

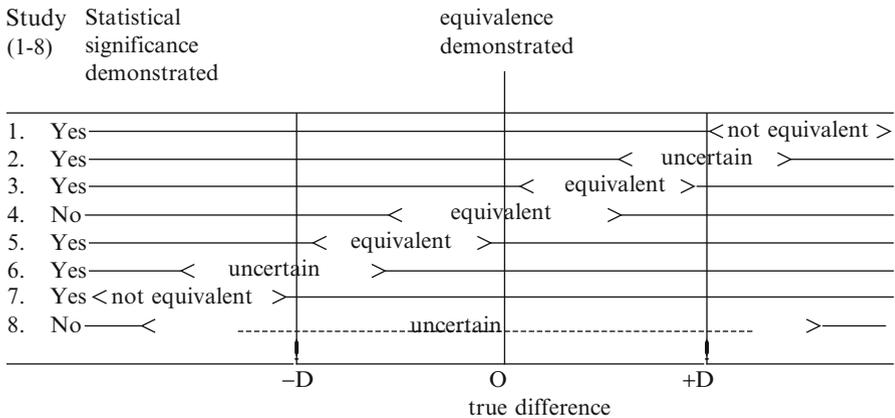
Chapter 14

Equivalence Testing Instead of Null-Hypothesis Testing

1 General Purpose

A negative study is not equal to an equivalent study. The former can not reject the null-hypothesis of no effect, while the latter assesses whether its 95 % confidence interval is between prior boundaries, defining an area of undisputed clinical relevance. Equivalence testing is important, if you expect a new treatment to be equally efficacious as the standard treatment. This new treatment may still be better suitable for practice, if it has fewer adverse effects or other ancillary advantages. For the purpose of equivalence testing we need to set boundaries of equivalence prior to the study. After the study we check whether the 95 % confidence interval of the study is

1. entirely within the boundaries (equivalence is demonstrated),
2. partly within (equivalence is unsure),
3. entirely without (equivalence is ruled out).



Then, the study is carried out, and the new and the standard treatment produce a mean reduction in blood pressure of 9.9 and 8.4 mm Hg (parallel-group study of 20 patients) with standard errors of 7.0 and 6.9 mm Hg.

$$\begin{aligned} \text{The mean difference} &= 9.9 \text{ minus } 8.4 \text{ mm Hg} \\ &= 1.5 \text{ mm Hg} \end{aligned}$$

The standard errors of the mean differences are 7.0 and 6.9 mm Hg

$$\begin{aligned} \text{The pooled standard error} &= \sqrt{(7.0^2 + 6.9^2)} \text{ mm Hg} \\ &= \sqrt{96.61} \text{ mm Hg} \\ &= 9.83 \text{ mm Hg} \end{aligned}$$

$$\begin{aligned} \text{The 95 \% confidence interval of this study} &= 1.5 \pm 2.0 \times 9.83 \text{ mm Hg} \\ &= \text{between } -18.16 \text{ and } +21.16 \text{ mm} \\ &\text{Hg} \end{aligned}$$

This result is entirely within the a priori defined boundary of equivalence, which means that equivalence is demonstrated in this study.

5 Conclusion

A negative study is not equal to an equivalent study. The former assesses the null-hypothesis of no effect, while the latter assesses, whether its 95 % confidence interval is between a priori defined boundaries, defining an area of undisputed clinical relevance. Equivalence testing is important, if you expect a new treatment to be equally efficacious as the standard treatment. This new treatment may still be better suitable for practice, if it has fewer adverse effects or other ancillary advantages. For the purpose of equivalence testing we need to set boundaries of equivalence prior to the study. The boundaries of equivalence must be in the protocol, and equivalence after the study has been completed is impossible. In an equivalence study, after the study has been completed, you should check, whether the 95 % confidence interval of the study is entirely within the a priori defined boundaries of equivalence. The boundaries have been defined on clinical, not statistical grounds.

In the current chapter, a study with continuous outcome data is used as an example. When studying binary outcome data, the result is often expressed as the proportion responders, e.g., 0.4 or 40 % responders. The calculation of the standard error with binary outcomes is explained in the Chap. 37. Briefly, with a proportion of 0.4 responders and a study sample size of 100, the standard error equals

$$\text{SE} = \sqrt{[(0.4 \times 0.6)/100]} = 0.049$$

$$\begin{aligned} \text{The confidence interval of this study} &= 0.4 \pm 1.960 \times 0.049 \\ &= \text{between } 0.304 \text{ and } 0.496 \end{aligned}$$

If the prior boundaries of equivalence were defined as being a proportion of responders between 0.25 and 0.50, then this study demonstrates the presence of equivalence.

6 Note

More background, theoretical and mathematical information of equivalence testing is given in *Statistics applied to clinical studies* 5th edition, Chap. 5, Springer Heidelberg Germany, from the same authors.