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. (details)

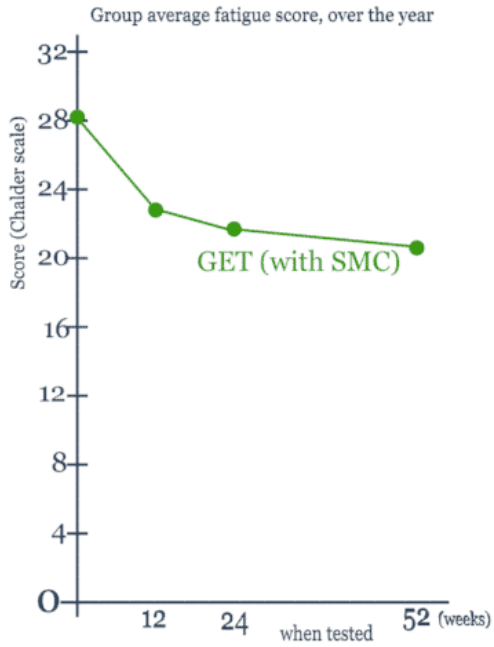
The PACE trial had 4 groups, each with around 160 patients: the members of one group had sessions of specialist medical care about 5 times over the year: the other three groups saw had specialist medical care 3 or 4 times, but also had an additional 12 to 15 sessions of Graded Exercise Therapy (GET), or of Cognitive Behaviour Therapy (CBT), or of Adaptive Pacing Therapy (APT). Their progress was tracked in a variety of ways and was measured at the start, after 12 weeks (half-way through), after 24 weeks (at the end of the courses) and after 52 weeks.

First let us look at the way that GET is claimed to moderately improve the fatigue levels of patients. We will focus on their assessment of fatigue using the continuous/Likert Chalder scale, which runs from 0 points to 33 points, where 33 points represents utter exhaustion (in the true sense of the term - not just feeling very tired). Patients were assessed at the start of the trial, after 12, after 24 and after 52 weeks. A drop in score therefore represents a drop in fatigue - an improvement in health.

This is the graph presented in the PACE report that shows the drop in fatigue levels of the GET group over the course of a year.

Doesn't it look impressive?

But the vertical scale should really go from 0 to 33, which is the full range of the Likert Chalder scores.



With the full vertical scale, the change looks less impressive, but it is still a small drop in fatigue, so may be worthwhile.

Remember that a score of 33 would be for someone who is at the maximum score of fatigue in this scale. Many patients would have got a score of 33.

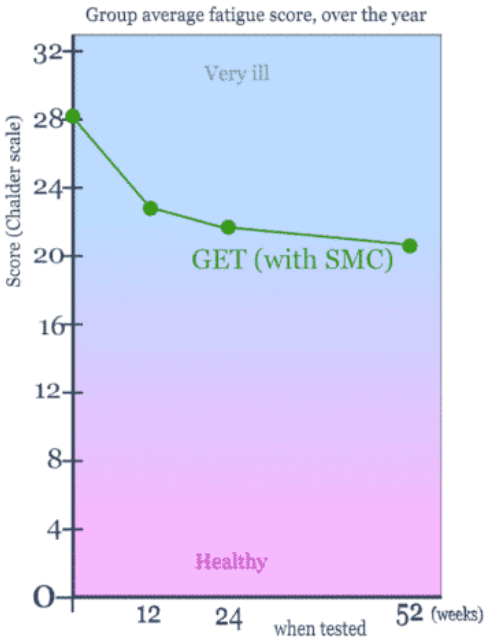
Remember also that patients must have been able to attend the centres and follow the therapies over the period of the trial: there will be patients with greater fatigue who were unable to participate in this trial.

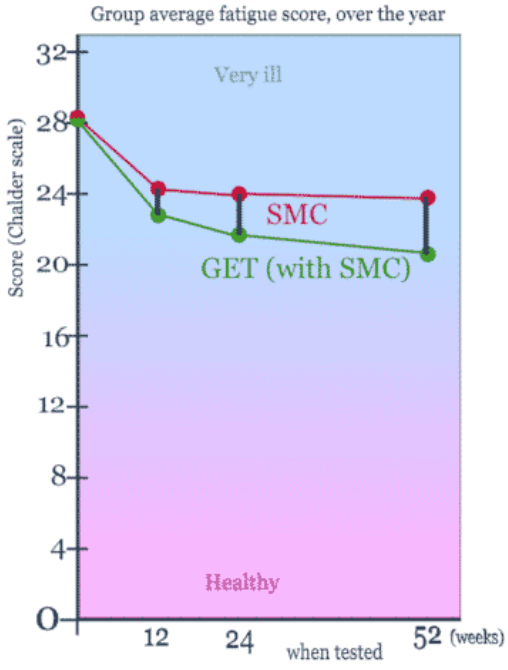
Here we shaded the graph: the colour blue for scores that represent people who are very ill, and pink for scores that represent people who are fit and well.

It is clear that the drop in fatigue level is not enough to lower it to a healthy score.

But this graph represents the average score of the patients who had Graded Exercise Therapy and Specialist Medical Care over the course of a year.

We need to compare it with the group that only had Specialist Medical Care.





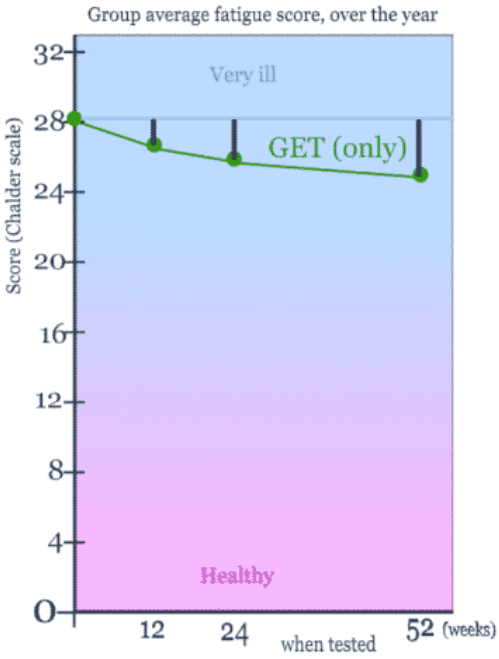
The line in red shows the drop in fatigue levels of the group that only had Specialist Medical Care.

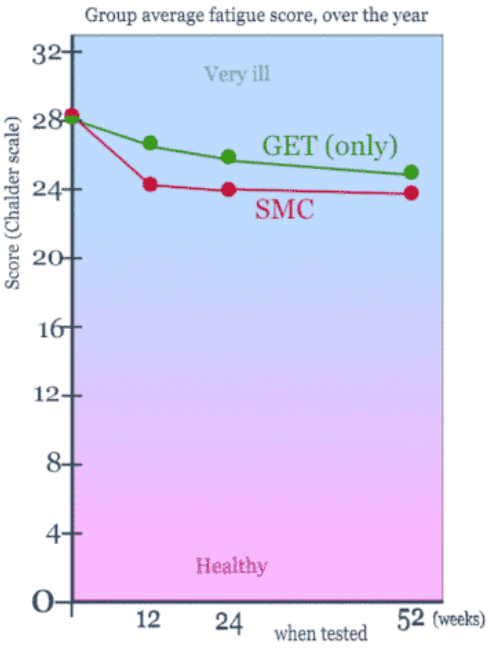
So if we are to consider the value of the Graded Exercise Therapy, we need to look at the difference between these lines (shown as black vertical lines).

We can regard this as the effect that GET has added on to the situation.

This is now the plot of that extra drop in fatigue that the Graded Exercise Therapy group saw.

This really isn't much, especially when we consider that it is the result of a whole year's commitment to graded exercise.





It is clear that adding a year's graded exercise has not even improved the fatigue by very much.

In fact, the very first drop in the red graph is greater – after just one or two sessions of Specialist Medical Care.

So why did the PACE report emphasise the improvement under Graded Exercise Therapy, but not the importance of having Specialist Medical Care (of the sort that is not usually given by a G.P.)?

Remember too that this (red) group had between 4 and 6 sessions with a specialist, which is more than most patients ever get.

Both of these graphs are flattening out. In the case of the red SMC line, the effect is obvious: the maximum improvement occurs in the first 12 weeks.

The green line, representing the average extra benefit produced by the GET sessions, a smaller benefit than the SMC produced, is also flattening out. It fits a classic exponential decay pattern, with a maximum benefit of a 3·5 point improvement after many years. There are similar graphs for GET and CBT for each of the other assessments used – section 3-Further Graphs shows them.

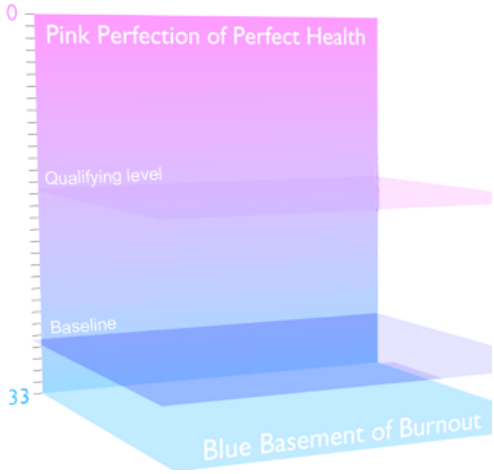
In most cases, the scores of the patients who only had Specialist Medical Care showed more than half of the overall improvement in the scores of the GET+SMC and of the CBT+SMC groups, and to use that combined score to reach targets gives a false impression of the effectiveness of GET and CBT. It is like setting the minimum height standard for entry into the police force at 5 feet 10 inches, then allowing candidates to stand on a box over three feet high. Interestingly the patient satisfaction with SMC (50%) was below the level of satisfaction with GET and CBT (82% or more), which suggests that patients overestimated the relative value of GET and CBT.

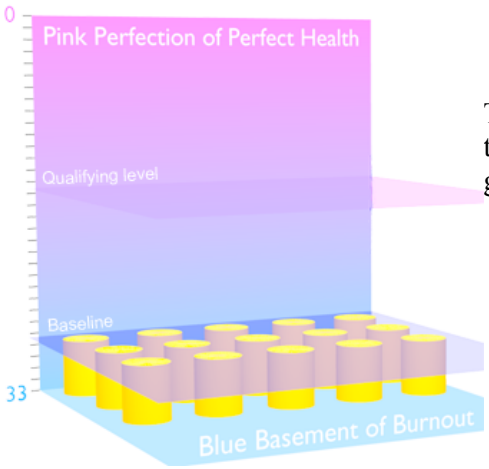
The next set of diagrams looks at the consequences of the average improvement being so low, and of the baseline score of the group (28·2) being so close to the end value of the scale (33). In this situation, it is very similar to the situation of describing the average salary of adults in the UK mentioned in conclusion 2 where a few high earners lift the average (the mean) to an unrepresentative value. We do not have access to the raw data, so the example here is an illustration of the situation rather than an actual example.

Here we have reversed the direction of the graph, so improving is shown by moving upwards. The blue and pink colours still show ill/well.

The level in blue marked as Baseline is the average score of the patients in the GET group at the start of the trial.

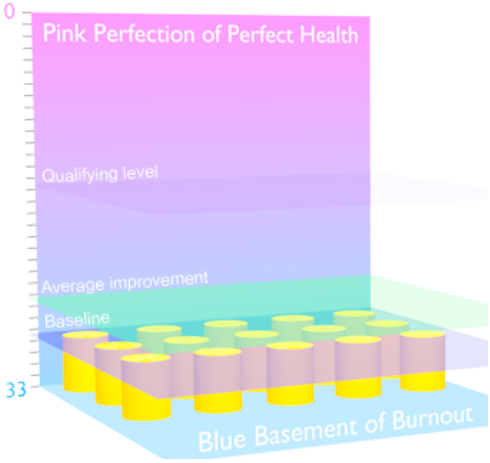
The pink "Qualifying" level is an estimate of the score needed to be judged to be ill enough to qualify for taking part in the trial. (It is not possible to be exact, as the scoring system was changed between determining entry qualifications and measuring progress.)

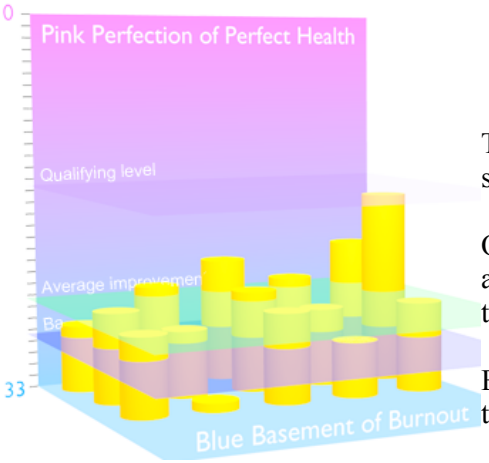




These fifteen yellow cylinders represent fifteen of the patients in the GET group whose scores were equal to the average in the group.

This new green level is the average improvement in the group brought about by following the graded exercise therapy for a year, minus the improvement measured in the group that only had specialist medical care.





These yellow cylinders show what could be happening to their scores, consistent with this weak average improvement.

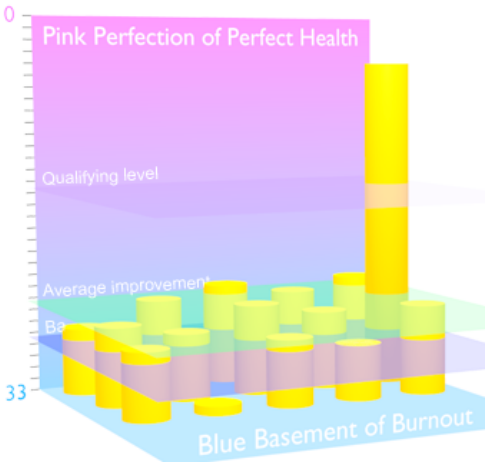
Only three of them have improved by much more than the average, and in order to balance their scores, many more of them have had to show only a tiny improvement, and one in fact has done worse.

Even the best of them though is still ill enough to qualify for the trial all over again, after a whole year of GET.

Here we show what happens if one patient improves enough so that he or she no longer is classed as ill enough to enrol in the trial. The low group average would mean that all of the others would have to have had shown very little improvement or actual deterioration to balance this.

This is the problem. Small average improvements in the whole group can be caused by just one person in 15 doing really well, masking the fact that the remainder hardly improve or deteriorate.

What we do not know is how many in the group, if any, showed a reasonable benefit. We do know that number must be small.



As you can see, with such a small average improvement, the consequence of just 2 in 15 patients making a reasonable amount of progress with GET is that the remaining 13 must have made very little progress, or even deteriorated. If just one patient in 15 were to have returned to "healthy", then the remaining 14 would have had to made very little progress or deteriorated in order to keep down the average improvement.

This situation is similar for the other measures, and for CBT. For that reason the case for applying GET or CBT to all patients with ME/CFS is neither an efficient use of resources, nor an effective use of patients' time and efforts.