Appendix: Statistical Analyses used in examples [posted as supplied by author]

Sample variances of the active and placebo arms (s_1^2 and s_2^2 respectively) were calculated by squaring sample standard deviations (s_1 and s_2) where these were reported. If standard deviations were not reported, estimates were derived from the standard error of the mean (using $s = SE*\sqrt{n}$) or 95% confidence limits (using $s = \sqrt{n}*(UCL – LCL)/3.92$).

An F statistic was calculated for the ratio of variances from the two groups to estimate the probability of the observed difference in variances arising by chance. Distribution curves, assuming normality, were constructed on the basis of estimated mean levels and standard deviations of change in outcome for the treatment and placebo groups.

Estimates of the magnitude of variability in treatment effect were calculated where appropriate. The mean treatment effect was calculated from the difference in mean change between the two groups. The variance in treatment effect was estimated from the difference in variances between the two groups (assuming a greater variance in the treated group). The square root of this was taken to estimate the standard deviation of treatment effect and distribution of effect for 95% of the population.

The method we have used relies on a number of assumptions. The data used should be approximately normally distributed and measurement variability should not be level dependent. If these assumptions do not hold on the natural scale it may be necessary to examine data transformed to another scale such as the log scale. The method also assumes there is no substantial correlation between treatment effect and the change that would have occurred without treatment. Provided the measurements are done over a short time frame (where it is not expected that patient's true mean value would change substantially in the absence of treatment) then this assumption will often appear reasonable.

The last assumption may be less tenable for change in measurement over longer periods of time where a patient's response to treatment will often be correlated with the change that would have occurred without treatment. Where this takes the form of a negative correlation, the treatment effect is greatest for those who would have progressed the most without treatment. If the negative correlation is strong, treatment may actually negate some of the natural variation in measurement, resulting in a lower variance of change in those on treatment. If the variance of change is sufficiently low in the treated group, monitoring may become unnecessary. In these cases the expected change for the individual can be predicted from the mean change found for the group on treatment.