. (2015)	CFQ (&)	SF-36 (&)				CGIS									WSAS			
Tummers et al. (2010)	CIS F										SIP, SF-36 PF							
Vos-Vromans et al. (2016)	CIS F (*)		SIP		SCL-90		SF-36 PCS, MCS					CAL				SES (*)	LSQ	PSCG (*), MAAS

(Continued)

Table 2. (Continued)

Name of the first author (year of publication)	Fatigue	Physical functioning	Functional impairment/ status	Distress and mood	Symptoms assessment	Change in health	Quality of life	Physiological assessment	Perceived exertion	Psychological well-being/ disorder	Disabilities	Causal attributions	Improvement	Anxiety and depression	Daily functioning	Self-efficacy/sense of control	Satisfaction (life/treatment)	No label/others
Wallman et al. (2004)	CFQ (*)							***THR (#), OU (#), RE (#), PO (#), WEE (#), RB (#), SCWT (*)***	BS (*)				CGIS	HADS (*)				
White et al. (2011)	CFQ (*)	SF-36 PF (*)				CGIS (*)					WSAS (*), ***6-MWT (^)***, JSDS (*), HADS (*), CFS symptoms (^), PEM (^), PCM						Non specific scale	Adverse events
White et al. (2013)	CFQ	SF-36 PF				CGIS												
Wiborg et al. (2015)	CIS F (*)	SF-36 PF (*)	SIP (*)	SCL-90 (*)														

FPR: Fatigue Problem Rating; CFQ: Chalder Fatigue Questionnaire; SF-36 PF: Short Form-36 physical functioning scale; WSAS: work and social adjustment scale; GHQ: General Health Questionnaire; BDI: Beck Depression Inventory; BQ: brief questionnaire; CGIS: clinical global impression (change) score; IA: illness attributions; LTGR: long-term goals rating; RF: relapse frequency; COS: course of symptoms; Emp. status: employment status; VAS: Visual Analogue Scale; HADS: Hospital Anxiety and Depression Scale; PSQI: Pittsburgh Sleep Quality Index; CQM: contraction of the quadriceps muscle; TWT: treadmill walking test; TCLC: Thumb Prick Capillary Lactate Concentrations test; BS: Borg scale; FSS: Fatigue Severity Scale; 6-MWT: 6 minute walk test; ass. rat.: assessor ratings; BAI: Beck Anxiety Inventory; BPI: Brief Pain Inventory; SES: Self-Efficacy Scale; PSS: Perceived Stress Scale; QOLS: Quality of Life Scale; POMS: Profile of Mood States; CFS: chronic fatigue syndrome; CDC-SC: Centers for Disease Control and Prevention based CFS symptoms; QOLI: Quality of Life inventory; IMQ: Illness Management Questionnaire; VO2 peak: maximum oxygen uptake; Exp. Gass: expired gas analysis; HR: heart rate; FIS: Fatigue Impact Scale; SHAQ P: Stanford Health Assessment Questionnaire Pain; W: Weakness; SF-36 T: Short Form-36 Total; SF-36 BP: Short Form-36 Bodily Pain; CM: comorbidities; SF-36 PCS, MCS: Short Form-36 physical component scores and mental health component scores; HUI: Health Utilities Index; CIS F: Checklist Individual Strength Fatigue; SIP: Sickness Impact Profile; SRS: Self Rating Scale; KPS: Karnofsky performance score; SCL-90: Symptom Checklist 90; CAL: Causal Attribution List; LSQ: Life Satisfaction Questionnaire; PSCG: Patient-specific Complaints and Goals questionnaire; MAAS: Mindfulness Attention Awareness Scale; THR: target heart rate; OU: oxygen uptake; RE: respiratory exchange; PO: power output; WEE: weekly energy expenditure; RB: resting blood pressure; SCWT: Stroop colour word test; JSDS: Jenkins Scale for Disturbed Sleep; PEM: post-exertional malaise; PCM: poor concentration and memory; FS: fatigue syndrome.

*Significant difference found compared to other/control group(s).

^Significant effect found compared to just one group (if more than one existed).

&Significant difference between baseline and final assessment within group.

#Increase over time from baseline to final assessment, significant difference between groups.

Subjective outcome measures in black; Objective measures are bold, italic and underlined.

also the statistical differences found per (subjective or objective) outcome.

In 10 studies, all of the outcome measures were subjective (Deale et al., 1997, 1998, 2001; Lopez et al., 2011; Nunez et al., 2011; Sharpe et al., 2015; Tummers et al., 2010; Vos-Vromans et al., 2016; White et al., 2013; Wiborg et al., 2015). In eight studies, outcome measures were both subjective and objective, though the proportion of subjective measures in these studies was larger (Fulcher and White, 1997; Jason et al., 2007; Moss-Morris et al., 2005; O'Dowd et al., 2006; Prins et al., 2001; Sharpe et al., 1996; Wallman et al., 2004; White et al., 2011).

In general, the interventions were found to be effective according the study results, as most studies ($n = 15$) found CBT and/or GET, either in the original form or an adapted therapy version, a positive treatment for ME/CFS patients. Three studies (Deale et al., 2001; Nunez et al., 2011; O'Dowd et al., 2006) did not consider CBT and/or GET directly suitable for treatment.

One study mentioned the occurrence of harmful/adverse events (White et al., 2011), where serious adverse events were uncommon. Moss-Morris et al. (2005) reported that physiological assessment tests were experienced as harmful to half of the sample. An intention-to-treat analysis showed minimal adverse events in Fulcher and White (1997).

Cut-off scores

For eight studies, explicit inclusion/entry criteria were specified for subjective measures that related to (physical) functioning and fatigue severity, providing a cut-off score of what was considered 'severe' fatigue or 'severe' or 'substantial' impairment (Prins et al., 2001; Sharpe et al., 1996, 2015; Tummers et al., 2010; Vos-Vromans et al., 2016; White et al., 2011, 2013; Wiborg et al., 2015). Three of these studies (Tummers et al., 2010; Vos-Vromans et al., 2016; Wiborg et al., 2015) did not provide an entry criteria/cut-off score for physical functioning (SF-36-PF scale) even though this was a primary outcome measure.

Seven studies operationalized an 'improvement' definition based on a combination of several different outcome measures and cut-off scores that patients had to meet to be classified as 'improved' (Deale et al., 1997, 1998, 2001; Sharpe et al., 2015; White et al., 2011, 2013; Wiborg et al., 2015) (Table 3). As to Recovery, three studies (Deale et al., 2001; White et al., 2013; Wiborg et al., 2015) operationalized 'recovery' as a combined definition of different outcome measures and cut-off scores that patients had to meet in order to be classified 'recovered' (Table 3).

Discussion

The main aim of this systematic review was to assess the methodological quality of studies on the effectiveness of GET and CBT for ME/CFS patients. Further, we have studied the effectiveness of CBT and GET for ME/CFS patients, analysed the cut-off scores used for outcome measures, identified whether and which harmful events are reported in the literature and evaluated the inclusion criteria of RCTs with regard to including patients with PEM.

Our results showed that the methodological quality of the included studies was evaluated to be relatively low, as all studies were found to score an unclear or high risk in three to six of the bias categories. Thus, robust evidence to support the effectiveness of CBT and GET is lacking. The included trials showed that GET and CBT mainly were not found to be effective treatments for CBT and GET.

In the choice of outcome measured on improvement and/or recovery, the included studies mainly relied on self-reported measures. Cut-off scores were not provided or not clearly defined. Almost none of the studies reported on the occurrence of harmful/unwanted events, making it difficult to analyse if CBT and GET lead to harmful events. Moreover, all studies included CFS patients using the Oxford or CDC criteria. These do not require having PEM as a symptom.

Our findings are similar to what was reported in literature earlier. Regarding the methodological

Table 3. 'Improvement' and 'recovery' operationalized as a combination of several outcome measures (cut-off scores).

Author(s)/outcome measures	Wiborg et al., 2015	Deale et al., 1997; Deale et al., 1998; Deale et al., 2001	White et al., 2011; White et al., 2013; Sharpe et al., 2015		
Definition of 'improvement' at entry level	'improvement'	'improvement' criterion	'improved' as published	'improved' as defined in PACE protocol	
Reliable change index CIS F	> 1.96				
CIS F	< 35 (Entry criteria: ⩾ 35 'severe fatigue')				
SIP 8	< 700 (Entry criteria: ⩾ 700 'substantial impairment')				
SF-36 PF	⩾ 65	Increase of ⩾ 50 per cent from baseline or ⩾ 83 score	Improvement: ⩾ 8 point increase (Entry criteria: ⩽ 65)	⩾ 75 or increase of ⩾ 50 per cent	
CFQ 11 items		< 4 no longer 'case' or no longer 'excessive fatigue' (bimodal scoring)	At least a 2-point decrease (Likert-type scoring) (Entry Criteria: ⩾ 6 bimodal scoring)	⩽ 3 or a 50-per-cent reduction in fatigue score (bimodal scoring)	
Author(s)/outcome measures	Wiborg et al., 2015	Deale et al., 2001	White et al., 2011	White et al., 2013	
Definition of 'recovery'	**Post hoc analysis** 'recovery'	'complete recovery'	**Post hoc**. '**normal range**'	'trail Recovery' (A) and 'clinical recovery' (A + B)	Recovery (had to meet all criteria) as mentioned in protocol
CIS F	< 27				
SIP 8	< 203				
SF-36 PF	⩾ 80	> 83 on SF General Health	60 or more	(A) 60 or more (Entry criteria: 65 or less)	85 or more
CFQ 11 items		< 4	18 or less (Likert-type scoring)	(A) 18 or less (Likert-type scoring (0,1,2,3) for the 11 items) (Entry criteria: 6 or more (bimodal score))	3 or less (bimodal scoring, 0, 0, 1, 1 for the 11 items)
CGIS				(A) Overall health self-rated as 'much better' or 'very much better' (score 1 or 2)	Overall health self-rated as 'very much better' (score 1)
'Caseness' criterion		No longer meeting CFS (Oxford) criteria		(A) No longer meeting the modified Oxford case definition of CFS. (B) plus no longer meeting the CDC criteria AND the London ME criteria.	No longer meeting the Oxford case definition of CFS plus no longer meeting the CDC criteria AND the London ME criteria.
Employment		Fulltime employment			

CIS F: Checklist Individual Strength Fatigue; SIP: Sickness Impact Profile; SF-36 PF: Short Form-36 physical functioning scale; CFQ: Chalder Fatigue Questionnaire; CGIS: clinical global impression change score; CFS: chronic fatigue syndrome; ME: myalgic encephalomyelitis; CDC: Centers for Disease Control and Prevention.

quality of the studies, criticism on deviations from study protocols, using subjective primary outcome measures and not using a ME-specific patient group, was brought forward earlier by Vink (2017) and Wilshire et al. (2018). Vink (2017) and Twisk and Corsius (2017) discussed that CBT and GET were ineffective. In addition, also in other studies, neither CBT nor GET brought improvement in several objective outcomes (Chalder et al., 2015; McCrone et al., 2012; Wiborg et al., 2010).

The majority of the included studies used self-reported outcome measures, which can enhance the likelihood of biased results as CBT and GET trials are non-blinded. The nature of the interventions could also introduce bias, as ME/CFS patients are taught to focus less on their symptoms and change their views about their illness and what perpetuates it, leading to fewer self-reported symptoms.

Moreover, in studies that stated an entry criteria/cut-off score for both relevant outcome measures and improvement or recovery, the difference between what was considered severe fatigue/severe impairment and improvement on subjective outcomes differed by just one point (Tummers et al., 2010; Wiborg et al., 2015). Several trials did not provide cut-off scores for outcomes, while they did provide an entry cut-off score (e.g. Prins et al., 2001; Tummers et al., 2010; Vos-Vromans et al., 2016). The definition of recovery based on subjective outcome measure and the trial being non-blinded may create problems such as a high expectancy regarding the effects of the treatment and changes in the symptom-reporting behaviour (Wilshire et al., 2017).

Regarding the reporting of harmful study events, it is challenging to conclude on how proportionate the occurrence of harmful/unwanted events is to the observed effects. Information on the occurrence of adverse/unwanted events is essential, as research conducted by Kindlon (2011) and Vink and Vink-Niese (2018) showed that over the past 20 years patients have consistently reported either no or adverse effects from these interventions. However, not reporting adverse events is typical for this field as psychotherapy trials generally report infrequently on adverse outcomes (Duggan et al., 2014; Jonsson et al. 2014; Vaughan et al., 2014). Concerning the inclusion of ME/CFS patients with or without PEM, the studies only included CFS patients using the Oxford or CDC criteria, making the study results not specifically applicable to ME/CFS patients with PEM. The Oxford and CDC criteria have been criticized for focusing on fatigue in general and for not including PEM, a key symptom of ME/CFS (Christley et al., 2012).

The strength of this review lies in the systematic approach with a validated tool to assess the methodological quality of the studies. A limitation of our review is that the definitions of CBT and GET, that were provided in the papers, were used even though a good distinction between the definitions is often not made.

The results of this review show that high-quality evidence on the effectiveness of CBT and GET for ME/CFS patients is needed. Future research should include well-defined patient populations, using a community-based participatory research approach, to increase patient involvement in determining priorities for research and care. In order to securely demonstrate the efficacy of CBT/GET within a non-blinded design, researchers need to show that self-reported improvements are supported by objectively measurable outcomes and report on the harms/adverse events.

Conclusion

The findings of this systematic review do not support the claim that CBT and GET are effective treatments for ME/CFS patients, due to methodological flaws and biases found in the studies that are investigated in this. This review highlights the need for more high-quality research to establish effective treatment options for well-defined patient groups using objective measurement instruments.

Declaration of Conflicting Interests

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

Funding

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

ORCID iD

SA Ahmed https://orcid.org/0000-0002-3452-262X

References

Centers for Disease Control and Prevention (CDC) (2018) Myalgic encephalomyelitis/chronic fatigue syndrome. Available at: https://www.cdc.gov/me-cfs/index.html (accessed 21 February 2018).

Chalder T, Goldsmith KA, White PD, et al. (2015) Rehabilitative therapies for chronic fatigue syndrome: A secondary mediation analysis of the PACE trial. *The Lancet Psychiatry* 2(2): 141–152.

Christley Y, Duffy T and Martin CR (2012) A review of the definitional criteria for chronic fatigue syndrome. *Journal of Evaluation in Clinical Practice* 18: 25–31.

Deale A, Chalder T and Wessely S (1998) Illness beliefs and treatment outcome in chronic fatigue syndrome. *Journal of Psychosomatic Research* 45(1): 77–83.

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

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

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

Duggan C, Parry G, McMurran M, et al. (2014) The recording of adverse events from psychological treatments in clinical trials: Evidence from a review of NIHR-funded trials. *Trials* 15: 335.

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

Fulcher YF and White PD (1997) Randomised controlled trial of graded exercise in patients with the chronic fatigue syndrome. *British Medical Journal* 314: 1647–1652.

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

Higgins JPT (2011) The cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 343: d5928.

Institute of Medicine (2015) *Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness*. Washington, DC: The National Academies Press.

Jason LA, Torres-Harding S, Friedberg F, et al. (2007) Non-pharmacologic interventions for CFS: A randomized trial. *Journal of Clinical Psychology in Medical Settings* 14: 275–296.

Jonsson U, Alaie I, Parling T, et al. (2014) Reporting of harms in randomized controlled trials of psychological interventions for mental and behavioral disorders: A review of current practice. *Contemporary Clinical Trials* 38(1): 1–8.

Kindlon T (2011) Reporting of harms associated with graded exercise therapy and cognitive behavioural therapy in myalgic encephalomyelitis/chronic fatigue syndrome. *Bulleting of the IACFS/ME* 19(2): 59–111.

Larun L, Brunberg KG, Odgaard-Jensen J, et al. (2017) Exercise therapy for chronic fatigue syndrome. *Cochrane Library* 4: CD003200.

Lopez C, Antoni M, Penedo F, et al. (2011) A pilot study of cognitive behavioral stress management effects on stress, quality of life, and symptoms in persons with chronic fatigue syndrome. *Journal of Psychosomatic Research* 70: 328–334.

McCrone P, Sharpe M, Chalder T, et al. (2012) Adaptive pacing, cognitive behaviour therapy, graded exercise, and specialist medical care for chronic fatigue syndrome: A cost-effectiveness analysis. *PLoS ONE* 7(8): e40808.

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

National Institute for Health and Care Excellence (NICE) (2017) NICE to begin review of its guidance on the diagnosis and treatment of CFS/ME. *NICE*, 20 September. Available at: https://www.nice.org.uk/news/article/nice-to-begin-review-of-its-guidance-on-the-diagnosis-and-treatment-of-cfs-me (accessed 21 February 2018).

Nunez M, Fernandes-Sola J, Nunez E, et al. (2011) Health-related quality of life in patients with chronic fatigue syndrome: Group cognitive behavioural therapy and graded exercise versus usual treatment. A randomised controlled trial with 1 year of follow-up. *Clinical Rheumatology* 30: 381–389.

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

Petrie KJ and Weinman J (2017) The PACE trial: It's time to broaden perceptions and move on. *Journal of Health Psychology* 22(9): 1198–1200.

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

Sharpe M, Archard L, Banatvala J, et al. (1991) Chronic fatigue syndrome: Guidelines for research. *Journal of the Royal Society of Medicine* 84: 118–121.

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

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

Smith ME, Haney E, McDonagh M, et al. (2015) Treatment of myalgic encephalomyellitus/chronic fatigue syndrome: A systematic review for a national institutes of health pathways to prevention workshop. *Annals of Internal Medicine* 162(12): 841–850.

The Health Council of the Netherlands (2018) *Myalgische Encefalomyelitis (ME)*. Available at: https://www.gezondheidsraad.nl/nl/taak-werkwijze/werkterrein/optimale-gezondheidszorg/mecvs (accessed 21 March 2018).

Tuller D (2017) Once again, the PACE authors respond to concerns with empty answers. *Journal of Health Psychology* 22(9): 1118–1122.

Tummers M, Knoop H and Bleijenberg G (2010) Effectiveness of stepped care for chronic fatigue syndrome: A randomized noninferiority trial. *Journal of Consulting and Clinical Psychology* 78(5): 724–731.

Twisk FNM and Corsius LAMM (2017) An analysis of Dutch hallmark studies confirms the outcome of the PACE trial: Cognitive behaviour therapy with a graded activity protocol is not effective for chronic fatigue syndrome and myalgic encephalomyelitis. *General Medicine Open* 1: 1–13.

Vaughan B, Goldstein MH, Alikakos M, et al. (2014) Frequency of reporting of adverse events in randomized controlled trials of psychotherapy vs. psychopharmacotherapy. *Comprehensive Psychiatry* 55(4): 849–855.

Vink M (2017) Assessment of individual PACE trial data: In myalgic encephalomyelitis/chronic fatigue syndrome, cognitive behavioral and graded exercise therapy are ineffective, do not lead to actual recovery and negative outcomes may be higher than reported. *Journal of Neurology and Neurobiology* 3(1): 10.16966/2379-7150.136.

Vink M and Vink-Niese A (2018) Graded exercise therapy for myalgic encephalomyelitis/chronic fatigue syndrome is not effective and unsafe: Re-analysis of a cochrane review. *Health Psychology Open* 5(2): 2055102918805187.

Vos-Vromans DC, Smeets RJ, Huijnen IP, et al. (2016) Multidisciplinary rehabilitation treatment versus cognitive behavioural therapy for patients with chronic fatigue syndrome: A randomized controlled trial. *Journal of Internal Medicine* 279: 268–282.

Wallman KE, Morton AR, Goodman C, et al. (2004) Randomised controlled trial of graded exercise in chronic fatigue syndrome. *The Medical Journal of Australia* 180: 444–448.

White PD, Goldsmith KA, Johnson AL, et al. (2011) Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): A randomised trial. *The Lancet* 377: 823–836.

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

Wiborg JF, Knoop H, Stulemeijer M, et al. (2010) How does cognitive behaviour therapy reduce fatigue in patients with chronic fatigue syndrome? The role of physical activity. *Psychological Medicine* 40(8): 1281–1287.

Wiborg JF, Van Bussel J, Van Dijk A, et al. (2015) Randomised controlled trial of cognitive behaviour therapy delivered in groups of patients with

chronic fatigue syndrome. *Psychotherapy and Psychosomatics* 84: 368–376.

Wilshire CE, Kindlon T, Courtney R, et al. (2018) Rethinking the treatment of chronic fatigue syndrome – a reanalysis and evaluation of findings from a recent major trial of graded exercise and CBT. *BMC Psychology* 6: 6.

Wilshire CE, Kindlon T, Matthees A, et al. (2017) Can patients with chronic fatigue syndrome really recover after graded exercise or cognitive behavioural therapy? A critical commentary and preliminary re-analysis of the PACE trial. *Fatigue: Biomedicine, Health & Behavior* 5(1): 43–56.