

# M.E. Analysis – Evaluating the results of the PACE study

a project supported by Phoenix Rising

## 2. Future studies should use additional measures of average and variation (more)

**If there was one billionaire** living in a small village of 1000 people, would it be fair to say that the average person in the village was a millionaire? If all the money was shared out equally, then each person in the village would be a millionaire (the usual way of working out an average). They would probably be quite happy with that, but it would be very unlikely to happen. So does it give a fair picture?

**This way of calculating an average**, by adding everything together and sharing it all out, is called the mean (or to be precise, the arithmetic mean). There is also a calculation that gives a measure of the variation between the values, and that is called the standard deviation.

**For adult males in the UK**, the mean (average) height is 5'10" with a standard deviation of 3". We add and subtract the standard deviation from the mean to get a range of 5'7" to 6'1" to give us an indication that two-thirds of men in the UK will be within those heights. That is because the distribution of heights is *Normal* - there are the same numbers and variations of height on either side of 5'10" (a less confusing term than *Normal* is *Gaussian*).

**Consider these numbers** - 10, 10, 10, 10, 10, 10, 10, 10, 10, and 0. The mean of these numbers is 9 and the standard deviation is 3.2. Suppose this was the number of fingers that each of ten people had. If we used the mean and standard deviation to give us an idea of average and range, we would say that on average the group had 9 fingers, and the normal range was from 6 to 12 (3.2 either side of 9).



But of course you have already dismissed this: you know that it would be daft to say that an average person has 9 fingers, and most people have between 6 and 12.

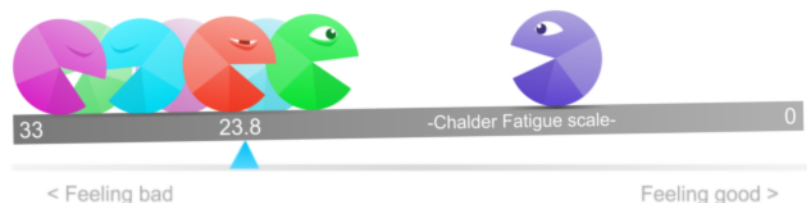
Using the mean and standard deviation here is wrong, because, unlike with men's heights, the distribution of values is very unbalanced.

**The data used in the PACE trial is also very unbalanced** (as it is in many other studies). In circumstances like that we have to use other methods to give a fair and clear picture. It isn't statistics that is at fault, but a poor choice of method.

**Regrettably, we tend to take statistics for granted**, especially when it comes to averages, but even averages are much more complex and potentially deceptive than would appear. When we talk about an average, we normally think of the arithmetic mean, but we do have other choices. The problem with the mean, and even more so with the standard deviation, is that they are greatly influenced by unusual values (such as the billionaire).

For what is known as Normal/Gaussian distributions (such as people's height, weight, IQ) the pattern is understood, and we know that if we add and subtract the standard deviation from the mean, we get a range of values for the middle two-thirds of people, as stated above for heights. But most collections of data do not fit this distribution. Many are very lopsided, or skewed, with very long tails on one side (such as income distribution - there is a large clump of people on typical wages, but a very very long tail of a small number of people getting very large incomes).

**This is a problem with the average/mean levels reported in the PACE trial.** The data is heavily skewed, and that moves the mean to an inappropriate level. The mean is like the balance point on a seesaw. If one of the arms is very short, and the other relatively long, it takes a lot of people on the short end to balance one out on the long end. If one patient improves by quite a lot, it takes many patients remaining as they are to keep the average down. We do not have access to the raw data, but the diagram below represents, in a simplified form, what is happening - it matches both the mean and the standard deviation of the results after 52 weeks for the group that had only specialist medical care. We know that some patients did well, but look what that entails for the others.



**If a distribution is well-balanced on either side,** then using the arithmetic mean as an average is fair and clear. But when distributions are skewed, or the measurements themselves are not fairly linear, other types of averages are more appropriate (and mathematicians have many others). The one that should also be quoted in situations like this is the median - the middle value - and this is done in good quality reports where distributions are skewed.

We are pleased to see that the recent study by Nacul et al. (mentioned in conclusion 1-details), undertaken as part of the ME/CFS Observatory Research Programme, and looking at the functional status of people with ME, specified both sets of averages in table 2 on the SF-36 scores: the medians (25.0) were consistently below the means (30.1), as is typical of a strongly skewed set of data.

Using the mean as an average in a situation like this consistently overstates the effectiveness of the treatments or therapies for the majority of the patients, but more importantly, it draws attention away from the clustering at the bottom end.

For a fuller explanation, please look at the 2-details page.