

The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study

ABSTRACT

The use of percentage change from baseline is not recommended in statistical analysis. Those who want to report this statistic should use a different technique, such as ANCOVA, and convert the results to corresponding percentage changes by using mean baseline scores.

INTRODUCTION

The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study

Introduction A randomized trial is often used to assess the continuous outcome at both the beginning and end of treatment, as demonstrated by studies on pravastatin for hypercholesterolemia, exercise and diet for obesity patients, and acupuncture for pain in athletes with shoulder injuries. In statistical analysis of trials that evaluate outcomes, there are four methods for including data in the analysis: using the baseline score alone and entering only the post-treatment score (as described in "POST"); analyzing the change from baseline either through absolute differences ("CHANGE") or through a percentage change to baseline ("FRACTION") via constructing — by means of least common mathematical approach—a regression model that adjusts the outcome of the treatment by the base score. Figure 1 shows how each method could be used to analyze the outcomes of these treatments with the same as an example showing how many people who had been treated In some trials, the outcome can be measured several times after treatment using a design known as "repeated measures". These trials can then be analysed using an area-under-curve or summary statistic such as momentum" or lma (repetitive measures analysis of variance) and more complex linear estimation methods. These methods are particularly useful when post-treatment scores show cyclic events (e.g. quality of life in late stage cancer patients) or when it is important to assess how treatment affects effects. By examining the statistical characteristics of each method, one can determine which one to use. An important factor to consider is that a good statistical method should have fewer false negatives () and be low enough to reduce the rate of errors. The statistical power (1-1) is used to measure statistical tests, with power usually being set at either 0.8 or 0.89, and data must be obtained in quantities such as the required number of evaluable patients. Economists describe an efficient method that requires less data to provide varying levels of data. The four methods POST, CHANGE, FACTION and ANCOVA have been the subject of research by statisticians for some time. In this paper, I will compare them using a hypothetical trial and demonstrate their statistical power with examples that are useful in clinical research.

CONCLUSION

The percentage change in outcome from baseline during a randomized trial provides clinically relevant information to both patients and clinicians, which is likely why researchers studying hot flashes or the effects of different chemotherapy regimens on quality of life report this statistic. Using percentage change from baseline is not an efficient method for reporting statistical data, as it does not correct for imbalance between groups at baseline and can result in a non-normally distributed statistic. Trialists

should use another technique, such as ANCOVA, to test significance and calculate confidence intervals. They should then convert their results into percentage difference by using both baseline (8000) and post-treatment scores. Results here restate earlier data that suggests previously reported data indicating that ANCOVA is the preferred method for analyzing results of trials in which baseline and post treatment measurements were taken, although other methods such as CHANGE or POST could be used instead, given that the assumptions under ANNOVC modeling do not appear to work (for instance, with small samples) and those where the correlation between baseline variables may be maintained through stratification; therefore there will always have been some level of agreement among groups for both outcomes, so either FRACTION or low.