M.E. Analysis - Evaluating the results of the PACE study



a project supported by Phoenix Rising

# 3. It has not been proven that we should continue to use GET and CBT with all ME/CFS patients. (Chalder Fatigue Scale)

# There are 11 questions in the Chalder Fatigue Scale, as used in the PACE trial:-

- Do you have problems with tiredness?
- Do you need to rest more?
- Do you feel sleepy or drowsy?
- Do you have problems starting things?
- Do you lack energy?
- Do your muscles have less strength?
- Do you feel weak?
- Do you have difficulty concentrating?
- Do you make slips of the tongue when speaking?
- Do you find it more difficult to find the correct word?
- How is your memory?

#### and asks if the factor is

- less (of a problem) than usual
- no more than usual
- more than usual
- much more than usual



In the PACE trial patients were asked to respond comparing their present state with when they were last well.

**Bimodal scoring** allocates 0, 0, 1 and 1 to each of these answers, which would give a score of 0 to someone who was perfectly healthy, down to a maximum of 11. In effect, Bimodal scoring simply counts how many of these factors are affected by the illness.

Continuous, or Likert scoring allocates 0, 1, 2, and 3 to these answers, with scores from 0 to 33.

**In our survey** (more details in section 6) many people commented on the difficulty of quantifying the difference between more and much more. A factor that would be considered by any normal healthy person to be severe can get even worse in a bad spell. They considered that at least one more level, extreme, scoring 4, would have been appropriate. There is a study by Goudsmit et al. looking at the limitations of both scoring methods.

# The PACE Trial

Originally, as shown in the protocol, the trial was to have used bimodal scoring, specifying that a positive outcome would be a 50% reduction in fatigue score, or a score of 3 or less. A score of 4 or more would show abnormal fatigue.

Later, after the trial had started, the scoring was changed from Bimodal to Likert/continuous, and a positive outcome was re-defined as a score of 18 or less. This change has led to unease among some scientists who regard a change of an agreed protocol as being acceptable only under very special circumstances.

### Difficulties with the Continuous Scoring

It puzzles us how any patient could say that their symptom had become less of a problem than it had been when they were fully healthy. That would mean that a score of zero for any component is very unlikely (although from other studies there are occasional scores which suggest that patients must have scored zero on some components, and we have found subsequently on our own survey that there are rare circumstances that permit this to happen). It would make more sense if patients had been asked to compare themselves with healthy friends of the same age, but this is not the case.

One of the entry conditions was that patients had to score 6 or more under the bimodal system - i.e. they had 6 or more factors that gave them difficulties. Subsequently 18 was used on the Likert scale as a boundary for returning to normal health.

In a relatively small study, looking at cortisol levels in CFS, the Likert scale was used to assess fatigue levels in 40 patients and 40 controls. The patients scored  $24 \cdot 4 \pm 2 \cdot 9$ , and the healthy controls  $7 \cdot 6 \pm 2 \cdot 3$  (which of course means that the healthy controls generally marked 3 or 4 of the items as being less of a problem than usual). A much larger Norwegian study gave a value of 11.2 as a score for healthy adults. Clearly, the use of a score of 18 to mark "normal function" is, at best, overgenerous, but the differences between the results of different studies prompts the question "Is the Likert scale fit for purpose?"

#### Start with the patient below:-

| Anne M Elass  | less than usual in<br>strength | no more than usual | more than usual | much more than usual | Score |
|---|--------------------------------|--------------------|-----------------|----------------------|-------|
| Do you have problems with tiredness?                    |                                |                    |                 | $\langle$            | 3     |
| Do you need to rest more?                               |                                |                    |                 | $\checkmark$         | 3     |
| Do you feel sleepy or drowsy?                           |                                |                    |                 | $\checkmark$         | 3     |
| Do you lack energy?                                     |                                |                    |                 | $\checkmark$         | 3     |
| Do you have trouble starting things?                    |                                | V                  |                 |                      | 1     |
| Do your muscles have less strength?                     |                                |                    | $\checkmark$    |                      | 2     |
| Do you feel weak?                                       |                                |                    |                 | V                    | 3     |
| Do you have difficulty concentrating"                   |                                |                    | $\checkmark$    |                      | 2     |
| Do you find it more difficult to find the correct word? |                                | $\checkmark$       |                 |                      | 1     |
| Do you make slips of the tongue when speaking?          |                                | $\checkmark$       |                 |                      | 1     |
| How is your memory?                                     |                                |                    | $\checkmark$    |                      | 2     |
|   |                                |                    |                 |                      | 24    |

Below is a series of graphs linking the Bimodal scoring system (which simply counts how many items gives problems) to the Likert scoring.

With a score of 24, and with 8 items giving problems, this yellow box is where our patient would be on the graph shown on the right.

If all eight of these problem areas scored 3 points each, her score would have been 27 – the worst possible Likert score for a person whose bimodal score is 8.

The best possible score for a person with a bimodal score of 8 is 16, which would involve three scores of zero.

That range is shown in yellow in the diagram below.





The diagram below shows the complete match of possible Bimodal with Likert scores. Some of them include possible scores of zero, and these have been coloured white.



In order to qualify for the trial, patients had to score 6 or more under the Bimodal scheme.

These would be to the left of the green line.

Originally, significant recovery was defined as halving the Bimodal score or better, which would be any patients ending up to the right of the black vertical line, and for the worst affected, ending up between the black and green lines (half of a score of 11 or 10 is 5).



Subsequently patients were defined as returning to normal functioning if their Likert score dropped to 18 or better (i.e. above the black horizontal line, shown in the diagram on the left).

The green shaded triangle shows where patients who would now be classified as returning to normal health on the later criteria, but who would not have reached the initial recovery target.

Any patient in the triangle to the left of the green line would simultaneously be both eligible to be included in the trial, but being defined as being within the range of normal health. A study published in 2008 by Goudsmit on 24 patients did measure both bimodal and Likert scores as well as whether they rated mild, moderate or severe with ME/CFS. They kindly let us have those scores: they are plotted on the graph below, and show just how complex the situation can be. We have added the results of our own survey to it (described in more detail in conclusion 6).



The overlap between mild and moderate, and between moderate and severe is surprising, as Goudsmit has noted before.

It is clear that there may be severe cases with scores equal to mild cases (here we have an example at 25).

Notice that five of the people in our survey would qualify for the original trial, as they are to left of the green line, but at the same time would be described as "recovered" - within the normal health scores - as they are above the horizontal black line. How did that possibility escape the notice of the authors and reviewers? It wasn't missed by Sarah Feehan who subsequently wrote online to The Lancet on behalf of the Liverpool ME Support Group.