Appropriate design and reporting of superiority, equivalence and non-inferiority clinical trials incorporating a benefit–risk assessment: the BRAINS study including expert workshop

Nikki Totton,[1*] Steven A Julious,[1] Elizabeth Coates,[1] Dyfrig A Hughes,[2] Jonathan A Cook,[3] Katie Biggs,[1] Catherine Hewitt,[4] Simon Day[5] and Andrew Cook[6]

[1]School of Health and Related Research, University of Sheffield, Sheffield, UK
[2]Centre for Health Economics and Medicines Evaluation, Bangor University, Bangor, UK
[3]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[4]York Trials Unit, Department of Health Sciences, University of York, York, UK
[5]Clinical Trials Consulting & Training Limited, Buckingham, UK
[6]Wessex Institute, University of Southampton, Southampton, UK

[*]Corresponding author n.v.totton@sheffield.ac.uk

Disclosure of interests of authors

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/BHQZ7691.

Primary conflicts of interest: Dyfrig A Hughes receives support from the Medical Research Council Trials Methodological Research Partnership (TMRP), reference number MR/S014357/1, and is a Health and Care Research Wales (H&CRW) Senior Research Leader, reference number 2021/22. Andrew Cook reports research contracts from the National Institute for Health and Care Research (NIHR) Public Health Research (PHR), Health Technology Assessment (HTA) and Invention for Innovation (i4i) programmes, payment for which was made to his employer. He also reports reimbursement for travel and provision of accommodation in March 2019 to attend a workshop on approaches to pancreatic surgery in Miami, FL, USA. Andrew Cook is also a member of the Data Monitoring Committee of the Diabetes Prevention – Long Term Multimethod Assessment 2 (DIPLOMA-2) study. Nikki Totton received financial support (paid to institution) for protected time on the project, as well as consumables and dissemination through conferences.

Published October 2023
DOI: 10.3310/BHQZ7691

Scientific summary

Appropriate design and reporting of superiority, equivalence and non-inferiority clinical trials incorporating a benefit–risk assessment: the BRAINS study including expert workshop

Health Technology Assessment 2023; Vol. 27: No. 20
DOI: 10.3310/BHQZ7691

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

Scientific summary

Background

Randomised controlled trials (RCTs) are the gold standard of health technology evaluation. They can be designed to assess the objective of superiority, equivalence or non-inferiority; the objective is determined by the selected research question and primary outcome. However, in practice, selecting the most appropriate outcome measure and, subsequently, trial design can be difficult.

In addition, some trials may have more than one outcome of importance to consider; this is particularly the case when using equivalence or non-inferiority designs, as a secondary superiority outcome is often important. This multidimensionality is not considered if the focus of a trial's success is a single primary outcome.

Benefit–risk (B–R) methods are used in the clinical trial regulatory setting to assess multiple outcomes and consider the trade-off of the benefits against the risks, but they are not regularly implemented in publicly funded UK trials.

Objectives

This guidance document aims to fill a knowledge gap, with a focus on publicly funded clinical trials throughout. This is undertaken by first providing recommendations on the most appropriate trial design to select and then identifying when a B–R method may be used within a trial to assess outcome trade-offs, whether this be qualitatively or quantitatively.

Methods

In this project, three key methodologies were used across three work packages (WPs) to elicit expert opinion on trial design and B–R methods. These were:

- WP1, a web-based survey of relevant researchers in the area. The survey was sent to researchers through appropriate mailing lists, as well as through known contacts and networks, for completion.
- WP2, a rapid review of current literature. The review built on previous reviews to find available B–R methods, but, in addition to previous work, we retrieved details on which of these could be implemented in publicly funded studies, as well as details on each method's strengths and weaknesses.
- WP3, a 2-day expert consensus workshop. Results from the survey and the review were presented to a group of experts ($n = 15$) who had been contacted because of their knowledge of and interest in the area. The nominal group technique was used to select items for inclusion in checklists. Open discussions, supported by the presented findings, were used to gain opinions on reasons for including B–R methods within a trial. Open discussions were thematically analysed to identify the key points discussed and support the results of the checklists.

Results

To aid researchers in selecting an appropriate trial design, a list of 19 factors to consider was created. This list relates to six different sections: population, intervention, comparator, outcomes, feasibility and

perspectives. Each factor is described in detail and the list of factors is used in conjunction with examples of trial designs to explain why a particular design might be chosen.

Once the appropriate trial design has been selected, considering a B–R method may be justified if one of six key reasons is present:

1. The success of the trial depends on more than one outcome.
2. Important outcomes within the trial are in competing directions (i.e. a health technology is anticipated to be better on one outcome but worse on another).
3. To allow patient preferences to be included and directly influence trial results.
4. To provide transparency on subjective recommendations from a trial.
5. To provide consistency in the approach to presenting results from a trial.
6. To synthesise multiple outcomes into a single metric.

In particular, the first reason (considering multiple outcomes) is often applicable to equivalence and non-inferiority trials, in which success also depends on the superiority of a secondary outcome; a description of this was flagged as important in the consensus workshop.

Owing to the range of B–R methods available, difficulties arise when discussing a globally recommended B–R assessment. To aid with this, the 92 methods that were identified from the literature review as potentially suitable for use in publicly funded RCTs were categorised into seven groups. The first group was an overarching framework that provided structure for a B–R assessment from start to finish. The most basic group, the second group, was the use of a narrative summary of the benefits and risks; although this is not an official method, it was considered important to provide such a summary at the end of a trial. The next two groups comprised a summary table with a specific structure to contain all relevant information, before moving on to methods that can quantitatively trade off multiple outcomes. The final three groups related to preference elicitation from stakeholders, uncertainty estimation and visualisations.

Importantly, these groupings show that quantitative methods are not always necessary or appropriate. However, this should not negate the need for a summary or narrative discussion of subjective interpretations. As it could be argued that favourable and unfavourable outcomes are present in all RCTs, a narrative summary of the benefits and risks would add value to the reports of all publicly funded trials.

It is noted that some methods require additional data to be collected, so the need for these added resources and expenses should be considered on a case-by-case basis. To provide some clarity on this, the methods groups were linked with different stages of the RCT, from design through to conclusion and dissemination. Group discussions identified the difficulty of additional work during the trial design stage, which is something that this breakdown of the methods at different stages hopes to address.

Finally, when a B–R method has been deemed important to use, whether that be a narrative summary or a quantitative trade-off method, the reporting should be transparent and consistent. A checklist of items to include when reporting the trial design in applications or protocols was created, as was a separate checklist for reporting the trial findings.

There are five pieces of information that are recommended when reporting the trial design:

1. a heading of ‘benefit–risk’
2. explicit use of the term ‘benefit–risk’
3. a plan for B–R assessment
4. the anticipated benefits and risks
5. discussion of the B–R balance with patient representatives.

Two further items are recommended when reporting the results of the trial:

6. a summary table of benefits and risks
7. reporting of quality-adjusted life-years in terms of B–R.

Conclusions

The advice and recommendations provided in this report aim to improve the selection of a trial design, understanding of B–R methods and when it is appropriate to include them and how to effectively report the use of the methods. This guidance is aimed at researchers designing publicly funded RCTs, such as those by the National Institute for Health and Care Research or the Medical Research Council, as well as those on the funding panels.

Using these recommendations will improve the design of RCTs, as well as the clarity and consistency of reporting the benefits and risks associated with the health technologies being evaluated.

Further research should focus on providing more detail on each of the different B–R methods and recommendations on how they can be fully integrated in publicly funded clinical trials. Practical barriers to this implementation could be assessed to ensure that the information presented in this report can be used most effectively.

Study registration

This study is registered as PROSPERO CRD42019144882.

Funding

Funded by the Medical Research Council UK and the National Institute for Health and Care Research as part of the Medical Research Council–National Institute for Health and Care Research Methodology Research programme.

Health Technology Assessment

HTA MRC MSAW

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.6

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2021 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), Embase (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

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

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

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 or, commissioned/managed through the MRC–NIHR Methodology Research Programme (MRP), 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

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

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.

This report

This issue of the Health Technology Assessment journal series contains a project commissioned by the MRC–NIHR Methodology Research Programme (MRP). MRP aims to improve efficiency, quality and impact across the entire spectrum of biomedical and health- related research. In addition to the MRC and NIHR funding partners, MRP takes into account the needs of other stakeholders including the devolved administrations, industry R&D, and regulatory/advisory agencies and other public bodies. MRP supports investigator-led methodology research from across the UK that maximises benefits for researchers, patients and the general population – improving the methods available to ensure health research, decisions and policy are built on the best possible evidence.

To improve availability and uptake of methodological innovation, MRC and NIHR jointly supported a series of workshops to develop guidance in specified areas of methodological controversy or uncertainty (Methodology State-of-the-Art Workshop Programme).

Workshops were commissioned by open calls for applications led by UK-based researchers. Workshop outputs are incorporated into this report, and MRC and NIHR endorse the methodological recommendations as state-of-the-art guidance at time of publication.

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 under a MRC–NIHR partnership. 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, the MRC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the MRC, the HTA programme or the Department of Health and Social Care

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland, and final files produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).