Psychological Medicine, 2004, **34**, 37–49. © 2004 Cambridge University Press
DOI: 10.1017/S0033291703001247 Printed in the United Kingdom

Is graded exercise better than cognitive behaviour therapy for fatigue? A UK randomized trial in primary care

L. RIDSDALE,[1] L. DARBISHIRE AND P. T. SEED

From the Departments of Neurology and General Practice, Guy's, King's and St Thomas's School of Medicine, London

ABSTRACT

Background. Patients frequently present with unexplained fatigue in primary care, but there have been few treatment trials in this context. We aimed to test cognitive behaviour therapy (CBT) and graded exercise therapy (GET) for patients presenting to their family doctor with fatigue. Secondly, we described the outcome for a cohort of patients who presented to the same doctors with fatigue, who received standard care, plus a booklet.

Method. This was a randomized trial, followed by a prospective cohort study. Twenty-two practices in SE England referred 144 patients aged 16 to 75 years with over 3 months of unexplained fatigue. Self-rated fatigue score, the hospital anxiety and depression rating scale, functional impairment, physical step-test performance and causal attributions were measured. In the trial six sessions of CBT or GET were randomly allocated.

Results. In the therapy groups the mean fatigue score decreased by 10 points (95% confidence interval (CI) = −25 to −15), with no significant difference between groups (mean difference = −1·3; CI = −3·9 to 1·3). Fewer patients attended for GET. At outcome one-half of patients had clinically important fatigue in both randomized groups, but patients in the group offered CBT were less anxious. Twenty-seven per cent of the patients met criteria for CFS at baseline. Only 25% of this subgroup recovered, compared to 60% of the subgroup that did not meet criteria for CFS.

Conclusions. Short courses of GET were not superior to CBT for patients consulting with fatigue of over 3 months in primary care. CBT was easier 'to sell'. Low recovery in the CFS subgroup suggests that brief treatment is too short.

INTRODUCTION

Interventions have been tested mainly for patients with unexplained fatigue in hospital specialist clinics. Most have been diagnosed with chronic fatigue syndrome (CFS). To be given this diagnosis patients must have fatigue of >6 months duration, that is disabling, and have additional symptoms such as headaches, myalgia and sleep difficulties (Fukuda *et al.* 1994). The treatments tested range from intravenous immunoglobulin therapy (Lloyd *et al.* 1990) to homeopathy (Awdry, 1996). Cognitive behavioural therapy (CBT) and graded exercise therapy (GET) seem particularly to have a beneficial effect (Whiting *et al.* 2001). Medically unexplained fatigue is commonly presented to doctors in primary care, but 75% of patients have fatigue that does not meet the criteria for CFS (Ridsdale *et al.* 2001*a*). Most of these patients are managed by their family doctor, with only 2% referred to specialists (Morrell, 1972). With usual medical care, 70–75% of patients in general practice report that their fatigue symptoms become chronic (Ridsdale *et al.* 1993; Whitehead & Campion, 2002).

[1] Address for correspondence: Dr Leone Ridsdale, Department of General Practice, Guy's, King's and St Thomas's School of Medicine, 5 Lambeth Walk, London SE11 6SP.

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

Despite the obvious difference between patients seen in specialist practice and general practice (Euba *et al.* 1996) and dangers in extrapolating from one context to the other (Raine *et al.* 2002), few treatments have been tested in primary care. One group was randomly offered an interview with a research nurse plus a booklet on CBT to patients who had been identified with fatigue in primary care, but who had not necessarily consulted their doctor for the problem (Chalder *et al.* 1997). At follow-up, 63% of the intervention group were no longer fatigued compared with 39% in the control group. Our group identified patients who presented to their doctor with >3 months of fatigue, and compared the effect of the offer of six sessions of counselling or CBT, provided by trained therapists. At outcome we found that 48% were no longer fatigued in the counselling group, compared to 47% in the CBT group (Ridsdale *et al.* 2001*a*). Whitehead & Campion (2002) offered to train family doctors to provide brief CBT for patients presenting with fatigue, but few doctors completed their training. One year later they found that when patients presented with >3 months of fatigue, 21% were no longer fatigued in the group whose GP was offered training, and 22% were no longer fatigued in the comparison group.

The Joint Royal Colleges (1996) and a UK Department of Health Report have proposed that treatments for patients with CFS should be tested and provided in primary care (Report of the Working Party on CFS/ME, 2002). They proposed that early diagnosis and management might reduce the burden of disability and referral to specialists. A systematic review had found that long courses (15 to 17 sessions) of CBT and GET are likely to be beneficial for CFS in specialist clinics (Whiting *et al.* 2001). However, developing and testing complex interventions is challenging (Campbell *et al.* 2000), and the review found there was little consensus on outcome measures used, or what constituted a clinically important effect size (Whiting *et al.* 2001).

There is little evidence-base to guide doctors on the management of patients in primary care; no standard treatment is provided and treatment is variable (Fitzgibbon *et al.* 1997). This is unfortunate for patients whose symptoms tend to become chronic. Lack of tested treatments in primary care may create greater demand, with long waits for specialist services. Outside specialist centres there are therapists available to provide graded exercise, and in UK primary care there is more access to physiotherapy than there is to therapists able to provide CBT.

In prior studies of patients in general practice presenting with fatigue we found that one-half to two-thirds believed their symptoms were physical (Ridsdale *et al.* 1994, 2001*a*). On the basis of our prior trial it was postulated that patients with physical attributions may respond better to physical therapy (Chalder *et al.* 2003). We reasoned that if a GET intervention for patients with fatigue was feasible and similar in its effect to CBT, more treatment options could be made available in primary care. Also we reasoned that as most patients with fatigue in primary care do not have CFS and do not get referred to specialists, their fatigue might respond to shorter courses of treatment.

Our main aim was to compare a short treatment of GET and the previously tested CBT, in primary care, by means of a randomized trial. We were not funded for a third arm to compare this with usual care. Following trial completion, we continued recruiting patients prospectively to a cohort group. This group received usual medical care, plus a booklet (Chalder, 1995) shown to have a better effect than usual medical care (Chalder *et al.* 1997). We present the outcome from this non-randomized, but prospectively recruited cohort group.

METHOD

Subjects

Patients aged 16 to 75 years attending their general practitioner (GP) with a main complaint of unexplained fatigue of >3 months duration were recruited from 22 general practices in London and South East England between January 1999 and June 2001. Total population size was 174 000 patients aged 16 to 75 years. The trial had multi-centre and local ethics committee approval.

Design – randomized trial

The main study was a phase II randomized trial in which individual CBT was compared with individual GET. Patients who met the trial eligibility inclusion criteria were asked to give

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

informed written consent. The inclusion criteria were: aged 16 to 75 years; complains of fatigue as a main or important problem; duration of fatigue symptoms for $\geqslant 3$ months; no recent change in drug regimen; normal full blood count, erythrocyte sedimentation rate and thyroid function test on entry or in the previous 6 months. The exclusion criteria were: patient unable to read English; concurrent physical problems, which in the judgement of the doctor have caused the fatigue symptoms; patient has asthma and/or ischaemic heart disease that would contraindicate a physical step-test; psychotic illness, organic brain syndrome, or substance dependency; current treatment from a psychiatrist, psychologist, community psychiatric nurse, physiotherapist, or exercise therapist. The trial statistician (P.S.) carried out stratified and blocked (size ten) randomization to four sets of random combinations. Results from a previous trial informed stratification by fatigue duration (more or less than 4 years) and fatigue severity (more or less than 25 on a fatigue scale) (Ridsdale *et al.* 2001*a*). Allocation was concealed in four series of consecutive, opaque, sealed envelopes and assigned by the research co-ordinator (L.D.) in the presence of the patient, following baseline assessment. Six cognitive behaviour therapists and five physiotherapists provided the active treatments, and were supervised. Therapists followed treatment manuals based on previously published protocols (Chalder, 1995; Fulcher & White, 1998). Treatment offered was six 45-min sessions over 12 weeks, with the first session used to assess and engage with the patient. Treatment was offered on the premises of each patient's doctor in primary care.

Cognitive behaviour therapy (*CBT*)

After an assessment, a rationale for treatment is provided. The treatment involves activity planning, homework, establishing a sleep routine and other cognitive interventions (Chalder *et al.* 1999). It is based on a model that distinguishes between precipitating and perpetuating factors, with the perpetuating factors becoming the focus of the intervention. The treatment ensures levels of activity and rest are both consistent and realistic given the patients' responsibilities. Sleep disturbance and negative beliefs regarding the symptom of fatigue, self-expectations or self-esteem are identified and patients are encouraged to challenge them in the conventional way. Specific lifestyle changes are encouraged if deemed appropriate and relapse prevention is addressed in the last two sessions.

Graded exercise therapy (*GET*)

Treatment is based on the principles of exercise prescription devised by the American College of Sports Medicine (American College of Sports Medicine, 2000), adapted to each patient's current physical capacity. It was developed from a GET protocol designed for patients with chronic fatigue syndrome in a specialist context (Fulcher & White, 1998). GET is structured and supervised activity management that aims for a gradual but progressive increase in aerobic activities, usually walking. Home exercise is programmed, with initial sessions lasting between 5 and 15 min at an intensity of 50% of the age-related estimated maximum heart rate. Patients are advised not to exceed the recommended exercise duration or intensity.

Design – prospective cohort

We were not funded to recruit a third 'control' arm in the trial. Recruitment to the two-arm trial was completed faster than expected. After recruitment was complete a further group was prospectively recruited from the same practices, using the same criteria for eligibility as described previously. This group was given a booklet on self-management of fatigue (Chalder, 1995) and continued with usual medical care. The booklet provides information about fatigue and the factors that contribute to its onset and maintenance, and describes a variety of cognitive and behavioural techniques for management.

Instruments

Assessments were made at baseline and at 3 and 8 months from baseline, with the mean of the two periods used as primary outcome. The primary outcome was fatigue measured with a fatigue scale scored using a Likert system (0, 1, 2, 3) with a maximum score of 33 (Chalder *et al.* 1993). Binary scoring (0, 0, 1, 1) was applied to the data to estimate fatigue 'caseness', with a cut-off of $\geqslant 4$ indicative of clinically significant fatigue. Chronic fatigue syndrome (CFS) status was determined, using the CDC criteria (Fukuda *et al.* 1994), by the research

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

associate (L.D.) at the baseline assessment. Criteria for this diagnosis include fatigue of a definite onset, a minimum duration of 6 months, with substantial functional impairment, and four or more additional symptoms from a list of eight. We measured secondary outcomes including anxiety and depression (HAD) (Zigmond & Snaith, 1983), degree of functional impairment in work, home, private and leisure activities (WASA) (not at all impaired = 0; very severely impaired = 8) (Ware *et al.* 1992) and illness attributions (physical = 1; psychological = 5) (Powell *et al.* 1990). A 1 min, four-count physical step-test was administered at baseline, with patients repeating it following completion of treatment at 3 months, and not at 8 months. Patients were asked whether they had previously been referred to a psychiatrist. The research associate extracted information on prior diagnoses and consultation frequency (measured over 8 months and then extrapolated to 1 year) from general practitioners' records.

Power calculations

Sample size calculation for randomized trial

We knew that other fatigue trialists had not used the same outcome measures, and where they had done so, the data have not been analysed in the same way (Whiting *et al.* 2001). In our previous trial clinician trialists reached a consensus that a difference of four points between one treatment and another at outcome (4/14·4 = 28%) would be clinically important (Ridsdale *et al.* 2001*b*). A smaller difference of $\leqslant$20% has been recommended particularly in drug trials when it is important and feasible to measure small differences between a tested treatment and a new drug with a similar effect (Williams, 1997). Preliminary data from our prior trial, which was in progress, suggested the mean fatigue after treatment was 14·4 (Likert scored) and standard deviation 7·4 (Ridsdale *et al.* 2001*a*). Using the same fatigue scale, with a correlation of 0·7, and one baseline and two outcome measures (Frison & Pocock, 1992), and using ANCOVA, we estimated 50 subjects per group would give a 90% power to detect a difference of 20% (2·88 units).

Sample size calculation for prospective cohort

The calculation for the prospectively recruited cohort was based on final data from the same trial (Ridsdale *et al.* 2001*a*). On the basis of this trial, with a standard deviation (S.D.) of 8·5 and a correlation of 0·5, we estimated that 40 patients in the cohort group and 100 in the combined treatment group would give a power of 90% to detect a four-point difference by ANCOVA, with a baseline and the average of two follow-up measures.

Statistical analysis

We based our analysis on data from all recruited patients when patients had completed measures at 3 and 8 months. Where data were only available at one follow-up, the single observed value replaced the average (five patients had missing data at 3 months and 18 had missing data at 8 months – nine patients had missing data at both 3 and 8 months). As a sensitivity analysis alternative approaches were applied: carrying the last value forward, Heckman regression (Heckman, 1979), and both single and multiple imputation (Little & Rubin, 1987; Greenland & Finkle, 1995). The main effect was the difference between mean scores for the two randomized groups using the Fatigue Scale, Likert scored (Chalder *et al.* 1993).

The main outcomes were analysed using the average of 3 and 8 month scores. For the randomized trial, analysis of outcomes was by analysis of covariance (ANCOVA) adjusted for baseline fatigue, fatigue duration and CFS status. For comparisons between therapy and the prospective cohort ANCOVA was adjusted for: baseline fatigue, fatigue duration, CFS status, gender, a history of anxiety/depression and past psychiatric referral.

We recognized the challenge, highlighted in a systematic review, of comparison and meta-analysis when data are reported differently (Whiting *et al.* 2001), and also report outcomes using criteria applied by other investigators using this fatigue scale. Chalder and colleagues (1993) reported a score of $\geqslant$4, using binary scoring, indicated 'case' levels of fatigue. Cleare and colleagues (1999) induced that a reduction in Likert score of 9 indicated a clinically significant reduction in fatigue, equating to a 30% fall in fatigue score produced by a standard course of CBT (16 weekly sessions in secondary care) (Sharpe *et al.* 1996). The attribution questionnaire was analysed using ordered logistic regression, a non-parametric method using the

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

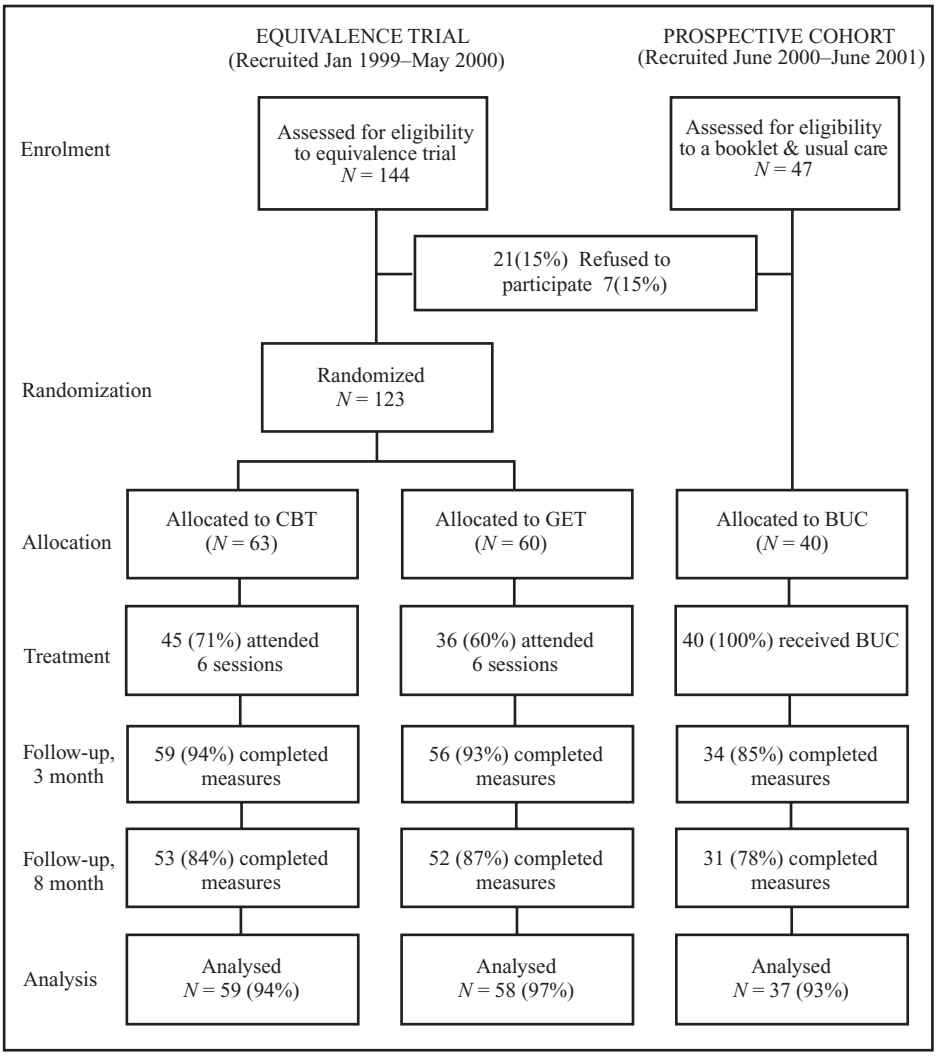


FIG. 1. Trial profile.

order of the values. This method of analysis is based on a category scale and not a binary score and so does not lend itself to the reporting of relative risks.

RESULTS

Baseline

GPs referred 144 patients for potential recruitment into the therapy trial (Fig. 1). All patients met Chalder's case criteria for fatigue (Chalder *et al*. 1993). Twenty-one (15%) refused to participate, and so 123 patients were randomized to active therapy. Patients in the two active therapy groups were similar, as shown in Table 1. Mean age was 40 and 68% were women, 65% were working and their mean fatigue score was 25. The median duration of fatigue was 11 months longer in the CBT group.

Forty-seven patients were referred for potential recruitment into the prospective cohort group. Seven (14%) refused to participate, and so 40 received a booklet and continued with usual care. Comparing the therapy trial groups to the cohort group, patients in the cohort group were significantly less fatigued with a mean

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

Table 1. *Clinical and social information by group: characteristics* (N (%), *unless stated otherwise*) *of the participants*

	Treatment trial		Cohort group, BUC ($N=40$) N (%)
	CBT ($N=63$) N (%)	GET ($N=60$) N (%)	
Age (years), mean (S.D.)	40 (12·3)	40 (10·8)	37 (11·6)
Women	43 (68)	41 (68)	31 (78)
Married/cohabiting	37 (59)	28 (47)	23 (58)
Employed (full or part time)	38 (60)	44 (73)	26 (65)
Fatigue score, mean (S.D.)	25·3 (4·7)	24·7 (5·5)	23·1 (5·2)*
Months of fatigue (duration), median (5–95%)	32 (4 to 224)	21 (3 to 93)	24 (3 to 274)
History of anxiety/depression	38 (60)	34 (57)	19 (48)
Past psychiatric referral	29 (46)	31 (52)	12 (30)†
1-min step-test score, mean (S.D.)	27·3 (8·1)	26·3 (5·1)	29·8 (5·5)‡
Consultations with GP in past 8 months, mean (S.D.)	6·0 (3·2)	6·9 (5·0)	5·6 (3·0)

CBT, Cognitive behavioural therapy; GET, graded exercise therapy; BUC, booklet and usual care.
* A lower score to combined active treatment group ($-4·4$, $-7·8$ to $-1·0$).
† Significantly less patients than in combined active treatment group ($\chi^2=4·3$, $P=0·04$).
‡ A significantly higher score than combined active treatment group (2·9, 0·6 to 5·3, $P=0·01$).

score of 23, and were less likely to have had a past psychiatric referral (30% *v.* 50%).

Baseline comparisons of patients with CFS, and fatigue that did not meet criteria for CFS, have been published (Darbishire *et al.* 2003; McCrone *et al.* 2003). Patients with CFS had higher mean scores on fatigue, functional impairment, anxiety, depression, and had more consultations, unemployment and informal care.

Treatment

The mean number of therapy sessions provided for all patients allocated to therapy was 4·6 (S.D. = 2·2). Patients who did not complete six sessions of therapy were significantly younger than those who did (36·5 *v.* 42·1; mean difference = 5·7; CI = 1·4 to 9·9), with no other significant differences found. Reasons for not completing therapy included: too busy ($N=13$); lack of faith in therapy ($N=8$); and feeling better ($N=4$). More patients (45/63, 71%) in the CBT group completed six sessions than in the GET group (36/60, 60%), (mean difference = 0·8; CI = 0·1 to 1·6). The difference is mainly due to seven patients not starting GET. More patients cited lack of faith in their allocated treatment as a reason for not starting or for stopping GET compared to CBT.

Two patients died before outcome, one from bronchopneumonia and the other committed suicide. The patient who committed suicide had a history of depression, but the depression score at the baseline was not particularly high (12/21). The patient was allocated CBT, and received four sessions. As the patient's mood deteriorated she was referred back to the doctor, and subsequently admitted to hospital under a section order. Death occurred following discharge from hospital.

Outcome

The randomized trial

Data for the intention to treat analysis were available for 117/123 patients, 59/63 in CBT and 58/60 in GET. Outcome scores are shown in Table 2. Mean fatigue scores fell significantly in both groups from baseline to outcome. Mean fatigue scores at outcome combining 3 and 8 month data were 13·9 (S.D. = 7·0) in the CBT group and 15·1 (S.D. = 8·1) in the GET group. After adjusting for baseline scores, the mean difference was $-1·3$ (CI = $-3·9$ to 1·3), with a small difference favouring CBT. Individual fatigue scores increased in 2/59 (3%) patients receiving CBT, and 7/58 (12%) receiving GET.

Data were also analysed 'per-protocol', including only patients who completed six sessions of therapy. Data were available for 80/123 patients, 45/63 in CBT and 36/60 in GET. Mean fatigue scores fell significantly in both groups from baseline to outcome. Mean fatigue scores at outcome combining 3 and 8 month data were 13·9 (S.D. = 6·3) in the CBT group and 15·1

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

Table 2. *Primary and secondary outcome scores**

	CBT Mean (S.D.)	GET Mean (S.D.)	BUC Mean (S.D.)
Fatigue			
Baseline	25·3 (4·7)	24·7 (5·5)	23·1 (5·2)
3 months	13·3 (8·8)	15·1 (9·0)	17·2 (8·7)
8 months	14·8 (7·9)†	15·3 (9·4)	17·9 (8·9)
Differences between groups‡		(CBT − GET) = −1·3; −3·9 to 1·3	
(mean difference; CI)		(Therapies − BUC) = −4·9; −7·7 to −2·2	
HAD − Depression			
Baseline	8·1 (3·2)	8·4 (4·0)	9·5 (3·8)
3 months	5·8 (3·8)	6·6 (4·6)†	9·3 (4·8)
8 months	6·6 (4·3)	6·8 (4·9)	8·5 (4·6)
Differences between groups‡		(CBT − GET) = −0·5; −1·7 to 0·7	
(mean difference; CI)		(Therapies − BUC) = −1·2; −2·4 to −0·1	
HAD − Anxiety			
Baseline	10·4 (4·4)	10·4 (4·7)	9·5 (3·8)
3 months	7·9 (4·4)	9·3 (4·8)†	9·3 (4·8)
8 months	8·5 (4·2)	9·8 (5·3)	8·5 (4·6)
Differences between groups‡		(CBT − GET) = −1·5; −2·6 to −0·4	
(mean difference; CI)		(Therapies − BUC) = −1·4; −2·6 to −0·3	
Functional impairment (WASA)			
Baseline	5·1 (1·8)	4·6 (2·2)	4·5 (1·7)
8 months	4·1 (1·9)	4·1 (2·4)	3·9 (2·1)
Differences between groups‡		(CBT − GET) = −0·3; −0·9 to 0·3	
(mean difference; CI)		(Therapies − BUC) = −0·3; −1·0 to 0·4	
Attributions			
Baseline	2·4 (0·9)	2·6 (1·1)	2·4 (0·9)
3 months	2·6 (0·9)†	2·4 (1·0)	2·6 (0·9)
8 months	2·9 (1·2)	2·3 (1·0)†	2·6 (1·0)
ORs between groups§		(CBT ÷ GET) = 4·0; 1·8 to 8·9	
(ORs; CI)		(Therapies ÷ BUC) = 0·8; 0·4 to 1·4	

* Numbers reported: baseline (CBT = 63, GET = 60, BUC = 40); 3-months follow-up (CBT = 59, GET = 56, BUC = 34); 8-months follow-up (CBT = 54, GET = 52, BUC = 31).

† Data missing for one patient.

‡ Comparison made from the average of 3 and 8 months, adjusting for baseline by ANCOVA.

§ Calculated using ordered logistic regression.

(S.D. = 7·8) in the GET group. After adjusting for baseline scores, the mean difference was −0·3 (CI = −2·7 to 3·3), with a small difference favouring CBT. For the sensitivity analysis the difference in outcome scores between the two therapy groups were calculated using an additional four methods (Heckman, 1979; Little & Rubin, 1987; Greenland & Finkle, 1995). When each of the results was compared with the main outcome, the most extreme difference was 0·3, less than half a scale point on the outcome measure.

Using the Chalder *et al.* (1993) case criteria for fatigue, the proportion who recovered was 30/59 (51%) following CBT, and 28/58 (48%) following GET. Using the Cleare *et al.* (1999) criteria, the proportion of patients who recovered was CBT 30/59 (51%) and GET 33/58 (57%).

The prospective cohort

Full data were available for 31/40 patients who received a booklet plus usual care (BUC). Mean fatigue scores at outcome combining 3 and 8 month data were 17·9 (S.D. = 8·1). Using the Chalder *et al.* (1993) case criteria for fatigue, the proportion that recovered was 14/37 (38%). Using the Cleare *et al.* (1999) criteria, the proportion of patients who recovered was 13/37 (35%). After adjusting for baseline scores, there was a significant difference between groups receiving therapy and a booklet plus usual care, favouring active therapy (mean difference = −4·9; CI = −7·7 to −2·2). Individual fatigue

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

scores increased in 10/37 (27%) of patients receiving BUC.

Secondary outcome measures

The randomized trial

Following treatment depression, anxiety, and functional impairment scores fell in both CBT and GET groups. At outcome anxiety scores were significantly lower in the CBT group compared to the group receiving GET (mean difference = −1·3; CI = −2·4 to −0·3) (Table 2). There were no other significant differences between groups at outcome. The frequency of consulting their doctor per year declined following both CBT (mean change in consultations = −2·2; CI = −3·3 to −1·1) and GET (mean change in consultations = −3·0; CI = −4·3 to −1·7), and there was no significant difference between groups at outcome (mean difference = −0·1; CI = −1·54 to 1·32).

The prospective cohort

Between baseline and outcome depression, anxiety and functional impairment scores did not change significantly in the BUC group. Depression, anxiety and functional impairment scores were significantly higher at outcome following BUC than following therapy (Table 2). The frequency of consulting their doctor per year declined following BUC (mean change in consultations = −1·3; CI = −2·0 to −0·5) and there was no significant difference between BUC and therapy groups at outcome (mean difference = −1·0; CI = −2·3 to 0·2).

The relationship between therapy and CFS status

On entry to the trial all patients met Chalder's case criteria for fatigue (Chalder *et al*. 1993), but only 36 (29%) patients conformed to CDC criteria (Fukuda *et al*. 1994) for CFS (CBT, $N = 15$; GET, $N = 21$). In the group meeting criteria for CFS, median fatigue duration at baseline was 30 months (5% to 95% = 6 to 109), mean fatigue score was 28·24 (S.D. = 4·54), 17 (47%) attributed their fatigue to physical causes and 6 (17%) were members of a self-help group. In the group that did not meet criteria for CFS, median fatigue duration at baseline was 24 months (5% to 95% = 3 to 161), mean fatigue score was 23·7 (S.D. = 4·7), 39 (45%) attributed their fatigue to physical causes and none were members of a self-help group.

In the group allocated to GET, patients meeting criteria for CFS attended significantly more sessions of therapy (5·2, S.D. = 1·6) compared with patients not meeting criteria for CFS (3·6, S.D. = 2·6; mean difference = 1·7, CI = 0·4 to 2·9). In the group allocated to CBT, patients meeting criteria for CFS attended more sessions of therapy (5·5, S.D. = 1·2) compared with patients not meeting criteria for CFS (4·8, S.D. = 2·0), but in this instance the difference was not significant (mean difference = 0·7, CI = −0·4 to 1·7).

For patients meeting criteria for CFS at baseline, fatigue scores decreased by 26% to 19·70 (S.D. = 7·40) (CBT, 17·53, S.D. = 6·78; GET, 20·05, S.D. = 7·90; mean difference = −4·29, CI = −9·4 to 0·8). For patients not meeting criteria for CFS at baseline, fatigue scores were decreased by 34% to 12·61 (S.D. = 6·81) (CBT, 12·60, S.D. = 6·64; GET, 12·61, S.D. = 7·11; mean difference = −0·4, CI = −3·4 to 2·6). Fatigue scores were higher at both baseline and outcome in the group meeting criteria for CFS compared to group not meeting criteria for CFS. The effect of treatment (CBT *v*. GET) was not significantly different in the two groups (CFS *v*. non-CFS) (Interaction test, adjusting for baseline fatigue and fatigue duration: $F(1, 110) = 0{\cdot}43$; $P = 0{\cdot}52$). The interaction is shown in Fig. 2.

At outcome, 25% ($N = 9/36$) of patients meeting criteria for CFS at baseline were no longer cases of fatigue (CBT = 5/15 (33%); GET = 4/21 (19%)). At outcome 60% ($N = 49/81$) of patients who did not meet criteria for CFS at baseline were no longer cases (CBT = 25/44 (57%); GET = 24/37 (65%)).

DISCUSSION

Limitations

No randomized comparison group

Our first trial was originally designed as a control trial of CBT (Ridsdale *et al*. 2001*a*). This protocol was changed before it was funded because our funder's referee argued that psychological interventions have a placebo response, and we needed to compare CBT with an active intervention providing an equivalent amount of

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

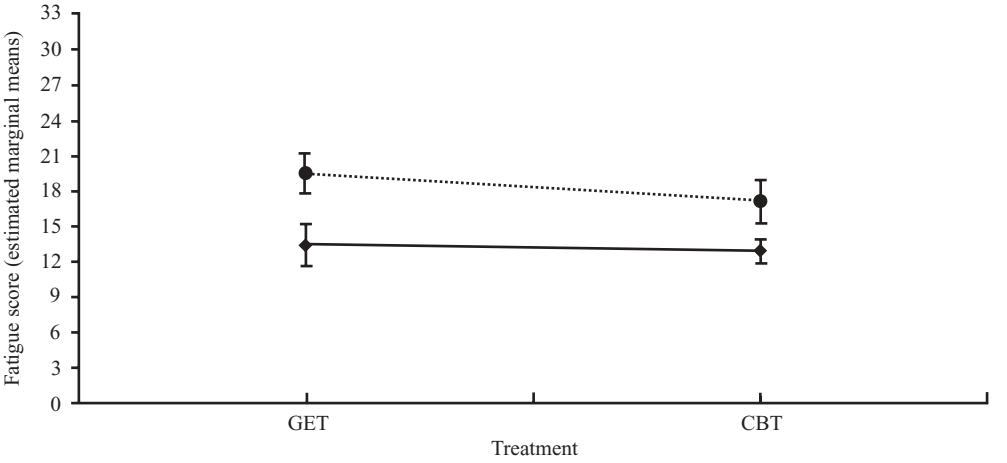


FIG. 2. Fatigue scores at outcome broken down by treatment group and CFS status (adjusted means and standard error bars). (●----●, CFS; ◆—◆, non-CFS.)

care. We introduced counselling as our comparison group because it was widely available in general practice, and because there was little evidence that it was effective (Corney, 1990). The outcome of the trial was that there was no significant difference between counselling and CBT (Ridsdale *et al.* 2001*a*). We agreed with our critics, Underwood & Eldridge (2001), that we could not demonstrate, except by use of historical comparison groups, that either therapy was better than nothing in the management of fatigue in primary care.

The trial reported here started before the end of the first trial (Ridsdale *et al.* 2001*a*). Once again we were not funded for a three-arm trial, which might have allowed us to compare two active treatments with usual care. We completed the two-arm trial of active therapies sooner than expected, and recognized that we could not compare their effect with usual care. In this context we decided to recruit from the same practices using the same criteria, a cohort group who were given a booklet plus usual care. We acknowledge this evidence is weaker than that derived from the randomized trial, as we cannot adjust for a different time-frame. We have reported the findings because the data are likely to be more similar than evidence from historical comparison groups. Evidence from historical comparison groups suggest that after usual care 24 to 31% of patients recover (Ridsdale *et al.* 1993; Ridsdale *et al.* 2001*a*; Whitehead & Campion, 2002). The result from our usual care plus booklet group is similar to this.

Lack of power

Whiting and colleagues (2001) drew attention to the challenge of systematically reviewing evidence from interventions for fatigue when trialists have not agreed on what should be the primary outcome, or how data from a commonly used measure should be reported. We used a fatigue measure, but using this measure the trial was under-powered to demonstrate equivalence of CBT and GET if a 20% difference is considered important. A *post hoc* power calculation is given here for clarification. The standard error of the treatment effect is 1·32 (CBT *v.* GET). The sample size calculation was based on a possible effect of 2·88 units, amounting (*post hoc*) to 2·17 standard errors. For a difference at the 5% level, the Z score is $2{\cdot}17-1{\cdot}96=0{\cdot}211$; giving a power of 56%. Using an effect size of 4 units (3·02 times the standard error), the *post hoc* power is 85%. So the study is adequately powered to answer the question regarding an effect of four points (and indeed has shown equivalence at this level), but further evidence would be needed to address the possibility of a difference of 2·88 units or less. This is illustrated in Fig. 3, which builds on Jones and colleagues (1996) excellent paper on equivalence trials. It is difficult to estimate the effect of complex

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

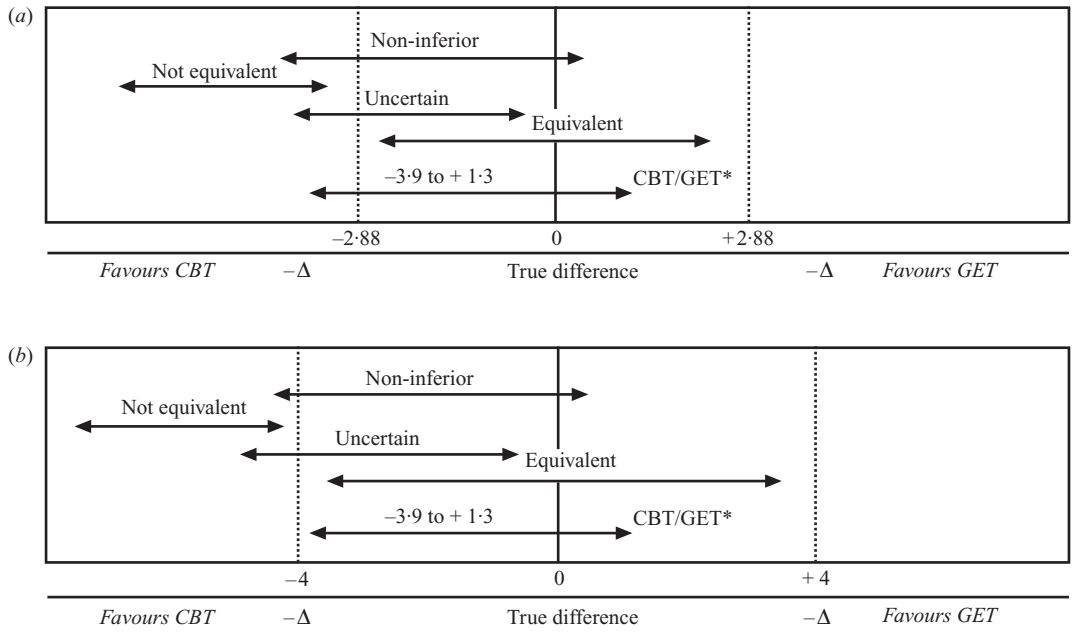


FIG. 3. Treatment effect: (*a*) assuming equivalence at 2·88 points; (*b*) assuming equivalence at 4 points. (* Based on (*a*) our results show non-superiority of GET compared with CBT; and based on (*b*) our results show equivalence.)

interventions, agree on outcomes, agree on the rating of existing measures and estimate important effect sizes. We submit that 20% was an ambitious difference for an exploratory trial of a complex intervention, and if a 20% difference is important, we have shown only that CBT is not inferior to GET.

Different treatments, similar outcomes

There is a consensus that we should test treatments for CFS in primary care (Joint Royal Colleges, 1996; Raine *et al.* 2002). Experts and patients also advocate the investigation of interventions for patients whose fatigue is of lesser duration and severity than CFS, in order to prevent disability (Report of the Working Party on CF/ME, 2002). In the trials that have shown significant differences, treatment has been compared with usual (i.e. less) care or with a placebo (i.e. relaxation) provided by a therapist trained in the active treatment, CBT or GET (Whiting *et al.* 2001). This is the second randomized trial of two active treatments, CBT *versus* GET. Our evidence supports the hypothesis that, for unexplained fatigue of $\geqslant$3 months in primary care, assuming a difference of 20% (2·88) is important, GET is not superior to CBT (Fig. 3*a*). If a difference of 4 points is clinically important, the effect of GET is equivalent to CBT (Fig. 3*b*). Our evidence also supports the general conjecture that the efficacy of bona fide treatments is roughly equivalent (Wampold *et al.* 1997).

Therapies may have different effects on attributions, but similar effects on fatigue

In primary care, at baseline, 50% to 69% of patients have physical attributions (Ridsdale *et al.* 1994, 2001*a*) compared to 65% to 78% when patients have been referred to specialists (Sharpe *et al.* 1996; Deale *et al.* 1997). There has been a debate about whether improvement in fatigue is associated with patients attributing their symptoms to psychological causes at baseline (Lawrie & Pelosi, 1996; Sharpe *et al.* 1996; Deale *et al.* 1997). Chalder *et al.* (2003) found that in primary care if patients have psychological attributions at baseline, they are likely to have a better outcome following psychological treatments than patients with more physical attributions. Differences in patients' beliefs at baseline may be a factor, which explains why

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

smaller doses of therapy reduce symptoms in primary care. However, it is interesting that following therapy, the group allocated to CBT shifted their views towards psychological attributions, while the group allocated to GET did not. Despite this divergence after treatment, both groups reported similar improvements in their fatigue symptoms. This evidence supports the hypothesis that improvement in fatigue and functioning may occur independently of changes in cognitions (Deale *et al.* 1997). We are exploring this with in-depth, qualitative research.

CBT was easier 'to sell'

In primary care one-half of patients think their fatigue has a physical cause. It is perhaps surprising, since CBT is a psychological intervention, that more patients randomized to CBT than GET started treatment and remained throughout (15%). Trials in secondary care have reported high dropout rates for GET also (Wearden *et al.* 1998). In future it will be useful to describe patient-preferences at baseline, or even match patients with different attributions to the therapies that better reflect their beliefs, in a patient-preference trial. Physiotherapy is more available in primary care, and if exercise therapy is to develop as a treatment, therapists may need to develop better strategies to engage and retain patient participation in treatment.

In the subgroup with CFS, brief therapy is too short

The subgroup analysis of patients with CFS was exploratory and too small to provide power, but the findings are consistent with evidence from trials in specialist settings and with our prior trial of CBT (Ridsdale *et al.* 2001*a*; Whiting *et al.* 2001). The fatigue score of the group receiving CBT was reduced by 34% and by 26% after GET. There was a trend for CBT to have a slight advantage, particularly in the group with CFS. CFS status did not alter the amount of response to treatment, but fatigue scores were significantly higher at baseline in the CFS group. So, notwithstanding an estimate of benefit, the offer of six sessions of therapy in primary care left three-quarters of the CFS subgroup still meeting case-criteria for clinical fatigue. For many patients with fatigue in primary care and particularly those with CFS, six sessions is too short. A longer duration of therapy should be tested in this setting, particularly for patients with CFS.

Conclusions

We conclude that in primary care, for patients with fatigue of $\geqslant$3 months, the effect of a brief intervention with CBT is not inferior to GET. Investigators testing complex interventions are faced with difficult choices about design, therapies, comparators, and outcomes (Campbell *et al.* 2000; Raine *et al.* 2002). In future, primary care studies need to include a randomized control group. Fewer patients took up the offer of GET, and exercise therapists need to learn why CBT is easier 'to sell'. Therapy with CBT and GET appeared to change patients' beliefs about the cause of their symptoms in opposite directions, but both groups got better. We are exploring this apparent paradox, using qualitative research. Patients with CFS responded to treatment. Their fatigue was more severe at baseline, and high 'caseness' at outcome, suggests they need longer therapy.

Decisions about which therapy to offer in practice will depend on the availability of these therapies in primary care, patients' preference, and the cost. We are reporting on the economic outcome of this trial separately (McCrone *et al.* 2004). The exigencies of research funding usually require measurement of outcome after a relatively short period, which we have done. Chronic fatigue is a relapsing-remitting condition, and outcomes need measuring in the long-term. We are following-up these patients, and will report on their long-term outcome in the future.

We acknowledge: the patients; the lead general practitioners – Drs J. Lee, M. Whyte-Venables, M. Ashworth, R. Byng, C. Cheal, G. Earnshaw, D. Abraham, A. Surawy, A. L. Hastie, D. Kay, K. A. Khan, H. Cohen, M.-A. Whitehead, C. Challecombe, C. McCarthy, J. Boxer, P. Bowskill, V. Thapar, M. A. Palmer, I. Bandara, K. Gordge, H. Knights and A. Stimmler; the therapists – G. Airey, P. Carroll, P. Cracknell, B. Gaimster, P. Garratt, E. Guyver, P. Musson, G. Pidd and K. Reilly; the therapist supervisors – Drs J. Bright and J. McCarthy; and many colleagues who provided feedback at different stages including – Drs P. White, P. McCrone and Professor R. Jones, and our anonymous referees.

The study was funded by the Linbury Trust.

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

REFERENCES

American College of Sports Medicine (2000). *Guidelines for Exercise Testing and Prescription, 6th edn*. Lipincott Williams & Wilkins: London.

Awdry, R. (1996). Homeopathy may help ME. *International Journal of Alternative & Complementary Medicine* **14**, 12–16.

Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D. & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *British Medical Journal* **321**, 694–696.

Chalder, T. (1995). *Coping with Chronic Fatigue*. Sheldon Press: London.

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D. & Wallace, E. P. (1993). Development of a fatigue scale. *Journal of Psychosomatic Research* **37**, 147–153.

Chalder, T., Wallace, P. & Wessely, S. (1997). Self-help treatment of chronic fatigue in the community: a randomised controlled trial. *British Journal of Health Psychology* **2**, 189–197.

Chalder, T., Deale, A. & Wessely, S. (1999). The treatment of chronic fatigue syndrome. In *Difficult Clinical Problems in Psychiatry* (ed. M. Rader and D. Naber), pp. 135–153. Martin Dunitz: London.

Chalder, T., Godfrey, E., Ridsdale, L., King, M. & Wessely, S. (2003). Predictors of outcome in a fatigued population following a randomized controlled trial. *Psychological Medicine* **33**, 283–287.

Cleare, A. J., Heap, E., Malhi, G. S., Wessely, S., O'Keane, V. & Miell, J. (1999). Low-dose hydrocortisone in chronic fatigue syndrome: a randomised crossover trial. *Lancet* **353**, 455–458.

Corney, R. (1990). Counselling in general practice – does it work? *Journal of the Royal Society of Medicine* **83**, 253–257.

Darbishire, L., Ridsdale, L. & Seed, P. T. (2003). Distinguishing patients with chronic fatigue from those with chronic fatigue syndrome: a diagnostic study. *British Journal of General Practice* **53**, 441–445.

Deale, A., Chalder, T., Marks, I. & Wessely, S. (1997). Cognitive behavior therapy for chronic fatigue syndrome: a randomised controlled trial. *American Journal of Psychiatry* **154**, 408–414.

Euba, R., Chalder, T., Deale, A. & Wessely, S. (1996). A comparison of the characteristics of chronic fatigue syndrome in primary and tertiary care. *British Journal of Psychiatry* **168**, 121–126.

Fitzgibbon, E. J., Murphy, D., O'Shea, K. & Kelleher, C. (1997). Chronic debilitating fatigue in Irish general practice: a survey of general practitioners' experience. *British Journal of General Practice* **47**, 618–622.

Frison, L. & Pocock, S. J. (1992). Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Statistics in Medicine* **11**, 1685–1704.

Fukuda, K., Straus, S. E., Hickie, I., Sharpe, M. C., Dobbins, J. G. & Komaroff, A. (1994). The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Annals of Internal Medicine* **121**, 953–959.

Fulcher, K. Y. & White, P. D. (1998). Chronic fatigue syndrome: a description of graded exercise treatment. *Physiotherapy* **84**, 223–226.

Greenland, S. & Finkle, W. D. (1995). A critical look at methods for handling missing covariates in epidemiologic regression analyses. *American Journal of Epidemiology* **142**, 255–264.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica* **47**, 153–162.

Joint Royal Colleges (1996). *Chronic Fatigue Syndrome: Report of the Royal Colleges of Physicians, Psychiatrists and General Practitioners*. Joint Royal Colleges: London.

Jones, B., Jarvis, P., Lewis, J. A. & Ebbutt, A. F. (1996). Trials to assess equivalence: the importance of rigorous methods. *British Medical Journal* **313**, 36–39.

Lawrie, S. M. & Pelosi, A. J. (1996). Cognitive behaviour therapy for the chronic fatigue syndrome. Essential elements of the treatment must be identified. *British Medical Journal* **312**, 1097.

Little, R. J. A. & Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. John Wiley: New York.

Lloyd, A., Hickie, I., Wakefield, D., Boughton, C. & Dwyer, J. (1990). A double-blind, placebo-controlled trial of intravenous immunoglobulin therapy in patients with chronic fatigue syndrome. *American Journal of Medicine* **89**, 561–568.

McCrone, P., Darbishire, L., Ridsdale, L. & Seed, P. T. (2003). The economic cost of chronic fatigue and chronic fatigue syndrome in UK primary care. *Psychological Medicine* **33**, 253–261.

McCrone, P., Darbishire, L., Ridsdale, L. & Seed, P. T. (2004). Cost effectiveness of cognitive behaviour therapy, graded exercise therapy, and usual care for patients with chronic fatigue in primary care. *Psychological Medicine* (submitted).

Morrell, D. C. (1972). Symptom interpretation in general practice. *Journal of the Royal College of General Practitioners* **22**, 297–309.

Powell, R., Dolan, R. & Wessely, S. (1990). Attributions and self-esteem in depression and chronic fatigue syndromes. *Journal of Psychosomatic Research* **34**, 665–673.

Raine, R., Haines, A., Sensky, T., Hutchings, A., Larkin, K. & Black, N. (2002). Systematic review of mental health interventions for patients with common somatic symptoms: can research evidence from secondary care be extrapolated to primary care? *British Medical Journal* **325**, 1082–1085.

Report of the Working Party on CFS/ME to the Chief Medical Officer for England and Wales (2002). *A Report of the CFS/ME Working Group*. Department of Health: London.

Ridsdale, L., Evans, A., Jerrett, W., Mandalia, S., Osler, K. & Vora, H. (1993). Patients with fatigue in general practice: a prospective study. *British Medical Journal* **307**, 103–106.

Ridsdale, L., Evans, A., Jerrett, W., Mandalia, S., Osler, K. & Vora, H. (1994). Patients who consult with tiredness: frequency of consultation, perceived causes of tiredness and its association with psychological distress. *British Journal of General Practice* **44**, 413–416.

Ridsdale, L., Godfrey, E., Chalder, T., Seed, P., King, M., Wallace, P., Wessely, S. & the Fatigue Trialists' Group. (2001*a*). Chronic fatigue in general practice: is counselling as good as cognitive behaviour therapy? A UK randomised trial. *British Journal of General Practice* **51**, 19–24.

Ridsdale, L., Godfrey, E. & Seed, P. (2001*b*). Chronic fatigue in general practice: authors reply. *British Journal of General Practice* **51**, 317–318.

Sharpe, M., Hawton, K., Simkin, S., Surawy, C., Hackmann, A., Klimes, I., Peto, T., Warrell, D. & Seagroatt, V. (1996). Cognitive behaviour therapy for the chronic fatigue syndrome: a randomised controlled trial. *British Medical Journal* **312**, 22–26.

Underwood, M. & Eldridge, S. (2001). Chronic fatigue in general practice. *British Journal of General Practice* **51**, 317–318.

Wampold, B. E., Moudin, G. W., Moody, M., Stich, F., Benson, K. & Ahn, H. (1997). A Meta-analysis of outcome studies comparing bona fide psychotherapies: empirically 'all must have prizes'. *Psychological Bulletin* **133**, 203–215.

Ware, J. E., Sherbourne, C. D. & Davies, A. R. (1992). Developing and testing the MOS 20-item short-form health survey: a general population approach. In *Measuring Functioning and Well-being: The Medical Outcomes Study Approach* (ed. A. L. Stewart), pp. 277–290. Duke University Press: Durham, NC.

Wearden, A. J., Morriss, R. K., Mullis, R., Strickland, P. L., Pearson, D. J., Appleby, L., Campbell, I. T. & Morris, J. A. (1998). Randomised, double-blind, placebo-controlled treatment

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press

trial of fluoxetine and graded exercise for chronic fatigue syndrome. *British Journal of Psychiatry* **172**, 485–490.

Whitehead, L. & Campion, P. D. (2002). Can general practitioners manage chronic fatigue syndrome? A controlled trial. *Journal of the Chronic Fatigue Syndrome* **10**, 55–64.

Whiting, P., Bagnall, A.-M., Sowden, A. J., Cornell, J. E., Mulrow, C. D. & Ramirez, G. (2001). Interventions for the treatment and management of chronic fatigue syndrome. *Journal of the American Medical Association* **286**, 1360–1368.

Williams, R. L. (1997). *Therapeutic Equivalence of Generic Drugs: Response to National Association of Boards of Pharmacy*. US Food and Drugs Administration: Centre for Drugs Evaluation and Research, USA.

Zigmond, A. S. & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica* **67**, 361–370.

https://doi.org/10.1017/S0033291703001247 Published online by Cambridge University Press