PACE trial: Replies to queries of Health Services and Public Health Research Board

"Randomised controlled trials of adequate size, using standard case definitions, eligibility criteria, and baseline and outcome assessments, could be used to evaluate one or more of the interventions which have been shown in one or more trials to have a benefit."

(Paragraph 12 (page 4) of summary of Draft Strategy Consultation Document of MRC CFS/ME Research Advisory Group, 17.12.2002).

Summary of responses

1. *Recruitment and retention*
   We have further developed strategies that will ensure optimal recruitment and retention of subjects in the trial. We are adopting a position of equipoise regarding the possible outcomes and therefore choice of treatments. On the basis of both our clinical and trial experiences and our discussions with Action for ME, we are confident that most potential participants will accept any of the treatments.

2. *Power analyses*
   Further power calculations have been done for a variety of possible outcomes and for both predictor and process variables. We have concluded that we should keep equal numbers of subjects in all four arms of the trial and have provided detailed justification for this decision.

3. *Outcome measures*
   We have radically reduced the number of outcome measures in order to both diminish the burden on subjects and maximise the amount of key outcome data collected.

Recruitment and retention

The Health Services and Public Health Research Board comments were:
"Further justification on the predicted recruitment and retention rates and how you would deal with potential differential drop out. In addition, the Board requested justification of the feasibility of randomisation to particular treatment arms, as it was considered likely that most patients would want APT."

(a) Predicted Recruitment

Summary response

We are aware that recruitment is a major concern in all trials. Based on our previous extensive experience (of six completed trials of treatments of CFS) and current referral data, we are confident that the planned recruitment is attainable. We also have contingency plans in case recruitment falls behind schedule.

Detailed response

We need 600 participating subjects, i.e. 200 per year over three years. All the collaborating centres are services that are under increasing pressure of referrals.

How many new referrals do we see?
The number of new CFS referrals seen in the previous year (2001) from each of the existing participating centres is as follows:

Barts Immunology	300
Barts Psychological Medicine	110
Kings College	350
Royal Free Hospital	270
Oxford	100
Edinburgh	110
Total	**1,240**

The total potentially available number of subjects is approximately 1,200 per year. Hence, if only 20 per cent of those new patients were eligible and willing to participate we would recruit approximately 240 patients per year, which is more than the minimum required. Our treatment arrangements will be flexible to avoid insufficient therapist time limiting recruitment in efficient recruiting centres. All centres will be encouraged to recruit above the minimum rate, and the recruitment rate will be closely monitored by both the Trial Management committee and the DMEC.

How many will be lost by application of inclusion and exclusion criteria?
We have chosen the broadest possible CFS criteria; the Oxford criteria. A detailed estimate of the effect of eligibility criteria and exclusions based on our existing service data is as follows. For every 100 new patients seen, 20 will not meet the operationalised Oxford diagnostic criteria, and 10 will be excluded for chronic somatisation disorder. The one subject at significant risk of self-harm will normally also have chronic somatisation disorder. This leaves 70 eligible subjects for every 100 assessed. No patient has been excluded in our previous trials because of language difficulties. To aid recruitment we will now include subjects who have previously received one of the three additional treatments.

What will we do if a centre fails to recruit?
The principal investigator and trial management committee and will discuss with the leader of the relevant centre the reasons for poor recruitment, and advise possible solutions, with defined weekly targets. We have two additional centres willing to be included, should a centre need to be replaced. These two centres are not already included in the trial because one has only just been formed and the other has just moved site. Adding another centre to the six already recruited would cost an additional £133,813 for research staff, £8,000 for travel, and £219,605 for NHS therapy costs (for a non-London centre), although these costs would partially be off-set by finishing the trial earlier. Replacing a centre would not involve these costs since resources would be transferred.

(b) Retention rates in the Trial

Summary response

We have both the confidence (derived from our previous experience of six comple
trials of treatments of CFS) and the strategies to retain the large majority of subje
in treatment. We also have strategies to minimise missing primary outcome data in
the event of patients dropping out of treatment (see below).

Detailed response

Retention refers to both completion of treatment and the obtaining of outcome data.

Will patients drop out of treatment?
Our previous three most relevant trials achieved a very high rate of treatment completion (only 0, 10 and 11 % were dropouts). This reflects the collaborative style of treatments employed. We will supply subjects with travel expenses for treatment sessions, which we found to be helpful in previous trials. If subjects drop out of treatment we will assess them at home in order to ascertain their clinical state (especially adverse outcome), obtain primary outcomes, and assess whether continuing trial treatment is possible, or what further treatment should be offered.

Might there be a differential drop-out from treatment?
Our previous trials had no differential drop-out from three different treatments. However if this does occur it will provide a useful description of practice, and it will not affect the analysis by intention to treat.

Will we obtain outcome data on treatment drop-outs?
Yes. Our three previous most relevant trials had very little missing outcome data and we obtained data on most treatment drop-outs. We believe this was because of the strategies we used to ensure this, which we have enhanced for the PACE trial.

(a) The research nurse in each centre will be selected and trained to establish a positive relationship with the trial subjects as an aim. This will be especially important in the UMC group. In addition to seeing them 4 times in 12 months, we will use techniques commonly employed in cohort studies, such as regular trial/CFS newsletters and access to a trial web-site in between these face to face meetings.

(b) In order to minimise missing data in primary outcomes we will: (i) invite all participants to attend clinic for assessment and supply travel expenses; (ii) offer to assess subjects at home if that is not possible; (iii) If this is not achieved we will use either postal or email questionnaires, supplemented by telephone calls if necessary.

Our ability to achieve very high response rates is shown by the trials we have published. Our consent form will specify these manners of assessment.

How will we manage missing data in the analysis?
The primary analysis will be pragmatic by intention to treat, thus comparing treatment policies in the four groups. In our previous trials we had less than 10 per cent missing primary outcome data. Secondary sensitivity analyses will be used to assess the robustness of conclusions about missing primary outcome data, involving imputation of all possible outcomes (Hollis, 2002).

A full Analysis Strategy, modelled on that successfully employed within the recently completed, MRC-funded, trial of Cognitive Behaviour Therapy in Bipolar Affective Disorder, will be developed before undertaking any analysis, and independently of the trial database.

(c) Feasibility of randomisation

Summary response

Those recruiting and randomising patients will maintain a position of equipoise and employ explanations that are consistent with this. Therefore, we do not anticipate a difficulty either in acceptability of the proposed treatments, with recruitment into the trial, or acceptance of randomisation. We make this statement based on having completed six trials of treatment for CFS.

Detailed response

All the participating clinicians regard all the four treatments as potentially effective and are of the view that most patients seen will accept randomisation if it is fully and openly explained. The justification for our equipoise is made in the appendix. Our experience indicates that most patients are initially nihilistic about treatment effectiveness and are willing to accept any recommended treatment, so long as it is appropriately explained and delivered.

Will there be a strong preference for APT?
Although a survey of members of a patient charity indicated a preference for pacing (Action for ME, 2001), this is not the view of the large majority of patients seen in our clinics, who are happy to accept either CBT or GET. Action for ME endorse this view.

Will patients not want to be randomised to CBT/GET?
Similarly, much of the reported dislike of CBT and GET by some members of patients' organisations is due to misconceptions. Once these therapies are fully explained as being collaborative and not imposed, and particularly that these therapies do not imply that the illness is psychiatric, the acceptance rate is very high.

Will patients not want UMC alone?
Usual medical care alone has the attraction to the subject of making fewer demands on them, while receiving medical supervision and treatment by a clinical specialist in CFS. PACE UMC will be a better standard of care than usual secondary care provided in most parts of the UK, since fatigue specialists and clinics are uncommon. UMC will also be more intensive and treatment orientated than previous trials have allowed. There is some evidence that UMC in specialist fatigue clinics can be reasonably effective (Lloyd et al, 1993). Subjects also have the offer of one of the other trial therapies after the trial. Based on our collective experience, we do not anticipate an issue with its acceptability.

Power analyses

The Health Services and Public Health Research Board comments were:
"A detailed statistical analysis of the power of the study to address the specific questions being posed. It was noted that the effect size for each arm was different, but that the sample size was the same. This raised concerns over the power calculations for the various comparisons you were proposing (e.g. APT vs UMC and CBT/GET vs APT). Members were not convinced that the study as proposed had sufficient power to adequately address the predictors of response. Therefore for each question being posed, including the predictors of response, the Board would like to see a table of the power of the study to address that question, including revision of the sample sizes for each arm if merited."

Summary response

We have given a detailed analysis of the power of the study to address each of the seven questions from the PACE protocol. We demonstrate that the trial has sufficient power to answer all the questions we have posed based on previous data and on the detection of clinically meaningful differences. We provide both scientific and practical reasons for having equal numbers in each arm of the trial. 150 subjects in each treatment group provide 90% power to detect a clinically significant difference of 20% between any two treatment groups (Machin et al, 1997).

Power analyses to compare efficacy

Question 1 (from original PACE protocol): "Are CBT and/or GET more effective than pacing in reducing both fatigue and disability?"
Question 2: "Is pacing more effective than usual medical care?"

Table 1, below shows the percentages improved by the two primary outcomes and our global change measure (the CGI) within each of the four treatment groups, based on previous studies or best estimates (see appendix: table and summaries of relevant studies).

Table 1: Best estimate of outcomes mainly based on published data

Primary Outcome	Treatment Group			
	UMC	APT	GET	CBT
% improved (Fatigue scale)	10	25	50	60
% improved (SF36)	5	25	70	65
% improved (CGI)	15	30	65	60

Since these do not vary widely within treatment groups we have used close variations of these conservative percentage improvements as the estimates for the power calculations.

Table 2: 90% power analysis for the primary outcome of fatigue

Selected % improved				Number of Subjects		
UMC	APT	GET	CBT	APT vs. UMC	GET vs. APT	CBT vs. APT
10	25	50	60	**133**	**77**	**40**
10	30	50	60	**82**	**124**	**56**
15	25	50	60	335	**77**	**40**
15	30	50	60	161	**124**	**56**

Table 2 shows the numbers of subjects *per treatment group* required to detect the corresponding difference with 90% power at P=0.05 (two-sided)(Machin et al, 1997). Numbers in bold are achieved with our projected recruitment of 150 subjects per group (inflated by 10% to allow for drop-outs). Despite some uncertainty in the response to both UMC and APT, each row of Table 2 indicates treatment groups of sufficient size to establish clinically meaningful differences between APT and both GET and CBT, as well as between APT and UMC with a 10% response to the latter. Any trend in the percentages improved from UMC to APT to GET / CBT will be evaluated formally by a test for linear trend in the log (odds).

Table 3: 90% power analysis for the primary outcome of disability

Selected % improved				**Number of Subjects**		
UMC	APT	GET	CBT	APT vs. UMC	GET vs. APT	CBT vs. APT
5	25	70	65	**65**	**24**	**31**
5	30	60	65	**47**	**56**	**41**
10	25	70	65	**133**	**24**	**31**
10	30	60	65	**82**	**56**	**41**

The figures in bold show that our figures of 150 per treatment group provide adequate power for all likely comparisons, including coping with drop-outs.

Keeping equal numbers in each arm of the trial
Equal allocation is the most efficient method, from a statistical point of view, for a given sample size where all the pair-wise comparisons are important. Our sample size of 150 in each treatment group allows for significant uncertainty in the percentage response to all treatments, in a multi-centre trial. It is a sufficient number to allow 90% power to detect a 20% difference in outcome between any two treatments (Machin et al, 1997). Equal numbers also have an advantage in a four-arm trial with some uncertainty of outcome, when measuring multiple outcomes. The only realistic alternative to this, on the basis of the power calculations in table 2, would be to reduce the numbers receiving CBT (and GET to a lesser extent), perhaps by combining CBT and GET. By doing this we would lose the power to examine differential predictors and processes of these treatments, as well as significantly reducing the power of cost-effective comparisons.

Importantly, randomisation to equally sized groups also avoids complicated explanations when consenting patients for recruitment, and better reflects the therapeutic equipoise held by the research team as a whole. This is essential for the trial to be considered a fair test of all treatments and thus be acceptable to both patients in general and the wider community.

Differential effects on primary outcomes

Question 4 of protocol: "Do different treatments have differential effects on outcomes (i.e. disability versus symptoms)?"

This question is of particular interest to Action for ME. We suspect that a differential effect is most likely with APT, with fatigue being more effectively treated than disability. Previous trials of CBT and GET suggest that both these treatments have similar effects on disability and fatigue (see appendix table). We expect that there will be no more than a 5 % difference between the two primary outcomes with CBT/GET combined. Assuming that 10 % of subjects would improve in terms of disability with

APT (double that seen with UMC: see appendix table) and 30 % would improve in fatigue with APT (see appendix table) gives a 20% difference in outcomes. With a power of 90 % and P = 0.05, 101 subjects would be required in each group (Machin et al, 1997). We will have 300 in the combined CBT/GET group and 150 in the APT group, which gives us sufficient power.

Predictive Factors

Question 3: "Are there differential predictors of response to CBT and GET (and does the mechanism of change differ)?"
Question 5: "What factors predict a favourable response to treatment in general and with specific treatments?"

The primary purpose of any RCT is to estimate comparative efficacy and cost-effectiveness of treatments, and power is calculated with this aim. Examination of prognostic factors is a secondary objective, the success of which is dependent upon the strength of association between these factors and outcome (here the percentage who improve), and among the factors themselves; this requires logistic regression modelling. Since there are no clearly identified prognostic factors for the response to CFS to different treatments, the modelling will be both exploratory and hypothesis driven, with variable selected on the basis of a combination of previous unreplicated predictors (Prins et al, 2001 & 2003; Bentall et al, 2002), clinical expertise and statistical significance. Baseline physical disability will be entered into later models. Subsequent modelling of secondary outcomes may also be examined but since these are mainly continuous measures, examination of power is restricted to the (much) less sensitive logistic regression model.

With four equal size treatment groups and estimated response percentages of 60, 50, 30, and 10, we expect an overall response of 40%. For binary risk factors, independent of treatment, the sample size to estimate an odds ratio (with 90% confidence) within a specified percentage (ε%) is conditioned on this overall response, and is dependant on both the sample distribution across the risk factor, and the response percentage within one of the two levels of the risk factor; formula (6.16), page 134, of Machin et al (1997) can be used. We estimate that a sample size of 600 is sufficient to estimate odds ratios of 1.3 (or greater) within 30%, with 50% response with the risk factor present, and the latter distributed as extreme as 3:1 within the sample. Further this sample size is sufficient to allow for modelling of multiple risk factors as determined by the variance inflation factor, $(1-\rho^2)$, where ρ, the multiple correlation coefficient between them, does not exceed 0.4.

For continuous risk factors, independent of treatment, we use Table II of Hsieh (1989), corresponding to 5% significance (one-tailed) and 80% power. With overall response of 40% (at the mean of the risk factor), 420 participants are required with an odds ratio (or greater) of 1.3 at one standard deviation above the mean. Again, with multiple risk factors, this number must be increased by the variance inflation factor.

Modelling of risk factors associated with treatment (treatment-covariate interactions) requires the multiple correlation of the treatment effect, risk factor, and their interaction to be taken into account. Even assuming this to be as high as 0.45, then $420/(1-0.45^2) = 526$, which is within our total sample size.

Pervasive inactivity power analysis
Prins et al (2003) found that pervasive inactivity predicted non-response to CBT at 8 months follow-up and provided data allowing a power analysis. None of the subjects who were pervasively passive responded to CBT, compared to 40% of the subjects who were moderately active. If we assume that 5% of pervasively passive subjects will respond in the PACE trial compared to 40% of moderately active subjects, only 28 subjects are required in the CBT group (Machin et al, 1997). Since GET is designed to reverse pervasive inactivity, we anticipate that significantly more will respond. If we therefore assume that 30% of the pervasively passive will respond to GET, we require 48 subjects in both CBT and GET, at 90% power and P = 0.05 (Machin et al, 1997).

Process of change

Question 3: "(Are there differential predictors of response to CBT and GET and) does the mechanism of change differ?"
Question 6: "What are the mechanisms of change with successful treatment?"

Summary response

We hypothesise that GET will improve fitness and walking distance more than CBT, whereas illness beliefs will change more with CBT than GET, and that increased activity is essential to successful treatment. The 10 week mid-therapy assessment will allow us to judge this, and we probably have sufficient power to do this, as suggested by the power analysis estimated for illness beliefs (see below).

Detailed response

No studies of mechanism of change with successful CFS treatments have been published. Power studies are therefore difficult. The most useful data, which gives some indication of likely power, involve changes in illness beliefs (Deale et al, 1998).

Illness beliefs
Three illness beliefs significantly changed with CBT. These were that "I should avoid exercise when tired, doing less helps fatigue, and exercise is harmful" (Deale et al, 1998). All three beliefs were held by 57% before CBT. Immediately after CBT, 33% of subjects still held the first and third beliefs and 22% held the second belief (Deale et al, 1998). This represents at least 24% changing their beliefs after CBT. We surmise that 20% will have changed their beliefs after 10 weeks of CBT compared to 5% with GET. With power = 90% and P = 0.05, 101 subjects will be required in both the CBT and GET groups to detect this difference (Machin et al, 1997).

Cost-effectiveness

Question 7: "What are the relative cost-effectiveness and cost-utility of these treatments."

The original proposal states that "Since costs are not expected to vary significantly between or within groups, the treatment determined number of 150 per arm is likely to find significant differences in cost-effectiveness". We are still of the view that costs will not vary significantly between groups. This view was reinforced by two recent studies that found, after standardisation for baseline factors, no significant cost differences in primary care settings (a) between counselling and CBT (Chisholm et

al, 2001), or (b) between CBT, GET and usual care (McCrone et al, 2003). The study will test for differences in cost-effectiveness and cost-utility (and not simply cost), so that if costs do indeed prove to be similar between treatments, the cost-effectiveness differences will come down to whether outcomes differ. A probable approach will be to combine costs and outcomes to compute net benefits. Power calculations are then not really relevant because the key consideration is the societal (or other) value attached to a unit change in outcome and the associated cost-effectiveness probability distribution (Briggs, 2001).

Outcome measures

The Board's comments:
"Further justification for the outcome measures you were intending to use, including consideration of their intensity, number and feasibility of being undertaken."

Summary response

We have either omitted or substituted eight measures from the trial in order both to concentrate on the essential measures, and to considerably reduce the measurement load on subjects.

Detailed response

Those measures being either deleted or substituted

1. Measurement of physical movement by Actiwatch
This will save 48 hours of subjects' time and commitment in all four interviews. We will substitute the six-minute walking distance test (Butland et al, 1982), which has been successfully used to objectively measure change in exercise tolerance in a previous trial (Sharpe et al, 1996).

2. The Queen's College three-minute step test
We are substituting the quicker and better tolerated self-paced step test, which has been shown to be acceptable to both unwell and elderly individuals, as well as correlating closely to more extended tests of fitness (Petrella et al, 2001). Because the test will be less of an effort for subjects, more subjects will complete the test and give useful data. A one-minute step test was shown to predict CFS in a previous study (White et al, 2001).

3. The EuroQOL (EQ-5D) Quality of life

4. Strength of the therapeutic alliance

5. All sub-scales of the SF-36, except the social and physical function sub-scales

6. The criteria for "fibromyalgia"

7 & 8. The kinesiophobic and illness beliefs questionnaires
We are substituting three simple questions regarding illness beliefs (see process measures, below), which Deale et al (1998) showed changed with CBT.

After these omissions/changes, we will have the following outcome measures left in the study:

Two primary outcome measures:

Chalder fatigue questionnaire 11 item scale (2 minutes to fill in)
The SF-36 physical function 10 item sub-scale (2 minutes to fill in)

Seven secondary outcome measures:

Client Service Receipt Inventory (CSRI) (15 minutes); the only measure of economic variables

Six-minute walking distance test (Butland et al, 1982); the only objective measure of physical function
SF-36 social function 2 item sub-scale (1 minute); the only measure of social function
The Hospital Anxiety and Depression scale (2 minutes), the only measure of mood
Clinical Global Impression (CGI) change score (1 minute), the only global measure of change
Likert scale scores of non-fatigue symptoms of CFS (Fukuda et al, 1994)(2 minutes); particularly requested by Action for ME in order to assess change in symptoms other than fatigue
Satisfaction with treatment scale (1 minute)

References

Action for ME. A survey of members of Action for ME. 2001, Action for ME, London.
Bentall RP et al (2002). Predictors of response to treatment for chronic fatigue syndrome. *Br J Psychiatry* 181, 248-52.
Briggs AH (2001). A Bayesian approach to stochastic cost-effectiveness analysis. *International Journal of Technology Assessment in Health Care* 17, 69-82.
Butland RJA et al (1982). Two-, six-, and 12-minute walking tests in respiratory disease. *BMJ* 284:1607-8.
Chisholm D, Godfrey E, Ridsdale L, Chalder T, King M, Seed P, Wallace P, Wessely S (2001) Chronic fatigue in general practice: economic evaluation of counselling versus cognitive behaviour therapy. *Br J Gen Pract*, 51, 15-18.
Deale A et al (1998). Illness beliefs and treatment outcome in chronic fatigue syndrome. *J Psychosom Res* 45, 77-83.
Fukuda K et al. The chronic fatigue syndrome: a comprehensive approach to its definition and study. *Ann Intern Med* 1994;121:953-959.
Fulcher KY & White PD (2000). Strength and physiological response to exercise in patients with the chronic fatigue syndrome. *J Neurol Neurosurg Psychiatry* 69:302-7.
Hollis S (2002). A graphical sensitivity analysis for clinical trials with non-ignorable missing binary outcomes. *Statistics in Medicine* 21, 3823-34.
Hsieh FY (1989). Sample size tables for logistic regression. *Statistics in Medicine* 8:795-802.
Lloyd A et al (1993). Immunologic and psychologic therapy for patients with chronic fatigue syndrome. Am J Med 94, 197-203.
Machin D et al (1997). *Sample Size Tables for Clinical Studies.* Blackwell Science, 2nd edition: Oxford.
McCrone P, Ridsdale L, Darbishire L, Seed P (2002). Cost-effectiveness of cognitive behavioural therapy, graded exercise and usual care for patients with chronic fatigue in primary care. Submitted for publication.
Petrella RJ et al (2001). A self-paced step test to predict aerobic fitness in older adults in the primary care clinic. *J Am Geriatr Soc* 49, 632-8.
Prins JB et al (2001). Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 357:841-7.
Prins JB et al (2003: in press). Cognitive behaviour therapy for chronic fatigue syndrome: predictors of treatment outcome. *Excerpta Medica International Congress Series.*
Sharpe M et al. Cognitive behaviour therapy for the chronic fatigue syndrome: a randomised controlled trial. *BMJ* 1996;312:22-26.
Wearden A et al (1998). A randomised, double-blind, placebo controlled treatment trial of fluoxetine and a graded exercise programme for chronic fatigue syndrome. *Br J Psychiatry* 172, 495-90.
White PD et al (2001). Predictions and associations of fatigue syndromes and mood disorders that occur after infectious mononucleosis. *Lancet* 358:1946-54.

APPENDIX

Table: Outcome values 1 year after treatment onset from previous trials
N.B. All analyses by intention to treat. Figures in bold italics (underlined) are estimates. Other figures are data taken from previous trials.

	UMC	APT	GET	CBT
Chalder scale mean (SD or 95% CI) scores at baseline and one year*	11 (10-11) > 10 (9-11)	***11 > 7***	11 (10-11) > 3 (2-5)	10 (1) > 4 (4)
% subjects improved by Chalder fatigue scale	***10,*** (6% at 6/12)	***30***	(18% at 6/12)	60
SF36 physical function (PF) mean (SD) scores at baseline and 1 year$				25 (19) > 72 (28)
% subjects much improved by SF 36 PF	6	***10***	69	63
% subjects "much improved" by self-rated CGI	12, 23	***30***	63, 70	50, 60, 63
% treatment drop-outs	3, 6	***10***	9, 17, 37	0, 10, 36

*Chalder fatigue scale scores: 11 = maximum, 3 or less = population normal energy
$ SF-36 physical function sub-scale scores: 0 is total disability, 90 = UK population mean functioning.

N.B. Although we have used these data, along with estimates of clinically important differences, for the purposes of our power analyses, we stress the uncertainty in much of these data, which are based on small, usually single centre studies undertaken by enthusiastic leaders for particular therapies.

Studies from which these data were derived:

COGNITIVE BEHAVIOUR THERAPY

Sharpe et al, 1996 RCT of CBT versus UMC
This study did not use the SF36 or the Chalder scale. By CGI at one year, 60% were much better after CBT versus 23% after UMC.
No subjects dropped out of CBT and 1/30 (3%) dropped out of UMC.

Deale et al, 1997 RCT of CBT versus relaxation
By Chalder fatigue scale, CBT subjects went from a mean (SD) of 10 (1) to 4 (4) compared to relaxation subjects who went from 9.5 (3) to 7 (4) at one year.
By SF36 PF sub-scale, CBT subjects went from a mean (SD) of 25.5 (19) to 72 (28) compared to relaxation subjects who went from 28 (27) to 38 (27).
By CGI 19/30 (63%) were much better compared to 8/30 (27%).
10% dropped out of CBT and 13% dropped out of relaxation.

Prins et al, 2001 RCT of CBT versus guided support and normal course
This study did not use the Chalder scale or SF36. By a different fatigue scale, 35% were significantly better by 14 months with CBT, compared with 13% with the control support group (SG) and 17% with normal course (NC). NC subjects were free to seek any treatment they wished, so we do not think it represents UMC and have not used this particular data. Significant functional improvement occurred in 49% after CBT, but only 19% with SG and 23% with NC. 50% felt "much better" after CBT compared with 15% after SG and 32% after NC. 36% dropped out of CBT compared to 28% after SG and 10% after NC.

GRADED EXERCISE THERAPY

Fulcher and White, 1997 RCT of GET versus relaxation and flexibility
Since this was a cross-over trial after initial treatment, we cannot use this trial for one year follow-up data, apart from noting that 35/56 (63%) subjects, who had either been randomised or crossed over to GET, rated themselves as much better at one year follow-up by CGI.
3/33 (9%) subjects dropped out of GET. 4/33 (12%) subjects dropped out of the control treatment.

Powell et al, 2001 RCT of education followed by GET versus UMC
GET: SF36 physical function (PF) score 69% improved to a score of 25/30. UMC: SF36 PF score 6% improved.
Using the full range of 10 – 30 scores on SF36 physical function scale (10 = maximum impairment, 30 = no impairment):
UMC went from a mean (95% CI) of 16 (15 to 17.5) to 17 (15 to 18) and GET went from 16 (15 to 17) to 25 (23 to 26), by one year.
Using the Chalder Fatigue Scale, UMC went from 11 (10 – 11) to 10 (9 – 11) and GET went from 10 (10 – 11) to 3 (2 – 5), by one year.
By the CGI, 12% were much better after UMC versus 70% after GET by one year.
6 % dropped out of UMC and 17 % dropped out of GET.

Wearden et al, 1998 RCT of GET
This trial was atypical by adopting a factorial design, with a drug treatment. It used a more strenuous exercise level from the start, with fewer therapy sessions, and relied on physiological evidence of improvement before increasing the intensity of GET, unlike the other two GET trials. We believe these factors were related to the higher drop-out rates and less positive outcomes and have therefore excluded these data for the purpose of our power analyses.
Using the original Chalder Fatigue Scale (range 0 – 14 (14 maximum fatigue)): GET brought about 18 % of subjects scoring lower than 4, and 6 % of UMC subjects met this criterion by 6 months. (NB a score of <4/14 is a more conservative criterion than the <4/11 that we will use in PACE.)
There were no significant changes in SF36 physical function scores between GET and the control arm.

37 % dropped out of GET (with or without fluoxetine). 22 % dropped out of the control group (with or without fluoxetine).

USUAL MEDICAL CARE

UMC in the two studies above were provided by the general practitioner (Sharpe et al, 1996) supplemented by an information booklet and once-off medical advice (Powell et al, 2001). The UMC planned in PACE is supplemented by fatigue specialist advice, follow-up and pharmacotherapy. We therefore believe that the outcomes will be better than those of the previous two studies, which was the case in usual specialist care in an Australian RCT (Lloyd et al, 1993). This supports our stance of equipoise between the four arms of the study. At the same time our power calculations have been based not only on previous studies, but also on what we believe a clinically significant difference will be.

ADAPTIVE PACING THERAPY

There are no published trials of adaptive pacing therapy (APT), but Friedberg and Krupp (1994) have published a non-randomised comparative trial of six weekly sessions of adaptive "cognitive behaviour therapy" compared to no treatment. Although called CBT the active therapy ensured that any lifestyle changes were compatible with activity limitations imposed by CFS. Therefore, we believe this is as near to a trial of APT that has been published. There were no drop-outs from 22 subjects, but also no significant differences in any outcomes. The therapy had only 6 sessions, rather than PACE trial's 14. This fact, when added to our clinical experience with APT, and the efficacy (26% and 27%) of our previous control treatments of relaxation and flexibility therapies, gives us our best estimate that the efficacy of APT will fall between that of CBT/GET and UMC.

END