Analysis and comment

Research methodology
Assessment of generalisability in trials of health interventions: suggested framework and systematic review
C Bonell, A Oakley, J Hargreaves, V Strange, R Rees

Most evaluations of new treatments use highly selected populations, making it difficult to decide whether they would work elsewhere. Systematic evaluation and reporting of applicability is required.

Randomised trials of health interventions generally describe outcomes among participants with little consideration of whether the effects can be generalised. However, generalisability cannot be assumed with either biomedical interventions or more complex social interventions. If their results are to be translatable into policy and practice decisions, trials must provide evidence about how relevant the interventions might be to other sites and populations. Such information is particularly crucial for resource poor settings.

Although CONSORT criteria for reporting randomised trials include assessment of generalisability, there is no mandated framework for empirically assessing and reporting this. We consider the factors affecting generalisability using examples from HIV and sexual health, examine how a sample of trials looked at generalisability, and suggest how to improve evaluation.

Can the intervention be delivered elsewhere?

Several factors affect whether an intervention can be delivered and received in other sites. Firstly, an intervention must be feasible. Providers will vary in their capacity to implement an intervention, as will institutions in being suitable places for an intervention. The presence of local “champions” may influence feasibility in a particular site. Some interventions require the existence of other health services—for example, services for treating sexually transmitted infections require microbiology laboratories to target the right patients. Interventions may also require adequacy in other sectors such as transport. Feasibility has a cost dimension: an unaffordable intervention lacks general feasibility.

Secondly, an intervention must achieve adequate coverage. This may depend on the overall comprehensiveness of health systems or on whether providers can reach people in other ways—for example, through outreach. Adequate coverage may be more difficult in some sites or sub-populations.

Finally, an intervention generally must be acceptable to be effective. Acceptability refers to participants’ assessment of their experience of an intervention and will influence whether recipients adhere to treatment plans, act on health advice, or return for follow-up. For example, condom promotion has proved acceptable and subsequently effective in urban Tanzania but not in rural regions. Acceptability will vary between populations as it depends on cultural norms and can have economic dimensions. For example, HIV voluntary counselling and testing services that require clients to attend clinics twice (first for testing and then for results) may be acceptable in high income settings but not low income settings because transport or opportunity costs are too great.

Factors relating to delivery of an intervention are best documented by embedding an evaluation of process in trials. The study collects quantitative and qualitative data on planning, delivery, and uptake and how context affects them.

Does the intervention meet recipients’ needs?

To be effective an intervention must meet recipients’ needs—that is, the recipients must have capacity to benefit from an intervention. Thus potential recipients of an intervention should have similar needs to those of the original study participants. Trial participants may be untypical of the general population even in the study site, let alone in other sites. Trials tend to under-represent certain groups, such as minority ethnic and low income groups, women, and older people, whose needs may differ from those of people included in trials. Trials should therefore describe the sociodemographic profile of participants and report the extent to which they are representative of the target population.

If the needs of future potential recipients differ from those of the study participants, interventions may not work in a new population or have to be adapted. For example, provision of antiretroviral drugs in low income countries, or to certain sub-populations may have to be accompanied by support to promote adherence in order to achieve similar outcomes to those achieved among trial participants.
This is also true of public health interventions. The extent to which a factor contributes to the incidence of a particular disease, and therefore needs intervention, varies across populations. For example, treating ulcerative sexually transmitted infections may have a significant effect on HIV incidence in an HIV epidemic localised within high risk groups but not in a more generalised epidemic. Assessing whether an intervention has met recipients' needs, or will meet those of future recipients, requires investigators to be explicit about the causal pathways through which an intervention is expected to act and to measure relevant pathway variables.

**Current assessment of generalisability**

We reviewed whether trials of HIV prevention targeting homosexually active men explored generalisability or factors affecting this. We obtained and examined all available evaluation reports of eight interventions that a recent systematic review reported to have rigorously evaluated outcomes. Two reviewers independently assessed whether the studies had empirically examined local factors affecting feasibility, coverage, and acceptability; evaluated process; assessed needs; and assessed the potential generalisability of interventions. Six of the eight trials had integral process evaluations, but only three of these collected quantitative and qualitative data on the planning, delivery, and receipt of the intervention (table 1). Only one process evaluation stated that consideration of generalisability was an aim. Six trials gave some information about participants' ethnicity (usually the proportion described as white). Seven trials provided data on educational level. None commented on the extent to which study samples were representative of the populations being targeted.

Only those trials incorporating process evaluations identified contextual factors influencing the feasibility, coverage, and acceptability of their intervention (table 2). Elford et al, for example, reported that recruitment and retention of peer educators to provide HIV prevention in gyms was extremely difficult because of educators' low confidence. Only one study reported on needs (table 2). Although other studies reported baseline sexual behaviour and sexual health related attitudes or knowledge of the target population or participants, the purpose was to check for baseline differences between intervention and comparison groups rather than to describe normative need.

Most of the studies speculated about the potential generalisability of their intervention to other sites but did not consider this empirically. Rosser et al, for example, wondered whether their intervention might prove more effective among populations with more risky sexual behaviour. The trials that examined contextual barriers and facilitators to delivering the intervention could make more considered assessments of generalisability. Two reports referred to sociological theory to hypothesise what contextual factors might have influenced the effect of the intervention in the study site compared with other sites. However, these trials both reported on interventions previously reported as effective in other contexts that were largely ineffective in their own sites. Therefore, rather than consider the scope for transferring the interventions to new sites, they (reasonably) considered the contextual reasons for failure of transfer.

**Systematic evaluation**

To make informed decisions about whether they should implement interventions, providers require more information than simply whether interventions are effective in original study sites. They need information on context and needs. However, most of the studies we looked at did not empirically examine generalisability. Phase III trials should be judged not only in terms of the designs and methods they use to exami-
This approach is compatible with a phased approach to intervention trials. Assessing generalisability in phase III should inform choice of sites for phase IV replicability research. However, such multi-site evaluations are unlikely unless funding for such work is increased.

Finally, systematic reviews should consider generalisability. Currently, many do not examine intervention process or context and do not comment on the potential for and limits to intervention effects being generalised to other settings and populations.

Contributors and sources: This article is based on an analysis of trials of HIV prevention for men who have sex with men that were identified in a systematic review. The authors have experience and expertise in primary evaluations and systematic reviews of public health interventions. They identified in a systematic review. The authors have experience and expertise in primary evaluations and systematic reviews of public health interventions.

Competing interests: None declared.


---

Table 2 Discussion of contextual factors and generalisability in eight studies of HIV prevention

<table>
<thead>
<tr>
<th>Outcome evaluation</th>
<th>Acceptability</th>
<th>Feasible delivery</th>
<th>Coverage</th>
<th>Local needs</th>
<th>Discussion of generalisability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dilley et al (2002)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Cost of intervention regarded as potential barrier to transfer</td>
</tr>
<tr>
<td>Elford et al (2001)</td>
<td>Educators thought intervention period insufficient to develop rapport</td>
<td>Recruitment and retention of educators difficult due to time, interest, and confidence problems. Time needed for planning was more than expected</td>
<td>Educators found it difficult to make contact (and discuss) with participants. Possibly related to social norms in UK, big city, or gym context</td>
<td>No</td>
<td>Intervention informed by US work and authors question its transfer to UK because of different norms</td>
</tr>
<tr>
<td>Flowers et al (2002)</td>
<td>Discussing sex in community sites went against local norms</td>
<td>Recruitment of popular peers difficult; educators found it difficult to discuss sex</td>
<td>No</td>
<td>No</td>
<td>Intervention informed by US work and authors question its transfer to UK because of different norms and needs</td>
</tr>
<tr>
<td>Gold and Rosenthal (1989)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Argues intervention is generalisable to other sites and populations because it is client centered</td>
</tr>
<tr>
<td>Imrie et al (2001)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Suggests intervention may be more effective in sites with higher rates of risk behaviour</td>
</tr>
<tr>
<td>Picciottos (2001)</td>
<td>No</td>
<td>More feasible to address some issues with minority ethnic than white men</td>
<td>More drop-outs among young and less educated men</td>
<td>No</td>
<td>Suggests prior rapport between health promoters and potential educators essential to recruitment</td>
</tr>
<tr>
<td>Rosser et al (2002)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Reported qualitative findings on sexual health needs from baseline interviews</td>
</tr>
<tr>
<td>Shepherd et al (1997)</td>
<td>No</td>
<td>Prior links between educators and health promoters enabled recruitment. Educators reluctant to address factors other than knowledge because of norms of what constitutes education</td>
<td>No</td>
<td>No</td>
<td>Suggests prior rapport between health promoters and potential educators essential to recruitment</td>
</tr>
</tbody>
</table>

---

Summary points

- Few randomised trials assess the generalisability of their results
- Such information is essential to decisions about adopting new interventions
- Trials should include evaluations of the feasibility, coverage, and acceptability of interventions
- They should also examine exactly for whom and what interventions are effective

---

ine outcomes but also how they assess generalisability. To enable this trials should:
- Include process evaluations as integral elements
- Develop evidence based theories about how intervention processes are influenced by context and how processes might differ if interventions are implemented in other sites
- Report the extent to which their participants are representative of the population being targeted
- Describe the prevalence of the needs being met by the intervention, informed by clear hypotheses about the intervention’s mechanism.

We believe that these elements are essential to comply with the existing CONSORT requirement to report on “clinical characteristics” of participants if clinical is interpreted as meaning need for health care. The most useful information on the potential for, as well as the barriers to, transfer of interventions comes from studies that compare an intervention in one site with similar interventions provided elsewhere, as in the study by Elford et al. Future phase III research might build on such work by setting out to examine interventions implemented across diverse contexts in multi-site studies. These would examine differential effects by site and explore contextual determinants of success to generate hypotheses for future research and guidelines for the implementation of interventions outside trials. This approach is compatible with a phased approach to intervention trials. Assessing generalisability in phase III should inform choice of sites for phase IV replicability research. However, such multi-site evaluations are unlikely unless funding for such work is increased.

Finally, systematic reviews should consider generalisability. Currently, many do not examine intervention process or context and do not comment on the potential for and limits to intervention effects being generalised to other settings and populations.
Confidentiality and consent in medical research
Balancing potential risks and benefits of using confidential data

Christina Davies, Rory Collins

Public health benefits arising from advances in medical research often rely on the use of personal data. How can we ensure that protecting patients’ interests does not unduly hamper scientific study?

Confidential medical information is used in almost every type of clinical and public health research. Different research scenarios raise different practical, ethical, and legal issues, and with these come the challenges of balancing the potential risks associated with the use of personal data against the potential benefits that might be gained from the research. We consider a strategy for explicitly reviewing the balance of these potential risks and benefits when planning research.

Effect of current legislation

Changes in the laws on data protection10–12 have had an important effect on training for medical research and on the design, costs, and feasibility of research projects. In many instances, this has improved the ways in which personal data are handled and protected the privacy of patients. There is, however, a general concern that varying interpretations of current legislation are stifling important research.13 Widespread uncertainty among professional bodies, hospital managers, ethics committees, clinicians, medical researchers, and the public may be producing disproportionate obstacles to the use of personal data when there is no genuine risk. In some instances, interpretations of legislation seem to have been driven less by careful consideration of the likelihood of real harm for individuals than by the desire to minimise the risk of criticism for organisations.

It needs just a few such decisions to impart an extra twist to the cycle of inefficiency in the use of public money for medical research. Clearly, research should conform to good practice, but it remains appropriate to consider whether over-interpretation of data protection legislation represents another real, albeit difficult to quantify, risk to the public.

Balancing risks and benefits

It is essential to achieve a rational view of the real risks and benefits of research using medical records and for any regulations to be drafted and interpreted appropriately. Risks and benefits can be presented from the perspectives both of safeguarding the interests of the participants in research and of pursuing the needs of future patients (rather than those whose data have been allowed to come to harm from research that uses information concerning them, particularly since it may be impossible to facilitate risk reduction among MSM: a pilot study. AIDS 2001; 15:251-61.


(Accepted 19 May 2006)