Use of methods for specifying the target difference in randomised controlled trial sample size calculations: Two surveys of trialists' practice

Jonathan A Cook, Jennifer M Hislop, Doug G Altman, Andrew H Briggs, Peter M Fayers, John D Norrie, Craig R Ramsay, Ian M Harvey, Luke D Vale and for the DELTA group

Clin Trials 2014 11: 300 originally published online 6 March 2014
DOI: 10.1177/1740774514521907

CLINICAL TRIALS ARTICLE Clinical Trials 2014; 11: 300–308

Use of methods for specifying the target difference in randomised controlled trial sample size calculations: Two surveys of trialists' practice

Jonathan A Cook[a,b], Jennifer M Hislop[c], Doug G Altman[b], Andrew H Briggs[d], Peter M Fayers[e,f], John D Norrie[a], Craig R Ramsay[a], Ian M Harvey[g], Luke D Vale[c]; for the DELTA group

Background Central to the design of a randomised controlled trial (RCT) is a calculation of the number of participants needed. This is typically achieved by specifying a target difference, which enables the trial to identify a difference of a particular magnitude should one exist. Seven methods have been proposed for formally determining what the target difference should be. However, in practice, it may be driven by convenience or some other informal basis. It is unclear how aware the trialist community is of these formal methods or whether they are used.

Purpose To determine current practice regarding the specification of the target difference by surveying trialists.

Methods Two surveys were conducted: (1) Members of the Society for Clinical Trials (SCT): participants were invited to complete an online survey through the society's email distribution list. Respondents were asked about their awareness, use of, and willingness to recommend methods; (2) Leading UK- and Ireland-based trialists: the survey was sent to UK Clinical Research Collaboration registered Clinical Trials Units, Medical Research Council UK Hubs for Trial Methodology Research, and the Research Design Services of the National Institute for Health Research. This survey also included questions about the most recent trial developed by the respondent's group.

Results *Survey 1*: Of the 1182 members on the SCT membership email distribution list, 180 responses were received (15%). Awareness of methods ranged from 69 (38%) for health economic methods to 162 (90%) for pilot study. Willingness to recommend among those who had used a particular method ranged from 56% for the opinion-seeking method to 89% for the review of evidence-base method. *Survey 2*: Of the 61 surveys sent out, 34 (56%) responses were received. Awareness of methods ranged from 33 (97%) for the review of evidence-base and pilot methods to 14 (41%) for the distribution method. The highest level of willingness to recommend among users was for the anchor method (87%). Based upon the most recent trial, the target difference was usually one viewed as important by a stakeholder group, mostly also viewed as a realistic difference given the interventions under evaluation, and sometimes one that led to an achievable sample size.

[a]Health Services Research Unit, University of Aberdeen, Aberdeen, UK, [b]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK, [c]Institute of Health & Society, Newcastle University, Newcastle upon Tyne, UK, [d]Health Economics and Health Technology Assessment, University of Glasgow, Glasgow, UK, [e]Population Health, University of Aberdeen, Aberdeen, UK, [f]Department of Cancer Research and Molecular, Norwegian University of Science and Technology, Trondheim, Norway, [g]Faculty of Medicine and Health Sciences, University of East Anglia, Norwich, UK

Author for correspondence: Jonathan A Cook, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Nuffield Orthopaedic Centre, Windmill Rd, Oxford, OX3 7LD, UK.
Email: jonathan.cook@ndorms.ox.ac.uk

10.1177/1740774514521907

Limitations The response rates achieved were relatively low despite the surveys being short, well presented, and having utilised reminders.

Conclusion Substantial variations in practice exist with awareness, use, and willingness to recommend methods varying substantially. The findings support the view that sample size calculation is a more complex process than would appear to be the case from trial reports and protocols. Guidance on approaches for sample size estimation may increase both awareness and use of appropriate formal methods.
Clinical Trials 2014; **11:** 300–308. http://ctj.sagepub.com

Introduction

Central to the design of a randomised controlled trial (RCT) is a calculation of the number of participants needed [1,2]. As part of this calculation, a target difference, the magnitude of a difference which is desired to be detected, is typically specified to provide reassurance that the trial will be informative and able to answer the research question posed with required statistical precision and certainty. It is important that an appropriate number of participants are sought as too large or small a study is arguable unethical, wasteful, and potentially misleading. Despite its critical role, the specification of the target difference has received little attention. A detailed systematic review and synthesis of the literature, which is reported elsewhere [3], identified seven methods for specifying a target difference (see Table 1). It is unclear the extent to which trialists are aware of these methods and whether they are used in practice when designing clinical trials. Reports of trial results and protocols will typically report the sample size calculation and the values assumed therein [1]. The process of how these values were determined (including the target difference) typically lacks detail, particularly in reports of trial results where there are space restrictions. Arguably, it is those with practical experience of designing RCTs who are best placed to provide advice about the use of such methods in their experience. Furthermore, there may be other methods or existing methods that are implemented in a way that has not been captured from the systematic review of the literature. To address this, we sought to assess the usage of methods among leading clinical trialists.

This was achieved using two related surveys: one of the membership of the International Society for Clinical Trials (SCT) [4] and one of the UK- and Ireland-based trialists [5–7]. The aim of the surveys was to evaluate current practice among clinical trialists, specifically, which methods respondents were aware of, used, and would be willing to recommend. While the two surveys were similar, the second, to UK- and Ireland-based trialists, was slightly more extensive (see below for details). This work was part of the Difference ELicitation in TriAls (DELTA) project: a study on target differences commissioned by the Medical Research Council, UK (MRC)/National Institute for Health Research (NIHR) Methodology Research Panel and is reported in detail elsewhere [3].

Methodology of the surveys

Survey 1: SCT membership

Members of the SCT were surveyed (sent 24 August 2011) [4]. The survey asked generic questions about the individual responding (position, affiliation, location, and whether they are currently involved in the design of RCTs), and about their group's awareness and usage of methods for determining the target difference, with the opportunity to suggest additional methods provided (see online Appendix 1). A brief summary of the seven previously identified methods was provided in the online form. Additionally, the respondents were asked whether they would be willing to recommend the use of any of the methods. Finally, an opportunity to comment on the issue was provided. Members received an email invitation sent *via* the society's email distribution list, inviting them to complete the online survey. The invitation included a brief introduction to this issue and the aim of the survey. The online survey was setup bespoke for this purpose by the University of Aberdeen's Health Services Research Unit (HSRU) Programming team. Once potential participants received the email, they were able to access the survey by clicking on the URL hyperlink provided. A generic reminder was sent to the entire study sample 1 week after the initial email invitation (it was not possible to tailor reminders to individuals).

Survey 2: UK- and Ireland-based trialists

The sample for the survey of UK- and Ireland-based trialists included three groupings who contribute to trial design: UK Clinical Research Collaboration

Table 1. Methods for specifying an important and/or realistic difference

Methods for specifying an important difference
• *Anchor*. The outcome of interest can be 'anchored' by using either a patient's or health professional's judgement to define an important difference. This may be achieved by comparing a patient's health before and after treatment and then linking this change to patients who showed improvement/deterioration using a more familiar outcome (for which either patients or health professionals more readily agree on what amount of change constitutes an important difference). Alternatively, a contrast between patients can be made to determine a meaningful difference.
• *Distribution*. Approaches that determine a value based upon distributional variation. A common approach is to use a value that is larger than the inherent imprecision in the measurement and therefore likely to represent a minimal level for a noticeable difference.
• *Health economic*. Approaches that use principles of economic evaluation. These typically involve including both resource cost and health outcomes and defining a threshold value for the cost of a unit of health effect that a decision-maker is willing to pay, to estimate the overall net-benefit of treatment. The net-benefit can be analysed in a standard frequentist framework or take the form of a (typically Bayesian) decision-theoretic value of information analysis.
• *Standardised effect size*. The magnitude of the effect on a standardised scale defines the value of the difference. For a continuous outcome, the standardised difference (most commonly expressed as Cohen's d 'effect size') can be used. Cohen's cut-offs of 0.2, 0.5, and 0.8 for small, medium, and large effects, respectively, are often used. Thus, a 'medium' effect corresponds simply to a change in the outcome of 0.5 standard deviations. Binary or survival (time-to-event) outcome metrics (e.g., an odds, risk, or hazard ratio) can be utilised in a similar manner, although no widely recognised cut-offs exist. Cohen's cut-offs approximate to odds ratios of 1.44, 2.48, and 4.27, respectively. Corresponding risk ratio values vary accordingly to the control group event proportion.
Methods for specifying a realistic difference
• *Pilot study*. A pilot (or preliminary) study may be carried out where there is little evidence, or even experience, to guide expectations and determine an appropriate target difference for the trial. In a similar manner, a phase II study could be used to inform a phase III study.
Methods for specifying an important and/or a realistic difference
• *Opinion-seeking*. The target difference can be based on opinions elicited from health professionals, patients, or others. Possible approaches include forming a panel of experts, surveying the membership of a professional or patient body or interviewing individuals. This elicitation process can be explicitly framed within a trial context.
• *Review of evidence base*. The target difference can be derived using current evidence on the research question. Ideally, this would be from a systematic review or meta-analysis of RCTs. In the absence of randomised evidence, evidence from observational studies could be used in a similar manner. An alternative approach is to undertake a review of studies in which an important difference was determined.

RCT: randomised controlled trial.

(UKCRC) registered Clinical Trials Units (CTUs) [6], Regional NIHR Research Design Service (RDS) offices in England [7], and the MRC UK Hubs for Trial Methodology Research [5] (as of 24 August 2011). One individual (typically the Director) from the CTUs, MRC Hubs, and RDS offices was invited to complete the survey (sent 24 August 2013). Where the same individual held a position with more than one entity, only one survey was sent and a response on behalf of the relevant groups requested. They were requested to forward the survey to the appropriate member of their group if they were not personally able to complete it.

In addition to the information collected in the SCT survey, this survey (see online Appendix 2) requested information about the most recent trial developed from the UK- and Ireland-based trialists sample. These details included the underlying basis adopted for the target difference (e.g., realistic difference or important difference) and any methods for determining the target difference used. Additionally, they were asked whether there is anything that would aid them in the design of RCTs and whether they would be happy to be contacted for further details. The initial request was personalised and sent by post and included an invitation letter, paper version of the survey, and description of the methods available for determining the difference. A paper reminder was sent 2 weeks from the initial notification of the survey. Following this, an additional (final) email reminder was sent after another week with an electronic invitation, version of the survey, and description of the methods.

Ethical review

The surveys were approved by the University of Aberdeen's College of Life Sciences and Medicine Ethics Review Board (CERB/2011/6/657). This project abided by the MRC's guidance on Good Research Practice and conformed to the University of Aberdeen's Research Governance Guidelines. We piloted the survey invitations and formats with members of the project team and local researchers. The responses to the online survey and submitted

Table 2. Respondent characteristics (survey 1, N = 180)

Characteristic	% of respondents (n)
Location[a]	
United States	71 (127)
United Kingdom	10 (18)
Canada	8 (15)
Other European country	6 (10)
Japan	2 (3)
Australia	2 (3)
African country	1 (2)
China	1 (1)
Profession	
Health professional	7 (13)
Statisticians/epidemiologists	85 (153)
Other scientist (e.g., ethicist or behavioural scientist)	1 (2)
Trial staff	4 (8)
Other	2 (4)
Institution[a]	
Academic institution	58 (103)
Contract research organisation	13 (23)
Governmental agency	9 (17)
Health-care provider	3 (6)
Private industry	13 (24)
Other	3 (6)
Currently involved in trial design	
Yes	90 (162)

[a]Based upon N = 179 as one respondent did not complete.

survey data are stored within a secure database on a secure server.

Data analysis

The surveys were analysed separately. The response rate was defined as the respective number of responding participants divided by the number of potential participants in the population. The statistical analysis was descriptive. Responses were summarised quantitatively or narratively as appropriate.

Results

Survey 1: SCT membership

Of the 1182 members on the SCT membership (519 of whom described themselves as statistician, epidemiologist, or principal investigator) email distribution list, 180 responses were received (15%). Respondent characteristics are given in Table 2. A total of 13 countries were represented, although over 75% were from North America (127 and 15 respondents from the United States and Canada, respectively). In all, 18 respondents were based in the United Kingdom. The vast majority of respondents were statisticians/epidemiologists, with only 13 respondents being health professionals. The affiliation of the majority was an academic institution with similar numbers from a contract research organisation, private industry, and a regulatory authority. Of the 180 respondents, 162 (90%) stated they were presently involved in trial design.

The responses regarding awareness, usage, and willingness to recommend methods are given in Table 3. Awareness of methods ranged from 69 (38%) for health economic methods to 162 (90%) for pilot study. No additional method was reported. As expected, usage was lower than awareness and ranged from 16 (9%) for the health economic to 133 (74%) for pilot study. The use of treatment selection methods and continual reassessment method and other adaptive models was highlighted, although these typically are based upon an arbitrary, though pre-specified, sample size [8].

The highest level of willingness to recommend was for review of evidence base (73%) and the

Table 3. Responses regarding awareness, usage, and willingness to recommend methods[a] (survey 1, N = 180)

Method	Aware of % (n)	Used % (n)	Recommend % (n)	Recommend if used % (n)
Anchor	43 (77)	33 (59)	30 (54)	71 (42)
Distribution	58 (104)	40 (72)	33 (60)	68 (49)
Health economic	38 (69)	9 (16)	16 (28)	69 (11)
Opinion-seeking	59 (106)	40 (72)	32 (58)	56 (40)
Pilot study	90 (162)	74 (133)	65 (117)	77 (103)
Review of evidence base	87 (156)	73 (132)	73 (132)	89 (118)
Standardised effect size	77 (138)	58 (104)	41 (73)	63 (65)
Other	0 (0)	0 (0)	0 (0)	NA
None	2 (3)	3 (6)	3 (6)	NA

NA: not applicable.
[a]Respondents were invited to select all that apply.

Table 4. Respondent characteristics (survey 2, N = 34)

Characteristics	% of respondents (n)
Representing	
CTU	52% of CTUs (25)
Hub	63% of Hubs (5)
RDS	80% of RDS offices (8)
Position	
Director	76 (26)
Statistician	9 (3)
Other	15 (5)
Intervention types in trial portfolio[a,b]	
Pharmacological	88 (29)
Non-pharmacological	97 (32)
Phase of trial in portfolio[a,b]	
I	3 (1)
II	73 (24)
III	94 (31)
IV	58 (19)
Clinical area in trial portfolio[a,b]	
Blood	15 (5)
Cancer	67 (22)
Cardiovascular	55 (18)
Dementias and neurodegenerative diseases	45 (15)
Diabetes	48 (16)
Ear	12 (4)
Eye	18 (6)
Genetics and congenital disease	18 (6)
Infection	36 (12)
Inflammatory and immune	21 (7)
Injuries and emergencies	30 (10)
Medication for children	42 (14)
Mental health	58 (19)
Metabolic and endocrine	24 (8)
Musculoskeletal	52 (17)
Neurological	33 (11)
Oral and gastrointestinal	42 (14)
Primary care	73 (24)
Renal	39 (13)
Reproductive heath	42 (14)
Respiratory	42 (14)
Skin	30 (10)
Stroke	52 (17)

CTU: Clinical Trials Unit; RDS: Research Design Service.
[a]Respondents were invited to select all that apply.
[b]Based upon N = 33 as one CTU had only recently started.

lowest was health economic (16%). Willingness to recommend was lower than awareness and usage for all methods except for the health economic method where the number willing to recommend this method although less than the number aware of it was more than usage (16% vs. 9%, respectively). Willingness to recommend among only those who had used a particular method was substantially higher than across all respondents ranging from 56% for opinion-seeking method to 89% for review of evidence-base method for specified methods.

Survey 2: UK- and Ireland-based trialists

Information on the groups represented is given in Table 4. Of the 61 surveys sent out, 34 (56%) responses were received representing: 25 (52%), 5 (63%), and 8 (80%) of CTUs, methodology trial Hubs, and RDS offices, respectively (some having more than one affiliation). Respondents were predominately directors of one of these institutions (26, 76%), with the remainder being a statistician (3, 9%) or others (5, 15%). The vast majority stated their group dealt with both pharmacological and non-pharmacological trials (29, 88%) and (32, 97%), respectively. With regard to RCT phases, the group's trial portfolio contained 24 (73%), 31 (94%), and 19 (58%) for phases II–IV, respectively. One (3%) reported also undertaking a phase I study. All clinical areas under the NIHR UK portfolio categorisation were represented with frequencies ranging from 4 (12%) for ear to 24 (73%) for primary care–related research.

The responses regarding awareness, usage, and willingness to recommend methods are given in Table 5. Awareness of methods ranged from 33 (97%) for the opinion-seeking and pilot study methods, to only 14 (41%) for distribution method. No other methods were suggested, and all stated that they had used at least one of the methods. The use of methods was substantially less than awareness except for the pilot study, review of evidence base, and standardised effect size methods. All respondents were aware of at least one of the different formal methods for determining the target difference. Almost all had used at least two methods (94%). One respondent stated their group had not used any of the methods as they had only recently formed. The highest level of willingness to recommend was for review of evidence base (76%) followed by standardised effect size (65%). The vast majority (88%) recommended more than one method. Two respondents (6%) stated that they would not recommend any method. The use of the anchor and review of evidence base in combination was suggested by one respondent. The difficulty in differentiating between some of the methods without full definitions was noted by another respondent.

Data on the most recent trial the respondent's group had been involved with is given in Tables 5 and 6. Based upon the most recent trial, all bar three groups used a formal method. In all, 21 (62%) stated using more than one formal method. Two respondents reported that their group had used alternative

Table 5. Responses regarding awareness, use, and willingness to recommend methods[a] (survey 2, N = 34)

Method	Aware of % (n)	Used % (n)	Recommend % (n)	Recommend if used % (n)	Most recent trial[b] % (n)
Anchor	65 (22)	44 (15)	47 (16)	87 (13)	18 (6)
Distribution	41 (14)	24 (8)	26 (9)	38 (3)	3 (1)
Health economic	62 (21)	24 (8)	32 (11)	63 (5)	3 (1)
Opinion-seeking	88 (30)	53 (18)	53 (18)	78 (14)	27 (9)
Pilot study	97 (33)	88 (30)	59 (20)	67 (20)	24 (8)
Review of evidence base	97 (33)	94 (32)	76 (26)	78 (25)	52 (17)
Standardised effect size	88 (30)	82 (28)	65 (22)	79 (22)	42 (14)
Other	0 (0)	0 (0)	0 (0)	NA	0 (0)
None	0 (0)	3 (1)	6 (2)	NA	9 (3)

NA: not applicable; CTU: Clinical Trials Unit.
[a]Respondents were invited to select all that apply.
[b]Based upon N = 33 as one CTU had only recently started.

Table 6. Most recent trial (survey 2, N = 33[a])

Most recent trial	% of respondents (n)
Primary outcome[b]	
Generic quality of life (e.g., EQ-5D)	12 (4)
Disease-specific quality of life (e.g., Oxford Knee Score)	21 (7)
Other patient reported outcome (non-quality-of-life measure)	6 (2)
Mortality	27 (9)
Clinical functional measure (e.g., forced expiratory volume (FEV))	33 (11)
Economic outcome (e.g., incremental cost per QALY)	6 (2)
Other	24 (8)
There was no primary outcome	3 (1)
What was the underlying principle(s) adopted in determining the difference?[b]	
A realistic difference given the interventions under evaluation	61 (20)
A difference which would led to an achievable sample size	33 (11)
A difference that would be viewed as important by a relevant stakeholder group (e.g., clinicians)	91 (30)
Other	6 (2)

CTU: Clinical Trials Unit; QALY: quality-adjusted life year.
[a]Based upon N = 33 as one CTU had only recently started.
[b]Respondents were invited to select all that apply.

informal methods: reverse engineering the study sample size to ensure the research cost fell 'within funding' range and basing it upon the lead clinical applicant's opinion although this was not formally elicited, respectively. The most common type of primary outcome in the most recent trial conducted by the respondent's group was a clinical function measure (33%) followed by mortality outcome (27%). Disease-specific (21%) and generic (12%) quality-of-life measures were also represented in multiple studies. A non-quality of life patient reported outcome and a health economic measure were each reported twice as being used as the primary outcome. Other outcome types were non-mortality time-to-event (6%), cardiovascular events (6%), weight-related outcomes (6%), length of stay, and violent events. In one case, there was no primary outcome and 11 (34%) had more than one primary outcome. There was no clear link between outcome type and methods used with a spread of methods used for the more common outcome types (quality of life, mortality, and clinical outcome measures).

The vast majority stated that the target difference was one that was viewed as important by a stakeholder group (91%). Just over half the respondents (61%) stated the basis for determining the target difference was to achieve a realistic difference given the interventions under evaluation. In all, 11 respondents (33%) stated that it was a difference that gave an achievable sample size. One considered what difference would be worthwhile detecting given the cost of the intervention and the other

considered what magnitude of a target difference (and hence size and cost of project) would likely to be funded. In all, 10 (30%), 7 (21%), and 2 respondents (6%), respectively, reported using two, three, and four bases of consideration when determining the target difference. Of the 19 respondents with two or more bases, 16 stated they sought both a realistic and an important differences. Two respondents used an addition basis for determining the target difference in combination with all three prespecified approaches (a difference that was realistic, important to stakeholder, and achievable): the cost of the intervention and the 'likelihood of securing funding'.

Discussion

The two surveys of current practice provided insight into the current practice among clinical trialists regarding specification of the target difference. To our knowledge, this has not been done before. Responses suggest that use of formal methods is greater than would appear the case from trial reports [1,9,10] or at least, the use of methods is higher for the type of RCTs that the second survey represents. The two surveys represented different stakeholders: an international society of people involved in clinical trials and leading UK- and Ireland-based trialists. Variations in awareness, use, and willingness to recommend between methods were substantial between respondents to the two surveys. There was some difference in the responses between the two groups which might be expected given the more heterogeneous composition of the SCT sample. Nevertheless, the findings support the view that sample size calculation is a more complex process than would appear to be the case from trial reports and protocols.

In both surveys, the awareness of formal methods was high for most methods although for some, notably the anchor, distribution, and health economic methods, it was substantially lower than for others. Awareness of the opinion-seeking method was lower in the SCT sample than among the UK- and Ireland-based trialists. This may reflect a greater focus among the SCT membership upon pharmacological trials conducted for regulatory approval where phase II trial results typically informs the phase III trial and the phase II sample size is influenced by convention/regulatory body expectations. The pattern of use of methods was broadly the same as awareness.

With regard to recommendations for use by researchers, the review of the evidence-base method consistently had the highest level of recommendation across the two surveys. In both surveys, the use of an informal approach such as 'reverse engineering' to suit expected recruitment and associated research cost was mentioned. Slightly more respondents were willing to recommend the health economic method than had actually used it (although in both cases, the number recommending was low compared with most of the other methods), perhaps reflecting the intuitive appeal of this approach and that cost considerations influence decisions about study design even when not explicitly stated. The lack of use of the health economic approaches may relate to its more recent development, lack of expertise, and resources necessary to conduct this approach.

Exploration of the basis for calculating the target difference was further explored in more depth in the second survey by referring to the last trial their group had been involved in. Details about the most recent trial conducted by respondents reflected a wide variety of outcome types, for which all of the methods were used to varying degrees. While the vast majority stated that the basis for the target difference was a difference viewed as important by a stakeholder group, most also used another basis and two used four separate bases (difference viewed as important to stakeholder, realistic, achievable, and either the cost of the intervention or consideration of funding) for determining the target difference. Furthermore, the majority of respondents stated they had used more than one method in determining the sample size calculation. Such complexity of considerations, in our experience, is rarely reported in the sample size section either in the trial report or the protocol. This view is supported by a review of sample size calculations in RCTs [1]. Clearer and more explicit reporting of the basis for determining the target difference, including any formal methods used, is needed.

Funders of RCTs naturally consider the associated research cost of the study and the feasibility of the proposal, and this was recognised by respondents to both surveys. The desirable precision may be prohibitively high or impractical to achieve. For respondents, the dilemma as to how best to balance the cost of a trial against the need for a large (and by extension an expensive) trial may be needed to answer the research question. This tension led more than one respondent to query how often the target difference is honestly chosen as opposed to being picked to make the sample size 'affordable'. Similarly, the need to choose a sample size that would allow a realistic recruitment rate was also raised. How cost and feasibility should be taken into account may be unclear to applicants and lead to a reluctance to be explicit about the practical considerations. In principle, these are the sort of considerations that the economic value of information methods appears capable of capturing while ensuring that the cost of data collection is less than the value generated by the RCT. Furthermore, such an approach could be tailored to a funder's perspective

and considerations. Nevertheless, while there is some awareness of these techniques, there is little evidence of their widespread use in practice.

The need for guidance *via* a central resource to clarify what is clinically important for common outcomes (one specifically regarding the patient perspective) and guidance on what methods are appropriate and in what situation is clear from the survey results. Such guidance, as some survey respondents themselves note, would need to recognise trial specifics (e.g., the phase of the trial) and dependent upon the type of outcome, the resources (time, expertise, etc.) available to implement methods, current knowledge of the clinical area, and also the proposed trial analysis (e.g., Bayesian [11] or update of a meta-analysis) [12].

The response rates achieved were relatively low despite the surveys being short and well presented and having utilised reminders. However, we view it as unlikely that non-response has led to unrepresentative findings and suspect the low response rate reflects the difficulty of achieving high rates in certain population groups [13] along with the lower response rate from email distribution lists, as was used in the SCT survey. A further factor may be the nature of the survey; it focused on RCT methodology, which made determining who should be invited to respond, and perhaps also completion, less than straightforward and required consideration of a portfolio of research projects. One respondent noted the difficulty in differentiating between some of the methods despite a brief description being provided. There may also have been some overlap in respondents between the two surveys. Nevertheless, the surveys provided insightful information about the practice and views of trialists from a range of backgrounds regarding determining the target difference.

The second survey included MRC Hubs for Trial Methodology and the 10 regional centres of the NIHR RDS in addition to CTUs. This reflects the varied way in which the design of RCTs is dealt with in the United Kingdom and Ireland. For example, in Scotland, CTUs typically design trials 'in-house' or have a very close affinity with a research group within their host institution. However, in England, the NIHR RDS may take on this role and advise on elements of trial design. Furthermore, across the United Kingdom and Ireland, the establishment of the MRC Hubs has altered the clinical trial landscape and are an additional grouping of trial expertise.

Conclusion

The awareness, use, and views of trialists regarding methods for specifying the target difference varied both in the United Kingdom and Ireland, and also internationally. Variations in practice exist, and a key requirement highlighted was the need for guidance documentation to inform the process. There would also appear to be a need for more transparent reporting of the considerations taken into account when determining the target difference.

Acknowledgements

We would like to thank the respondents of the surveys for their submissions. We would also like to thank Mark Forrest and Lara Kemp for help with producing the surveys, and Marion Campbell, Ian Ford, Dean Fergusson, and Adrian Grant for serving on the project advisory group which provided guidance on the DELTA project's conduct and interpretation of findings. Views expressed are those of the authors alone. Additionally, we would like to thank the journal's editors and reviewers for constructive comments which have improved the manuscript.

DELTA Group: Jonathan A. Cook, Jennifer M. Hislop, Doug G. Altman, Andrew H. Briggs, Peter M. Fayers, John D. Norrie, Craig R. Ramsay, Ian M. Harvey, Ian Ford, Brian Buckley, Kirsten Harrild, Temitope E. Adewuyi, Cynthia Fraser, Dean Fergusson, and Luke D. Vale.

Funding

This research was funded by the Medical Research Council (MRC)/National Institute for Health Research (NIHR) Methodology Research Panel which is jointly funded by the MRC (reference no. G0902147) and Health Technology Assessment (HTA) (project number 06/98/01). Jonathan Cook held MRC Medical Research Council UK training (reference no. G0601938) and methodology (reference no. G1002292) fellowships while this research was undertaken. The Health Services Research Unit, Institute of Applied Health Sciences (University of Aberdeen), is core-funded by the Chief Scientist Office of the Scottish Government Health and Social Care Directorates.

Conflict of interest

None declared.

References

1. **Charles P, Giraudeau B, Dechartres A**, *et al*. Reporting of sample size calculation in randomised controlled trials: Review. *BMJ* 2009; **338**: b1732.
2. **Julious SA**. *Sample sizes for clinical trials*. CRC Press, Boca Raton, FL, 2010.
3. **Cook JA, Hislop J, Adewuyi TE**, *et al*. Assessing methods to specify the targeted difference for a randomised

controlled trial – DELTA (Difference ELicitation in TriAls) review. *Health Technol Asses*, in press.
4. **Society for Clinical Trials, Philadelphia, PA**. Available at: http://www.sctweb.org/ (2011, accessed January 2014).
5. **MRC Network of Hubs for Trials Methodology Research**. Medical Research Council, London. Available at: http://www.methodologyhubs.mrc.ac.uk/ (2010, accessed January 2014).
6. **UKCRC Registered Clinical Trials Units**. NIHR Clinical Research Network Coordinating Centre, Leeds. Available at: http://www.ukcrc-ctu.org.uk/ (2010, accessed January 2014).
7. **NIHR Research Design Service**. National Institute for Health Research, London. Available at: http://www.nihr.ac.uk/infrastructure/Pages/infrastructure_research_design_services.aspx (2011, accessed January 2014).
8. **Friedman LM, Furberg CD, DeMets DL**. *Fundamentals of Clinical Trials*. Springer, New York, 2010.
9. **Gayet-Ageron A, Agoritsas T, Combescure C**, *et al*. What differences are detected by superiority trials or ruled out by noninferiority trials? A cross-sectional study on a random sample of two-hundred two-arms parallel group randomized clinical trials. *BMC Med Res Methodol* 2010; **10**: 93.
10. **Goudie AC, Sutton AJ, Jones D**, *et al*. Empirical assessment suggests that existing evidence could be used more fully in designing randomized controlled trials. *J Clin Epidemiol* 2010; **63**: 983–91.
11. **Spiegelhalter DJ, Abrams KR, Myles JP**. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. John Wiley & Sons, Chichester, 2003.
12. **Sutton AJ, Cooper NJ, Jones DR**. Evidence synthesis as the key to more coherent and efficient research. *BMC Med Res Methodol* 2009; **9**: 29.
13. **Cook JV, Dickinson HO, Eccles MP**. Response rates in postal surveys of healthcare professionals between 1996 and 2005: An observational study. *BMC Health Serv Res* 2009; **9**: 160.