

How to design a randomised controlled trial

P. Brocklehurst*¹ and Z. Hoare²

In brief

Explains the PICO statement and the basics of randomisation.

Readers will understand the basics of statistical testing, sample size calculations and the basics of bias.

Readers will also understand the role of Clinical Trials Units and why they are important when considering an experimental design

This practical paper explains how to design an randomised controlled trial (RCT) for those who have little prior knowledge of the topic. It covers the basics of randomisation, statistical testing, sample size calculations, bias and the role of Clinical Trial Units.

Introduction

Randomised controlled trials (RCTs) are the workhorse of evidence-based healthcare and the only research design that can demonstrate causality, that is, that an intervention causes a direct change in a clinical outcome. Although they can be complex, the idea at its simplest is to 'create two identical systems into one of which a new component [the intervention] is introduced'.¹ Observations are then made of outcome differences that occur between experimental and control conditions and 'should a change occur, it is attributed to the one difference between them' (Figs 1 and 2).¹ This paper aims to explain how to design an RCT for those who have little prior knowledge of the topic and specifically explore the following areas:

- The PICO statement
- Randomisation
- Trial design
- Statistical testing
- Sample size calculations
- Bias
- Clinical Trials Units.

¹Director of NWORD Clinical Trials Unit and Honorary Consultant in Dental Public Health; ²Principal Statistician, NWORD Clinical Trials Unit

*Correspondence to: Professor Paul Brocklehurst
Email: p.brocklehurst@bangor.ac.uk

Refereed Paper. Accepted 3 February 2017
DOI: 10.1038/sj.bdj.2017.411

Before you start

Whether you are designing an individually randomised, cluster randomised, stepped wedge or adaptive trial, your research question always returns to your PICO statement. Precision in defining a research question is a key skill; the more precise the research question is, the easier it is to design your study. The PICO statement divides a research question into four basic parts: the patient/population (who are you intending to conduct the study with/on), the intervention, the control and the outcome measure.

P = Population

The first step in developing a well-built question is to identify the patient problem or describe the population group of interest. When identifying the P in PICO it is helpful to ask yourself how you would describe this population group to another person. What are the important characteristics of the group? For example, this could be children, or more specifically young children under the age of five years of age. For example, in the Northern Ireland Caries Prevention In Practice (NIC-PIP) trial, the eligible population was

Fig. 1 The experimental approach to evaluation

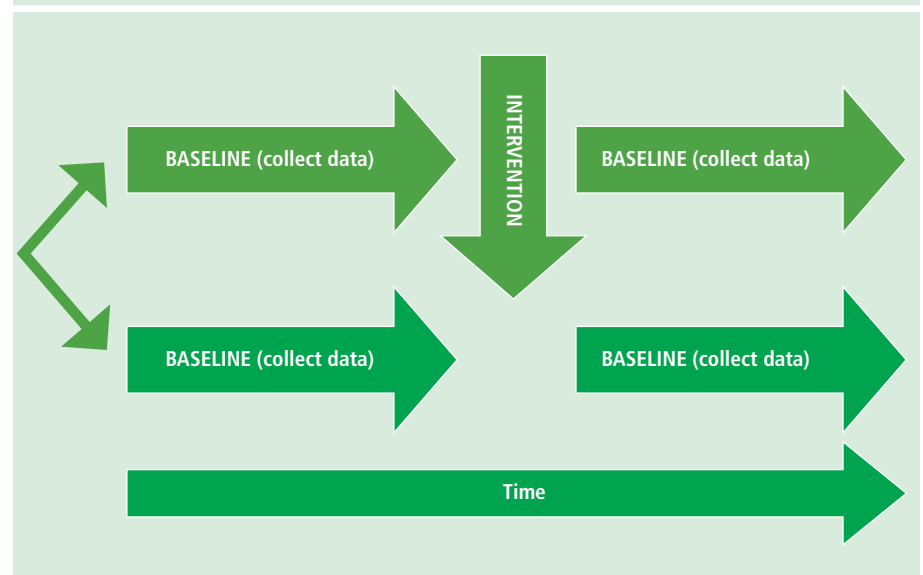


Table 1 Examples of PICO statements for recent/ongoing trials conducted in a primary care environment

| Trial | Population | Intervention | Control/Comparator | Clinical primary outcome measure | Follow-up period |
|----------|--|---|--------------------------------|---|------------------|
| iQuaD | Regularly attending NHS dentate adult patients who have periodontal health, gingivitis or moderate periodontitis | Personalized oral hygiene advice | Standard oral hygiene advice | Gingival inflammation/bleeding on probing at the gingival margin AND oral hygiene self-efficacy | Three years |
| NIC-PIP | Children aged 2–3 years, who were caries free at baseline | Fluoride varnish, a toothbrush and a 50 ml tube of adult toothpaste and standardized evidence-based prevention advice provided at 6-monthly intervals | Prevention advice alone | Proportion of children who converted from caries-free to caries-active states | Three years |
| INTERVAL | Regularly attending NHS dentate adult patients | Risk based recall interval based on the patient’s risk of dental disease AND 24 month recall interval | 6 month recall interval | Gingival bleeding AND quality of life (OHIP-14) | Four years |
| FiCTION | Children aged 3–7 years | Conventional caries management with best practice prevention AND biological management of caries with best practice prevention | Best practice prevention alone | Incidence of either pain or infection related to dental caries | Three years |

caries free children aged between two and three years of age (Table 1).²

I = Intervention

Identifying the intervention is the second step in the PICO process. It is important to identify what you plan to do to your population. This may include the use of a specific diagnostic test, treatment, adjunctive therapy, medication or a recommendation to use a product or procedure. When thinking about conducting a randomised controlled trial, the intervention would be the health technology that you intend to test experimentally. In NIC-PIP, this was the delivery of a preventive regime in line with *Delivering better oral health*. In the iQuaD trial, the intervention was personalised oral hygiene advice (Table 1).³ Some RCTs also employ more than one active arm simultaneously. For example, the FiCTION trial utilises two active arms: conventional caries management with best practice prevention and biological management of caries with best practice prevention (Table 1).⁴

I = Control or comparison

The control or comparison is the third step to take in building your PICO question. This represents the alternative you are planning to compare the intervention to. This can take a number of forms. For example, it could be no active intervention. This is classically known as the ‘control group’ in an RCT. For example, you might be comparing a new high fluoride toothpaste to prevent dental caries. However, it could equally refer to testing a new intervention compared to existing treatment or two

different types of intervention like Atraumatic Restorative Treatment *versus* the Hall Technique, in what is called a ‘head-to-head’.⁵

O = Outcome

Determining the primary outcome measure (POM) is the final step in building the PICO question and one of the most important as it has ramifications on how you statistically test for differences between the intervention and the control/comparator. It specifies what you would expect to see, should the intervention be successful. It is important to decide here whether your POM would be measured using a continuous variable or an ordinal one. The difference between these two types of variables is that a continuous variable describes outcomes that are measured on a scale, like height or weight, whereas ordinal variables are categorical in nature and as the name suggests, can be placed in order. For example, if a person is asked about their feelings towards their dental care and the available responses are unsatisfied, neutral or satisfied, this would be an ordinal variable.

Another key aspect to specify when thinking about your POM is its time to expression, that is, how quickly you would expect to see your result. Time to expression has a critical influence on the duration of the trial (and thereby cost) and will obviously vary with the type of disease under investigation. For example, trials evaluating interventions for gingivitis will have a much shorter duration compared to caries trials. In the FiCTION trial, the research question is ‘what is the clinical and cost effectiveness of restoration caries in

primary teeth, compared to no treatment?’ Here, the POM is the incidence of either pain or infection related to dental caries and the follow-up period is three years (Table 1).

Randomisation

When a new RCT is being planned, researchers are said to be in equipoise. This means that we are uncertain whether the new treatment being experimentally tested actually produces a benefit for the participant. This is an ethical position. If we already have evidence that a new treatment is better than another, we should be giving this treatment to the patient already and if we know there is no difference or that the new treatment is harmful, we shouldn’t be offering it all to the patient. Consolidated Standards of Reporting Trials (CONSORT) 2010 states ‘ideally, participants should be assigned to comparison groups in the trial on the basis of a chance (random) process characterised by unpredictability.’⁶ The requirement is there for a reason. Randomisation of the participants is crucial because it allows the principles of statistical theory to stand and as such allows a thorough analysis of the trial data without bias.

So, how do we randomise? Surely putting participants into random groups is as simple as tossing a coin? This is randomisation in its simplest form but in many cases it results in an unbalanced sample. For example, in a small trial of say 50 participants, tossing a fair coin 50 times would result in a 25:25 split only 7.95% of the time!

There are many different types of randomisation. Tossing a coin or using a random number

Fig. 2 Trials in context

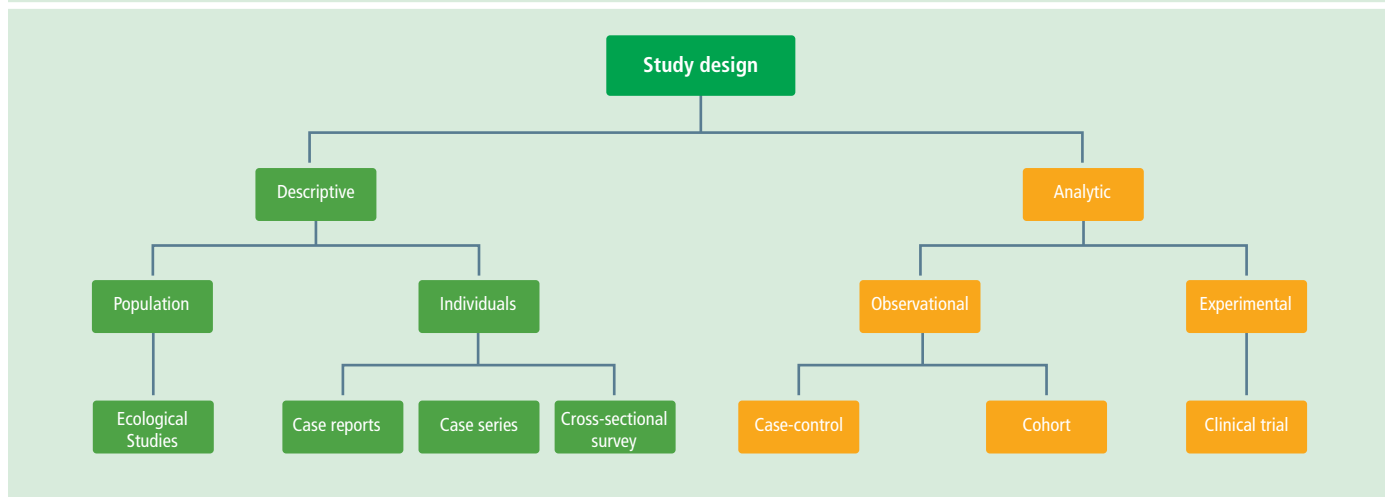


table are examples of simple randomisation. Restricted randomisation uses methods to control the imbalance between the groups; generating a random list AAABBBABABAABB allows participants to be allocated as they arrive to the next treatment on the list. With the list here we know that at the sixth, eighth, tenth and fourteenth participants we have balance in allocation. Stratified randomisation allows us to account for and control certain characteristics within the population of participants such as gender or age (factors that might confound the final effect). It is recommended that stratification should be used sparingly and only in those characteristics that you think would potentially affect your outcome.

When choosing a randomisation method it is important to determine whether the method can accommodate enough treatment groups. For example, tossing a coin would be difficult to implement for a trial with three arms. It is also important to determine how predictable the method is. A deterministic algorithm (not considered randomisation) would allow you to be able to predict what treatment would be allocated next. A static random element would mean that each allocation is made with a pre-determined probability (tossing a coin gives a 50:50 chance of either treatment being assigned). A dynamic element adjusts the probability of being allocated to a treatment based on what has already been allocated in the trial so far. This is the basis of the North Wales Organisation for Randomised Trials in Health's (NORTH) remote randomisation system.⁷

Other considerations include, can the method accommodate for stratification variables and if so, how many? Can the method handle an unequal allocation ratio?

Is thresholding used (that is, maximum level of accepted imbalance)? Can the method be implemented sequentially that is, as the patients walk through your door? Is the method complex? Is the method suitable for cluster randomisation? Decisions like these mean that often a Clinical Trials Unit is needed in the design and planning of your trial. Further reading can be found here.^{8,9}

Designing your trial

There are different phases of RCTs. Phase I trials are described as 'first into person', whilst Phase II are slightly larger trials that commonly determine efficacy that is, does the intervention work or not. Phase III trials take this a step further and determine effectiveness that is, does the intervention produce health benefits in the real world. This section will focus on Phase III trial designs. Again, it is important that you consult a statistician at this stage: 'to call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of'.¹⁰

Feasibility or pilot or full trial

A key question to ask when designing a trial is do you have all the information to inform all the parameters needed. A feasibility study helps you determine whether a definitive trial is 'feasible'. This type of study is often not randomised. This is because the intervention under study is commonly under development and the study plan or intervention itself would change before a definitive study is started. The important outcomes of a feasibility study will be things like the ability to recruit participants,

ability to retain participants for the length of time required, the suitability of the proposed outcome measures, and willingness of the participants/clinicians to be involved.

Pilot studies can be thought of as a small version of the definitive trial, that is, your intervention is now established, but there is still some uncertainty about whether your definitive trial can run as planned. A pilot study will assess all the features that a feasibility study does and often the terms are used interchangeably. Further reading on the issue of terminology can be found here.¹¹⁻¹³ With a stable intervention to test, enough information about the difference in the POM you expect (known as the effect size), enough understanding about the time to expression of your chosen POM, confidence about the feasibility of running the trial *as designed*, then it may be time to design a definitive trial of the intended intervention.

Randomised or non-randomised

A RCT is seen as the 'gold standard' of trial design but there are some situations where randomisation is not possible. Uncontrolled or non-randomised trials are used when randomisation is not possible or is unethical. The results of non-randomised or uncontrolled trials may be considered less reliable as there is an increased risk for errors affecting the outcome of the trial. An example of a non-randomised trial is Lam *et al.*'s (2010) study of mental-health first aid training.¹⁴ The 'SPIRIT 2013 Statement' provides recommendations for a minimum set of scientific, ethical, and administrative elements that should be addressed in a clinical trial protocol.¹⁵ It is worth remembering that a non-randomised trial will have to be analysed and reported

Table 2 Statistical considerations when determining a sample size

| # | Question |
|---|--|
| 1 | How large do you expect the difference to be between the intervention and control group to be? Is this clinically significant that is, a difference of 1mm in probing depth may be statistically significant, but is it really important in practice? If the effect size is small, the sample size will be large and vice versa. |
| 2 | How much variation do you expect to see in the difference in the POM between the intervention and control arm? If the variability is small then the sample size will be small, whereas if the variability is large then the sample size will be large. |
| 3 | What direction will the effect be? For example, do you think your intervention will produce a positive or negative difference in your POM, compared to the control (or will it be indifferent). |
| 4 | Are you going to record the POM at baseline and at the end of your trial in the same participants (known as 'repeated measures') or in a different sample. The former is common with clinical interventions, whilst the latter is more common with public health interventions at a population level, where it is impractical to measure the same group of people at the end of the trial compared to at baseline. The latter requires a larger sample size. |
| 5 | Are the same number of participants required in both the intervention and control arm, or are more participants are required in one group compared to the other. Unequal allocation to groups will require a larger sample size than equal allocation of groups. |

Table 3 Non-statistical considerations when determining a sample size

| # | Question |
|---|--|
| 1 | Can you recruit the number of participants required from the sample size calculation. This can be problematic, particularly with rare disorders. |
| 2 | Is it 'do-able' in terms of costs. The cost of the study needs to be considered and whether there will be sufficient funds to conduct a study with the number of participants required. |
| 3 | How long does your study need to run for? This is linked to the 'time to expression' of your POM. It is also influenced by how quickly you can recruit the numbers of participants that meet the eligibility criteria of your trial. |
| 4 | Do you need to run a pilot/feasibility study first? Do you have enough information to reliably calculate a sample size? |
| 5 | How many participants might withdraw from the study before the study is completed (attrition of the study). This depends on the study and the population of participants. As an example, if the study is in an elderly population then you would need to consider that participants may pass away before the study is completed. Remember there is a difference between withdrawing from a 'treatment' and withdrawing from the study and data should be collected on all participants as long as possible regardless of their level of participation. |

very differently if randomisation is not a key component. All trials should be reported to CONSORT standards and it worth keeping these guidelines in mind during the design process.¹⁰

A key question with RCTs is whether the randomisation is at an individual level or at a cluster level. The individually randomised parallel group design is typically seen as the standard RCT design and remains the favoured approach by funders. An example of an individually randomised parallel group design in dentistry is the FICTION trial.⁵ This is appropriate *IF* the intervention is to be delivered to an individual and there is no possibility of contamination. However, this is not always possible. For example, if a community based oral health prevention programme was being delivered in a school, it would be difficult to undertake the intervention on one child and not affect another child. The environment in the school and the teachers that undertook

the intervention would find it difficult not to influence a child in the control group. In such cases, cluster randomisation would be used. In this example, schools would be the unit of randomisation, not the individual (so a whole school would initiate the intervention and a whole school wouldn't). Two examples of cluster randomised trials in dentistry are NIC-PIP and IQuaD.^{2,3}

Another design of definitive trial is a stepped wedge trial. Here, the unit of randomisation is not the individual or the cluster, but time. This might sound complex, but all this means is that everyone gets the intervention, but just not yet. These designs are not commonly used, but have the advantage that they map on well to a policy roll-out, where everyone will get the intervention. As a result, there are ethical advantages to the approach. For example, if a new practice prevention tool from Public Health England was going to be introduced across England and could be rolled

out on a staggered basis, this approach could be undertaken. At specific time-points in the trial (the 'step'), participating practices that are not yet adopting the intervention (who are currently acting as the control) come 'on-line'. The down-side of stepped-wedge trials is that as each new wave of practices adopt the intervention, all the recruited practices in the trial have to have the POM measured again. Another disadvantage of this design is that you must have sufficient time between each step for the disease of interest to express itself. If this was a gingivitis measure, then this would not be too problematic, but if the researchers were examining the impact of the preventive intervention on dental caries in adults, this would mean that the trial would be very long. One example of a dental stepped wedge trial is the SOCLE-II trial.¹⁶ Here, researchers are exploring whether enhanced oral healthcare or usual oral healthcare is the most effective for people in stroke care settings. Rather than randomising the participants into the two groups, they are rolling out the intervention (enhanced oral healthcare), one ward at a time. More information on stepped-wedge trials can be found here.^{17,18}

Demonstrating success

A statistical hypothesis in a trial describes what the researcher expects to see happen to their chosen POM as their intervention is applied to the intervention arm. This assumption may or may not be true. The null hypothesis in a superiority trial assumes that changes in the POM result from chance and that there is no difference between the intervention and the control arm. The alternate hypothesis assumes that changes are influenced by some non-random cause that is, the intervention the researcher has introduced has worked!

Individually, cluster and stepped-wedged designs commonly test a directional hypothesis that the new intervention produces a health benefit compared to the control or an existing intervention ('head-to-head'). These are termed 'superiority designs' and from a statistical perspective, test whether your point estimate (mean if the POM is measured using a continuous variable) lies above or below the 95% confidence interval (CI) in the control arm. However, sometimes our question is whether a new treatment is as good as another treatment or meets a certain standard. Trials that explore these issues are known as 'equivalence' or 'non-inferiority'

designs. ‘Equivalence’ trials determine whether the value for the POM in both arms is not statistically different that is, that the 95% CI of the difference of the two groups lies within an acceptable margin (the equivalence margin). ‘Non-inferiority’ designs test the difference between two arms and test whether the new intervention is not unacceptably worse than the other that is, that the lower end 95% CI of the difference in POM does not extend below the pre-defined non-inferiority margin. More reading can be found here.¹⁹

Statistical tests commonly quote the ‘p’ value, which describes the statistical significance of the results. The p-value is the probability of obtaining an effect at least as extreme as the one observed assuming the null is true, therefore the lower the p-value, the more likely that there is a significant difference between the scores. It is generally accepted that any POM being tested is statistically significant if the p-value is below 0.05. If this is the case, the null hypothesis can be rejected. However, it is worth knowing that the more times you test something, the more likely by chance you are going to find something statistically significant. In the cases of multiple testing, consideration should be given to adjusting the level of significance.

Participants

There are three elements to a sample size calculation: the p-value, power and the effect size. The p-value is set commonly at 0.05 (as highlighted above). Power is the probability that you will see an effect *IF* an effect is there to be seen. Sometimes we don’t see a statistically significant effect because quite simply, no effect exists. However, sometimes an effect is there, but we don’t have the numbers to see it (this is called being under-powered). For RCTs, we set this probability of detecting an effect *IF* an effect is there to be seen at 90%. The only element that varies in the power calculation is the effect size that is, this is the main element that the researcher needs to determine in consultation with a statistician. An effect size is a point estimate of the measure of the strength of effect standardised by the variability of the measure that is, the expected difference in your POM between the intervention and control arm (for example, if your POM is measured using a continuous variable, this would commonly be the mean with the variability represented by the standard deviation).

Table 1 and Table 2 provide you with the details that a statistician would need to know

before a sample size can be calculated. Table 3 also provides some of the important non-statistical considerations in trial design. More reading can be found here.²⁰

What about bias?

Common to all trial designs is the need to reduce bias. A bias is a systematic error and can operate in either direction: under or over-estimating the true intervention effect. Bias is caused by flaws in the design of the study and so is not the same as imprecision, which is a random error. Selection bias refers to systematic differences between the intervention and control arm caused by differences in baseline characteristics. This should be removed if the randomisation process was effective. Ensuring that participants are blind to their allocation (where possible) reduces the risk that knowledge of which intervention was received, rather than the intervention itself, affects the outcome. This is called allocation concealment. Detection bias (or ascertainment bias) refers to systematic differences produced by differences in how outcomes are determined. Blinding of outcome assessors may reduce the risk that knowledge of the intervention, rather than the intervention itself, affects the POM. Blinding of outcome assessors is especially important when subjective POMs are used, for example, how nervous were you during your dental treatment?

Attrition bias describes the systematic differences between the intervention and control arm caused by withdrawals from the trial that is, when the participants do not want to take part anymore. This can skew the numbers and mix of participants in each arm. It may also tell you that your trial is not socially acceptable! Reporting bias (or publication bias) refers to systematic differences caused by researchers and journals only reporting positive effects of intervention.²¹ This can be seen in the pharmaceutical industry where negative results about the effects of a particular drug can get hidden.²²

Clinical Trials Units

Clinical Trials Units (CTUs) are ‘specialist units which have been set up with a specific remit to design, conduct, analyse and publish clinical trials and other well-designed studies’ (<https://youtu.be/QvGaGEHgwXg>).²³ Commonly, they have a number of functional areas:

- Statistical support (pre, per and post-trial)
- Trial management

- Quality assurance
- Information technology.

CTUs have expertise in the co-ordination of trials, particularly those that involve Investigational Medicinal Products, where compliance with the Medicines Health Regulatory Authority is critical to discharge the expectations in the ‘UK Medicines for Human Use (Clinical Trials) Regulations’.²⁴ Some also provide specialist statistical advice for clinicians. For example, although NWRTH has over £18 million of trials on its portfolio from across the United Kingdom, it is also part-funded by the Welsh Government to provide the Research Design and Conduct Service (<http://nworth-ctu.bangor.ac.uk/research-support-service/index.php.en>).

Most CTUs, but not all, are registered with the United Kingdom Clinical Research Collaboration (UKCRC) and many specialise in specific areas, like Clinical Trials of Investigative Medicinal Products (drug trials). NWRTH has a traditional strength in pragmatic trials and trials of complex interventions. Methodologically, they link with initiatives like TrialForge (<http://www.trialforge.org>) and work to understand how to ‘make trials work’ (see <http://nworth-ctu.bangor.ac.uk/trials.php>).

When preparing for a grant application, researchers are encouraged to approach CTUs early to get help on designing their project. The National Institute of Health Research sees CTUs ‘as an important component of any research application and funded project’ and you are expected to inform them whether you have contacted a CTU in any grant application. They also provide a useful schematic of the necessary steps to take when planning a definitive trial (<http://www.ct-toolkit.ac.uk>).

In summary

This paper has explored the key design elements of RCTs. Although there are significant challenges encountered when designing such complex studies, thinking through each component described above will provide clarity and hopefully encourage more GDPs to get involved in research.²⁵ This is important as there is an increasing need for high-quality evidence from primary care settings to guide the delivery of future healthcare.

1. Pawson R. *The science of evaluation*. London: Sage Publications, 2013.
2. Tickle M, O’Neill C, Donaldson M *et al*. A randomised controlled trial to measure the effects and costs of a dental caries prevention regime for young children

- attending primary care dental services: the Northern Ireland Caries Prevention In Practice (NIC-PIP) trial. *Health Technol Assess* 2016; **20**: 1–96.
3. Clarkson J E, Ramsay C R, Averley P *et al*. IQaD dental trial; improving the quality of dentistry: a multicentre randomised controlled trial comparing oral hygiene advice and periodontal instrumentation for the prevention and management of periodontal disease in dentate adults attending dental primary care. *BMC Oral Health* 2013; **13**: 58.
 4. Innes N P T, Clarkson J E, Speed C, Douglas G V, Maguire A, FICTION Trial Collaboration. The FICTION dental trial protocol filling children's teeth: indicated or not? *BMC Oral Health* 2013; **13**: 25.
 5. Hesse D, de Araujo M P, Olegário I C, Innes N, Raggio D P, Bonifácio C C. Atraumatic restorative treatment compared to the Hall Technique for occluso-proximal cavities in primary molars: study protocol for a randomized controlled trial. *Trials* 2016; **17**: 169.
 6. Schulz K F, Altman D G, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c332.
 7. Bangor University. NWORTH remote randomisation system. Available at <http://nworth-ctu.bangor.ac.uk/randomisation/index.php.en> (accessed January 2017).
 8. Hoare Z, Whitaker C J, Whitaker R. Introduction to a generalized method for adaptive randomization in trials. *Trials* 2013; **14**: 19.
 9. Russell D, Hoare Z S J, Whitaker R, Whitaker C J, Russell I T. Generalized method for adaptive randomization in clinical trials. *Stat Med* 2011; **30**: 922–934.
 10. Fisher R A. Presidential Address by Professor RA Fisher. *Sankhyā. Ind J Statistics* 1938; **4**: 14–17.
 11. Arain M, Campbell M J, Cooper C L, Lancaster G A. What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Med Res Methodol* 2010; **10**: 67.
 12. Lancaster G A. Pilot and feasibility studies come of age! *Pilot Feasibility Stud* 2015; **1**: 1.
 13. Eldridge S M, Lancaster G A, Campbell M J *et al*. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. *PLoS ONE* 2016; **11**: e0150205.
 14. Lam A Y, Jorm A F, Wong D F. Mental health first aid training for the Chinese community in Melbourne, Australia: effects on knowledge about and attitudes toward people with mental illness. *Int J Mental Health Systems* 2010; **4**: 18.
 15. Chan A W, Tetzlaff J M, Altman D G. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. *Ann Intern Med* 2013; **158**: 200–207.
 16. Brady M C, Stott D, Weir C J *et al*. Clinical and cost effectiveness of enhanced oral healthcare in stroke care settings (SOCLE II): a pilot, stepped wedge, cluster randomized, controlled trial protocol. *Int J Stroke* 2015; **10**: 979–84.
 17. Brown C A, Lilford R J. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol* 2006; **6**: 54.
 18. Woods B, Russell I. Randomisation and chance-based designs in social care research. Available at <http://sscr.nihr.ac.uk/PDF/MR/MR17.pdf> (accessed January 2017).
 19. Schumi J, Wittes J T. Through the looking glass: understanding non-inferiority. *Trials* 2011; **12**: 106.
 20. Ellis P D. *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge: Cambridge University Press, 2010.
 21. Chan A W, Altman D G. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005; **330**: 753.
 22. Goldacre B. *Bad science*. Harper Collins: London, 2008.
 23. UKCRC . Clinical Trials Units. Available at <http://www.ukcrctu.org.uk/?page=CTURole> (accessed January 2017).
 24. Legislation.gov.UK Medicines for Human Use (Clinical Trials) Regulations. Available at <http://www.legislation.gov.uk/uksi/2004/1031/contents/made> (accessed January 2017).
 25. Martin-Kerry J M, Lamont T J, Keightley A *et al*. Practical considerations for conducting dental clinical trials in primary care. *Br Dent J* 2015; **218**: 629–634.