Generalising the results of trials to clinical practice

J Paul Seale, Val J Gebski and Anthony C Keech

Randomised controlled trials should be the basis for developing clinical guidelines and for decisions about individual patient management. They should also inform public health policy. However, their capacity to fulfil these roles will depend on how closely a trial's participants reflect the general population of patients with the disorder that has been investigated. The extent to which a trial's findings are relevant to the broader population of patients with the disorder is referred to as the trial's generalisability, or external validity. The CONSORT statement refers to generalisability under Item 21 (Box 1).[1] Well-written reports should discuss the various factors that influence the generalisability of the trial's findings. Julian and Pocock have proposed a checklist of questions to assist with this assessment (see Box 2).[2]

To determine the generalisability of a trial's findings, several aspects require scrutiny.

1 CONSORT checklist of items to report when reporting a randomised trial.[1]

Section and topic	Item no.	Descriptor
Discussion		
Generalisability	21	Generalisability (external validity) of the trial findings.

2 Concepts covered in Julian and Pocock's criteria for assessing generalisability[2]

- Representativeness of patients for the condition in practice.
- Proportion of eligible patients participating.
- Conformity of the treatments and background care (doses, durations, follow-up period, etc) to standard practice patterns.
- Consistency of measured outcomes with conclusions drawn.
- Appropriate balance of surrogate and clinical outcomes.
- Reliability of evidence on efficacy and safety findings.
- Coverage of all relevant outcomes (adverse events and side-effects).
- Consideration of the study findings in the context of other available evidence.

Is the patient population representative of the broad target group?

To make this assessment, it is firstly necessary to examine the inclusion and exclusion criteria for the trial. These criteria determine the characteristics of the potential participants. They are particularly important for trials assessing new drugs, because patients with any significant degree of renal or hepatic impairment, or any significant comorbidity, are often excluded. Excluding such participants may result in a trial population that represents only a subsection of the broader population with the disorder for which the drug may be indicated.

Secondly, the baseline data should describe the population that participated in the trial. Demographic variables, age range, as well as clinical data such as blood pressure, staging of disease and any listed comorbidities, will help readers decide whether the trial population closely resembles the patient population (or individual patient) for which a decision about management is required. In some trials, the entry criteria are considerably broader than the population actually recruited. This discrepancy will only be evident if adequate baseline data are presented. For example, in a study of combination chemotherapy in malignant breast cancer, no radiotherapy was allowed in the protocol.[3] Results were reported on the basis of the extent of lymph node involvement — 0–3 nodes or 4 or more — and readers might assume that the findings of the trial would apply to participants with many (more than 10) involved nodes. However, less than 8% of patients with more than 10 involved nodes were included in the study, as clinicians referred these higher-risk patients for radiotherapy rather than enrolling them in the trial.[4] The original trial report did not tell readers that the group with more than 4 involved nodes actually comprised patients with primarily 4 to 10 involved nodes.[5]

Participant flow diagram

These diagrams are useful for assessing generalisability of trials. If properly completed, flow diagrams will indicate the number of participants:

- screened for participation;
- with the condition of interest;
- classed as ineligible (on the basis of exclusion criteria); and
- who did not elect to participate.

If the population randomly allocated to groups within the trial represents only a small proportion of those with the condition of interest and assessed for eligibility, it is probable that the generalisability of the findings of the study will be limited. Large trials assessing warfarin therapy for atrial fibrillation have enrolled only about 15% of those who were potentially eligible, and this substantially limits the generalisability of their results.[6,7]

Where eligible patients who entered randomised trials have been compared with those who were eligible but did not participate, differences have emerged. In a study of therapy for temporomandibular disorders, 18 eligible patients did not consent and 60 were randomly allocated to trial arms.[8] The 18 patients who did not consent reported more pain than those who participated, perhaps restricting the findings of the trial to those with milder pain. Differences were also evident between

University of Sydney, NSW.
J Paul Seale, PhD, FRACP, FRCP, Professor of Clinical Pharmacology; Val J Gebski, MStat, Associate Professor, and Principal Research Fellow; Anthony C Keech, FRACP, MClinEpi, Deputy Director, NHMRC Clinical Trials Centre.
Reprints will not be available from the authors. Correspondence: Professor J P Seale, Department of Pharmacology, Blackburn Building, University of Sydney, Sydney, NSW 2006.
enquiry@ctc.usyd.edu.au

enrolled and unenrolled patients in the Thrombolysis in Myocardial Infarction (TIMI 9) trial.[9] The TIMI 9 registry prospectively evaluated patients with ST-segment-elevation myocardial infarction. There were no exclusion criteria for the registry, but there were exclusion criteria for the randomised trial. Patients in the registry, but not enrolled in the trial, had higher baseline risk for adverse outcomes.

Screening logs

Screening logs list the numbers of individuals screened, eligible and enrolled, as well as reasons for not enrolling eligible patients. They thus allow readers to judge whether there are differences between patients who were and were not enrolled in the trial. If the two populations are similar, the generalisability of the trial is increased. A template of the typical information collected in the screening log is presented in Box 3. This information should be limited to the most important characteristics of the relevant population to minimise the burden on trial staff collecting the data. Screening logs also describe the range of participants with the disease being seen at each investigation site, and the patterns of care of these people. For those deemed ineligible, the criteria excluding them from the study are documented, providing further information as to the generalisability of the intervention to this cohort.[10] The results of the study will apply more to subjects excluded because they were not available for follow-up or were just outside the age range than to those with concomitant disease.

3 Screening log template

Site identification/name:_____

Date of visit:_____

Subject initials:_____

Clinician initials:_____

Age:_____

Sex:_____

Is the subject recruited into the study?yes / no

If no: then:

a) Main reason for exclusion:_____

_____

_____

Reasons for exclusion:

1.Subject refusal

2.Clinician refusal (with possible reasons)

3.Ineligible (specify which inclusion criteria are not met and which exclusion criteria exist)

4.Language difficulties

5.Other

b) Treatment actually given:_____

_____

_____

Comorbidities

In some trials, during random allocation, patients may be stratified by comorbidities regarded as potential confounders. In others, particularly in clinical trials of new drugs, comorbidities may be exclusion criteria. If these comorbidities are relatively common, the exclusion criteria will significantly limit the generalisability of the trial outcomes. This is an important issue in drug trials, as comorbidities are often exclusion criteria. When the drug is registered for use, the listed indication is often relatively broad, so it is necessary to scrutinise the clinical trials section of the product information to obtain a clearer picture of the patient population which was studied. If a patient for whom the drug is being considered has one of the comorbidities which was an exclusion criterion, whether the drug will be efficacious or safe is unknown. Such a dilemma exists with the "statin" lipid-lowering drugs. Over 160 000 patients have participated in trials of statins, but almost all of these trials have excluded patients with significant renal or hepatic disease.[11-13]

If some patients have characteristics which were not reported in the trial's population, it is conceivable that the trial's results are not relevant to these patients.

Subgroup analyses

As adolescents and pregnant women are not usually included in trials, most clinical trials are of limited relevance to these groups. If there are a priori reasons to expect differences between subgroups, the trial may have stratified participants by these potential confounders. The findings of the trial will be most generalisable if the benefits are evident in each subgroup of the trial, as well as across the entire study population.[14] In clinical trials with large numbers of patients or events, it is possible to have reliable subgroup analyses which may help prescribers to relate the trial's findings more closely to patients for whom they are trying to select appropriate therapies.

Conclusions

The main purpose of conducting randomised clinical trials is to identify improvements in clinical care. Ideally, the findings of trials should apply to a wider population than those included in the trial. It is therefore vital that every effort is made to have a broad selection of patients from the population of interest to minimise selection bias.

Numerous exclusion criteria will restrict the patient population and progressively diminish the generalisability of the findings of the intervention under evaluation. It is the responsibility of those reporting trials to include aspects of generalisability when discussing their findings. It is the task of those responsible for treating patients, producing clinical guidelines and formulating public health policy to carefully assess the generalisability of clinical trials before applying their findings.[15]

Acknowledgements

We thank Rhana Pike for expert assistance in preparation of this manuscript.

Competing interests

None identified.

References

1 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001; 357: 1191-1194.

2 Julian DG, Pocock SJ. Interpreting a trial report. In: Pitt B, Julian D, Pocock S, editors. Clinical trials in cardiology. London: Saunders; 1997.

3 Bonadonna G, Brusamolino E, Valagussa P, et al. Combination chemotherapy as an adjuvant treatment in operable breast cancer. *N Engl J Med* 1976; 294: 405-410.

4 Bonadonna G, Valagussa P. Chemotherapy of breast cancer: current views and results. *Int J Radiat Oncol Biol Phys* 1983, 9: 279-297.

5 Langlands AO, Tiver KW. Adjuvant chemotherapy in breast cancer [letter]. *Int J Radiat Oncol Biol Phys* 1984, 10: 943-946.

6 Evans A, Kalra L. Are the results of randomized controlled trials on anticoagulation in patients with atrial fibrillation generalizable to clinical practice? *Arch Intern Med* 2001; 161: 1443-1447.

7 Stroke Prevention in Atrial Fibrillation (SPAF) Investigators. Stroke Prevention in Atrial Fibrillation study: final results. *Circulation* 1991; 84: 527-539.

8 Yuasa H, Kurita K, Westesson P-L. External validity of a randomized clinical trial of temporomandibular disorders: analysis of the patients who refused to participate. *Br J Oral Maxillofacial Surg* 2003; 41: 129-131.

9 Bahit MC, Cannon CP, Antman M, et al. Direct comparison of characteristics, treatment, and outcomes of patients enrolled versus patients not enrolled in a clinical trial at centers participating in the TIMI 9 trial and TIMI 9 registry. *Am Heart J* 2003; 145: 109-117.

10 Moore DAJ, Goodall RL, Ives NJ, et al. How generalizable are the results of large randomized controlled trials of retroviral therapy? *HIV Med* 2000; 1: 149-154.

11 Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994; 344: 1383-1389.

12 LIPID Study Group. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. *N Engl J Med* 1998; 339: 1347-1357.

13 Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 4963 people with diabetes: a randomised placebo-controlled trial. *Lancet* 2003; 361: 2005-2016.

14 Lord SJ, Gebski VJ, Keech AC. Multiple analyses in clinical trials: sound science or data dredging? *Med J Aust* 2004; 181: 452-454.

15 Horton R. Common sense and figures: the rhetoric of validity in medicine: Bradford Hill Memorial Lecture 1999. *Stat Med* 2000; 19: 3149-3164.

(Received 11 Oct 2004, accepted 12 Oct 2004) ❑