

Guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome (GETSET): a pragmatic randomised controlled trial



Lucy V Clark, Francesca Pesola, Janice M Thomas, Mario Vergara-Williamson, Michelle Beynon, Peter D White



Summary

Background Graded exercise therapy is an effective and safe treatment for chronic fatigue syndrome, but it is therapist intensive and availability is limited. We aimed to test the efficacy and safety of graded exercise delivered as guided self-help.

Methods In this pragmatic randomised controlled trial, we recruited adult patients (18 years and older) who met the UK National Institute for Health and Care Excellence criteria for chronic fatigue syndrome from two secondary-care clinics in the UK. Patients were randomly assigned to receive specialist medical care (SMC) alone (control group) or SMC with additional guided graded exercise self-help (GES). Block randomisation (randomly varying block sizes) was done at the level of the individual with a computer-generated sequence and was stratified by centre, depression score, and severity of physical disability. Patients and physiotherapists were necessarily unmasked from intervention assignment; the statistician was masked from intervention assignment. SMC was delivered by specialist doctors but was not standardised; GES consisted of a self-help booklet describing a six-step graded exercise programme that would take roughly 12 weeks to complete, and up to four guidance sessions with a physiotherapist over 8 weeks (maximum 90 min in total). Primary outcomes were fatigue (measured by the Chalder Fatigue Questionnaire) and physical function (assessed by the Short Form-36 physical function subscale); both were self-rated by patients at 12 weeks after randomisation and analysed in all randomised patients with outcome data at follow-up (ie, by modified intention to treat). We recorded adverse events, including serious adverse reactions to trial interventions. We used multiple linear regression analysis to compare SMC with GES, adjusting for baseline and stratification factors. This trial is registered at ISRCTN, number ISRCTN22975026.

Findings Between May 15, 2012, and Dec 24, 2014, we recruited 211 eligible patients, of whom 107 were assigned to the GES group and 104 to the control group. At 12 weeks, compared with the control group, mean fatigue score was 19·1 (SD 7·6) in the GES group and 22·9 (6·9) in the control group (adjusted difference -4·2 points, 95% CI -6·1 to -2·3, $p < 0·0001$; effect size 0·53) and mean physical function score was 55·7 (23·3) in the GES group and 50·8 (25·3) in the control group (adjusted difference 6·3 points, 1·8 to 10·8, $p = 0·006$; 0·20). No serious adverse reactions were recorded and other safety measures did not differ between the groups, after allowing for missing data.

Interpretation GES is a safe intervention that might reduce fatigue and, to a lesser extent, physical disability for patients with chronic fatigue syndrome. These findings need confirmation and extension to other health-care settings.

Funding UK National Institute for Health Research Research for Patient Benefit Programme and the Sue Estermann Fund.

Copyright © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license.

Introduction

Chronic fatigue syndrome is characterised by chronic, disabling fatigue in the absence of an alternative diagnosis.¹ Myalgic encephalomyelitis is thought by some to be the same disorder, whereas others consider it a different illness.¹ The prevalence of chronic fatigue syndrome in the UK and US population varies between 0·2% and 2·6%,¹ depending on definition, and the prognosis is poor if untreated, with a median of 5% of patients recovering.²

The UK National Institute for Health and Care Excellence (NICE) recommends graded exercise therapy (GET) and cognitive behavioural therapy (CBT) as treatments.³ However, patient organisations in the UK have reported, on

the basis of survey results, that GET might at times be harmful and have recommended different management approaches such as pacing (ie, living within the limits imposed by the illness).^{4,5} In a 2016 Cochrane review of eight randomised controlled trials,⁶ exercise therapy was more effective in improving both fatigue and physical functioning than either passive treatments or no treatment. GET delivered by specialist therapists, as used in these trials, is intensive and expensive, with up to 15 sessions required over a 3–6 month period.^{6,7} Furthermore, access to clinics that provide these interventions might be unavailable.⁸

Lancet 2017; 390: 363–73

Published Online

June 22, 2017

[http://dx.doi.org/10.1016/S0140-6736\(16\)32589-2](http://dx.doi.org/10.1016/S0140-6736(16)32589-2)

S0140-6736(16)32589-2

See [Comment](#) page 335

Centre for Psychiatry (L V Clark PhD, M Beynon MSc, Prof P D White MD), Centre for Cancer Prevention (F Pesola PhD), and Vice Principal (Health) Offices (J M Thomas MSc), Wolfson Institute of Preventive Medicine, Bart's and the London School of Medicine and Dentistry, Queen Mary University, London, UK; and Chronic Fatigue Syndrome/Myalgic Encephalopathy Service, Kent and Medway National Health Service and Social Care Partnership Trust, Maidstone, Kent, UK (M Vergara-Williamson ClinPsyD)

Correspondence to:

Dr Lucy Clark, Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Bart's and the London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK
l.clark@qmul.ac.uk

Research in context

Evidence before this study

We searched PubMed, PsychINFO, and the Cochrane Library from database inception until Aug 1, 2016, without language restrictions, for full reports of randomised controlled trials, systematic reviews, and meta-analyses using the search terms “chronic fatigue syndrome”, “myalgic encephal*”, “self-help”, “self-management”, “self-care”, and “self-instruction”.

We excluded trials of adolescents, education, and group interventions. Chalder and colleagues found that a self-help approach based on cognitive behavioural therapy (CBT) benefited patients with fatigue in the community, but not all patients in that study met the criteria for chronic fatigue syndrome. After excluding studies in which participants had unexplained fatigue but were not diagnosed with chronic fatigue syndrome, we identified two randomised controlled trials. Knoop and colleagues found that “guided self-instruction”, provided by a cognitive behavioural therapist using a booklet, improved fatigue and disability more than being on a waiting list. Tummers and colleagues found that the

same approach delivered by psychiatric nurses improved fatigue, but not social or physical functioning, more than being on a waiting list. We found no trials of self-help management for chronic fatigue syndrome based on guided exercise therapy principles.

Added value of this study

In GETSET, a guided self-help management approach, when added to specialist medical care, improved symptomatic fatigue and physical functioning more than did specialist medical care alone. This was also the case in two subgroups of patients who met the Oxford criteria and US Centers for Disease Control and Prevention criteria for chronic fatigue syndrome. The intervention was at least as safe as specialist medical care alone.

Implications of all the available evidence

Guided exercise self-help might be useful as an initial management approach for patients with chronic fatigue syndrome, but confirmation and extension to other health-care settings are needed.

Chalder and colleagues⁹ found that a self-help approach based on CBT benefited patients with fatigue in the community, but not all patients in this study met the criteria for chronic fatigue syndrome. Two trials^{10,11} of self-help management approaches for chronic fatigue syndrome based on CBT principles have been done, but self-help approaches based on GET principles have not been tested in clinical trials so far. We designed the graded exercise therapy guided self-help trial (GETSET)¹² to compare the efficacy and safety of guided graded exercise self-help (GES) added to specialist medical care (SMC) with SMC alone. We hypothesised that GES plus SMC would be more efficacious in improving fatigue and physical function than, and as safe as, SMC alone.

Methods

Study design

GETSET was an open-label, pragmatic randomised controlled trial done at two UK National Health Service (NHS) secondary-care clinics for chronic fatigue syndrome in central London and Kent, a more rural area. The study was approved by the UK National Research Ethics Service Committee London—London Bridge (reference 11/LO/1572) on the Nov 23, 2011. We followed ethical guidance from the Medical Research Council Good Clinical Practice Guide, which recommends adherence to the 1996 version of the Declaration of Helsinki. The study protocol has been published.¹² At the time of publication, approvals were being sought to make suitably anonymised data available for external researchers to request through the Yale University Open Data Access (YODA) Project, which is prepared to share these data. The YODA Project is an independent organisation that advocates for the responsible sharing of

clinical research data. Data from available clinical trials are shared through the YODA Project with registered users with approved proposals for scientific research.

Participants

We recruited adult patients (aged 18 years and older) attending these clinics who were diagnosed with chronic fatigue syndrome and placed on a waiting list for therapy. Several diagnostic criteria exist for chronic fatigue syndrome.^{3,13,14} In this trial, we recruited participants who met the NICE criteria,³ which are used by NHS clinicians. The NICE criteria require at least 4 months of clinically evaluated, unexplained, persistent, or relapsing fatigue with a definite onset that has resulted in a substantial reduction in activity and that is characterised by postexertional malaise or fatigue, or both.³ They also require at least one of ten related symptoms: difficulty sleeping, headaches, cognitive dysfunction, general malaise or flu-like symptoms, painful lymph nodes, sore throat, physical or mental exertion making symptoms worse, dizziness or nausea, palpitations, or multisite muscle or joint pain without evidence of inflammation.³ Medical assessment by the clinic doctor included history, physical, and mental state examinations, and laboratory tests, as recommended by NICE,³ before trial entry to exclude alternative diagnoses. Participants were also assessed by the US Centers for Disease Control and Prevention (CDC) criteria¹³ and the Oxford criteria¹⁴ for chronic fatigue syndrome. The structured clinical interview from the Diagnostic and Statistical Manual of Mental Disorders IV was used to diagnose both exclusionary and allowable comorbid psychiatric disorders.¹⁵

Patients were excluded if they were younger than 18 years, had current suicidal thoughts or comorbid

For the Medical Research Council Good Clinical Practice Guide see <http://c.yocdn.com/sites/www.tmn.ac.uk/resource/collection/476F14CA-DAE9-4144-BEA0>

For the Yale University Open Data Access Project see <http://yoda.yale.edu/>

psychiatric conditions requiring exclusion, had read the GES guide previously, had already received GET at one of the trial clinics, were unable to speak or read English adequately, or had physical contraindications to exercise.¹² Participants provided written informed consent before entry into the trial.

Randomisation and masking

Participants were randomly allocated to receive GES plus SMC (GES group) or SMC only (control group) by the King's College London Clinical Trials Unit, after baseline consent and assessment, independently of the trial team. Block randomisation (randomly varying block sizes) was done at the level of the individual with a computer-generated sequence and was stratified by centre, depression score according to the Hospital Anxiety and Depression Scale (HADS, ≤ 10 or ≥ 11),¹⁶ and severity of disability according to the Short-Form 36 physical function subscale (SF-36 PF, ≤ 40 and ≥ 45).¹⁷ After notification of intervention allocation, the trial manager informed the participant. As with any trial with therapist and participant involvement, neither could be masked from intervention allocation. The trial manager was responsible for arranging remote randomisation and informing the participants and therapists of intervention allocation, and therefore could not be masked from intervention allocation. The trial statistician, Trial Steering Committee, and the Data Monitoring and Ethics Committee were masked from intervention allocation.

Procedures

Before randomisation, all patients had at least one SMC consultation, delivered by doctors with specialist experience in chronic fatigue syndrome. SMC could involve prescriptions or advice regarding medication, as indicated for symptoms or comorbid conditions such as insomnia, pain, or depressive illness. Although not routinely scheduled during the trial, further SMC sessions were available after randomisation for patients who required it, but it was not a standardised intervention.

In addition to SMC being available, participants in the GES group were given a copy of a self-help booklet¹⁸ describing a six-step programme of graded exercise self-management. The GES booklet was based on the approach of GET developed for the PACE trial⁷ and on NICE recommendations,³ and was iteratively developed with pilot patients with chronic fatigue syndrome. Participants were encouraged to use the six steps described in the booklet: stabilising a daily routine, starting regular stretching, deciding on a physical activity goal and choosing a type of activity with which to start, setting a physical activity baseline, increasing the duration of physical activity and finally the intensity.

The most commonly chosen exercise was walking. Importantly, if a participant found that their symptoms increased after an incremental change in their activity, they were advised to maintain their activity at the same

level for longer than a week, until symptoms had settled, before considering another incremental increase.

In the first session, which lasted 30 min, a physiotherapist provided guidance on following the booklet and answered any questions from participants. This first session occurred within 5 working days of randomisation and was delivered face to face, by Skype, or by telephone. Up to three further appointments of 20 min each were offered over 8 weeks, and participants could choose to receive the guidance either by Skype or by telephone. Two experienced physiotherapists were trained to support the participants in using the booklet, but they were explicitly told not to provide therapy. During each session, they inquired about progress and answered any questions, with a focus on moving forward to the next step. They recognised achievements and provided feedback to participants on their efforts, with the aim of increasing motivation and self-efficacy. Mid-way through the guidance intervention, the physiotherapist discussed setbacks with the participant. A therapy leader trained the two physiotherapists until they were deemed competent and then provided regular individual supervision. The physiotherapists followed a manual when delivering the guidance, and all participant guidance sessions were audio-recorded for supervision, feedback, and monitoring of treatment integrity. If a participant could not be contacted by telephone or Skype, an email was sent to re-engage them.

Participants were assessed at baseline and 12 weeks after randomisation, roughly 4 weeks after GES guidance had ended. Participants were also followed up at 12 months after randomisation, and these results will be published in a separate paper. The trial manager undertook the baseline assessments either face to face in the clinic or via telephone or Skype. Follow-up questionnaires were sent by mail with freepost envelopes for return. If the follow-up questionnaires were not returned after reminders, the questionnaire was completed in the waiting room before the first clinic therapy appointment. All measures were self-rated by the participants.

The number of contacts with the physiotherapist during the trial, as well as the mode and duration of each contact, were recorded. At 12 weeks, participants were also asked in a questionnaire how satisfied they were with the help they had received during the study; the seven-point scale was condensed into four categories: dissatisfied (very or moderately dissatisfied), minimum (slightly dissatisfied, neither, or slightly satisfied), satisfied (moderately or very satisfied), and "did not receive help". The physiotherapists rated participants on their health (with clinical global impression [CGI]), adherence to their GES programme on a five-point scale (completely, very well, moderately well, slightly, and not at all), and their acceptance of the therapy model (also completely, very well, moderately well, slightly, and not at all). Information was retrieved from NHS service computerised attendances to determine how many SMC sessions participants had received during the trial.

For more on the manual followed by physiotherapists see <http://www.wolfson.qmul.ac.uk/images/pdfs/getset/GETSET%20therapists%20manual%20with%20appendices.pdf>

For the GES booklet see <http://www.wolfson.qmul.ac.uk/images/pdfs/getset/GET%20guide%20booklet%20version%201%2022062010.pdf>

Outcomes

The two primary outcomes were fatigue—measured by the Chalder Fatigue Questionnaire¹⁹ (CFQ; Likert scoring 0, 1, 2, and 3; range 0–33; highest score is most fatigue)—and physical function, measured by SF-36 PF (range 0–100; highest score is best function) 12 weeks after randomisation.¹⁷ These questionnaires are valid and reliable measures, and have been used in previous trials.^{6,7} The original protocol had only one primary outcome measure, the SF-36 PF. However, when some eligible participants were found to have high SF-36 PF scores at randomisation (because of their illness affecting cognitive or social functions but not physical function), we decided to also include fatigue, using the CFQ, as a co-primary outcome. This decision was made mid-way through trial recruitment (on June 20, 2013, after recruitment of 99 [47%] patients), before any outcome data had been examined, and was approved by the Research Ethics Committee, the Data Monitoring and Ethics Committee, and the Trial Steering Committee.¹²

The main secondary outcome was self-rated CGI score to assess change from baseline in overall health (CGI-health) and in chronic fatigue syndrome (CGI-CFS). These seven-point scales were condensed into three categories: negative change (very much worse or much worse), minimum change (a little worse, no change, or a little better), and positive change (much better or very much better) for analyses.²⁰ Other secondary outcomes were anxiety and depression (measured by the Hospital Anxiety and Depression Scale),¹⁶ overall health (measured by Patient Health Questionnaire-13 [PHQ-13]),²¹ quality of life (measured by the Euroqol Questionnaire [EQ-5D]),²² global functioning (measured by the Work And Social Adjustment Scale [WSAS]),²³ and physical activity (measured by the International Physical Activity Questionnaire [IPAQ]).²⁴ We removed two symptoms (menstrual cramps and dyspareunia) from PHQ-15 to create PHQ-13 because these symptoms were not reported by patients of both sexes.²¹ The IPAQ categorises patients as having undertaken a low, moderate, or high level of activity in the previous week.²⁵ Quality of life, (measured by EQ-5D) and economic measures will be reported separately in another paper.

Safety outcomes were non-serious adverse events, serious adverse events, serious adverse reactions to trial interventions, serious deterioration, and active withdrawal from intervention.¹² Adverse events were self-reported by participants in the follow-up questionnaire in response to specific questions regarding any new health problem since the start of the trial. Adverse events were considered serious if they involved death, hospital admission, increased severe and persistent disability, or self-harm, or when they were life-threatening or required an intervention to prevent one of these events. SMC clinicians and GES physiotherapists reported possible serious adverse events and reactions to the centre leader, and appropriate action was taken.¹² Two suitably qualified scrutineers who were masked from intervention

allocation reviewed all such adverse events, independently from the trial team, to establish whether there were any further serious adverse events. Disagreements were resolved by consensus between the two scrutineers. They were then unmasked from intervention allocation to establish if any serious adverse events were serious adverse reactions to trial interventions. Serious deterioration in health was defined as any of the following outcomes: a decrease in SF-36 PF score of 10 points or more between baseline and 12 weeks;¹⁷ scores of “much worse” or “very much worse” on the participant-rated CGI-health score;²⁰ or active withdrawal from the intervention because of worsening.

Statistical analysis

Our original sample size calculation was based on use of the SF-36 PF as our primary outcome measure, but the significance level was reduced from 5% to 2·5% to accommodate two primary outcomes. In a previous trial of chronic fatigue syndrome and myalgic encephalomyelitis in secondary care (n=641),⁷ mean SF-36 PF scores were 37 (SD 15) at baseline and 48 (21) after 12 weeks of practitioner-led GET (ie, an 11-point increase).⁷ On the basis of these findings⁷ and our estimate that GES would be less effective than therapist-delivered GET, our sample size calculations were based on the assumption that a mean difference of 8 (SD 18) points between study groups at 12 week follow-up would be a clinically meaningful difference on the SF-36 PF scale. Thus, assuming a significance level of $\alpha=2\cdot5\%$ and a power of 80%, we required a minimum of 98 participants in each group. This sample size was upwardly adjusted to allow for loss to follow-up and other adherence issues. Based on results from previous trials^{6,10} we expected a roughly 10% loss to trial follow-up; therefore, we aimed to recruit 109 patients to each group (ie, a total of 218). Based on results from a previous trial of GET,⁷ in which the difference between baseline and 12 weeks was 5·4 points on the CFQ,⁷ we assumed that a mean difference between intervention groups of 3 (SD 6) points would represent a clinically meaningful difference at 12 week follow-up. Hence, assuming a significance level of $\alpha=2\cdot5\%$ and a power of 80%, we required a total of 154 patients, which was within the 218 already planned.

We assessed data distribution of continuous measures, and normality of the data and regression residuals were explored. If the data were not normally distributed, they were transformed. Items with missing data on the primary and secondary outcome variables at baseline and follow-up were imputed with mean replacement (prorating) when less than 20% of item responses per scale were missing.

In the primary analysis, we compared results from both the SF-36 PF and the CFQ at 12 weeks between study groups, adjusted with multivariable linear regression analyses. All randomised participants with outcome data at follow-up (ie, the modified intention-to-treat population) were included in analyses, regardless of any departure

from the allocated intervention. We adjusted models for study group, baseline measure of the outcome, and stratification factors (ie, study centre, depression score [high vs low], and baseline physical function [high vs low]). We estimated effect sizes using Cohen's *d*, calculated by the adjusted mean difference between the two groups divided by the pooled SD at follow-up.

In the secondary analysis, we used the χ^2 analysis to assess the differences between study groups in the number of participants who achieved a clinically meaningful improvement on the SF-36 PF (ie, an 8 point increase) and the CFQ (ie, a 3-point decrease). We dichotomised the outcome measures using their corresponding clinically useful criteria and assessed the association between study group and clinically meaningful change score using a χ^2 test. For the two CGI scales, we compared the proportions of participants reporting a positive change, minimal change, or negative change across study groups using ordinal logistic regression, adjusting for our stratification variables.

To do a strict intention-to-treat analysis when more than 20% of data from any questionnaire of any participant was missing for either the main outcomes or covariate variables, we imputed missing data using multiple imputations by chained equations. Since multiple imputations rely on the assumption that data are missing at random, we included the prorated baseline outcome measures, covariates, and auxiliary variables associated with the outcomes and their so-called missingness in our imputation model (ie, IPAQ and PHQ-13). Inclusion of auxiliary variables related to missingness has been found to produce more accurate estimates;²⁶ therefore, we included the following auxiliary variables in the model: ethnic origin (white vs other), education level, employment status, and national patient organisation membership.

Following recent guidelines,²⁷ we also did a sensitivity analysis that took into account the partially nested design of the study to assess the potential effect of the so-called therapist effect on our results. We did a further sensitivity analysis to adjust for potential confounders, including sex, age, employment status, ethnic origin, education level, anxiety, baseline IPAQ, and stratification factors. We did additional sensitivity analyses in the per-protocol population to investigate the robustness of the conclusions of the primary analysis, following departures from the randomised intervention policies. We first excluded participants in the GES group who did not attend any of the four guided support sessions, and then excluded participants in the control group who reported using a GET self-help approach. We also did a subgroup analyses in patients meeting the CDC¹³ or Oxford¹⁴ criteria for chronic fatigue syndrome, using regression analyses as implemented in our primary intention-to-treat analysis. This analysis was done to assess whether results were consistent regardless of the definition criteria implemented. We assessed whether the effect in the GES group was moderated by baseline levels of depression or

severity of physical function (eg, our stratification factors). These subgroup analyses were exploratory.²⁸

We used χ^2 tests, or Fisher's exact test when appropriate, to describe differences in proportions of participant satisfaction between intervention groups and any differences in proportions with serious adverse events, serious adverse reactions to trial interventions, or serious deterioration (one or more vs none) between study groups. 97 patients from the GES group and 101 patients from the control group were included in the safety analysis.

Finally, we did regression analyses, adjusted for baseline scores and stratification factors, for several secondary outcomes to assess the effect of study group on overall health and physical functioning (ie, WSAS, HADS, PHQ-13, and IPAQ). Because of the ordinal

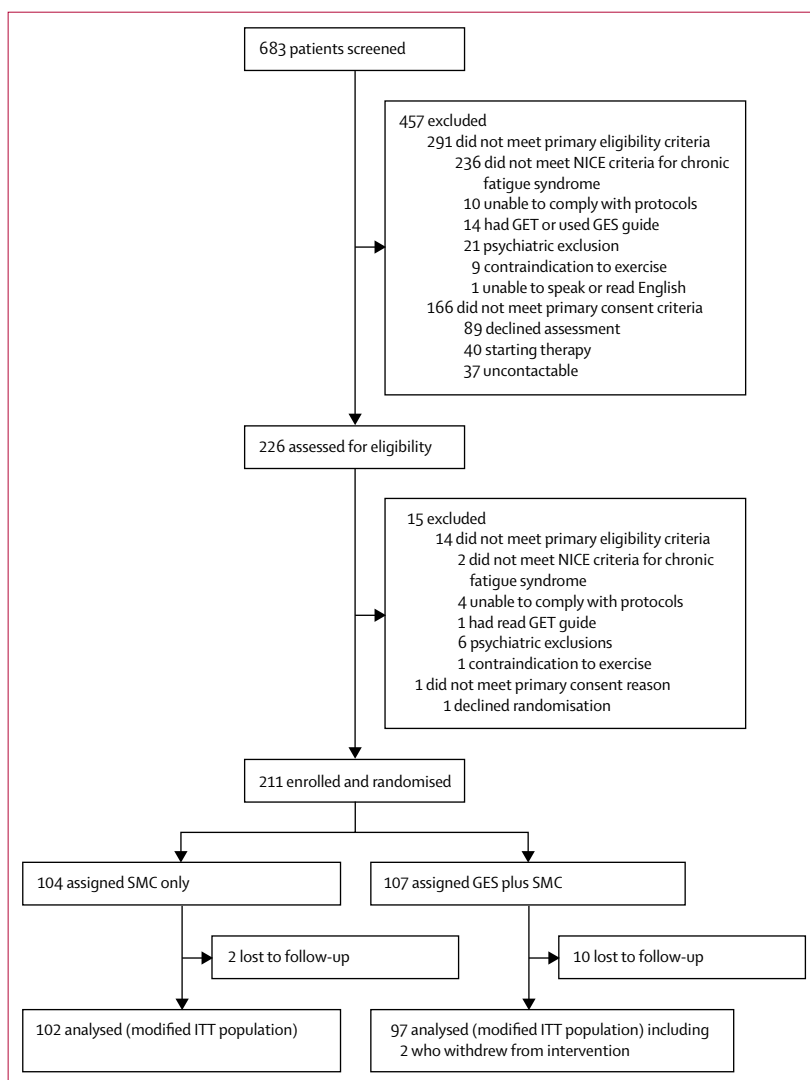


Figure: Trial profile

GET=graded exercise therapy. GES=guided graded exercise self-help. ITT=intention-to-treat. NICE=UK National Institute for Health and Care Excellence. SMC=specialist medical care.

	GES group (n=107)	Control group (n=104)
Age (years)	38.1 (11.1)	38.7 (12.7)
Female	88 (82%)	79 (76%)
White	94 (88%)	94 (90%)
CDC criteria ¹²	73 (68%)	77 (74%)
Oxford criteria ¹³	83 (78%)	87 (84%)
Current major depressive disorder	10 (9%)	11 (11%)
Duration of illness (months)	46 (23–114)	42 (25–99)
Physical activity (min per week)	120 (30–360)	185 (75–570)

Data are mean (SD), n (%), or median (IQR). GES=guided graded exercise self-help. CDC=US Centers for Disease Control and Prevention.

Table 1: Baseline characteristics

	GES group
Mode and attendance (n=107)	
First guidance session*	
Face to face	43 (40%)
Skype or telephone	61 (57%)
Number of guidance sessions attended	
Four	80 (75%)
Three	14 (13%)
Two	7 (7%)
One	3 (3%) [†]
Zero	3 (3%) [‡]
Satisfaction with GES (n=96)[§]	
Moderately or very satisfied	82 (85%)
Minimally satisfied or dissatisfied	10 (10%)
Moderately or very dissatisfied	1 (1%)
Reported not receiving help	3 (3%)

Data are n (%). GES=guided graded exercise self-help. *Three participants did not attend. [†]Two patients actively withdrew from the intervention after attending one session, stating family commitments; one patient was not contactable after the first session. [‡]One patient reported a family bereavement, one reported grief, and one was not contactable. [§]The remaining 11 patients did not answer the question (n=1) or have data available at short-term follow-up (n=10).

Table 2: Mode and attendance of, and satisfaction with, GES

nature of the IPAQ and CGI scales, we used ordinal regressions. Before conducting ordinal regressions, we checked for the proportional odds assumption using a Brent test. The two CGI scales met the assumption, but the IPAQ did not. Therefore, when analysing IPAQ results, we implemented a more flexible ordered logistic regression, which allows for partial proportional odds using the Stata command “gologit2”.²⁹ All statistical analyses were done with Stata version 13 and SPSS version 22.

This trial is registered at ISRCTN, number ISRCTN22975026.

Role of the funding source

The funders and sponsors of the study had no role in study design, data collection, data analysis, data interpretation,

or writing of the report. LVC, FP, and PDW had access to all the data. All authors commented on drafts and approved the final report. The corresponding author had final responsibility for the decision to submit for publication.

Results

Between May 15, 2012, and Dec 24, 2014, we screened 683 patients, of whom 226 (33%) were assessed for eligibility and 211 (93%) were subsequently enrolled (figure). The most common reason for exclusion at initial clinician screening was failure to meet the NICE criteria for chronic fatigue syndrome (236 participants). 90 (24%) of 378 patients who were clinically eligible declined assessment or randomisation, although only 11 (3%) of these patients declined because they explicitly did not want GES. 104 patients were randomly allocated to the control group and 107 to the GES group. Two (2%) patients in the control group and ten (9%) patients in the GES group were lost to follow up ($p=0.03$; figure). Two (2%) participants in the GES groups and none from the control group actively withdrew from the intervention; this was not significantly associated with study group ($p=0.50$).

Baseline characteristics are shown in table 1. Most participants in the GES group attended four guidance sessions with a physiotherapist, and most reported satisfaction with GES (table 2).

12 week outcome data collection was completed on April 30, 2015. Follow-up questionnaires were returned after a mean of 92 days (SD 31) from randomisation. On the SF-36 PF, we had one participant (in the GES group) with one item missing at baseline and five (one in GES group and four in control group) with one item missing at follow-up. For the CFQ, we had five participants (two in GES group and three in control group) with one or two items missing at baseline and three (two in GES group and one in control group) with one item missing at follow-up. By prorating values for these participants, our modified intention-to-treat population included 102 participants in the control group and 97 in the GES group. After adjusting for baseline scores and stratification factors, participants who received GES scored 4.2 points (95% CI 2.3–6.1; $p<0.0001$) lower on the CFQ scale and 6.3 points (1.8–10.8; $p=0.006$) higher on the SF-36 PF scale than the control group at 12 week follow-up (table 3). The absolute effect sizes were 0.53 (0.24 to 0.81) for fatigue and 0.20 (–0.08 to 0.48) for physical function. Twice as many participants in the GES group improved by more than 8 points on the SF-36 PF scale compared with the control group, and more participants in the GES group improved by at least 3 points on the CFQ compared with the control group (table 3). 33 (34%) of the 97 participants in the GES group improved by both at least 8 points on SF-36 PF and at least 3 points on the CFQ, compared with 14 (14%) of 102 in the control group ($\chi^2 11.4$; $p=0.001$).

Strict intention-to-treat analysis, with imputed missing information included (n=211), confirmed the findings of the main analysis for the primary outcomes (table 4).

Adjustment for clustering at the physiotherapist level did not modify the results (intraclass correlation coefficient 0), and neither did limiting the analysis to participants who met either the Oxford criteria or the CDC criteria (table 4), nor adjusting for age, sex, ethnic origin, study centre, education level, employment status, high SF-36 PF, high anxiety, high depression, and baseline IPAQ scores (low, medium, or high).

Non-serious adverse events were reported by about a quarter of participants in both study groups, with no significant difference between groups (table 5). Serious adverse events were uncommon (a participant attended Accident and Emergency [A&E] department after falling and damaging an arm; no fracture was found, and they were discharged; a participant attended A&E after twisting a knee, a damaged cartilage was diagnosed in the knee, and they were discharged; and a participant was admitted to hospital overnight for numbness in the right arm and leg, a neurologist assessed them and they were discharged the next day), and no serious adverse reactions were reported in either group. The difference in serious deterioration was not significant between study groups (table 5). Around a quarter of participants in both groups had deteriorations of 10 or more points in physical functioning, but more participants in the control group reported worsening in both CGI-health and CGI-CFS (table 5). As a post-hoc sensitivity analysis, if we assume that all 13 participants without these follow-up data would have reported worsening, the two groups would not differ in terms of CGI-health (11 in GES group vs 11 in control group; χ^2 0.01, $p=0.94$) and CGI-CFS (10 vs 12; χ^2 0.27, $p=0.60$).

At 12 weeks, participants in the GES group were more likely to have had a positive change in both overall health and chronic fatigue syndrome on the CGI scale than did those after SMC alone (table 6). Seven (7%) patients in the control group reported a negative change for both overall health and chronic fatigue syndrome, whereas only one (1%) patient in the GES group reported a negative change for overall health and none reported a negative change in chronic fatigue syndrome. In view of the small cell sizes, we did a post-hoc sensitivity analysis using 10 000 Monte Carlo permutations to ascertain whether results might have been due to chance, and results from this analysis confirmed our findings ($p=0.0002$, 95% CI 0.00002–0.001 for overall health; $p=0.0013$, 0.0006–0.002 for chronic fatigue syndrome).

At 12 weeks, participants in the GES groups had better outcomes than did participants in the control group for work and social adjustment scores, depression, and anxiety, but not for general physical symptoms (table 7). Similarly, participants in the GES group were more likely to score high, rather than moderate or low, on physical activity (IPAQ) than those in the control group (table 7).

We found a significant effect for the interaction between study group and baseline physical functioning on the SF-36 PF scores at 12 weeks (mean difference -10.7 ,

	Fatigue (CFQ)		Physical function (SF-36 PF)	
	GES group	Control group	GES group	Control group
Baseline				
n	107	104	107	104
Mean score (SD)	26.3 (4.8)	26.0 (4.6)	47.3 (22.2)	50.1 (22.6)
12 weeks				
n	97	102	97	102
Mean score (SD)	19.1 (7.6)	22.9 (6.9)	55.7 (23.3)	50.8 (25.3)
Mean difference compared with control group (95% CI)*	-4.2 (-6.1 to -2.3)	..	6.3 (1.8 to 10.8)	..
p value	<0.0001	..	0.006	..
Number improved from baseline†	62 (64%)	45 (44%)	44 (45%)	22 (22%)
Pearson χ^2	7.8	..	12.7	..
p value	0.005	..	<0.0001	..

CFQ=Chalder Fatigue Questionnaire. SF-36 PF=Short-Form 36 physical function subscale. GES=guided graded exercise self-help. *Adjusted for baseline, study centre, high SF-36 PF (≥ 45), and high depression (≥ 11). †By ≥ 3 points on the CFQ and ≥ 8 points on SF-36 PF.

Table 3: Primary outcomes

	Fatigue (CFQ)	Physical function (SF-36 PF)
Strict intention-to-treat analysis		
n	211	211
Mean difference compared with control group (95% CI)	-4.0 (-5.9 to -2.1)	6.3 (1.8 to 10.9)
p value	<0.0001	0.006
Therapist effect		
n	199	199
Mean difference compared with control group (95% CI)	-4.1 (-5.4 to -2.8)	6.3 (2.7 to 9.9)
p value	<0.0001	0.001
Met CDC criteria		
n	138	141
Mean difference compared with control group (95% CI)	-4.1 (-6.5 to -1.7)	6.3 (1.1 to 11.6)
p value	0.001	0.019
Met Oxford criteria		
n	156	159
Mean difference compared with control group (95% CI)	-3.5 (-5.7 to -1.3)	5.6 (0.8 to 10.4)
p value	0.002	0.024
Adjustment for covariates*		
n	199	199
Mean difference compared with control group (95% CI)	-4.3 (-6.3 to -2.4)	6.9 (2.2 to 11.6)
p value	<0.0001	0.004

CFQ=Chalder Fatigue Questionnaire. SF-36 PF=Short-Form 36 physical function subscale. CDC=US Centers For Disease Control and Prevention. *Adjusted for age, sex, ethnic origin, study centre, education level, employment status, high SF-36 PF (≥ 45), high depression (≥ 11), high anxiety (≥ 11), and baseline International Physical Activity Questionnaire Scores (low, medium, or high).

Table 4: Sensitivity analyses

95% CI -19.7 to -1.6 ; $p=0.02$). After adjustment for covariates, patients in the GES group who had worse baseline physical functioning (ie, low SF-36 PF score) had

	GES group (n=97)	Control group (n=101)	Difference between groups
Participants reporting non-serious adverse events	27 (28%)	23 (23%)	$\chi^2=0.67$; $p=0.41$
Serious adverse events	1 (1%)	2 (2%)	..
Serious adverse reactions	0	0	..
Serious deterioration			
Composite*	20 (21%)	30 (30%)	$\chi^2=2.2$; $p=0.14$
Physical functioning reduction*	20 (21%)	25 (25%)	$\chi^2=0.48$; $p=0.49$
Worsened CGI-health	1 (1%)	8 (8%)	$p=0.04^\dagger$
Worsened CGI-CFS	0	9 (9%)	$p=0.003^\dagger$
Withdrawal due to worsening	0	0	..

Data are number of participants reporting these events. Data are n (%) unless otherwise stated. No unexpected serious adverse reactions were suspected. GES=guided graded exercise self-help. CFS=chronic fatigue syndrome. CGI=clinical global impression. SF-36 PF=Short-Form 36 physical function subscale. *Reduction in SF-36 PF score by ≥ 10 points, scores of "much worse" or "very much worse" on participant-rated CGI-health, or active withdrawal due to worsening. † Fisher's exact test.

Table 5: Safety outcomes

	GES group (n=97)	Control group (n=101)
Overall health		
Positive change	17 (18%)	5 (5%)
Minimum change	79 (81%)	88 (87%)
Negative change	1 (1%)	8 (8%)
Odds ratio* (95% CI)	4.8 (1.9-12.4)	..
p value	0.001	..
Chronic fatigue syndrome		
Positive change	14 (14%)	6 (6%)
Minimum change	83 (86%)	86 (85%)
Negative change	0	9 (9%)
Odds ratio* (95% CI)	4.4 (1.7-12.2)	..
p value	0.002	..

Positive change was defined as "much better" or "very much better"; minimal change was defined as "a little better", "a little worse", or "no change"; and negative change was defined as "much worse" or "very much worse". GES=guided graded exercise self-help. CGI=clinical global impression. *Positive change vs negative and minimum changes, across trial groups (adjusted).

Table 6: Participant-rated change in CGI from baseline

a higher mean score at follow-up (56.9, SE 3.3) than their counterparts in the control group (44.2, 3.3). Baseline physical function score also predicted the effect of the study group on both CGI-health score (odds ratio 0.1, 95% CI 0.02-0.64; $p=0.01$) and CGI-CFS score (0.1, 0.02-0.63; $p=0.01$). Participants in the GES group with poor baseline physical functioning were more likely to report improvement at follow-up (eight [20%] reported improvement in CGI-CFS and nine [23%] in CGI-health) whereas this was not observed in their counterparts in the control group (zero patients for both CGI-CFS and CGI-health). Physical functioning did not moderate the effect of study group on fatigue (CFQ) as an outcome (mean difference -3.3, 95% CI -7.3 to 0.7; $p=0.1$). Depression did not moderate the effect of study group on the main

outcomes (SF36: -7.1 [95% CI -16.7 to 2.5], $p=0.2$; CFQ: -3.3 [-7.3 to 0.72], $p=0.1$; CGI-health: 2.7 [0.5 to 14.3], $p=0.2$; and CGI-CFS: 1.9 [0.4 to 10.4], $p=0.5$).

Three participants in the GES group did not attend any of the four guidance sessions, and were excluded in the per-protocol analysis. Exclusion of these participants did not alter the results for SF-36 PF (mean difference 6.5, 95% CI 2.0 to 11.0; $p=0.005$) or CFS (-4.1, -6.0 to -2.2; $p<0.0001$). No patient in the control group reported using the GES guide at the 12 week follow-up.

Participants in the control group were no more likely to receive SMC sessions during the trial than those in the GES group. In the GES group, 28 (26%) of 107 patients attended one session, two (2%) attended two sessions, and none attended three sessions; in the control group, 28 (27%) of 104 patients attended one session, two (2%) attended two sessions, and two (2%) attended three sessions ($\chi^2 2.126$; $p=0.54$).

Physiotherapist-rated data on adherence to GES were available for 104 (97%) participants offered GES; the remainder did not attend enough therapy sessions to be rated. The physiotherapists reported that 43 participants (42%) adhered to GES completely or very well, 31 (30%) moderately well, and 30 (29%) slightly or not at all.

Discussion

When added to SMC, GES significantly improved fatigue and physical functioning compared with SMC alone, and our findings held for patients meeting different diagnostic criteria for chronic fatigue syndrome. Significantly more participants exceeded predefined clinically meaningful changes for fatigue, physical functioning, and both after GES plus SMC than after SMC. In the GES group, a similar proportion of participants improved by a clinically meaningful amount on both primary outcomes (34%) and scored themselves in the high range of physical activity (IPAQ; 30%) at follow-up, which provides some support for these thresholds. The greatest improvements in physical functioning after GES occurred in those with more physical disability. The mean differences in outcome between study groups exceeded a clinically meaningful threshold for fatigue but not for physical functioning. Since 34% (33 of 97) of participants in the GES group and 14% (14 of 102) in the control group improved on both primary outcomes, five patients would need to be treated for one to benefit from GES. Using a higher threshold of improvement, 18% (17 of 97) rated their overall health as "much better" or "very much better" after GES. All but one secondary outcome (PHQ) showed similarly significant differences. Our sensitivity analyses supported the main findings, indicating that the results are robust. Of participants who remained in follow-up, most (85%) were satisfied with the support they had received with GES.

We were unable to compare two measures of the safety of the interventions because too few serious adverse events occurred, and no serious adverse reactions were reported in either study group. Significantly more

participants were lost to follow-up in the GES group than in the control group; a sensitivity analysis, assuming participants who withdrew would report worsening, found no significant difference in worsening between study groups. Two participants actively withdrew from GES, but neither reported their withdrawal as being due to the intervention causing them harm. The safety outcomes in this trial were consistent with those of physiotherapist-delivered GET⁶—namely, no differences between GES and the control interventions.

GETSET is the first trial of GES for chronic fatigue syndrome to be published, but results from other trials of self-help approaches based on cognitive behavioural interventions suggest that behaviour-based self-help approaches might be helpful.^{10,11} In one of these studies,¹⁰ the difference in SF-36 PF between intervention groups (5.7 points) was similar to our finding (6.3 points), and results from another study¹¹ showed no difference in functioning but that minimal CBT-based self-management significantly decreased fatigue compared with patients placed on a waiting list. As expected, a meta-analysis⁶ of therapist-delivered, more intensive GET showed larger effects, with a 13.1 point difference in SF-36 PF at the end of the intervention compared with comparison interventions. In our trial, the difference in CFQ between groups (4.2 points) was at the upper limit of the 95% CI of the mean difference reported in the Cochrane review⁶ of therapist-delivered GET (2.82 points, 95% CI 1.57–4.07), whereas the difference in SF-36 PF (6.3 points) was lower than that in the Cochrane review (13.10 points, 1.98–24.22).

To our knowledge, GETSET is the largest trial of self-help interventions for chronic fatigue syndrome, and the findings are strengthened by the small number of dropouts (12 [6%] of 211) from trial follow-up. We were able to offer GES sessions face to face and by telephone and Skype; remote guidance was popular with participants, mitigating the fatiguing effects of travel for the intervention. Primary outcomes were self-reported by participants, thus preventing observer bias. Adverse events were categorised by independent scrutineers, and the safety outcomes were self-rated. Our finding that GES was more useful in those with worse physical functioning is reassuring and has been reported previously,¹¹ but further exploration is necessary because it might be related to a ceiling effect in those with good physical functioning at baseline. This ceiling effect might also explain the relatively smaller difference in the effect size for physical function, which would reduce the overall difference between study groups.

Our trial has limitations. The results apply only to patients with chronic fatigue syndrome who are diagnosed in secondary care and referred for therapy. SMC was not standardised and we did not record its content. More participants were lost to follow-up in the GES group, although results of the sensitivity analysis were reassuring because serious deterioration was not different between

	GES group (n=97)	Control group (n=102)
Work and social adjustment scale		
Completed questionnaires at 12 weeks	97 (91%)	100 (96%)
Mean score at baseline	26.0 (7.48)	26.4 (7.0)
Mean score at 12 weeks	23.4 (8.6)	25.4 (8.3)
Mean difference at 12 weeks*	-1.9 (-3.7 to -0.2)	..
p value	0.033	..
HADS depression scale		
Completed questionnaires at 12 weeks	97 (91%)	101 (97%)
Mean score at baseline	9.0 (3.9)	8.8 (4.1)
Mean score at 12 weeks	7.6 (4.0)	8.6 (4.3)
Mean difference at 12 weeks*	-1.2 (-1.9 to -0.4)	..
p value	0.002	..
HADS anxiety scale		
Completed questionnaires at 12 weeks	97 (91%)	101 (97%)
Mean score at baseline	8.6 (4.7)	8.7 (4.7)
Mean score at 12 weeks	7.4 (4.3)	8.6 (4.7)
Mean difference at 12 weeks*	-1.1 (-2.0 to -0.3)	..
p value	0.006	..
PHQ-13		
Completed questionnaires at 12 weeks	97 (91%)	102 (98%)
Mean score at baseline	11.8 (4.1)	12.0 (4.3)
Mean score at 12 weeks	11.1 (4.3)	12.1 (4.4)
Mean difference at 12 weeks*	-0.9 (-1.8 to 0.1)	..
p value	0.07	..
IPAQ		
Completed questionnaires at 12 weeks	97 (91%)	99 (95%)
Baseline result		
Low	62 (64%)	49 (50%)
Moderate	32 (33%)	31 (31%)
High	3 (3%)	19 (19%)
12 week result		
Low	33 (34%)	46 (47%)
Moderate	35 (36%)	33 (33%)
High	29 (30%)	20 (20%)
Odds ratio† (95% CI)	3.2 (1.8 to 5.8)	..
p value	<0.0001	..
Data are n (%), mean (SD), or mean difference compared with the control group (95% CI) unless otherwise stated. GES=guided graded exercise self-help. HADS=Hospital Anxiety And Depression Scale. PHQ-13=Patient Health Questionnaire-13. IPAQ=International Physical Activity Questionnaire. *Comparisons were from the final adjusted models including study centre and stratification factors. †High vs moderate and low combined.		

Table 7: Secondary outcomes

the groups. We had only two physiotherapist clusters, and so estimates of the intraclass correlation coefficient might be inaccurate. 27 (25%) of 107 participants in the GES group received three or fewer sessions of guidance, which might reflect disengagement with the intervention; our

qualitative study will address this. All outcomes were self-rated, which might lead to bias by expectation, although the effects of such self-rating are uncertain because of the mixed perception of GET.³⁰ We did not measure any objective outcomes, such as actigraphy, which might have tested the validity of our self-rated measures of physical activity. The relative absence of intervention by diagnostic subgroup interactions might be related to the study not being powered to detect all interactions, including some that might have been clinically important. The clinical diagnosis of chronic fatigue syndrome is not straightforward, but we believe that our subgroup findings of similar efficacy in those meeting the Oxford and CDC criteria provide confidence that our findings are generalisable. It is important to note that this trial was not designed to test causative factors in chronic fatigue syndrome, and the relative efficacy of a behavioural intervention does not imply that chronic fatigue syndrome is caused by psychological factors.

We suggest that these findings show that a guided self-help intervention, when added to SMC, is a moderately effective intervention for fatigue, but has less effect on physical functioning, for people with chronic fatigue syndrome waiting for clinic therapy. These findings need confirmation in other health-care settings and testing of delivery by other health-care professionals.

Contributors

PDW was the principal investigator. LVC was the trial manager. PDW and LVC conceived the study and participated in trial design and coordination. MB entered and checked the data. LVC, PDW, and FP drafted the report. JMT and FP designed the statistical method and FP did the analyses. MV-W participated in trial design and coordination as centre leader. All authors contributed to the final report.

Declaration of interests

PDW reports grants from the UK National Institute of Health Research and the Sue Estermann Fund during the conduct of the study, and personal fees from Swiss Re-insurance company outside the submitted work; he is an appointed member of the Independent Medical Experts' Group, a non-governmental body, which advises the UK Ministry of Defence about its Armed Forces Compensation Scheme. PDW provided unpaid advice to the UK Department for Work and Pensions until 2015. All other authors declare no competing interests.

Acknowledgments

This paper presents independent research funded by the UK National Institute for Health Research (NIHR) under its Research for Patient Benefit (RfPB) Programme (grant reference number PB-PG-0610-22060). The views expressed are those of the authors and not necessarily those of the UK National Health Service (NHS), the NIHR, or the UK Department of Health. This study was also funded by the Sue Estermann Fund, a donor advised fund managed by The London Community Foundation. This study was supported by the UK Clinical Research Collaboration-registered King's Clinical Trials Unit at King's Health Partners, which is part funded by the NIHR Biomedical Research Centre for Mental Health at South London and Maudsley NHS Foundation Trust, King's College London, and the NIHR Evaluation, Trials and Studies Coordinating Centre. The members of our Trial Steering Committee were Christopher Williams (Chair), a patient charity representative, Alastair Miller, Rona Moss-Morris, and a patient representative. The members of our Data Monitoring and Ethics Committee were Astrid Fletcher (Chair), Charlotte Feinmann, and Irwin Nazareth. Hiroko Akagi and Vikki McKeever scrutinised our adverse event data. The physiotherapists were Jenny McClure and Emily Tims. We thank the referring clinicians at Barts Health NHS Trust, East London NHS Foundation Trust, and Kent and Medway NHS and Social Care Partnership

Trust: Julius Bourke, Maurice Clancy, Wendy Hedgecock, Elizabeth Jones, Maurice Murphy, Ewa Okon-Rocha, Areti Pavlidou, Kris Ray, and Elisabeth Wilmetts.

References

- Prins JB, van der Meer JW, Bleijenberg G. Chronic fatigue syndrome. *Lancet* 2006; **367**: 346–55.
- Cairns R, Hotopf M. A systematic review describing the prognosis of chronic fatigue syndrome. *Occup Med (Lond)* 2005; **55**: 20–31.
- Clinical Guideline [CG53]. Chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy): diagnosis and management chronic fatigue syndrome. London: National Institute for Health and Care Excellence. <http://guidance.nice.org.uk/CG53> (accessed July 7, 2015).
- Action for ME. ME 2008: what progress? Action for ME, 2008.
- ME Association. Managing my M.E. What people with ME/chronic fatigue syndrome and their carers want from the UK's health and social services. <http://www.meassociation.org.uk/2010/05/managing-my-me-me-association-publish-results-of-huge-survey-report/> (accessed July 7, 2015).
- Larun L, Brurberg KG, Odgaard-Jensen J, et al. Exercise therapy for chronic fatigue syndrome. *Cochrane Database Syst Rev* 2016; **2**: CD003200.
- White PD, Goldsmith KA, Johnson AL, et al. Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial. *Lancet* 2011; **377**: 823–36.
- Collin SM, Sterne JA, Hollingworth W, et al. Equity of access to specialist chronic fatigue syndrome (chronic fatigue syndrome/ME) services in England (2008–2010): a national survey and cross-sectional study. *BMJ Open* 2012; **2**: e001417.
- Chalder T, Wallace P, Wessely S. Self-help treatment of chronic fatigue in the community: a randomized trial. *Br J Health Psychol* 1997; **2**: 189–97.
- Knoop H, van der Meer JWM, Bleijenberg G. Guided self-instructions for people with chronic fatigue syndrome. *Br J Psychiatry* 2008; **193**: 340–41.
- Tummers M, Knoop H, van Dam A, Bleijenberg G. Implementing a minimal intervention for chronic fatigue syndrome in a mental health centre: a randomised controlled trial. *Psychol Med* 2012; **42**: 2205–15.
- Clark LV, McCrone P, Ridge D, et al. Graded exercise therapy guided self-help trial for patients with chronic fatigue syndrome (GETSET): protocol for a randomized controlled trial and interview study. *JMIR Res Protoc* 2016; **5**: e70.
- Reeves WC, Lloyd A, Vernon SD, et al. Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. *BMC Health Serv Res* 2003; **3**: 2.
- Sharpe MC, Archard LC, Banatvala JE, et al. A report—chronic fatigue syndrome. *J R Soc Med* 1991; **84**: 118–21.
- First MB, Spitzer RL, Gibbon M, et al. Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition with psychotic screen (SCID-I/P W/ PSY SCREEN). New York: Biometrics Research, New York State Psychiatric Institute, 2002.
- Zigmond A, Snaith R. The hospital anxiety and depression scale. *Act Psychiatr Scand* 1983; **87**: 361–70.
- McHorney CA, Ware JE, Raczek AE. The MOS 36 item short form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993; **31**: 247–53.
- Clark LV, ed. Graded exercise therapy: a self-help guide for those with chronic fatigue syndrome/myalgic encephalomyelitis. London: Bart's and the London NHS Trust, 2011. <http://www.wolfson.qmul.ac.uk/images/pdfs/getset/GET%20guide%20booklet%20version%201%2022062010.pdf> (accessed June 13, 2017).
- Chalder T, Berelowitz G, Hirsch S, et al. Development of a fatigue scale. *J Psychosom Res* 1993; **37**: 147–53.
- Guy W. ECDEU assessment manual for psychopharmacology. Rockville, MD: National Institute of Mental Health 1976: 218–22.
- Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* 2002; **64**: 258–66.

- 22 The EuroQOL. The EuroQOL—a new facility for the measurement of health-related quality of life. *Health Policy* 1990; **16**: 199–208.
- 23 Mundt JC, Marks IM, Shear K, et al. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *Br J Psychiatry* 2002; **180**: 461–64.
- 24 Craig CL, Marshall AL, Sjoström M, et al. International Physical Activity Questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003; **35**: 1381–95.
- 25 The IPAQ group. IPAQ scoring protocol. International Physical Activity Questionnaire. <https://sites.google.com/site/theipaq/scoring-protocol> (accessed June 13, 2017).
- 26 Johnson DR, Young R. Towards best practices in analysing datasets with missing data: comparisons and recommendations. *J Marriage Fam* 2011; **73**: 926–45.
- 27 Baldwin SA, Bauer DJ, Stice E, Rohde P. Evaluating models for partially clustered designs. *Psychol Methods* 2011; **16**: 149–65.
- 28 Committee for Proprietary Medicinal Products (CPMP). Points to consider on adjustment for baseline covariates. London: European Agency for the Evaluation of Medicinal Products, 2003. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003639.pdf (accessed June 13, 2017).
- 29 Williams R. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata J* 2006; **6**: 58–82.
- 30 M.E. Time to deliver: Initial findings of Action for M.E.'s 2014 survey. Bristol: @ Action for ME, 2014. <https://www.actionforme.org.uk/uploads/pdfs/me-time-to-deliver-survey-report.pdf> (accessed June 13, 2017).