Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial

P D White, K A Goldsmith, A L Johnson, L Potts, R Walwyn, J C DeCesare, H L Baber, M Burgess, L V Clark, D L Cox, J Bavinton, B J Angus, G Murphy, M Murphy, H O'Dowd, D Wilks, P McCrone, T Chalder*, M Sharpe*, on behalf of the PACE trial management group†

Summary

Background Trial findings show cognitive behaviour therapy (CBT) and graded exercise therapy (GET) can be effective treatments for chronic fatigue syndrome, but patients' organisations have reported that these treatments can be harmful and favour pacing and specialist health care. We aimed to assess effectiveness and safety of all four treatments.

Methods In our parallel-group randomised trial, patients meeting Oxford criteria for chronic fatigue syndrome were recruited from six secondary-care clinics in the UK and randomly allocated by computer-generated sequence to receive specialist medical care (SMC) alone or with adaptive pacing therapy (APT), CBT, or GET. Primary outcomes were fatigue (measured by Chalder fatigue questionnaire score) and physical function (measured by short form-36 subscale score) up to 52 weeks after randomisation, and safety was assessed primarily by recording all serious adverse events, including serious adverse reactions to trial treatments. Primary outcomes were rated by participants, who were necessarily unmasked to treatment assignment; the statistician was masked to treatment assignment for the analysis of primary outcomes. We used longitudinal regression models to compare SMC alone with other treatments, APT with CBT, and APT with GET. The final analysis included all participants for whom we had data for primary outcomes. This trial is registered at http://isrctn.org, number ISRCTN54285094.
The Intervention

**Intervention**
PACE is a multicentre randomised controlled trial. The group assignment is parallel group.
1. Standardised Specialist Medical Care alone (SSMC) - manual guided advice from a secondary care clinic specialist in chronic fatigue
2. Standardised Specialist Medical Care plus adaptive pacing therapy (APT)
3. Standardised Specialist Medical Care plus graded exercise therapy (GET)
4. Standardised Specialist Medical Care plus cognitive behaviour therapy (CBT)

There is no masking as the supplementary treatments being trialled are delivered by therapists and maintaining any blind would be very difficult. Even though treatment allocation is not blinded, staff are encouraged not to discuss randomisations or any subject that might inadvertently lead to bias.

1. SSMC alone
2. SSMC + APT (Adaptive Pacing Therapy)
3. SSMC + GET (Graded Exercise Therapy)
4. SSMC + CBT (Cognitive Behaviour Therapy)

Ref: ISRCTN54285094 registry. | DOI 10.1186/ISRCTN54285094 | A randomised controlled trial of adaptive pacing, cognitive behaviour therapy, and graded exercise, as supplements to standardised specialist medical care versus standardised specialist medical care alone for patients with the chronic fatigue syndrome/myalgic encephalomyelitis or encephalopathy (Last edited 3/11/2015).
### Study Design

**Methods**

**Study design and participants**

PACE was a parallel, four group, multicentre, randomised trial, with outcomes assessed up to 52 weeks after randomisation for patients with chronic fatigue syndrome. We recruited 641 participants from consecutive new outpatients attending six specialist chronic fatigue syndrome clinics in the UK National Health Service between March 18, 2005, and Nov 28, 2008, and completed outcome data collection in January, 2010.

---

### Selection Bias

- Unblinded. Revealing the study aims to the participant biases the odds ratio (OR) either up or down.
- People interested are more likely to consent, **biasing the OR up**.
- Non-stratified: selection may **not represent the population of interest**.
- The prevalence of the intervention (exposure) is **under-estimated & OR over-estimated** if the participant drops out or refuses randomisation.
- Level of confounders **not mentioned** biases OR either up or down.

---

### Randomised Controlled Trial

- An RCT needs a control group to negate major selection biases ie. probable risk between treatment & controls.
- RCT without a control group is a **different study protocol**.
- Randomisation isn't about "fairness" in allocation.

### Parallel

If the intervention doesn't work, participants should be offered a different therapy instead of waiting 52 weeks eg. cross-modal therapy arms.

### Consecutive new outpatients

Consecutive new outpatients can't be randomised concurrently (or parallel) as opposed to a database of existing patients with CFS.

### 52 weeks

Non-evidence based treatment endpoints don't set realistic milestones for trial completion.

---

- **Preregistration**: Minimises p-hacking & opens up pre-trial corrections for a better study design.
- **Stratified groups**: Groups participants by previous treatment condition, anticipates cross-modal.
- **Correct study design**: Cohort study (non-randomised trial) participant chooses treatment option.
- **Site specific assessment**: Recruit PIs who are passionate and committed to the study aims.
- **Evidence-based endpoints**: Sets a realistic timeframe for trial recruitment, follow-up & completion.
**PACE Clinical Protocol**

**Initial screening for eligibility – visit 0 (clinic doctor)**

New referrals to the outpatient clinics may be received from GPs or any other appropriate medical practitioner. Each clinic doctor will ensure that all consecutive new outpatients with a clinical diagnosis of CFS/ME are considered for the trial (i.e. if thought to be eligible they are told about the trial). Each centre leader will keep a trial log-book of every new chronic fatigue outpatient referral. This log book will detail each patient seen, whether or not they were referred for the trial and the reasons if not.

Where the patient is thought to be suitable by the clinic doctor (with a CFQ score of 6 or above and an SF-36 score of 65 or below), and the patient agrees to be assessed for eligibility, the clinic doctor will forward the patient’s contact details to the RN. The clinic doctor will give the patient the trial Participant Information Sheet. The RN will contact the patient to arrange the first research visit (visit 1).

- **GP referrals:** GPs shouldn’t be making referrals to a trial, but to a specialist who can treat CFS appropriate to the patient’s medical history. Specialists eg. psychologists, physiotherapists, geneticists as PI (onsite principal investigators) could offer new patients an option to join the trial because it is the best adjuvant option for the patient, proceed with screening, fulfill the eligibility criteria, then have the external nurse coordinator from the principal site go through the consent process.

- **Duplicate screening leads to biases.** Participants are already exposed to screening questions. Duplicate screening primes and validates other measures opposed to validating the eligibility criteria. Randomisation isn’t appropriate in this instance, unless the patient is experiencing a recurrence or have exhausted all available treatment options specific to their condition.

- **Keeping a log book breaches confidentiality, raises ethical issues, and opens up a can of biases.** If a patient doesn’t go through the consent process- no information should be logged. They aren’t part of the trial. A log book is only appropriate if participants are recruited from a pre-existing database that abide by privacy laws.

**An RCT study is possible for a CFS population subset**, who have exhausted all available treatment options and there is evidence (case studies) to suggest the intervention as an adjuvant therapy is unlikely to cause considerable harm. eg. A patient diagnosed with ME undergoing unsuccessful drug therapy consents to an adjuvant therapy trial, and is informed about randomisation and the harms associated with undertaking adjuvant therapy.

**The benefits and harms of the intervention** should be thoroughly researched for all eligible pre-existing medical conditions, and stratified into groups under the one PI (specialist for a pre-existing condition eg. ME) who then experiment with multi-modal adjuvant therapies eg. measures CBT for 8 weeks, then pacing 8 weeks, if CBT response is better, then try CBT for 12 weeks, eliminate pacing, try CBT and GET at the same time etc. as opposed to complying with one treatment arm for 52 weeks.
Selection Bias

Eligibility Criteria

Randomisation

Baseline Measures

Exclusion Criteria

Eligibility Scales
## Eligibility Criteria

3158 screened for eligibility: 2260 excluded
- 139 unable to comply with protocol: protocol exposed prior to randomisation: blinded; selection bias; misclassification biases.
- 71 contraindication to trial treatment: unspecified confounder.
- 46 psychiatric exclusions: unspecified exclusion validity criteria.
- 533 did not meet primary consent criteria: significant loss from inappropriate study design.
- 46 doctors declined patient's randomisation: investigator led bias.
- 29 unrecorded: missing data.

257 excluded (baseline measure discrepancies unspecified)
- 67 no current Oxford diagnosis of CFS yet was diagnosed previously. Investigator led bias can lead to differential misclassification of outcome and bias results towards or away from the null.
- 52 unable to comply with the protocol yet wasn't a problem previously: unspecified exclusion validity criteria; investigator led bias.
- 16 physical function score > 65 yet satisfied previously: unspecified exclusion validity criteria; investigator led bias.
- 13 contraindication to trial treatment: unspecified confounder.
- 12 psychiatric exclusion yet not excluded prior: investigator led bias.
- 4 bimodal fatigue score < 6 yet not excluded prior: investigator led bias.
- 1 unable to speak or read English adequately but could speak or read English previously: investigator led bias.
- 39 patients declined further assessment: unspecified reason, potential protocol exposure prior to randomisation: investigator led bias.
- 28 patients declined randomisation yet consented previously: investigator led bias.
- 12 unrecorded: missing data.
Randomisation Technique

Database programmer: institution or software unspecified.

Dubious randomisation technique: "straightforward"; "computer-generated probabilistic minimisation."

In RCTs, participants, therapists and doctors are masked to treatment allocation so long as the control group undergoes a placebo version of the same routine. If this is impractical- a different study design is recommended.

Given the variable differences in treatment eg. self reports vs CBT- its unlikely the statistician was masked to allocation but perhaps to randomisation undertaken by a database programmer (although its unusual to have a computer programmer in a clinical setting).
## Baseline Measures

<table>
<thead>
<tr>
<th>Demographic Data</th>
<th>Adaptive pacing therapy (n=159)</th>
<th>Cognitive behaviour therapy (n=161)</th>
<th>Graded exercise therapy (n=160)</th>
<th>Specialist medical care alone (n=160)</th>
<th>Overall (n=640)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>39 (11)</td>
<td>39 (12)</td>
<td>39 (12)</td>
<td>37 (11)</td>
<td>38 (12)</td>
</tr>
<tr>
<td>Female</td>
<td>121 (76%)</td>
<td>129 (80%)</td>
<td>123 (77%)</td>
<td>122 (76%)</td>
<td>1495 (77%)</td>
</tr>
<tr>
<td>White</td>
<td>146 (92%)</td>
<td>151 (94%)</td>
<td>148 (93%)</td>
<td>150 (94%)</td>
<td>595 (93%)</td>
</tr>
<tr>
<td>Any ME group membership</td>
<td>31 (19%)</td>
<td>26 (16%)</td>
<td>25 (16%)</td>
<td>23 (14%)</td>
<td>105 (16%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clinical Data</th>
<th>Adaptive pacing therapy (n=159)</th>
<th>Cognitive behaviour therapy (n=161)</th>
<th>Graded exercise therapy (n=160)</th>
<th>Specialist medical care alone (n=160)</th>
<th>Overall (n=640)</th>
</tr>
</thead>
</table>
| International CFS criteria:  
As randomised | 99 (62%)                        | 100 (62%)                          | 98 (61%)                        | 100 (63%)                           | 397 (62%)       |
| Actual          | 107 (67%)                       | 106 (66%)                          | 106 (66%)                       | 108 (68%)                           | 427 (67%)       |
| London CFS criteria:  
As randomised | 89 (56%)                        | 90 (56%)                           | 89 (56%)                        | 89 (56%)                            | 357 (56%)       |
| Actual          | 81 (51%)                        | 84 (52%)                           | 84 (53%)                        | 80 (50%)                            | 329 (51%)       |
| Any depressive disorder:  
As randomised | 55 (35%)                        | 55 (34%)                           | 54 (34%)                        | 55 (34%)                            | 219 (34%)       |
| Actual          | 54 (34%)                        | 52 (32%)                           | 54 (34%)                        | 53 (33%)                            | 213 (33%)       |
| Any psychiatric disorder:  
As randomised | 75 (47%)                        | 75 (47%)                           | 73 (46%)                        | 77 (48%)                            | 300 (47%)       |
| Duration of illness (months) | 33 (16-69)                  | 36 (16-104)                        | 35 (18-67)                      | 25 (15-57)                         | 32 (16-68)      |
| Body-mass index (kg/m²) | 25.9 (5.5)                    | 25.4 (5.2)                         | 25.5 (4.6)                      | 25.1 (4.5)                          | 25.5 (5.0)      |

Data are mean (SD), n (%), or median (IQR). ME=myalgic encephalomyelitis. CFS=chronic fatigue syndrome. *Psychiatric disorders included any depressive disorder and any anxiety disorder, including phobias, obsessive-compulsive disorder, and post-traumatic stress disorder.

**Table 1: Baseline demographics and clinical characteristics**

No evidence presented on CFS being prevalent in this baseline demographic: white (93%); female (77%); 1 in 5 chance (20%) a ME group member. **Conflict of Interest.**
Exclusion Criteria

Participant exclusion criteria

1. All potential participants will be screened for medical exclusions, by history and physical examination. Appropriate investigations will be undertaken by either the referring doctor or the centre doctors (checked by the RN). Patients with a relevant alternative medical diagnosis will be excluded. Investigations will be those recommended by the Royal Colleges’ Report on CFS/ME and the CMO’s working group report. These results will be collated by the RN, and will have been undertaken within six months of the baseline assessment.

2. The Research Nurse (RN) will use a standardised psychiatric interview (the Structured Clinical Interview for DSM-IV - SCID), under supervision by a participating centre PI or nominated deputy, to exclude those who are at significant risk of self-harm and those with psychiatric exclusions listed in the Oxford diagnostic criteria for CFS.

3. Patients who are considered by the RN in discussion with their centre leader to be unable to do one or more of the trial therapies or to complete all trial measures or for whom participation in the PACE trial would be inappropriate to their clinical needs (e.g. someone with significant post traumatic stress disorder or borderline personality disorder).

4. Patients who have previously received one of the trial treatments before from a centre participating in PACE (rather than any secondary care clinic for Chronic Fatigue Syndrome) and received a course of any of the supplementary therapies of CBT, GET or pacing therapy from a therapist will be excluded from taking part in the trial, or of advice from a PACE doctor that is judged to have been similar to SSMC (changed from ‘Patients who have previously attended a specialist fatigue clinic and received a course of any of the supplementary therapies of CBT, GET or pacing therapy from a therapist will be excluded from taking part in the trial’ in April 2006).

Screening by different research officers (referring doctor, centre doctor, RN) for the same thing can lead to different measures (differential effects) and bias results either towards or away from the null hypothesis.

Beginning treatment at a maximum of six months from baseline assessment is considered negligent unless justified.

The Research Nurse shouldn't be making psychiatric assessments on those who are at significant risk of self-harm. It's unclear whether the RN (Registered Nurse? Enrolled Nurse?) received specialist training under the supervision of an accredited psychiatrist or counselor.
Eligibility Scales

Participant inclusion criteria

Consent:
1. Both participant and clinician agree that randomisation is acceptable
2. The participant has given written informed consent

Eligibility:
3. The participant meets operationalised Oxford research diagnostic criteria for CFS
4. The participant's Chalder Fatigue Questionnaire score is 6 or more
5. The participant's SF-36 physical function sub-scale score is 65 or less (changed from '60 or less' in April 2006)
6. The participant will be aged at least 18 years old, either sex

The inclusion criteria doesn't specify the research inclusion criteria eg. criteria set by the second screening.

The benefits of randomisation outweighing the risks hasn't been justified with evidence for a population with baseline characteristics. eg. are these treatments accessible and affordable in the region for those diagnosed with CFS? Has it shown any benefit? Is it sustainable? Does randomisation justify the patient's treatment choice? Are the treatment arms equally beneficial?

Oxford research diagnostic criteria for CFS (1991): is overshadowed with newer CFS diagnostic criteria eg. Canadian 2003 criteria. It's unclear why the former was picked over the latter.

Chalder Fatigue Questionnaire (1993): was originally a 14-item scale for fatigue participants recruited from general practice. Then later 11-item (CSFQ 11) from Morriss, Wearden & Mullis (1998) who validated the questionnaire on CFS patients undertaking treatment in aerobic exercise and fluoxetine (antidepressant). It's unclear why the Chalder Fatigue Questionnaire was included in the eligibility criteria if fluoxetine wasn't, or why the questionnaire wasn't validated for other inclusions or exclusions before testing on control groups or standardising validations across treatment arms= BEFORE the trial ie. a pilot study among those diagnosed with CSF to set the trial's evidence-based inclusion and exclusion (eligibility) criteria.

We used continuous scores for primary outcomes to allow a more straightforward interpretation of the individual outcomes, instead of the originally planned composite measures (50% change or meeting a threshold score).\textsuperscript{10,30} We prorated primary outcomes scales only when there were at most two items per scale missing (nine participants for Chalder fatigue questionnaire and 11 for short form-36). Prorating involved calculating the mean value of the item scores present and replacing the missing values with that score.

Changing from eg. bimodal to continuous scores is a different patient reported outcomes instrument. It moots null baseline measures. It's unclear why the instrument was altered, unless the scores didn't meet the 50% change and threshold score. In this instance, the instruments used for measuring the outcomes likely turned out to be invalid and unreliable.

Refining the inclusion and exclusion criteria in the pilot study before constructing the trial's protocol, and arriving at an expert consensus in validating the instrument on a specific CFS cohort, or constructing a new patient reported outcome instrument, ensures crises like these are averted.
SSMC (Standardised Specialist Medical Care) alone

**Treatment:** Sessions with an experienced CFS doctor. Participants were given leaflets on the condition and given medication for symptomatic conditions eg. insomnia, pain, mood.

**Sessions attended:** Average 5 (3-6) over 52 weeks.

**Compliance:** Face-to-face sessions.

**Lack of sessions:** the same amount of therapy sessions should be offered like the other treatment arms. On average 5 sessions compared to 16 sessions sets a qualitative difference in treatment satisfaction and outcomes.

**Confounders:** CFS specialists are likely to treat CFS and other forms of therapy such as physiotherapy for those who break a limb, diary entries for mental well being, encourage physical activity, prescribe medications with drowsy side effects etc. and should be controlled- to not bias the analysis.

**Control condition:** for the SSMC to be the control condition, current evidence of no difference in treatment efficacy on CFS participants should be offered for the 15 treatment sessions.
SSMC + Adaptive Pacing Therapy

Treatment: 1. Sessions with an experienced CFS doctor. Participants were given leaflets on the condition and given medication for symptomatic conditions eg. insomnia, pain, mood. And...
2. Self-reports. Pilot manual- endorsed by Westcare and Action for ME administered by an occupational therapist: help participant to plan and pace activity and provide the best conditions for natural recovery.

Theory: Envelope theory of chronic fatigue syndrome: participants’ perceived energy envelopes.

Evidence: No empirical evidence.

Sessions attended: Average 13 (12-15) over 52 weeks + 3 (3-4) SSMC.

Compliance: Self-reports and face-to-face sessions.

Conflicts of Interest: 1. ME pilot groups are for ME conditions and don’t sample randomly the population of interest. 2. Occupational therapists are specialists and are unlikely to comply with just trial regimens.

No empirical evidence: empirical evidence on self-reports (including endpoints) can be developed into a patient reported outcome instrument with a self-report component specific to the study. This ensures scientific validity and reliability that can be assessed, quantified, and empirically deduced.
SSMC + Graded Exercise Therapy

**Treatment:** 1. Sessions with an experienced CFS doctor. Participants were given leaflets on the condition and given medication for symptomatic conditions eg. insomnia, pain, mood. And...
2. Exercise therapy delivered by a physiotherapist and one exercise physiologist based on manuals used in previous GET RCTs.

**Theory:** Exercise intolerance theories of chronic fatigue syndrome-theory assumes reversible physiological changes of deconditioning and avoidance of activity perpetuates further inactivity.

**Evidence:** Previous GET RCTs on CFS patients.

**Sessions attended:** Average 13 (12-14) over 52 weeks + 3 (3-4) SSMC.

**Compliance:** Face-to-face sessions.

**Lack of physiology tests** eg. weight loss, to validate a participant was undertaking eg. 30 minutes of light exercise five times a week.

**Lack of a patient reported outcome instrument** to account for neutral responses (eg. I walked...); account for investigator effects; and deduce reported outcomes into empirical contexts.
SSMC + Cognitive Behaviour Therapy

**Treatment:** 1. Sessions with an experienced CFS doctor. Participants were given leaflets on the condition and given medication for symptomatic conditions eg. insomnia, pain, mood. And...
2. Therapy sessions with a clinical psychologist or a nurse therapist.

**Theory:** Fear avoidance theory of chronic fatigue syndrome as being reversible and that cognitive responses and behavioral responses can perpetuate fatigue.

**Evidence:** Therapy manuals in previous CBT trials for CFS.

**Sessions attended:** Average 14 (12-15) over 52 weeks + 3 (3-4) SSMC.

**Compliance:** Face-to-face sessions.

---

The timeframe (endpoints) and recommended frequency in CBT for therapeutic benefit is unclear. Timeframes and achieved milestones in efficacy, should be set by the clinical psychologists' assessment of cohort (n=161) baseline characteristics- based on CBT evidence of timeframe and efficacy for a population with CFS, or therapeutic successes of CBT CFS in case studies. Milestones for each participant are compared to projected cohort milestones, and treatment regimens are adjusted accordingly to keep inline with cohort milestones. That way, therapeutic outcomes can be standardised, and can assist or validate trial outcomes.
# Primary Outcomes

**Primary outcome measures**

1. Is APT and SSMC more effective than SSMC alone in reducing (i) fatigue, (ii) disability, or (iii) both? **no, both**
2. Is CBT and SSMC more effective than APT and SSMC in reducing (i) fatigue, (ii) disability, or (iii) both? **yes, both**
3. Is GET and SSMC more effective than APT and SSMC in reducing (i) fatigue, (ii) disability, or (iii) both? **yes**
4. Are the active rehabilitation therapies (of either CBT or GET) more effective than the adaptive approach of APT when each is added to SSMC, in reducing fatigue, in reducing physical disability? **yes**
5. What are the relative cost-effectiveness and cost-utility of these treatments?

<table>
<thead>
<tr>
<th>1. APT + SSMC vs SSMC alone.</th>
<th>2. CBT + SSMC vs APT + SSMC</th>
<th>3. GET + SSMC vs APT + SSMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue: $p = 0.99$; $p &gt; 0.05$</td>
<td>Fatigue: $p = 0.0136$; $p &lt; 0.05$</td>
<td>Fatigue: $p = 0.0013$; $p &lt; 0.05$</td>
</tr>
<tr>
<td>Physical function: $p = 0.89$; $p &gt; 0.05$</td>
<td>Physical function: $p = 0.0342$; $p &lt; 0.05$</td>
<td>Physical function: $p = 0.0025$; $p &lt; 0.05$</td>
</tr>
</tbody>
</table>

### Statistical Method - Inappropriate

Mean differences don't reflect outliers or direction of skews, as opposed to a regression analysis that can meaningfully reflect robust findings within and between individual scores.

Mean scores are generally taken when there are too many outliers and there is no noticeable pattern in the data. In this instance, a rank method eg. Kruskal Wallis has been added to rank scores, and possibly curb outliers towards a recognisable pattern. This is often known as "p-hacking" because it gives the statistician control in manipulating the outcomes in favour of a hypothesis.

Mean differences including SSMC scores, assume SSMC to be the control condition (when it wasn't). In this instance SSMC should be taken out of the analysis. It's unclear which measurement scales and data were collected in the SSMC condition, aside from treating adverse events that weren't included in the outcomes.
Statistical Method

**Power calculations (reduce type I & II errors)**

The recommended minimum sample size in each trial arm for an **RCT is n = 200; total n = 800**

An example of why n = 200 should be used for each trial arm can be seen with risk tables eg. 2x2 contingency tables. When calculating the Odds Ratio, Relative Risk etc. each quadrant has a minimum sample size of n = 50, and each quadrant determines power rather than each group size.

**Incorrect sample sizes**

For RCT should be n = 800

**Dubious power calculations**

No evidence APT is the least effective therapy.

**Incorrect study design**

For n = 150 per treatment arm, a different study design (not RCT) eg. cohort (without randomisation) is preferrable.
statistical analysis plan was finalised, including changes to the original protocol, and was approved by the trial steering committee and the data monitoring and ethics committee before outcome data were examined.

It's unclear which committee members approved or which amendments were made or the relevance in this claim. There is also no mention of where the data was stored and with whom before it was examined.
It's unclear which committee members approved or which amendments were made or the relevance in this claim. There is also no mention of where the data was stored and with whom before it was examined.
Avoiding measurement biases
(RCTs & Cohort studies)

- **Measurement BIAS:** Poor measurement of the exposure, outcome or confounding factor(s) biasing frequency or effect.

- **Measurement ERROR:** Systematic error in the estimate of frequency or effect due to measurement errors or misclassification of:
  - 1. Study factor(s)
  - 2. Outcome factor(s)
  - 3. Confounding factor(s)

---

1) Exposure and confounders (usually) are measured blind to outcome.
2) **ERROR** in outcome measurement is the major risk. Ensure outcome measured is **BLIND** to exposure/intervention: blinding, double blinding, placebo controls in RCTs.
3) Minimise **CHANCE** by addressing error.
4) Cluster RCT: Groups of subjects randomised, so further randomisations potentially avoids measurement biases.
5) **NON-DIFFERENTIAL MISCLASSIFICATION** for OUTCOME: Underestimate effects of exposure on outcome. **RR biased towards null.** Error is **SAME** in E & UE.
6) **DIFFERENTIAL MISCLASSIFICATION** for OUTCOME: Measure with error and outcome is **DIFFERENT** in exposed and unexposed subjects. Unpredictable bias, might overestimate or underestimate the effect of exposure. **RR biased either toward or away from null.**

**RCT and COHORT** measure outcomes **OBJECTIVELY** and **BLIND** to exposure status. Blinding will help prevent differential misclassification of outcome.
## Misclassification effects

<table>
<thead>
<tr>
<th>Misclassification (Differential)</th>
<th>Misclassification (Non-differential) – INFORMATION BIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differential misclassification of exposure (case-control study) or outcome (cohort study) can bias the measure of effect <strong>towards OR away from the null.</strong></td>
<td>Identifying an error is non-differential. Bias towards null.</td>
</tr>
<tr>
<td>Differential RR &gt; 1 (not lower). If RR = 1.6 – can’t tell which way the error goes, towards null or not, bias which way don’t know.</td>
<td>Bias towards null. Risk difference = 0. RR = 1. No effect.</td>
</tr>
<tr>
<td>Differential error – no explicit error. Likely to overestimate? Often can’t tell.</td>
<td>Two study groups, one more like the other.</td>
</tr>
<tr>
<td><em>Over-counting</em> outcomes in <em>exposed</em> group but not the unexposed group. Bias away from null.</td>
<td>Non-differential misclassification of exposure or outcome (nearly) <strong>ALWAYS biases the measure of effect towards the null.</strong> (confident)</td>
</tr>
<tr>
<td><em>Under-counting</em> outcomes in <em>exposed</em> group but not in unexposed. Bias towards null.</td>
<td></td>
</tr>
<tr>
<td>Under-counting outcomes in <em>unexposed</em> group but not the exposed group. Bias away from null.</td>
<td></td>
</tr>
</tbody>
</table>

Any misclassification of confounders can bias the measure of effect towards OR away from the null.
# Measurement Error Types

<table>
<thead>
<tr>
<th>SYSTEMATIC ERROR</th>
<th>RANDOM ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Error repeated systematically with all subjects in study. Causes BIAS. E.g. Scales broken.</td>
<td>- Feature of randomness in world we live in. These are NOT big biases.</td>
</tr>
<tr>
<td>- People not wanting to own up to a systematic fault.</td>
<td>- Leads to <a href="#">more variability</a> in measurement without a recognizable pattern. E.g. weight can <a href="#">under/over estimate</a>.</td>
</tr>
<tr>
<td></td>
<td>- Many biological variables vary day to day in random way. E.g. High BP on day of measurement.</td>
</tr>
</tbody>
</table>
Measurement Biases

Self-reports: Differences in reporting patterns (over and under reporting) can lead to differential effects and bias the estimate of effect in either direction.

Dubious allocation concealment: It's unclear when and how often computer-generated randomisation occurred.

Unblinded: Blinding for some onsite investigators and not for others is the same as some aware of what's going on and some not. This situation is still regarded as unblinded to the outcome. Influences information reporting and collection (differential misclassification of the outcome). Bias measurement of effect in either direction.

Research officers: No indication research officers were trained in each respective trial arm, and measures of good agreement (reliability) and valid measures were used across specialists and across multi-centers. This can lead to different measures (differential effect) and bias the result in either direction.

Accuracy of information: Can't be verified (non-differential effect). Outcomes could be unrelated to what's been tested. No validation of self-report or evidence of disagreement in self-report (unclear which is correct). This biases the outcome towards the null hypothesis.

Different clinics: Difference between baseline cutoffs (determined by one center) and endpoints (determined by another). This difference is assumed the same across groups (non-differential effects) and biases the outcome towards the null.
Confounders

**Lack of information:** Presence of exposure may influence answers to the patient's history (differential effects) leading to *measurement error in confounders* and bias in either direction.

**Confounder levels may change over time:** *Behaviour modification.* Baseline measures may not reflect actual levels of confounding throughout the study (differential effects) bias either towards or away from the null.

**Inaccurate past measurements:** Can't verify without a gold standard or a reference standard (non-differential effects) bias towards null- *underestimates risks arising from the exposure.*

**Information about other variables:** *Lack of data on potential confounders* during the screening process (non-differential effects) can bias either towards or away from the null.
PACE Bias Timeline

Registered study design: RCT (highest evidence)
ISRCTN54285094
Funder requirements.

8 March 2007
Intervention problems
- Varied administration; compliance issues.
- Complaints; escalating conflict of interest.

18 March 2005
Published protocol
Journal: BMC Neurol.
Protocol problems.

28 Nov 2008
Data collection errors
- Missing data.
- Logbook problems.
- Compliance problems.

18 Feb 2011
Data analysis errors
- Sampling errors.
- Inconclusive results.
- Altered patient outcome measures eg. continuous
- Data trimming; P-HACKING.

2013 onwards
Published outcomes
Journal: Lancet
Outcome problems.

2015 onwards
Data requests denied
- Missing follow-up data.
- Consulting ME groups (trial participants?) for publication
  (Conflict of Interest).

Further publications
- Exposes further methods problems.
- Exposes further COIs.
- Exposes further statistical problems
eg. statisticians still apply RCT
design analyses and other
inappropriate analyses. Escalating
FILE DRAWER PROBLEM.

Recruitment

Trial participants exposed to the protocol.
- Screening problems: sign up frenzy.
- Altered eligibility requirements.
- Randomisation and stratification problems.
- Conflict of Interest eg. ME groups share experiences.

Criticisms: ME groups, academics
- Trial participants speak out.
- Data requests denied.
PACE Trial: In Summary

- Even though RCTs are considered a high form of scientific evidence- its inappropriate for most interventions.
- Systems for data storage, encryption, collection and analyses in part determine seamless operation, reduce compliance issues, missing data, help with auditing, and effect the outcomes across treatment centers.
- Patient groups eg. ME shouldn’t be involved in designing intervention manuals. This may prompt trial sign ups from those diagnosed with ME leading to skewed samplings, COIs, selection and measurement biases.
- Patient reported outcome instruments should be validated on a pilot group with pre-determined inclusion and exclusion criteria BEFORE the trial- so the inclusion and exclusion criteria can be refined, the correct population sample is recruited, and to safeguard from potential biases and limitations arising in the trial.
- A new patient reported outcome instrument should be designed and itemised appropriately if present PROMs aren’t sufficient for answering the hypotheses across treatment arms.
- Duplicate screenings and duplicate samplings can lead to selection and measurement biases. Statistical methods eg. bootstrapping can randomise re-samplings if duplicate screenings/samplings are unavoidable.
- Publishing the trial protocol before or during the trial, exposes all treatment arms and prime participants and unblinds those involved with expectations eg. ADT evidence is weak, therefore ADT doesn’t work.
- Each treatment arm should have evidence-based endpoints in administering the intervention in a timeline and frequency most beneficial for that particular population. If there is no specific evidence, then program interventions eg. rehab guidelines, health boot camps, could be used as evidence in extracting endpoints.
- If quantitative data don’t conform to the parameters set for statistical analyses, then qualitative methods eg. grounded tools on self-reports can be used to discuss the results, and the inclusion and exclusion criteria refined for future trials.
- Further participant characteristics could be explored in follow-up sessions for stratifying groups by prevalence of non-CMF (comorbid) conditions in designing future trials- including physiological measures eg. hormonal, endocrine, weight, diet, genes, EEG to aid in selection, monitoring, validate or refute outcome data.
What next?

ACCESS TO DATA - Precisely map and trace problems to the center, researcher, participant.

RECTIFY or RETRACT findings.