TRIALS

Specifying the target difference in the primary outcome for a randomised controlled trial: guidance for researchers

Jonathan A Cook[1,2*], Jenni Hislop[3], Douglas G Altman[1], Peter Fayers[4,5], Andrew H Briggs[6], Craig R Ramsay[2], John D Norrie[7], Ian M Harvey[8], Brian Buckley[9], Dean Fergusson[10], Ian Ford[11], Luke D Vale[3] and for the DELTA group

Abstract

Background: Central to the design of a randomised controlled trial is the calculation of the number of participants needed. This is typically achieved by specifying a target difference and calculating the corresponding sample size, which provides reassurance that the trial will have the required statistical power (at the planned statistical significance level) to identify whether a difference of a particular magnitude exists. Beyond pure statistical or scientific concerns, it is ethically imperative that an appropriate number of participants should be recruited. Despite the critical role of the target difference for the primary outcome in the design of randomised controlled trials, its determination has received surprisingly little attention. This article provides guidance on the specification of the target difference for the primary outcome in a sample size calculation for a two parallel group randomised controlled trial with a superiority question.

Methods: This work was part of the DELTA (Difference ELicitation in TriAls) project. Draft guidance was developed by the project steering and advisory groups utilising the results of the systematic review and surveys.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, containing recommendations and reporting items for a trial protocol and report. The guidance and was subsequently drafted and circulated for further comment before finalisation.

Results: Guidance on specification of a target difference in the primary outcome for a two group parallel randomised controlled trial was produced. Additionally, a list of reporting items for protocols and trial reports was generated.

Conclusions: Specification of the target difference for the primary outcome is a key component of a randomized controlled trial sample size calculation. There is a need for better justification of the target difference and reporting of its specification.

Keywords: Target difference, clinically important difference, sample size, randomised controlled trial, guidance

* Correspondence: jonathan.cook@ndorms.ox.ac.uk
[1]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Botnar Research Centre, Nuffield Orthopaedic Centre, Windmill Road, Oxford OX3 7LD, UK
[2]Health Services Research Unit, University of Aberdeen, Health Sciences Building, Foresthill, Aberdeen AB25 2ZD, UK
Full list of author information is available at the end of the article

Background

Well-conducted randomised controlled trials (RCTs) are widely viewed as providing the optimal evidence on the relative performance of competing healthcare interventions [1,2]. However, simply detecting any statistical difference in the effectiveness of interventions may not be sufficient or useful; if the interventions differ to a degree or in a manner that is of little consequence in patient, clinical or economic (or other meaningful) terms, then the interventions might be considered not to be different. If RCTs are to produce useful information that can help patients, clinicians and planners make decisions about health care, it is essential that they are designed to achieve this. This is typically achieved by specifying a target difference for a primary outcome as part of a sample size calculation, which provides reassurance that the trial will have the specified statistical power to identify whether a difference of a particular magnitude exists. Beyond purely statistical or scientific concerns, the sample size calculation has financial and ethical implications. Failing to recruit sufficient participants to be able to confidently detect a relevant difference between interventions may be viewed as an inefficient use of finite research resources, while recruiting substantially more than are needed risks exposing participants to unnecessary experimentation [3].

Given these considerations, determining an appropriate sample size is of critical importance. Surprisingly, little practical advice is available on specifying the target difference of the chosen primary outcome, which as noted above is a key component of the sample size calculation. A comprehensive systematic review of the literature identified methods for determining the target difference that are available and surveys have shown these methods are in use [4,5]. Nevertheless, uncertainty regarding the magnitude of the target difference when designing the trial will lead to uncertainty regarding the interpretation of the results, even when the trial is otherwise successfully conducted [6,7].

This article aims to provide practical guidance primarily for researchers involved in determining the sample size for an RCT and, in particular, the specification of the target difference in the primary outcome. It is also relevant to those who are involved in commissioning and publishing such studies. We provide guidance on the choice of the primary outcome, specification of the target difference and a brief summary of available methods that can be used to inform its specification and reporting. Additionally, two sets of reporting items, one for a trial protocol and the other a report of the trial findings in a peer reviewed biomedical journal, are also proposed and examples provided. A comprehensive systematic review and discussion of the individual methods for specifying a target difference has been reported elsewhere [4,5]. The focus of this guidance is upon what might be termed the conventional, or standard, approach to an RCT sample size calculation: a standalone trial utilising the conventional statistical framework for sample size calculation and primarily for superiority trials (those where the difference to be detected is specified). The key issues considered are relevant to other RCT designs and analysis approaches though implementation may differ. We note that the conventional approach to sample size calculation is not without its limitations and alternatives have been proposed [8], nevertheless it continues to be the most widely adopted approach [1,9].

The conventional approach to the sample size calculation for a two parallel group RCT is as follows:

1. The RCT is conceived as a standalone definitive study (a study that is designed to provide a meaningful answer on its own);
2. It addresses a superiority question evaluating evidence of a difference (in either direction);
3. Adoption of a two parallel-group RCT design (typically 1:1 allocation);
4. Application of the Neyman-Pearson framework to calculate the sample size [2,10-12]. This requires specification of: the primary outcome for which the required sample size is to be calculated; the target difference (specification varies according to outcome type); statistical parameters (significance level and power) and other component(s) of the sample size calculation (such as standard deviation (SD)).

Methods

Development of the guidance

This work was part of the DELTA (Difference ELicitation in TriAls) project, a study on target differences commissioned by the Medical Research Council/National Institute for Health Research Methodology Research Panel (MRC/NIHR) in the United Kingdom. It comprised three interlinking components: a comprehensive systematic review of methods for specifying the target difference, two surveys of current practice amongst clinical trialists and generation of structured guidance. This article is an abridged version of this guidance and other components of the project which have been reported in full elsewhere [4]. DELTA was undertaken by a collaborative group in which the majority of members have extensive experience of the design and conduct of RCTs (both as investigators and as independent committee members) and have conducted methodological research related to RCTs (such as quality-of-life measurement, statistical methodology, reporting, surgical trials and economic evaluation). 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 a list of recommendations and reporting items for a trial protocol and report. Both the structure and main recommendations were agreed at this meeting. The guidance was subsequently drafted and circulated for further comment before finalisation. No ethical approval was needed for this research.

Scope of the guidance

This guidance is based upon the conventional approach to a sample size calculation, though it should be applicable to most RCTs [1,9]. However, other approaches, for example trials with an explicitly Bayesian analysis framework, will require adaptation of the reporting items. It focuses upon guidance for a trial with a 'superiority' question; one which seeks evidence of a difference between intervention groups. Although this guidance is primarily aimed at researchers, it is also relevant for publishers, funders and commissioners of research.

Results

Abridged guidance is given below.

Choosing the primary outcome

In the conventional approach to the sample size calculation for an RCT, a single outcome is usually chosen to be the primary measure upon which the sample size calculation is based (in some cases more than one primary outcome may be appropriate) [2,10,13]. The specification of a primary outcome performs a number of functions in terms of trial design, but it is clearly a pragmatic simplification to aid the design, interpretation and use of RCT findings. Through the corresponding sample size calculation and specification of the target difference, it clarifies what the study aims to identify, and the statistical power and precision with which this can be achieved. Stating the primary outcome in the study protocol also helps prevents undue over-interpretation arising from testing multiple outcomes and selective outcome reporting bias, whereby authors report only statistically significant (on possibly clinically irrelevant) outcomes or change the primary focus of the study to match a statistically significance finding. Additionally, it helps clarify the initial basis upon which to judge the study findings. This is particularly important in presence of a 'negative' result, where the result does not meet the criteria for statistical significance (typically 5%). In all cases, focus should be upon the confidence interval as well as the point estimate, where a justifiable target difference can guide the interpretation. However, such justification of the target difference is often lacking in trial reports [1,6]. Calculating (or reverse engineering) the magnitude of a difference that can be detected at conventional levels of statistical significance and power (typically two-sided 5% and 80%, respectively), given a sample size which is believed to feasible, is often performed in

practice for a selection of key outcomes before determining the primary outcome. Nevertheless, it is important to report the final sample size calculation, including the chosen primary outcome, the target difference and any justification of the value chosen, in as robust and transparent a fashion as possible to allow others to judge the basis of the calculation.

Specifying the target difference

The specification of the target difference in an RCT sample size calculation has received surprisingly little discussion in the literature. For a superiority trial, it is the difference in the primary outcome value that the study is designed to detect reliably [2,10,13]. There are two main bases for specifying the target difference: a difference considered to be 'important' (for example, by a stakeholder group such as health professionals or patients), and a 'realistic difference' based upon current evidence (for example, seeking the best available estimates in the literature through some form of knowledge synthesis).

It has been argued that a target difference should always meet both of these criteria [14]. The desire to be able to consider an (clinically) important difference can be viewed as a middle ground between ignoring the consequences of the treatment decision and a full assessment of the benefits, harms and costs of an intervention against the alternatives, which seeks to ensure that any harms and costs are incurred for a good reason. Focusing on a benefit (or harm) of the most important outcome is a natural and intuitive, if imperfect, way to guide a decision. A large body of literature exists on defining a clinically important difference, though not in the context of an RCT sample size calculation [15-17]. The most common general approach is the minimal clinically important difference (MCID). This has been defined as 'the smallest difference …. which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management', or more simply as 'minimum difference that is important to a patient' [17]. Many variants on this basic approach exist [18,19]. In the context of specifying a target difference for a typical two parallel-group trial, the focus is on a difference at the group level, between two groups of different participants. This contrasts with the vast majority of the MCID (and related) literature, which focuses overwhelmingly upon within-patient change and whether an important difference can be said to have occurred [15-17]. An alternative approach is to consider all relevant issues, including the consequences of decision-making, whereby a difference of any magnitude can be viewed as important and therefore a study's size (and implicitly the target difference) is determined by reference to resource implications [20,21]. Whatever definition is used, estimation of an

important difference is not without its challenges and limitations [22,23].

The other main basis for a target difference is to specify a realistic difference; there is, for example, little point in setting as the target difference one that is so large that it cannot plausibly exist. If a systematic review of RCTs on the research question is available, it can be used to specify what difference is supported by current evidence. In essence, a realistic difference makes no claim regarding its clinical importance or otherwise. However, where a realistic difference is used, consideration of the importance of the difference is needed if the study findings are intended to inform clinical, patient or policy decisions. For some outcomes, the importance may be very clear (for example, mortality), whereas for others (especially quality of life and surrogate outcomes) further explanation is needed. Recruitment, study management and finance will naturally come into play when determining the sample size of a study. However, such considerations do not negate concerns about what is a realistic and/or important difference.

For a superiority trial it is generally accepted that the target difference should be a clinically important difference [2,10-12] or 'at least as large as the MCID [minimum clinically important difference]' [24]. The target difference in a conventional sample size calculation is not the minimum difference that can be statistically detected; statistical significance alone is not a sufficient consideration for attributing importance to a difference [2,12].

The target difference is specified differently depending upon the type of primary outcome. For a continuous outcome, this target difference on either the original or standardised scale is often referred to as the 'effect size'. Strictly speaking, this value alone does not fully (uniquely) specify the target difference; the assumed variability of the outcome (standard deviation) is also needed to convert the effect size between the original and standardised scales. For a binary outcome, the target difference will be conditional on the control group event proportion. To uniquely specify the sample size, the target difference and the control group event proportion are needed, which together imply a unique pair of absolute and relative target

Table 1 Methods for specifying an important and/or realistic difference [5]

Name	Description	Difference specified
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 participants who showed improvement and/or 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). An outcome can be anchored to another which more is known about. Contrasts between patients (such as individuals with varying severity of a disease) can also be used to determine a meaningful difference.	Important
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.	Important
Health economic	Approaches that use principles of economic evaluation. These typically include both resource cost and health outcomes, and define 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 frequentist framework or take the form of a (typically Bayesian) decision-theoretic value of information analysis. Due to difficulties in implementing a value of information analysis simpler heuristic frameworks, also based on the principles of economic evaluation, have been proposed.	Important
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.	Important and/or realistic
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 two study could be used to inform a phase three study, though this would need to take account of methodological differences (such as inclusion criteria and outcomes), which should be reflected in the target difference.	Realistic
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.	Important and/or realistic
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 cutoff values 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 SDs. Binary or survival (time-to-event) outcome metrics (such as an odds, risk or hazard ratio) can be utilised in a similar manner, though no widely recognised cutoff values exist. Cohen's cutoff values approximate odds ratios of 1.44, 2.48 and 4.27, respectively. Corresponding risk ratio values vary according to the control group event proportion.	Important

Table 2 Reporting items for the protocol and report of a two parallel group superiority trial

Item no.	Item description	Protocol	Report
1	State any divergence from the conventional approach.	√	√
2	State the primary outcome (and any other outcome which the study sample size calculation is based upon), or state why there is not one.	√	√
3	Reference the formula and/or simulation approach if the standard binary, continuous or survival outcome formulas are not used [2,11]. The primary analysis should be stated in the statistical analysis section.	√	
4	State the values used for statistical parameters (such as significance level and power).	√	√
5	State the underlying basis for determining the target difference:	√	
	a. an important difference as judged by a stakeholder,		
	b. a realistic difference based upon current knowledge or		
	c. both an important and realistic difference.		
6	Express the target difference according to the outcome type:	√	√
	a. Binary: state the target difference as an absolute and/or relative effect, along with the intervention and control group proportions. If both an absolute and a relative difference are provided, clarify if either takes primacy in terms of the sample size calculation.		
	b. Continuous: state the target mean difference on the natural scale, the common SD and standardised effect size (mean difference/SD). It is preferable to also provide the anticipated control group mean even though it is not required for the sample size calculation.		
	c. Time-to-event (survival): state the target difference as an absolute and/or relative difference, provide the control group event proportion, along with the intervention and control group survival distributions. Additionally, the planned length of follow-up should be stated along with the assumed accrual pattern. If both an absolute and a relative difference are provided, clarify if either takes primacy in terms of the sample size calculation.		
7	Explain the choice of target difference: specify and reference any formal method used or relevant previous research.	√	
8	State the sample size based upon the assumptions specified above (for a time-to-event outcome, the number of events required should also be stated). If any factors are incorporated which alter the required sample size (such as allowance for loss-to-follow-up) they should also be specified along with the final sample size.	√	√
9	Reference the trial protocol		√

differences. Similarly, survival outcomes require the control group proportion or survival distribution and length of follow-up period to be stated, in addition to the target difference. This is necessary as the sample size required is sensitive to both the absolute level and the relative difference. Despite this, it is not uncommon for only one or the other to be specifically stated in trial reports.

Seven methods for specifying the target difference have been identified [4] which can be used to inform the choice of target difference: anchor, distribution, health economic, opinion-seeking, pilot study, review of the evidence base and standardised effect size (see Table 1 for a brief summary and elsewhere for a summary of the literature assessment of the use of each method [5]).

Table 3 Reworked example RCT protocol sample size calculation sections

Primary outcome type	Example text
Binary	Men After Prostate Surgery (MAPS) radical prostatectomy trial [28]: The primary outcome is continence. The sample size was based upon a target difference of 15% absolute difference (85 versus 70%). This magnitude of a target difference was determined as both a realistic and important difference from discussion between clinicians and the project management group, and inspection of the proportion of urinary incontinence in the trials included in the Cochrane systematic review [29]. The control group proportion is also based upon the observed proportion in the RCTs in this review. Setting the statistical significance to the two-sided 5% level and seeking 90% power, 174 participants per group are required; 348 in total.
Continuous	Full-thickness macular hole and Internal Limiting Membrane peeling Study (FILMS) [7]: The primary outcome is ETDRS (Early Treatment Diabetic Retinopathy Study) distance visual acuity. A target difference of a mean difference of five letters, with a common standard deviation (SD) of 12, was assumed as five letters is equivalent to one line on a visual acuity chart and is viewed as an important difference by patients and clinicians. The SD value was based upon two previous studies; one RCT and one observational comparative study [30,31]. This target difference is equivalent to a standardised effect size of 0.42. Setting the statistical significance to the two-sided 5% level and seeking 90% power, 123 participants per group are required; 246 in total.
Time to event	Arterial Revascularisation Trial (ART) [32]: The primary outcome is all-cause mortality. The sample size was based upon a target difference of 5% in 10-year mortality with a control group mortality of 25%. Both the difference and control group mortality proportions are realistic based upon a systematic review of observational (cohort) studies [33]. Setting the statistical significance to the two-sided 5% level and seeking 90% power, 1464 participants per group are required; 2,928 participants (651 events) in total. Participants will be followed for 10 years.

Reporting the sample size calculation and target difference

The assumptions made in the sample size calculation should be clearly specified. All inputs should be clearly stated so that the calculation can be replicated. It is recommended that trial protocols clearly and fully state the sample size calculations, including where the approach taken differs from the conventional approach (for example, the adoption of a Bayesian framework instead of a frequentist approach), statistical parameters and the target difference, with justification for the choice of values. Due to space restrictions in many publications the main trial paper is likely to contain less detail. A minimum set of items for the main trial results paper along with full specification in the trial protocol is recommended below in Table 2. These are more extensive lists of reporting items building upon the Consolidated Standards for Reporting Trials (CONSORT) including the 2010 version) and Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statements, which provide guidance on reporting the sample size calculation, but not explicitly how to report the target difference and its justification [25-27] Examples for the three most common outcome types are provided in Table 3.

Discussion

The RCT is widely considered to be the best method for comparing the effectiveness of health interventions [1]. Determining the target difference is a key element of an RCT design. Improved standards in both RCT sample size calculations and reporting of these calculations would aid health professionals, patients, researchers and funders in judging the strength of the available evidence and would ensure better use of scarce resources. While no single method provides a perfect solution to a difficult question, we have provided practical guidance for researchers on sample size calculation with reference to specifying the target difference and how this should be reported in trial protocols and reports. To our knowledge, no alternative guidance exists. Although our examples and framing are from a medical context, the issues are relevant to social care, animal and other non-medical research as well. Further research into the implementation, practicality and consequence of using alternative methods for specifying the target difference (such as health economic and opinion-seeking), and exploration of the justification of some methods (such as the standardised effect size method, where the magnitude of the effect is used to infer the important of a difference) is needed.

Conclusions

Specification of the target difference for the primary outcome is a key component of an RCT sample size calculation.

There is a need for better justification of the target difference and for corresponding reporting of its specification. Raising the standard of RCT sample size calculations would aid health professionals, patients, researchers and funders in judging the strength of the evidence and would ensure better use of scarce resources.

Author details

[1]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Botnar Research Centre, Nuffield Orthopaedic Centre, Windmill Road, Oxford OX3 7LD, UK. [2]Health Services Research Unit, University of Aberdeen, Health Sciences Building, Foresthill, Aberdeen AB25 2ZD, UK. [3]Institute of Health and Society, Newcastle University, The Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne NE2 4AX, UK. [4]Population Health, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen AB25 2ZD, UK. [5]Department of Cancer Research and Molecular, Norwegian University of Science and Technology, Mailbox 8905, Trondheim N-7491, Norway. [6]Health Economics and Health Technology Assessment, University of Glasgow, 1 Lilybank Gardens, Glasgow G12 8RZ, UK. [7]Centre for Healthcare Randomised Trials (CHaRT), University of Aberdeen, Health Sciences Building, Aberdeen AB25 2ZD, UK. [8]Faculty of Medicine and Health Sciences, University of East Anglia, Elizabeth Fry Building, Norwich Research Park, Norwich NR4 7TJ, UK. [9]National University of Ireland, University Road, Galway, Ireland. [10]Clinical Epidemiology Program, Ottawa Hospital Research Institute, 725 Parkdale Avenue, Ottawa, ON K1Y 4E9, Canada. [11]Robertson Centre for Biostatistics, University of Glasgow, Boyd Orr Building, University Avenue, Glasgow G12 8QQ, UK.

References

1. Charles P, Giraudeau B, Dechartres A, Baron G, Ravaud P. Reporting of sample size calculation in randomised controlled trials: review. BMJ. 2009;338:b1732.
2. Julious S. Sample Sizes for Clinical Trials. Boca Raton, FL: Chapman and Hall/CRC Press; 2010.
3. McDonald A, Knight RC, Campbell MK, Entwistle VA, Grant AM, Cook JA, et al. What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies. Trials. 2006;7:7.
4. Cook JA, Hislop J, Adewuyi TE, Harrild K, Altman DG, Ramsay CR, et al. Assessing methods to specify the targeted difference for a randomised controlled trial – DELTA (Difference ELicitation in TriAls) review. Health Technol Assess. 2014;18:28.
5. Hislop J, Adewuyi T, Vale LD, Harrild K, Fraser C, Gurung T, et al. Methods for specifying the target difference in a randomised controlled trial: the Difference ELicitation in TriAls (DELTA) systematic review. PLoS Med. 2014;11:e1001645.
6. Hellum C, Johnsen LG, Storheim K, Nygaard OP, Brox JI, Rossvoll I, et al. Surgery with disc prosthesis versus rehabilitation in patients with low back pain and degenerative disc: two year follow-up of randomised study. BMJ. 2011;342:d2786.
7. Lois N, Burr J, Norrie J, Vale L, Cook J, McDonald A, et al. Internal limiting membrane peeling versus no peeling for idiopathic full-thickness macular hole: a pragmatic randomized controlled trial. Invest Ophthalmol Vis Sci. 2011;52:1586–92.
8. Bacchetti P. Current sample size conventions: flaws, harms, and alternatives. BMC Med. 2010;8:17.
9. Clark T, Berger U, Mansmann U. Sample size determinations in original research protocols for randomised clinical trials submitted to UK research ethics committees: review. BMJ. 2013;346:f1135.
10. Friedman LM, Furberg CD, DeMets DL. Fundamentals of Clinical Trials. New York: Springer; 2010.
11. Matthews JN. Introduction to Randomized Controlled Clinical Trials. London: Taylor & Francis; 2006.
12. Peace KE, Chen DG. Clinical Trial Methodology. London: Chapman & Hall; 2010.
13. Pocock SJ. Clinical Trials: A Practical Approach. Chichester: Wiley & Co; 1983.
14. Fayers PM, Cuschieri A, Fielding J, Craven J, Uscinska B, Freedman L. Sample size calculation for clinical trials: the impact of clinician beliefs. Br J Cancer. 2000;82:213–9.
15. Copay AG, Subach BR, Glassman SD, Polly J, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J. 2007;7:541–6.
16. Wells G, Beaton D, Shea B, Boers M, Simon L, Strand V, et al. Minimal clinically important differences: Review of methods. J Rheumatol. 2001;28:406–12.
17. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MICD): A literature review and directions for future research. Curr Opin Rheumatol. 2002;14:109–14.
18. Hays RD, Woolley JM. The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? Pharmacoeconomics. 2000;18:419–23.
19. Barrett B, Brown D, Mundt M, Brown R. Sufficiently important difference: expanding the framework of clinical significance. Med Decis Making. 2005;25:250–61.
20. Willan AR, Eckermann S. Optimal clinical trial design using value of information methods with imperfect implementation. Health Econ. 2010;19:549–61.
21. Kikuchi T, Pezeshk H, Gittins J. A Bayesian cost-benefit approach to the determination of sample size in clinical trials. Stat Med. 2008;27:68–82.
22. Blanton H, Jaccard J. Arbitrary metrics in psychology. Am Psychol. 2006;61:27–41.
23. Carragee EJ. The rise and fall of the "minimum clinically important difference". Spine J. 2010;10:283–4.
24. Van TM, Malmivaara A, Hayden J, Koes B. Statistical significance versus clinical importance: trials on exercise therapy for chronic low back pain as example. Spine. 2007;32:1785–90.
25. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med. 2001;134:663–94.
26. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
27. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gotzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158:200–7.
28. Glazener C, Boachie C, Buckley B, Cochran C, Dorey G, Grant A, et al. Urinary incontinence in men after formal one-to-one pelvic-floor muscle training following radical prostatectomy or transurethral resection of the prostate (MAPS): two parallel randomised controlled trials. Lancet. 2011;378:328–37.
29. Hunter KF, Moore KN, Glazener CM. Conservative management for postprostatectomy urinary incontinence. Cochrane Database Syst Rev. 2007;2:CD001843.
30. Brooks Jr HL. Macular hole surgery with and without internal limiting membrane peeling. Ophthalmology. 2000;107:1939–48.
31. Paques M, Chastang C, Mathis A, Sahel J, Massin P, Dosquet C, et al. Effect of autologous platelet concentrate in surgery for idiopathic macular hole: results of a multicenter, double-masked, randomized trial. Platelets in Macular Hole Surgery Group. Ophthalmology. 1999;106:932–8.
32. Taggart DP, Lees B, Gray A, Altman DG, Flather M, Channon K, et al. Protocol for the Arterial Revascularisation Trial (ART). A randomised trial to compare survival following bilateral versus single internal mammary grafting in coronary revascularisation. Trials. 2006;7:7.
33. Taggart DP, D'Amico R, Altman DG. Effect of arterial revascularisation on survival: a systematic review of studies comparing bilateral and single internal mammary arteries. Lancet. 2001;358:870–5.