SMALL TRIALS

From the Therapy chapter for the 3rd edition of Clinical Epidemiology, by DL Sackett
17 April 2004 (day 108)

A team of us, led by Brian Haynes, asked the question: "Among hypertensive steel workers who are neither compliant with antihypertensive drugs, nor at goal blood pressure six months after starting treatment, could a compound behavioral intervention (teaching them how to measure their own blood pressures, asking them to chart their own blood pressures and pill-taking, teaching them how to "tailor" their pill-taking to their daily habits and rituals, and checking their performance every two weeks and rewarding positive changes), provided by a high school graduate with no health professional training, compared to an "attention placebo," improve their compliance and blood pressure control over the next six months?". Only 38 patients entered our trial.[1] We allocated them to the experimental and control groups by concealed minimization (that considered each patients' blood pressure level, compliance, and prior exposure to instruction about their hypertension). Despite their small number (20 experimental and 18 control patients), they were enough to show, with great confidence, that the combined outcome of increased compliance and decreased diastolic blood pressure was far more likely to occur among experimental patients (70% vs.11%; p=0.001).

Most of the examples in this chapter come from the large, long, complicated RCTs that occupy most of my time. This section has the purpose of showing that small RCTs are not only possible, but also often sufficient to generate highly confident answers to questions about efficacy.

Small RCTs can serve three useful purposes[2], as summarized in Table 3-02S-1.

Table 3-02S-1. Useful purposes served by small trials

	Purpose served	Explanation
1	Providing a definitive answer to a question about efficacy.	The important number in an RCT is the number of events, not the number of patients.
2	Suggesting that the "expert" emperor has no clothes.	A small, indeterminate RCT can still be big enough to create uncertainty.
3	Supplying randomized evidence for meta-analysis.	Larger numbers generate smaller confidence intervals.

Their first purpose follows from the fact that the important number in an RCT is the number of events, not the number of patients. As a result, small trials can be large enough to provide confident, definitive answers to important questions. The record for smallness, as far as I know, is held by a cross-over trial of beta blockers for angina pectoris conducted three decades ago. As recalled by Richard Doll[3], it required only 12 patients to generate a p-value of 0.003[4]. Those days are gone forever in angina trials, because the growing array of efficacious interventions has sent control event rates (CERs) plummeting. For example, a recent multicenter trial of an intravenous beta blocker vs. "standard care," despite recruiting nine times as many patients, failed to achieve a definitive answer[5].

There is an only semi-facetious moral here for budding trialists. Pick a primitive specialty (where they've never validated any of their interventions) for your first RCT. Then, test its most promising (but unproven) intervention against placebo for its ability to improve short-term outcomes[1]. Specialties with a long tradition of performing RCTs are victims of their own progress. They must replace placebos with treatments they've validated in prior RCTs. This

[1] Of course, it must be both clinically sensible and ethical to perform this trial.

translates to decreases in control event rates (CERs) and absolute risk reductions (ARRs). As a result, they need to perform ever larger RCTs to confidently show that this year's treatment is superior to last year's, even when their relative risk reductions are identical.

The second useful purpose served by small trials is to reveal that the "expert" emperor may have no clothes: that this year's experts are as full of baloney as last century's. Even small, indeterminate trials (with huge confidence intervals) can prompt us to question conventional but untested therapeutic wisdom. Consider the internal mammary artery. Surgeons now use it with considerable success as a conduit in coronary artery bypass grafting for intractable angina pectoris. Forty years ago, with that same objective in mind, they tied it off! This was because, believe it or not, they found that the simple ligation of the internal mammary artery (a very safe operation, easily performed using local anesthesia) was followed by both subjective and objective improvement of angina. Pain lessened, and electrocardiograms and treadmill tests improved in more than three-quarters of the patients who underwent the procedure[6]. No wonder, then, that this operation rapidly gained popularity throughout Europe and North America.

Then R.G. Fish and his colleagues suspected that its apparent efficacy might be a "placebo effect" (if this is a new term for you, visit page xx for a definition). So, they told a group of preoperative patients that the procedure was experimental, had no physiologic basis, and was of uncertain benefit. Only 20% (rather than 75%) of their patients improved[7]. Their results generated enough uncertainty in the cardiology community to conduct two simultaneous small RCTs of internal mammary artery ligation. In these trials, the operation proceeded to the point of placing a loose ligature around the isolated internal mammary artery. Then, they opened a sealed envelope that told them to tie off the arteries of a random half of patients, and to leave the other half alone. On follow-up, the sham-operated patients fared as well as those whose arteries were tied off. The "expert" emperors' nether parts were exposed to all, and the procedure was rapidly abandoned.

Although conducted to answer an important clinical question, and used to justify randomized trials of coronary bypass two decades later, these two' 'negative" trials contained a grand sum of just 35 patients. To confidently exclude even the huge relative risk improvement of 50% in symptoms and function, however, they should have enrolled four times as many. Although they were too small to be conclusive in terms of formal statistics, these small trials changed the course of cardiac surgery. They not only forced the rethinking of the biologic rationale for the operation, but also led to its abandonment.

Third, small trials, even when individually inconclusive, can serve as the basis for convincingly conclusive overviews and meta-analyses. For example, infusing albumin in the attempt to save the lives of critically ill burn patients was "standard" practice for decades. When three small RCTs (among 14, 79, and 70 patients, respectively) generated conflicting evidence on whether this practice actually saved or cost lives, they were combined in a meta-analysis[8]. This meta-analysis generated a relative risk of dying following albumin infusion of 2.4 (95% confidence interval from 1.1 to 5.2) in burn patients. Moreover, when combined with a meta-analysis of 28 other small RCTs performed in patients with hypovolemia or hypoproteinemia, it emerged that for every 20 patients treated for one of these conditions with albumin infusions, one more patient would die. This startling meta-analysis of 31 trials with an average size of just 51 patients had major repercussions among clinicians (many of whom stopped using albumin in these patients) and regulators (some of whom cautioned against its further use in them). Moreover, it led to the design and initiation of a large albumin trial with a target sample size of 7000 (over 4 times as many patients as all previous trials combined).

Some ethicists and trialists hold that it is inappropriate to embark on an RCT when you know that it is too small to generate a confident answer about efficacy. I disagree with this view, and discuss it, along with other special ethical issues in RCTs, on page xx. However, to justify small trials on the basis of later inclusion in systematic reviews, they must be registered at their inceptions so that later meta-analysts can find them.

REFERENCES

[1] Haynes RB, Sackett DL, Gibson ES, Taylor DW, Hackett BC, Roberts RS, Johnson AL. Improvement of medication compliance in uncontrolled hypertension. Lancet 1976;1:1265-8.

[2] Sackett DL, Cook DJ. Can we learn anything from small trials? Ann NY Acad Sci 1993;703:25-31.

[3] Doll R. Discussion. Ann NY Acad Sci 1993;703:31.

[4] Prichard BNC, Dickinson CJ, Alleyne GAO, Hurst P, Hill ID, Rosenheim ML, Laurence DR. Report of clinical trial from medical unit and MRC Statistical Unit, University College Hospital Medical School, London. BMJ 1963;2:1226-7.

[5] Mitchell RG, Stoddard MF, Ben-Yehuda O, Aggarwal KB, Allenby KS, Trillo RA, Loyd R, Chang CT, Labovitz AJ. Esmolol in acute ischemic syndromes. Am Heart J. 2002;144:E9.

[6] Barsamian EM. The rise and fall of internal mammary ligation in the treatment of angina pectoris and the lessons learned. In" Costs, Risks and Benefits of Surgery. Bunker JP, Barnes BA, Mosteller F, editors. New York: Oxford University Press, 1977. Pp212-20.

[7] Fish RG, Crymes TP, Lovell MG. Internal-mammary-artery ligation for angina pectoris. Its failure to produce relief. N Engl J Med 1958; 259:418-20.

[8] The Albumin Reviewers (Alderson P, Bunn F, Lefebvre C, Li Wan Po A, Li L, Roberts I, Schierhout G). Human albumin solution for resuscitation and volume expansion in critically ill patients (Cochrane Review). In: *The Cochrane Library*, Issue 4, 2003. Chichester, UK: John Wiley & Sons, Ltd.