

Randomised controlled efficacy studies have been positioned at the top of the hierarchy of study designs, yet often their results have limited general applicability. Should they be knocked off their pedestal?

Randomised studies – useful for whom?

Medical treatment is greatly influenced by the results from double-blind, randomised and placebo-controlled efficacy studies that have been designed to determine whether new drugs or therapies are effective and the scope of any treatment effects. These studies often have a stringent experimental design, strict inclusion criteria and selected end points that make their generalisability (external validity) to clinical practice low (1).

Participants in randomised controlled efficacy studies are often quite dissimilar to those patients who actually seek medical help. The former are often recruited through the mass media (i.e. they are not seeking treatment) and are subsequently carefully selected using a number of inclusion and exclusion criteria. Often, more than half of them will be deemed unsuitable for inclusion or randomisation (1). Moreover, participants have a relatively low socioeconomic status. I myself regularly receive emails from desperate Americans who want to participate in one or more of our clinical studies (found at www.clinicaltrials.org) because they cannot afford obesity surgery.

Globally there are major differences in the organisation of the health services, treatment objectives, research management, funding and recruitment of patients to studies. Accordingly, the results from relatively well implemented studies cannot always be generalised to other countries or continents. Clinical intervention studies, the results of which depend on clinicians' level of competence and skills as well as less clinically relevant endpoints or surrogate endpoints, may also undermine the generalisability of randomised efficacy studies (1, 2).

The results from a relatively small randomised controlled single-centre study of 150 overweight persons with Type 2 diabetes were recently published in the prestigious *New England Journal of Medicine* (3). The research group compared the effects of three different therapies: gastric bypass, longitudinal ventricular resection (gastric sleeve) and intensive medical treatment with regard to the primary surrogate endpoint $HbA_{1c} \leq 6.0\%$. The participants had a body mass index (BMI) ranging from 27 kg/m² to 43 kg/m² and were recruited from either the Cleveland clinic or through the media. The external validity of this study is weakened by a high risk of selection bias, a lower threshold value for the primary surrogate endpoint than that which is recommended by international and national guidelines (Norway: $HbA_{1c} \leq 7.0\%$), as well as inclusion of participants with a BMI < 35 kg/m² (lower than that which is recommended by international guidelines for obesity surgery). One may also ask whether intellectual and financial conflicts of interest for the main author may have had an effect on the treatment (experimenter bias) (4). Consequently, the results of this study have limited relevance for Norwegian clinical practice.

The knowledge base upon which good patient treatment rests should not be restricted to findings from randomised controlled efficacy studies, but should also include results from other types of studies with higher external validity (5, 6). A randomised study with a more practically useful (pragmatic) design can serve as a good example of the latter. As early as 1967, Schwartz and Lellouch discussed the differences between so-called explanatory efficacy studies and pragmatic studies, which compare the effects of various treatment

methods in clinical practice (7). Pragmatic randomised studies are undertaken to help decision makers, clinicians and patients choose among alternative treatment methods. More emphasis is given to clinically relevant endpoints and fewer exclusion criteria are used.

Nor should we underestimate the value of non-randomised clinical studies and purely comparative observational cohort studies (6). American authorities have therefore encouraged more widespread use of comparative effectiveness research (CER) (6). This type of research compares the advantages and disadvantages of various treatment methods and strategies in a «real world» setting, for example in purely observational studies. However, researchers, sponsors, ethics committees and authors of guidelines continue to take insufficient account of the importance of high-quality observational studies, while disregarding the low external validity of many randomised efficacy studies (1). This notwithstanding, it should be emphasised that even though a study design adapted to the real world may produce a higher external validity, this may come at the cost of lower internal validity, which may also weaken the generalisability of the results (5).

Decision makers, clinicians and patients should thus be critical towards the results from randomised, controlled efficacy studies, and should also take into account the results from more pragmatically oriented studies before a treatment method is selected. Randomised controlled efficacy studies remain necessary, but are not in themselves sufficient for the provision of good patient treatment (8). Many studies already include elements from both types of study design.

Jøran Hjelmesæth
joran.hjelmeseth@siv.no

Jøran Hjelmesæth (born 1957), dr. med. and specialist in internal medicine and renal diseases. He heads both the multidisciplinary outpatient clinic and research group at the Morbid Obesity Centre, Vestfold Hospital Trust. He is also Professor II at the Department of Endocrinology, Morbid Obesity and Preventive Medicine, the Medical Clinic, Institute of Clinical Medicine, University of Oslo.

The author has completed the ICMJE form and declares no conflicts of interest.

References

1. Rothwell PM. External validity of randomised controlled trials: «to whom do the results of this trial apply?». *Lancet* 2005; 365: 82–93.
2. Lassen K, Høye A, Myrmet T. Randomised trials in surgery: the burden of evidence. *Rev Recent Clin Trials* 2012; 7: 244–8.
3. Schauer PR, Bhatt DL, Kirwan JP et al. Bariatric surgery versus intensive medical therapy for diabetes—3-year outcomes. *N Engl J Med* 2014; 370: 2002–13.
4. Cleveland Clinic. Staff display. http://my.clevelandclinic.org/staff_directory/staff_display.aspx?doctorid=6010 [12.9.2014].
5. Zwarenstein M, Treweek S, Gagnier JJ et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008; 337: a2390.
6. Dreyer NA. Making observational studies count: shaping the future of comparative effectiveness research. *Epidemiology* 2011; 22: 295–7.
7. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967; 20: 637–48.
8. Brean A. Hva er en medisinsk sannhet? *Tidsskr Nor Legeforen* 2013; 133: 381.