

Editorial

Equivalence and Non-inferiority Trials of CAM

Edzard Ernst

Complementary Medicine, Peninsula Medical School, Universities of Exeter and Plymouth, Exeter, UK

For several reasons, the value of placebo-controlled trials has often been disputed (not only) in complementary/alternative medicine (CAM) (1). Many clinicians feel that giving placebos to suffering patients is unethical. In fact, the Declaration of Helsinki advocates placebo-controlled trials only for conditions for which no therapy of proven efficacy exists (2). Other frequently cited reasons against the use of placebos in controlled clinical trials include the notions that patients find them hard to accept, that the placebo effect is an important contributor to the overall therapeutic effect, which should be cultivated rather than eliminated, and that placebo effects interact in a complex way with specific therapeutic effects, which renders the entire concept of the placebo-controlled trial an unscientific over-simplification (1). At the very minimum, placebo-controlled studies can be difficult to conduct, and it is therefore obvious that researchers should look for other types of methodology.

Two such options that potentially still retain major design features (namely blinding and randomization to a comparator treatment) of rigorous clinical trials are the non-inferiority trial (NIT) and the equivalence trial (ET). NITs test the hypothesis that one therapy is not worse than another, while ETs are aimed at finding out whether one treatment is neither better nor worse than another. In CAM, as in many other areas of medicine, it is often relevant to ask, is therapy X as good (i.e. efficacious) as therapy Y (i.e. another treatment used for the same condition)? For instance, in palliative cancer care it may be much more relevant to know whether massage therapy yields the same benefit in terms of quality of life as does aromatherapy, compared to determining whether massage is different from a placebo intervention. NITs and ETs are designed to answer such questions and are therefore potentially useful—in CAM and other areas of medical research.

At first glance, NITs and ETs look very much like conventional studies. They have a comparison group, can be randomized and even double-blind. Yet there is one crucial difference: while conventional studies aim to test whether there is a difference between the experimental and the control treatment, NITs and ETs aim to test whether both interventions yield the same (equivalent) result (3).

Despite the fact that such studies seem ideally suited to answer many research questions in CAM, they are associated

with several major drawbacks. First, the sample size of a typical NIT or ET has to be substantially larger than that of a typical conventional trial. For a conventional superiority trial, one needs to define what is clinically different (e.g. a difference of 5 points on a scale to arrive at an estimated sample size). For a NIT or an ET, one has to define what is clinically non-superior or equivalent. Common sense says that this must be less (e.g. plus or minus 2 points on the scale), thus the sample size of a typical NIT/ET needs to be larger. Analysis is different too: NITs and ETs require confidence intervals. One also needs to be quite clear whether one is doing a strict ET, i.e. no better and no worse, or a NIT, i.e. no worse. The implications can be considerable, particularly in CAM where research funding is usually at a premium (4). For example, while a conventional (placebo-controlled) study with, say, 200 patients is (at least sometimes) fundable, one with 400 patients very rarely is!

Second, equivalence or non-superiority only makes sense if the comparator treatment is of proven efficacy. Worse, one must make sure that, in the trial, one gives the comparator treatment its best chance of success. This usually means using it under the same conditions as those in which it was originally shown to be effective, and on a similar population. In other words, doses, inclusion and exclusion criteria etc must be the same as in the original trials. This can be a problem for CAM where one often wants to treat mild to moderate symptoms, whereas in a typical superiority trial one might opt to exclude mild to moderate symptoms.

If the comparator treatment is not of proven efficacy, establishing equivalence or non-inferiority between the experimental and the two treatments can be interpreted in at least two (dramatically different) ways: treatment A is equally effective as B, or it is equally ineffective as B. In the latter case, the result would obviously be less than informative (5). Important implications arise here for CAM: situations where a therapeutic option of proven efficacy exists are few and far between. The usual CAM scenario is a setting where several CAM therapies are used but none are of proven efficacy. Some would even insist that lack of proof of efficacy is a hallmark of CAM (6), in which case NITs or ETs of one form of CAM versus another hardly seem to be a good idea. NITs and ETs of one form of CAM versus an orthodox therapy, on the other hand, may encounter some of the above-listed

problems typical of placebo-controlled studies, e.g. reservations of participating patients.

And what about trials in which one active treatment A is tested against another active therapy B but which are *not* designed as NITs or ETs? In recent years such trials have become more and more frequent in CAM, but what do they show? If A turns out to be superior to B, then this is precisely what they show. However, more often than not, A is not significantly different from B and the authors use language suggesting that this means they are 'similar' or 'equal' or 'equipotent'. This terminology is seriously misleading. All that such a trial shows is that the result was inconclusive and thus there was no difference: in common English the difference seems minute; in terms of statistics it is, however, considerable.

In conclusion, NITs and ETs seem to be an attractive alternative to conventional, placebo-controlled, superiority trials of CAM. We are therefore likely to see more and more of such investigations. However, it is crucial to ensure that they are properly designed and that their conclusions are correct.

Studies incorrectly posing as NITs or ETs can be seriously misleading.

Acknowledgment

I am grateful to Anna Hart for her comments on this manuscript.

References

1. Heusser P. Probleme von Studiendesigns mit Randomisation, Verblindung und Placebogabe. *Forsch Komplementarmed* 1999;6:89–102.
2. World Medical Association. The Declaration of Helsinki (Doc 17.c). 2002.
3. Jadad AR. Randomised controlled trials. A user's guide. London, UK: BMJ Books 1998.
4. Ernst E. Funding research into complementary medicine: the situation in Britain. *Complement Ther Med* 1999;7:250–3.
5. Yueh B. On equivalence trials and alternative medicine. *Arch Otolaryngol Head Neck Surg* 2003;129:403–4.
6. Eisenberg DM, et al. Unconventional medicine in the United States. *N Engl J Med* 1993;328:246–52.