Managing the resource demands of a large sample size in clinical trials: can you succeed with fewer subjects?

Anthony C Keech and Val Gebski

IN PLANNING CLINICAL TRIALS, it is common to find that the calculated sample size[1] (Item 7 of the CONSORT checklist; Box 1) is too large for available resources. Strategies to determine whether the trial question(s) can be answered with fewer subjects are needed. These include:

■ focusing on higher-risk subjects;
■ using a run-in phase before randomisation;
■ "expanding" the primary study endpoint; or
■ running the trial for a longer period, with an event-based, rather than a calendar-based, stopping rule.

Choosing subjects with higher risk

If the subjects in a trial have a very low risk of the condition that the intervention is hypothesised to prevent, the trial,

NHMRC Clinical Trials Centre, University of Sydney, Camperdown, NSW.
Anthony C Keech, FRACP, MScEpid, Deputy Director; **Val Gebski,** MStat, Principal Research Fellow.
Reprints will not be available from the authors. Correspondence: Associate Professor Anthony C Keech, NHMRC Clinical Trials Centre, University of Sydney, Locked Bag 77, Camperdown, NSW 1450. enquiry@ctc.usyd.edu.au

---

1: CONSORT checklist of items to include when reporting a trial[1]

Section and topic	Item no.	Descriptor
Methods Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules

---

regardless of sample size, will not prove the value or otherwise of the intervention. For example, in the "Finnish Businessmen's Study", the efficacy of a multifactorial risk-factor intervention to prevent cardiovascular death among middle-aged men could not be proven, as only five such deaths had accrued at the end of the scheduled follow-up.[2] The proof required from trials relies on demonstrable differences in event counts between the intervention and control groups, and whether this difference could reasonably have occurred by chance alone. It matters little how many subjects produced these event counts — the evidence rests in the main with the event counts themselves and the size of the difference between them.

Consequently, if the calculated sample size of a proposed clinical trial is larger than feasible, limiting the subjects to

those in a higher-risk category should be considered. In the Finnish Businessmen's Study, it might have been better to recruit only men with prior heart disease, with four to eight times the risk of those in the primary prevention category. Similarly, in trials to prevent cancer recurrence after initial therapy, focusing on individuals with above-average risk of recurrence would require a smaller sample size. At times, however, the cost and feasibility advantage of using a lower sample size might be outweighed by the extra time and effort needed to identify high-risk individuals. This might occur especially where the features determining higher risk are not clinical characteristics, but are based on medical testing. In Box 2, a comparison of two possible trials shows that Trial B, with a similar study power, is more feasible and presumably less costly than Trial A.

Maximising study power through better compliance — use of a "run-in" design

In a clinical trial design, a "run-in" phase can reduce the required sample size.[3] Subjects who are entering a long-term trial are asked to take the study medication(s) for a period before randomisation. Individuals who lose interest early on (potential "drop-outs") can then be excluded before random allocation. Similarly, any subjects who feel they may have an indication to receive the intervention treatment (potential "drop-ins") can also withdraw before randomisation. This potentially lowers rates of anticipated non-compliance to allocated treatment during a trial, resulting in a smaller required sample size. As the calculated sample size is exquisitely sensitive to compliance, this procedure can be of major benefit (Box 3). (Once randomised, these participants would generally be included in an intention-to-treat analysis[4] and only dilute the apparent effect of the intervention, boosting the sample size needed and/or follow-up duration.)

Run-in phases can use either placebo or active therapy, and are usually single blind (ie, only the study staff are aware of the nature of the medication). A placebo run-in allows trial staff to be sure that reported side effects are not caused by treatment (colouring agents and excipients in placebos can occasionally cause reactions), whereas an active run-in can identify and exclude individuals who may be unable to tolerate the medication being tested in a long-term trial. In the US Physicians Study (testing the value of aspirin to prevent coronary death and β-carotene to prevent cancer), a placebo run-in phase allowed a trial of 22 000 doctors to deliver comparable results to a trial requiring 33 000 doctors, assuming that doctors who withdrew during the run-in period would otherwise have stopped taking the study medication soon after randomisation.[3,5] Whether excluding any potential trial subjects in this way will reduce the generalisability of the ultimate trial results needs to be carefully considered.

Choosing a different endpoint to limit the sample size

If a more frequently occurring endpoint can be substituted, with the same biologically anticipated effects of treatment,

2: Comparison of two possible trials — Trial A, with lower-risk subjects, and Trial B, with higher-risk subjects — to determine the value of the same treatment hypothesised to reduce events by 25% (ie, relative risk [RR]=0.75) during follow-up*

Treatment group	Trial A – lower risk (n=2000)		Trial B – higher risk (n=1000)	
	Control	Active	Control	Active
Number of subjects	1000	1000	500	500
Proposed RR with treatment		0.75		0.75
Expected event rate	20%	15%	40%	30%
Expected number of events	200	150	200	150
Study power at 2P = 0.05	82%		90%	

*The number of events, rather than the number of subjects, principally determines the power of the study, although the number of subjects determines in part the reliability of each event count and of the difference.

3: Possible effect of a "run-in" design on the sample size of a randomised trial

Trial scenario	Required sample size
A: 100% compliance in both trial arms	400
B: Average of 80% compliance in active arm (ie, 20% drop-outs at study mid-point)	625
C: Half (10%) of the average long-term non-compliers (drop-outs) instead withdraw during run-in phase before randomisation	494
D: Average of 80% compliance in both study arms (ie, 20% drop-ins plus 20% drop-outs)	1110
E: Half the average long-term non-compliers (10% drop-outs plus 10% drop-ins) instead withdraw during run-in phase before randomisation	625
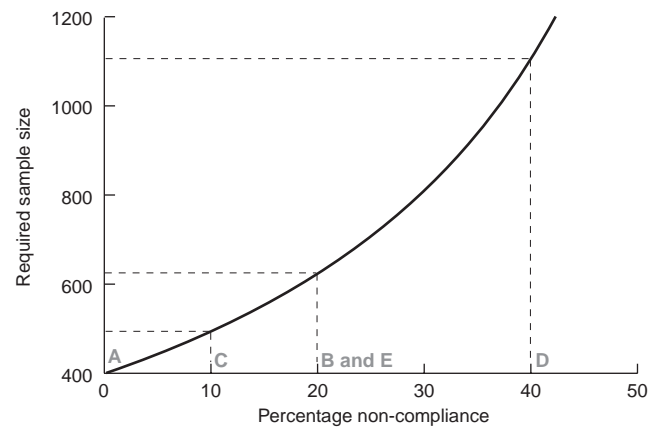
then the required sample size will fall accordingly. For example, while trials of lowering cholesterol level to reduce total mortality over 5 years may require, say, 12 000 patients, similar trials to reduce coronary mortality only (which cause a fall in total mortality) may only need 8000 patients, depending on the proportion of deaths due to coronary causes. Furthermore, trials designed to reduce the

combined endpoint of coronary death plus non-fatal myocardial infarction may require perhaps 4000 patients, with even fewer required for trials designed to reduce all vascular events (all cardiovascular deaths plus non-fatal myocardial infarction plus non-fatal stroke plus any revascularisation procedure). Of course, in the above example, as the endpoint becomes broader, "softer" clinical outcomes are included (ie, some outcomes, such as a decision to send a patient for a revascularisation procedure, may be more subjectively based, and even influenced by a patient's treatment, including the study treatment, if blinding has failed).

The decision as to the choice of the primary endpoint in trials should be made in consultation with the clinicians who will ultimately use the trial's outcomes in practice. Selection of the endpoint must ensure that sufficient information is available to determine whether the new treatment should be applied in clinical practice.[1]

In any case, tracking (which is blinded to study treatment) of the risk profile of subjects randomised into a clinical trial should occur during recruitment, as well as monitoring during follow-up (also blinded) of the event rates in the entire cohort to allow consideration of
- a possible increase (or, rarely, decrease) in the target sample size before the end of recruitment;
- a change in the primary outcome of the study; and
- extending the scheduled follow-up period to yield more events.

Whenever possible, it is important to specify a stopping rule in the study protocol, based on accrued numbers of events rather than a calendar date, to allow a trial to continue without major disruption when trial outcome risks are lower than expected.

Buying extra science for little extra cost — substudies in large clinical trials

Once a study outline has been finalised, formal consideration should be given to substudies nested within the larger trial. The use of surrogate outcomes offers the opportunity to answer questions of related interest, or to explore the mechanism of the treatment effect[6] in ways which might otherwise be prohibitively costly (ie, setting up substudies as separate enterprises). For example, in a study of the effects of lipid-lowering therapy on coronary death and stroke in many thousands of subjects with prior cardiovascular disease, substudies exploring the effects of treatment on (i) the measured progression of coronary atherosclerosis (using serial coronary angiography), (ii) the progression of carotid intima media thickness, (iii) the change in brachial vascular reactivity (using serial ultrasound examinations), or (iv) endothelial vasoactive peptide levels, may have sufficient

---

4: Checklist for managing sample size demands in clinical trials

- Determine the risk profile of the intended population of interest. Can a subpopulation at higher risk readily be found?
- Determine whether the study design will accommodate either a placebo or an active run-in phase? Establish clear guidelines on whether to randomise potentially non-compliant subjects.
- Determine the clinically justifiable power for the particular trial.
- Adjust the calculated sample size for the expected level of non-compliance with treatment.
- If the event rates are small, identify potential outcomes which may provide alternative endpoint(s) for which the event rate is much larger.
- Ensure that the risk profile of subjects is monitored blinded during recruitment as well as the event rate during follow-up.
- Where possible, base stopping rules on the number of events rather than the duration of follow-up.
- Identify related questions which may be investigated using surrogate outcomes on a subpopulation of randomised subjects.

---

power with only several hundred subjects each. For each substudy, the resources needed for subject identification and recruitment, running trial clinics and follow-up are already largely covered by the main trial infrastructure, resulting in extremely cost-effective research opportunities.

Conclusion

A number of strategies can help to ensure that clinical trials research can be done within limited budgets and by smaller-scale collaborations (Box 4). Care must be taken, however, to deliver results that are still meaningful to clinicians, and have a low risk of false-negative conclusions. As always, seeking professional advice can help to ensure success.

References

1. Kirby A, Gebski V, Keech AC. Determining the sample size in a clinical trial. *Med J Aust* 2002; 177: 256-257.
2. Miettinen TA, Huttunen JK, Naukkarinen V, et al. Multifactorial primary prevention of cardiovascular diseases in middle-aged men. Risk factor changes, incidence and mortality. *JAMA* 1985; 254: 2097-2102.
3. Pablos-Méndez A, Barr G, Shea S. Run-in periods in randomized trials. Implications for the application of results in clinical practice. *JAMA* 1998; 279: 222-225.
4. Gebski V, Marschner I, Keech AC. Specifying objectives and outcomes for clinical trials. *Med J Aust* 2002; 176: 491-492.
5. Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the on-going Physicians' Health Study. *N Engl J Med* 1989; 321: 129-135.
6. Simes RJ, Topol EJ, Holms DR Jr, et al. Link between the angiographic substudy and mortality outcomes in a large randomized trial of myocardial reperfusion. Importance of early and complete infarct artery reperfusion. GUSTO-I Investigators. *Circulation* 1995; 91: 1923-1928.

❑