Assessing methods to specify the target difference for a randomised controlled trial: DELTA (Difference ELicitation in TriAls) review

Jonathan A Cook,[1]* Jennifer Hislop,[1] Temitope E Adewuyi,[1] Kirsten Harrild,[2] Douglas G Altman,[3] Craig R Ramsay,[1] Cynthia Fraser,[1] Brian Buckley,[4] Peter Fayers,[5] Ian Harvey,[6] Andrew H Briggs,[7] John D Norrie,[1] Dean Fergusson,[8] Ian Ford[9] and Luke D Vale[10]

[1]Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[2]Medical Statistics Team, University of Aberdeen, Aberdeen, UK
[3]Centre for Statistics in Medicine, University of Oxford, Oxford, UK
[4]Department of General Practice, National University of Ireland, Galway, Ireland
[5]Population Health, University of Aberdeen, Aberdeen, UK
[6]Faculty of Medicine and Health Sciences, University of East Anglia, Norwich, UK
[7]Health Economics and Health Technology Assessment, University of Glasgow, Glasgow, UK
[8]Ottawa Hospital Research Institute, Ontario, Canada
[9]Robertson Centre for Biostatistics, University of Glasgow, Glasgow, UK
[10]Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK

*Corresponding author

Published May 2014
DOI: 10.3310/hta18280

Scientific summary

Assessing methods to specify the target difference for a RCT: DELTA review

Health Technology Assessment 2014; Vol. 18: No. 28
DOI: 10.3310/hta18280

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

The randomised controlled trial (RCT) is widely considered to be the gold standard study for comparing the effectiveness of health interventions. Central to the design and validity of a RCT is a calculation of the number of participants needed (the sample size). This provides reassurance that the trial will identify a difference of a particular magnitude if such a difference exists. The value used to determine the sample size can be considered the 'target difference'. From both a scientific and an ethical standpoint, selecting an appropriate target difference is of crucial importance. Specifying too small a target difference could be a wasteful (and unethical) use of data and resources. Conversely, too large a target difference could lead to an important difference being easily overlooked because the study is too small. Furthermore, an undersized study may not usefully contribute to the knowledge base and could potentially have a detrimental impact on decision-making.

Determination of the target difference, as opposed to statistical approaches to calculating the sample size, has been greatly neglected. A variety of approaches have been proposed for formally specifying what an important difference should be [such as the 'minimal clinically important difference (MCID)'], although the current state of the evidence is unclear, particularly with regard to informing RCT design by specifying a target difference.

Aim

The aim was to provide an overview of the current evidence on methods for specifying the target difference in a RCT sample size calculation.

Objectives

- To conduct a systematic review of methods for specifying a target difference.
- To evaluate current practice by surveying triallists.
- To develop guidance on specifying the target difference for a RCT.
- To identify future research needs.

Methods

The study comprised three interlinked components.

Systematic review of methods for specifying a target difference for a randomised controlled trial

A comprehensive search of both biomedical and some non-biomedical databases was undertaken. Additionally, clinical trial textbooks and guidelines were reviewed. To be included, a study had to report a formal method that could potentially be used to specify a target difference. The biomedical and social science databases searched were MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, EMBASE, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Methodology Register, PsycINFO, Science Citation Index, EconLit, Education Resources Information Center (ERIC) and Scopus for in-press publications. All were searched from 1966 or the earliest date of the database coverage and searches were undertaken between November 2010 and January 2011.

Identification of triallists' current practice

This involved two surveys:

- Members of the Society for Clinical Trials (SCT) were sent an invitation (followed by a reminder) to complete an online survey through the society's email distribution list. Respondents were asked about their awareness and use of, and willingness to recommend, methods for determining a target difference in a RCT.
- Survey of leading UK- and Ireland-based triallists. The survey was sent to UK Clinical Research Collaboration (UKCRC)-registered Clinical Trials Units (CTUs), Medical Research Council (MRC) UK Hubs for Trials Methodology Research and National Institute for Health Research (NIHR) Research Design Services (RDS). One response per triallist group was invited. In addition to the information collected in the SCT survey, this survey included questions on the approach used for the most recent trial developed. The initial request was personalised and sent by post, followed by two reminders.

Production of guidance on specifying the target difference for a randomised controlled trial

The draft guidance was developed by the project steering and advisory groups utilising the results of the systematic review and surveys. Findings were circulated and presented to members of the combined group at a face-to-face meeting along with a proposed outline of the guidance document structure and list of recommendations. Both the structure and main recommendations were agreed at this meeting. The guidance was subsequently drafted and circulated for further comment.

Results

Systematic review of methods for specifying a target difference for a randomised controlled trial

The search identified 11,485 potentially relevant studies, of which 1434 were selected for full-text assessment, with 777 included in the review. Fifteen clinical trial textbooks and the International Conference on Harmonisation (ICH) tripartite guidelines were also reviewed. Seven methods were identified – anchor, distribution, health economic, opinion-seeking, pilot study, review of evidence base (RoEB) and standardised effect size (SES) – each with important variations. The most frequently identified methods used to determine an important difference were the anchor, distribution and SES methods. No new methods were identified by this review beyond the seven pre-identified methods described earlier; however, substantial variations in the implementation of each method were detected. It is critical when specifying a target difference to decide whether the focus is to determine an important and/or a realistic difference as the appropriate methods vary accordingly. Some methods for determining an important difference within an observational study are not appropriate for specifying a target difference in a RCT (e.g. statistical hypothesis testing approach). Multiple methods for determining an important difference were used in some studies although the combinations varied, as did the extent to which results were triangulated.

Identification of triallists' current practice

The two surveys regarding formal methods to determine the target difference in a RCT provided insight into current practice among clinical triallists.

Society for Clinical Trials survey

Of the 1182 members on the SCT membership email distribution list, 180 responses were received (15%). Awareness ranged from 69 (38%) for the health economic method to 162 (90%) for the pilot study method. Usage was lower than awareness and ranged from 16 (9%) for the health economic method to 133 (74%) for the pilot study method. The highest level of willingness to recommend was for the RoEB method ($n = 132$, 73%) and the lowest was for the health economic method ($n = 28$, 16%). Willingness to recommend among those who had used a particular method was substantially higher than across all

respondents: the lowest level was for the opinion-seeking method ($n = 40$, 56%) and the highest level was for the RoEB method ($n = 118$, 89%).

UK- and Ireland-based triallist survey

Of the 61 surveys sent out, 34 (56%) responses were received. Awareness of methods ranged from 97% ($n = 33$) for the RoEB and pilot methods to only 41% ($n = 14$) for the distribution method. All respondents were aware of at least one of the different formal methods for determining the target difference. Usage ranged from 24% ($n = 8$) for both the distribution and health economic methods to 94% ($n = 32$) for the RoEB method. Usage was substantially less than awareness for all methods except for the pilot study, RoEB and SES methods. The highest level of willingness to recommend was for the RoEB method (76%, $n = 26$) followed by the SES method (65%, $n = 22$), with the distribution method having the lowest level of willingness to recommend (26%, $n = 9$). Based on the most recent trial ($n = 33$), all bar three groups (91%, $n = 30$) used a formal method. The vast majority (91%, $n = 30$) stated that the target difference was one that was viewed as important by a stakeholder group. Just over half (61%, $n = 20$) stated that the basis for determining the target difference was to achieve a realistic difference given the interventions under evaluation.

Guidance on specifying the target difference in a randomised controlled trial

Guidance was developed for specifying the target difference in a RCT. Additionally, guidance on reporting the sample size calculation was developed which includes a minimum set of items for reporting the specification of the target difference in the trial protocol and main results paper. A minimum set of items for reporting the specification of the target difference in the trial protocol and main results paper was developed.

Conclusions

The specification of the target difference is a key component of a RCT design. There is a clear need for greater use of formal methods to determine the target difference and for better reporting of its specification. Although no single method provides a perfect solution to a difficult question, methods are available to inform specification of the target difference and should be used whenever feasible. Raising the standard of RCT sample size calculations and the corresponding reporting of them would aid health professionals, patients, researchers and funders in judging the strength of the evidence and ensure better use of scarce resources.

Further research priorities

1. A comprehensive review of observed effects in different clinical areas, populations and outcomes is needed to assess the generalisability of the Cohen's interpretation for continuous outcomes, and to provide guidance for binary and survival (time-to-event) measures. To achieve this, an accessible database of SESs should be set up and maintained. This would aid the prioritisation of research and help researchers, funders, patients and health-care professional assess the impact of interventions.
2. Prospective comparison of formal methods for specifying the target difference is needed in the design of RCTs to assess the relative impact of different methods.
3. Practical use of the health economic approach is needed; the possibility of developing a decision model structure that reflects the view of a particular funder (e.g. the Health Technology Assessment programme) and incorporates all relevant aspects, should be explored.
4. Further exploration of the implementation of the opinion-seeking approach in particular is needed. The reliability of a suggested target difference that would lead to a change in practice should be explored. Additionally, the impact of eliciting the opinion of different stakeholders should also be evaluated.

5. The value of the pilot study for estimating parameters (e.g. control group event proportion) for a definitive study should be further explored by comparing pilot study estimates with the resultant definitive trial results.
6. Qualitative research on the process of specifying a target difference in the context of developing a RCT should be carried out to explore the determining factors and interplay of influences.

Funding

The Medical Research Council UK and the National Institute for Health Research Joint Methodology Research Programme.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Five-year impact factor: 5.596

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (http://www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) 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 reviewers 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.

HTA programme

The 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 journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.hta.ac.uk/

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 06/98/01. The contractual start date was in October 2010. The draft report began editorial review in April 2012 and was accepted for publication in August 2012. 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 reviewers 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.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).