Using equivalence designs to improve methodological rigor in medical education trials

Martin G Tolsgaard¹ & Charlotte Ringsted²

Editor – Experimental studies comparing the effects of one type of training against those of no training are often deemed to be of little value to the development of theory.¹ Few such papers are now published in leading journals dedicated to medical education. However, comparing different types of educational intervention with one another stands a greater chance of both providing empirical evidence of their relative effectiveness and strengthening the underlying theoretical frameworks that might yield even better ways of training. Furthermore, in such comparisons no trainees are restricted from training, a situation that might be considered unethical in contexts in which the effectiveness of training is well established.² Examples of research comparing different types of intervention include studies on peer teaching versus teaching provided by faculty staff,³ dyad versus single-learner training⁴ and high-versus low-fidelity simulation training.⁵ These studies often use superiority designs to test whether or not one type of training is better than another. Unfortunately, when these studies fail to demonstrate statistically significant differences, equivalence between the interventions is often concluded. Equivalence – or non-inferiority – cannot be established based on negative results. Rather, it requires certain methodological considerations that are not accounted for in a superiority trial.

An equivalence trial requires a predefined educationally or clinically relevant maximum difference, also called delta, below which the new and existing type of training can be regarded as equivalent. Sample size calculations should be based on delta values, which often results in samples larger than those in corresponding superiority trials. Finally, to assess equivalence, the researcher compares the confidence interval for any observed difference in outcome with the delta value, rather than simply comparing the mean values of the interventions being examined. Hence, a failed superiority study cannot be converted to an equivalence trial as delta values must be determined a priori, and equivalence should not be inferred from negative results, which may just as well be explained by small sample sizes or inadequate measurement instruments. An extension of the CONSORT statement has now been developed for reporting equivalence trials⁶ and current literature provides several useful methodological guidelines for conducting this type of research.⁷

The adoption of equivalence trials in medical education research would undoubtedly improve the scientific rigor of many comparative studies in which the objective is to assess the effectiveness or efficiency of new types of training that may yield learning outcomes equivalent to those of existing methods of training but at lower costs, with less time, better adherence to training, higher feasibility or greater user satisfaction. To date, only a very limited number of equivalence trials have been published over the last decade in the field of medical education⁸,⁹ despite both theoretical and methodological support for this approach.

REFERENCES


