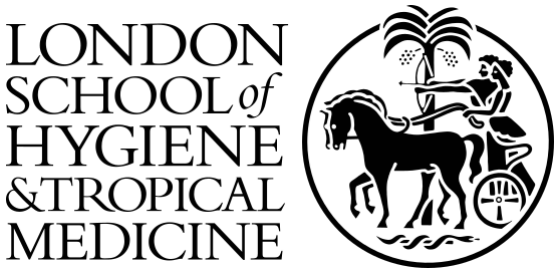


LSHTM Research Online

Bonell, C; Oakley, A; Hargreaves, J; Strange, V; Rees, R; (2006) Assessment of generalisability in trials of health interventions: suggested framework and systematic review. BMJ, 333 (7563). pp. 346-9. ISSN 1468-5833 DOI: https://doi.org/10.1136/bmj.333.7563.346

Downloaded from: http://researchonline.lshtm.ac.uk/11611/

DOI: https://doi.org/10.1136/bmj.333.7563.346

Usage Guidlines:

Please refer to usage guidelines at http://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: Creative Commons Attribution Non-commercial http://creativecommons.org/licenses/by-nc/3.0/

https://researchonline.lshtm.ac.uk

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

London School of Hygiene and Tropical Medicine, London WC1E 7HT
C Bonell *senior lecturer*
J Hargreaves *lecturer*
Social Science Research Unit, Institute of Education, University of London, London
A Oakley *professor*
V Strange *research officer*
R Rees *research officer*
Correspondence to: C Bonell chris.bonell@lshtm.ac.uk

BMJ 2006;333:346–9

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.[w1] 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.[1 w2] Such information is particularly crucial for resource poor settings.[2]

Although CONSORT criteria for reporting randomised trials include assessment of generalisability,[3] a framework for empirically assessing and reporting this is lacking. 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,[w3] as will institutions in being suitable places for an intervention.[w4] The presence of local "champions" may influence feasibility in a particular site.[4] Some interventions require the existence of other health services[4]—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.[4] For example, condom promotion has proved acceptable and subsequently effective in urban Tanzania but not in rural regions.[w5] 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.[w3]

Factors relating to delivery of an intervention are best documented by embedding an evaluation of process in trials.[5] 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 underrepresent 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.[6] 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.[w6]

References w1-w10 are on bmj.com

Table 1 Interventions and process evaluation in eight studies of HIV prevention

Outcome study	Additional process study	Intervention	Site	Any positive effects?	Methods of process evaluation	Stated rationale for process evaluation
Dilley et al (2002)[8]	None	Cognitive behavioural HIV prevention counselling	San Francisco HIV clinic	Yes	Adherence to treatment. Some sessions taped but data not reported	Assess consistency, completeness, and adherence to intervention guidelines
Elford et al (2001)[9]	Elford et al (2002),[10] Elford et al (2000)[11]	HIV peer education	London gyms	No	Survey of men's awareness, contact with and perceived usefulness of work. Interviews and group discussions with providers and stakeholders and documents about costs, planning, and delivery	Explore feasibility, practical constraints, transferability, and cultural adaptation
Flowers et al (2002)[12]	Flowers et al (1999)[13]	HIV peer education, gay specific genitourinary medicine service, and telephone advice line	Gay venues and clinic, Glasgow	No (intention to treat); yes (treatment analysis)	Observation of intervention. Survey of men's awareness and contact with work and its acceptability. Provider diaries, interviews and group discussions, and documents about planning and delivery	None
Gold and Rosenthal (1998)[14]	None	Face to face HIV prevention discussion versus posters	Homes, Melbourne and Sydney	No	Survey of men's views on usefulness of interventions	None
Imrie et al (2001)[15]	None	Cognitive behavioural HIV prevention workshop	Sexual health clinic, London	No	None	None
Picciano et al (2001)[16]	None	HIV telephone counselling	US	Yes	Counsellor reports of contents and ratings of each session	None
Rosser et al (2002)[17]	None	HIV education seminars	US university	Yes	None	None
Shepherd et al (1997)[18]	Shepherd et al (1999)[19]	HIV peer education	UK homes and other informal sites	Yes	Group discussions and interviews with educators, including drop-outs, about intervention training, reach, acceptability, and delivery	To examine how peer education can be undertaken with gay men not yet involved in the gay community

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.[w1] 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.[7] 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,[8–14 16 18] but only three of these collected quantitative and qualitative data on the planning, delivery, and receipt of the intervention (table 1).[9 12 18] Only one process evaluation stated that consideration of generalisability was an aim.[10] Six trials gave some information about participants' ethnicity (usually the proportion described as white).[8 9 15–18] Seven trials provided data on educational level.[8 9 12 14–17] 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.[10]

Only one study reported on needs (table 2).[19] Although other studies reported baseline sexual behaviour[8 9 12 14–17] or sexual health related attitudes or knowledge[8 12 16] 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.[17] 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.[10 12] However, these trials both reported on interventions previously reported as effective in other contexts[w7] 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 exam-

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

Outcome evaluation	Empirical examination of:				Discussion of generalisability
	Acceptability	Feasible delivery	Coverage	Local needs	
Dilley et al (2002)[8]	No	No	No	No	Cost of intervention regarded as potential barrier to transfer
Elford et al (2001)[9]	Educators thought intervention period insufficient to develop rapport	Recruitment and retention of educators difficult due to time, interest, and confidence problems. Time needed for planning was more than expected	Educators found it difficult to make contact (and discuss sex) with participants. Possibly related to social norms in UK, big city, or gym context	No	Intervention informed by US work and authors question its transfer to UK because of different norms
Flowers et al (2002)[12]	Discussing sex in community sites went against local norms	Recruitment of popular peers difficult; educators found it difficult to discuss sex	No	No	Intervention informed by US work and authors question its transfer to UK because of different norms and needs
Gold and Rosenthal (1998)[14]	No	No	No	No	No
Imrie et al (2001)[15]	No	No	No	No	No
Picciano et al (2001)[16]	No	More feasible to address some issues with minority ethnic than white men	More drop-outs among young and less educated men	No	Argues intervention is generalisable to other sites and populations because it is client centred
Rosser et al (2002)[17]	No	No	No	No	Suggests intervention may be more effective in sites with higher rates of risk behaviour
Shepherd et al (1997)[18]	No	Prior links between educators and health promoters enabled recruitment. Educators reluctant to address factors other than knowledge because of norms about what constitutes education	No	Reported qualitative findings on sexual health needs from baseline interviews	Suggests prior rapport between health promoters and potential educators essential to recruitment. Also that educators need longer involvement to address factors other than knowledge

ine outcomes[3] but also how they assess generalisability. To enable this trials should:

- Include process evaluations as integral elements[5]
- Develop evidence based theories about how intervention processes are influenced by context[w8] and how processes might differ if interventions are implemented in other sites[w9]
- Report the extent to which their participants are representative of the population being targeted[6]
- 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 intervention.

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.[10] Future phase III research might build on such work by setting out to examine interventions implemented across diverse contexts in multisite 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.[w9] 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.[20] 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.[w10]

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

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 and the integrated analysis of outcome and process data. All authors contributed to the conception and design of the analysis presented and to analysis and interpretation of the studies reviewed. All contributed to drafting and revising the intellectual content of the article. CB is the guarantor.

Competing interests: None declared.

1 Pawson R, Greenhalgh T, Harvey G, Walshe K. Realistic review—a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy* 2005;10(suppl 1):21-34.
2 Chinnock P, Siegfried N, Clarke M. Is evidence-based medicine relevant to the developing world? *PLoS Med* 2005;2(5):e107.
3 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191-4.
4 Kraft JM, Mezoff JS, Sogolow E, Spink Neumann M, Thomas PA. A technology transfer model for effective HIV/AIDS interventions: science and practice. *AIDS Educ Prev* 2000;12(suppl A):7-20.
5 Oakley A, Strange V, Bonell C, Allen E, Stephenson J, RIPPLE Study Team. Process evaluation in randomised controlled trials of complex interventions. *BMJ* 2006;332:413-6.
6 Bartlett C, Doyal L, Ebrahim S, Davey P, Backmann M, Egger M, et al. The causes and effects of socio-demographic exclusions from clinical trials. *Health Technol Assess* 2005;9(38).
7 Rees R, Harden A, Thomas J, Oliver S, Kavanagh J, Burchett H, et al. *HIV health promotion and men who have sex with men (MSM): a systematic review of research relevant to the development and implementation of effective and appropriate interventions.* London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, 2004. http://eppi.ioe.ac.uk/EPPIWebContent/hp/reports/MSM/MSM_HIV_2004.pdf (accessed 21 Jun 2006).
8 Dilley JW, Woods WJ, Sabatino J, Lihatsh T, Adler B, Casey S, et al. Changing sexual behavior among gay male repeat testers for HIV. *J Acquir Immune Defic Syndr* 2002;30:177-86.
9 Elford J, Bolding G, Sherr L. Peer education has no significant impact on HIV risk behaviours among gay men in London. *AIDS* 2001;15:535-8.

10 Elford J, Sherr L, Bolding G, Serle F, Maguire M. Peer-led HIV prevention among gay men in London: process evaluation. *AIDS Care* 2002;14: 351-60.
11 Elford J, Sherr L, Bolding G, Maguire M, Serle F. Peer-led HIV prevention among gay men in London (the 4 gym project): intervention and evaluation. In: Watson J, Platt S, eds. *Researching health promotion.* New York: Routledge, 2000:207-30.
12 Flowers P, Hart GJ, Williamson LM, Frankis JS, Der GJ. Does bar-based, peer-led sexual health promotion have a community-level effect amongst gay men in Scotland? *Int J STD AIDS* 2002;13:102-8.
13 Flowers P, Hart G. Everyone on the scene is so cliquey. In: Aggleton P, Hart GJ, Davies P, eds. *Families and communities responding to AIDS.* London: UCL Press, 1999:83-98.
14 Gold RS, Rosenthal DA. Examining self-justifications for unsafe sex as a technique of AIDS education: the importance of personal relevance. *Int J STD AIDS* 1998;9:208-13.
15 Imrie J, Stephenson J, Cowan F, Wanigaratne S, Billington AJP, Copas A, et al. A cognitive behavioural intervention to reduce sexually transmitted infections among gay men: randomised trial. *BMJ* 2001;322:1451-6.
16 Picciano JF, Roffman RA, Kalichman SC, Rutledge SE, Berghuis JP. A telephone based brief intervention using motivational enhancement to facilitate risk reduction among MSM: a pilot study. *AIDS Behav* 2001; 5:251-61.
17 Rosser BRS, Bochting BO, Rugg DL, Robinson BE, Ross MW, Bauer GR, et al. A randomized controlled intervention trial of a sexual health approach to long-term HIV risk reduction for men who have sex with men: effects of the intervention on unsafe sexual behaviour. *AIDS Educ Prev* 2002;14(suppl A):69-71.
18 Shepherd J, Weare K, Turner G. Peer-led sexual health promotion with young gay and bisexual men: results of the HAPEER project. *Health Educ* 1997;6:204-12.
19 Shepherd J, Turner G, Weare K. A new method of peer-led HIV prevention with gay and bisexual men. In: Aggleton P, Hart GJ, Davies P, eds. *Families and communities responding to AIDS.* London: UCL Press, 1999:163-84.
20 Campbell M, Fitzpatrick R, Haines A, Kinmouth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;321:694-6.

(Accepted 19 May 2006)

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?

Editorial by Souhami

This article is the last in a four part series building on a recent Medical Research Council initiative relating to use of personal information in medical research

MRC/BHF/Cancer Research UK Clinical Trial Service Unit and Epidemiological Studies Unit, University of Oxford, Oxford OX3 7LF
Christina Davies *senior research fellow*
Rory Collins *British Heart Foundation professor of medicine and epidemiology*
Correspondence to: R Collins secretary@ctsu.ox.ac.uk

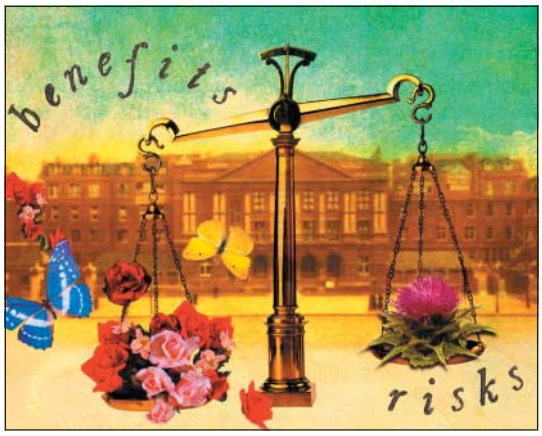
BMJ 2006;333:349–51

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 protection[1–3] 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.[4] 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 not 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.



SARAH PERKINS

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 patients and the wider public for evidence on which to base healthcare decisions.[5] Individuals should not be allowed to come to harm from research that uses information concerning them, particularly since it may be future patients (rather than those whose data have been used) who benefit from such research. There is, however, little evidence that serious harm has been caused by the use of confidential records in medical research.[4]

When designing a research project using confidential data, researchers should consider the ways in which the data are to be used and the measures to be taken to protect confidentiality. They should assess the likelihood of any harm being caused to individuals and the value of