

## ENDGAMES

## STATISTICAL QUESTION

## Superiority trials

Philip Sedgwick *senior lecturer in medical statistics*

Section of Medical and Healthcare Education, St George's, University of London, Tooting, London, UK

Researchers investigated the effectiveness in primary care of topical intranasal corticosteroids in children aged 4 to 11 years with bilateral otitis media with effusion. A randomised double blind, placebo controlled superiority trial was performed.<sup>1</sup>

The primary end point was the proportion of children cured of bilateral otitis media with effusion in one or both ears at one month. The proposed smallest effect of clinical interest was a difference of 15%. A total of 105 children were allocated to receive nasal mometasone furoate 50 µg, given once daily into each nostril for three months, and 112 were allocated to placebo spray.

At one month 40.6% of the children in the topical steroid group and 44.9% in the placebo group were cured. The difference between the steroid and placebo groups in the percentage cured was not significant (difference favouring placebo 4.3% (95% confidence interval -9.3% to 18.1%; P=0.55)).

Which one of the following statements best describes the alternative hypothesis for the statistical test of steroid spray compared with placebo in the primary end point?

- In the population, the proportion of children cured is greater for steroid spray than for placebo.
- In the population, the proportion of children cured is less for steroid spray than for placebo.
- In the population, the proportion of children cured is equal between steroid spray and placebo.
- In the population, the proportion of children cured with steroid spray is not equal to placebo.

## Answer

Answer *d* is the best description of the alternative hypothesis. The aim of the superiority trial was to establish whether steroid treatment was superior to placebo in effectiveness or whether placebo was superior. In particular, the sample size for the trial was calculated so that superiority between treatments would be shown if a difference in cure rate of at least 15% was obtained. The predetermined margin of 15% is called the smallest effect of clinical interest.<sup>2,3</sup> The smallest effect of clinical interest represents the smallest difference between treatments in

percentage cured that needed to be observed for one treatment to be considered clinically more effective than the other. Statistical hypothesis testing was used to test for superiority, with a null and alternative hypothesis as described in a previous Statistical Question.<sup>4</sup>

In the above superiority trial the null hypothesis incorporated the traditional starting position of equipoise. The null hypothesis states that, in the population of children aged 4-11 years where the sample was obtained, there is no difference between steroid and placebo treatments in the proportion of children cured of bilateral otitis media with effusion in one or both ears at one month. The trial's aim was to establish whether the results supported the null hypothesis or provided evidence of a difference between treatment arms, as specified by the alternative hypothesis.

The alternative hypothesis states that in the population of children aged 4-11 years where the sample was obtained there is a difference between the steroid and placebo treatments in the proportion of children cured of bilateral otitis media with effusion in one or both ears at one month. No direction is specified, and the alternative hypothesis is termed two sided: the proportion cured by the steroid spray could be less or greater than that cured by placebo.

Despite expectations before the trial that nasal steroid spray would be more effective than placebo, a two sided alternative hypothesis was chosen because results may not be as expected.<sup>5</sup> Indeed the cure rate for placebo was greater than for the steroid treatment by a difference of 4.3%. It is not obvious why this occurred. However, the P value for the test of the statistical hypotheses was 0.55, and hence there was not a statistically significant difference between cure rates. This lack of statistical significance at the 5% level was confirmed by the 95% confidence interval for the population difference in cure rates that straddled zero (-9.3% to 18.1%).

Traditionally, clinical trials have been performed as superiority trials—that is, they have attempted to establish whether a difference existed between a new treatment and the standard treatment or placebo. As medical treatments and drugs are developed it is becoming increasingly difficult to show that new

products are sufficiently more effective than the standard regimen. This has led to the development of other types of trials, including non-inferiority and equivalence trials. These trials will be discussed in future Endgames.

Competing interests: None declared.

- 1 Williamson I, Bengt S, Barton S, Petrou S, Letley L, Fasey N, et al. Topical intranasal corticosteroids in 4-11 year old children with persistent bilateral otitis media with effusion in primary care: double blind randomised placebo controlled trial. *BMJ* 2010;340:b4984.
- 2 Sedgwick P. Sample size calculations I. *BMJ* 2010;340:c3104.
- 3 Sedgwick P. Sample size calculations II. *BMJ* 2010;340:c3299.
- 4 Sedgwick P. Statistical hypothesis testing. *BMJ* 2010;340:c2059.
- 5 Sedgwick P. One sided and two sided hypothesis tests. *BMJ* 2010;340:c2458.

Cite this as: [BMJ 2011;342:d2981](https://doi.org/10.1136/bmj.d2981)