

Randomised controlled trials for policy interventions: a review of reviews and meta-regression

S Oliver,^{1*} AM Bagnall,² J Thomas,¹
J Shepherd,³ A Sowden,² I White,⁴
J Dinnes,³ R Rees,¹ J Colquitt,³ K Oliver¹
and Z Garrett¹

¹Social Science Research Unit, Institute of Education, University of London, UK

²Centre for Reviews and Dissemination, University of York, UK

³Southampton Health Technology Assessments Centre, Wessex Institute for Health Research & Development, University of Southampton, UK

⁴Medical Research Council Biostatistics Unit, Institute of Public Health, Cambridge, UK

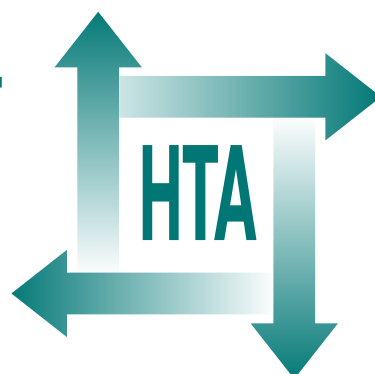
*Corresponding author



Executive summary

Health Technology Assessment 2010; Vol. 14: No. 16
DOI: 10.3310/hta14160

Health Technology Assessment
NIHR HTA programme
www.hta.ac.uk





Executive summary

Background

While the randomised controlled trial (RCT) is generally regarded as the design of choice for assessing the effects of health care, within the social sciences there is considerable debate about the relative suitability of RCTs and non-randomised studies (NRSs) for evaluating public policy interventions.

Objectives

To determine whether RCTs provide the same effect size and variance as NRSs of similar policy interventions, and whether these findings can be explained by other associated factors.

Methods

This study employed four approaches:

1. Resampling studies: comparing controlled trials that are identical in all respects other than the use of randomisation by 'breaking' the randomisation in a trial to create smaller non-randomised trials and smaller randomised trials by resampling randomised and non-randomised comparisons from the data.
2. Replication studies: comparing randomised and non-randomised arms of controlled trials mounted simultaneously in the field.
3. Investigating comparable 'field' studies: controlled trials drawn from systematic reviews that include both randomised and non-randomised studies. These include structured narrative reviews and sensitivity analyses within meta-analyses.
4. Meta-epidemiology: investigating associations between randomisation and effect size using a pool of more diverse randomised and non-randomised studies within broadly similar areas. These more diverse studies can be drawn from across reviews addressing different questions, or from broad sections of literature.

This study sought earlier reports of all four approaches and conducted new analyses for three of these approaches (1, 3 and 4 above) across a range of public policy sectors. The new analyses

were strengthened by testing pre-specified associations supported by carefully argued hypotheses. Data were drawn from: two RCTs of policy interventions for resampling studies; comparable studies drawn from systematic reviews of health promotion and of transition for youths with disabilities; and a systematic search for prior work. The search strategy comprising free text terms for RCT and non-randomised studies (e.g. non-experimental, pseudorandom, semi-random) was applied to 14 electronic bibliographic databases spanning health, education, social policy and social science in June and July 2004 [Applied Social Sciences Index and Abstracts (ASSIA), Australian Education Index (AEI), British Education Index (BEI), CareData, Dissertation Abstracts, EconLIT, Educational Resources Information Centre (ERIC), International Bibliography of the Sociological Sciences (IBSS), ISI Proceedings: Social Sciences and Humanities, PAIS International (Public Affairs Information Service), PsycINFO, SIGLE (System for Information on Grey Literature in Europe), Social Science Citation Index (SSCI), Sociological Abstracts]. This was supplemented by citation searching for key authors, contacting review authors and searching key internet sites.

For investigating comparable field studies, and the meta-regression, studies were coded for characteristics of the population, policy intervention and evaluation. Differences in effect sizes between studies were investigated using random-effects meta-regression to allow for unexplained heterogeneity between studies as well as the known uncertainty in estimated effect sizes (measured by their standard errors). Associations between different characteristics of the studies and whether or not they employed randomisation were measured using chi-squared tests.

Results

Reviews of methodological studies and empirical reviews

Prior methodological reviews included a review of within-study comparisons of randomised and non-randomised participants, six single meta-analyses and one review of meta-analyses. Between them these covered interventions for preventing juvenile delinquency, treatment of alcohol abuse,

and other psychological, mental health or health-care interventions. These studies investigated whether randomisation influenced effect sizes. Most also investigated the influence of other variables or modifiers of effect such as population, sample size, attrition, intervention, type of control group and publication status. The results suggest that effect sizes from RCTs and non-randomised controlled trials (nRCTs) may indeed differ in some circumstances and that these differences may well be associated with factors confounded with design. Inter-relationships among variables make it difficult to determine the likely impact of any one factor.

A systematic review of meta-analyses of existing reviews comparing effects from RCTs and nRCTs found that the effect sizes were similar in five reviews, dissimilar in eight reviews, and mixed in three. Most reviews appeared to ignore the variability associated with effect size. Considerable variation in the studies pooled within reviews, in terms of population, intervention, outcome and other methodological details, makes it difficult to separate the potential effect of random assignment from the potential effects of all the other variables.

Resampling studies

Re-analysis of data from two trials suggests that nRCTs can give the same answers as RCTs. This was a tightly controlled examination in which the only factor that was different between the RCTs and nRCTs was randomisation.

Comparable 'field' studies and meta-epidemiology

In the examination of trials sampled from systematic reviews we found considerable variation, with RCTs producing smaller effect sizes than nRCTs in systematic reviews conducted at the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) (using within review comparisons and meta-epidemiology) and larger effect sizes than nRCTs in the studies reviewed by Colorado State University (using meta-epidemiology alone).

Investigation of potential confounding factors in the EPPI-Centre reviews suggests that RCTs have smaller effect sizes, even though their sample sizes tend to be smaller with participants allocated individually (both attributes associated to some extent with effect size) and their theoretical frameworks more readily apparent. Other attributes commonly associated with quality were not associated with randomisation or effect size:

attrition rates, time to follow-up or quality of reporting.

Conclusions

From the resampling studies we have no evidence that the absence of randomisation directly influences the effect size of policy interventions in a systematic way. At the level of individual studies, non-randomised trials may lead to different effect sizes, but this is unpredictable. Many of the examples reviewed and the new analyses in the current study reveal that randomisation is indeed associated with changes in effect sizes of policy interventions in field trials. Despite extensive analysis, we have identified no consistent explanations for these differences.

Recommendations for research

1. Policy evaluations should adopt randomised designs whenever possible.
2. Policy evaluations should also adopt other standard procedures for minimising bias and conducting high-quality assessment of effects of intervention, particularly blinded allocation of either individuals or groups and the avoidance of small sample sizes.
3. Feasibility studies of randomising geographical areas, communities and regions should be carried out for evaluating policy interventions in a range of sectors, implemented within interventions, communities and across regions.
4. Feasibility studies of blinded allocation should be carried out for policy interventions in a range of sectors, implemented within interventions, communities and across regions.
5. Clear descriptions should be included in systematic reviews of how judgements of equivalence (or otherwise) have been reached when comparing the effects found in randomised and non-randomised studies of policy interventions.
6. Research is required into the reasons for choosing randomisation or not, particularly in the presence and absence of an explicit collective plan of action.

Publication

Oliver S, Bagnall AM, Thomas J, Shepherd J, Sowden A, White I, *et al.* Randomised controlled trials for policy interventions: a review of reviews and meta-regression. *Health Technol Assess* 2010;**14**(16).

NIHR Health Technology Assessment programme

The Health Technology Assessment (HTA) programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needed in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this issue of the journal was commissioned by the National Coordinating Centre for Research Methodology (NCCRM), and was formally transferred to the HTA programme in April 2007 under the newly established NIHR Methodology Panel. The HTA programme project number is 06/90/22. The contractual start date was in April 2004. The draft report began editorial review in January 2009 and was accepted for publication in March 2009. The commissioning brief was devised by the NCCRM who specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

Editor-in-Chief: Professor Tom Walley CBE

Series Editors: Dr Martin Ashton-Key, Dr Aileen Clarke, Professor Chris Hyde,
Dr Tom Marshall, Dr John Powell, Dr Rob Riemsma and Professor Ken Stein

ISSN 1366-5278

© 2010 Queen's Printer and Controller of HMSO

This journal may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to: NETSCC, Health Technology Assessment, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.

Printed on acid-free paper in the UK by the Charlesworth Group.