

How to quickly critically appraise a randomized control trial

Will Carroll

Abstract

Being able to critically appraise scientific papers is an important skill. A thorough and methodical approach is necessary, particularly for peer-review purposes. However, it is also important to be able to do this quickly and efficiently. In some circumstances, students and trainees may be asked to perform a review on an unseen paper as part of an assessment. With apologies to my erstwhile teacher, David Sackett, this brief article offers a framework which will allow a fairly complete review of a randomized controlled trial in 15–20 min. This framework requires a reviewer to assess the Validity, Importance, Applicability, GReatness of benefit and Acceptability of any study and is encompassed by the acronym VIAGRA.

Keywords assessment tools; evidence-based medicine; randomized controlled trial; RCT

Introduction: ignore the introduction (and discussion)

<epigraph>The great tragedy of science - the slaying of a beautiful hypothesis by an ugly fact.

<epigraph source>Thomas Huxley

Reviewing scientific papers is surprisingly difficult. There are many possible reasons for this, but I believe that there are three main barriers to reviewing well.

The first is that even with the best of intentions most clinicians simply do not have the time. As a journal editor with a passion for research, I perhaps should be setting a good example. However, I readily admit that my office is littered with journals (and some books) still in their original wrappings. The second is that we are all approach any topic or study with a degree of bias. This is difficult to avoid. I know that I am more likely to be less critical about a paper that confirms my view of the world and I am sure this tendency affects other reviewers too. Thirdly, we are swayed by good narratives. As Huxley noted in the 19th century, good science often kills beautiful hypotheses. Therefore beware the narratives that exist in your own experience and that of the authors. Whilst placing research into context is helpful once the study has been completely assessed, the introduction and discussion sections of any paper which draw out the authors (and reviewers) narratives on the topic these should be disregarded completely in the first instance.

Having limited your initial decision about a paper to the methods and results section it then helps to have a structured

Will Carroll MD FRCPCH, Consultant Paediatrician, University Hospitals of the North Midlands and Keele University, UK. Conflict of interest: none declared.

VIAGRA: making you a more potent reviewer

Validity – Were the methods valid? Follow an ABCDE approach

Assignment and accounting
Blindness
Controls
Demographics
Equality of treatment

Importance – Was the study important? How common is it? Are the effects clinically meaningful?

Applicability – Can this research be applied to your patients or a wider group?

GReatness (of benefit) – What is the number needed to treat (or harm)?

Acceptability – Is it ethical? Would it be acceptable in other health-care settings? Who funded it?

Box 1

approach. My own approach is given by the acronym VIAGRA (see [Box 1](#)).

Assessing validity: an ABCDE approach

Being able to adequately (rather than completely) review a paper quickly requires a memorable approach. In determining whether a study is valid it is useful to first of all determine whether the research question was adequately defined. On first reading therefore I try to re-frame any study using a PICO format.

Who were the *patients*, what was the *intervention*, what was the *comparison* and what was the primary *outcome*? If I cannot do this, then the question being addressed is not adequately focussed (or I am in need of more coffee). To then critically appraise it I then run through five key questions, as follows.

Assignment and accounting

How were participants assigned to their groups and were they all accounted for at the end of the study?

If this is a randomized controlled trial then the assignment should be random and this should be clearly described. Historically, reviewers and researchers have worried a lot about the randomization process as this can lead to bias being introduced. There is probably not a 'perfect' solution to randomization. However, clinicians should be at least one step removed from randomization wherever possible. Otherwise we know that there is a temptation to enter participants into active treatment groups. Clinicians and participants should have the allocation sequence concealed from them.

Once individuals are assigned to a group it is crucial to track their progress through the study and carefully check to see whether everyone entered into the study is accounted for. In many better studies there is a flow diagram explaining this ([Figure 1](#)).

Blinding

Blinding is an important part of any trial. It is not always possible but if it is not done, or done poorly then it can lead to bias as researchers or participants will know which treatment arm they are in.

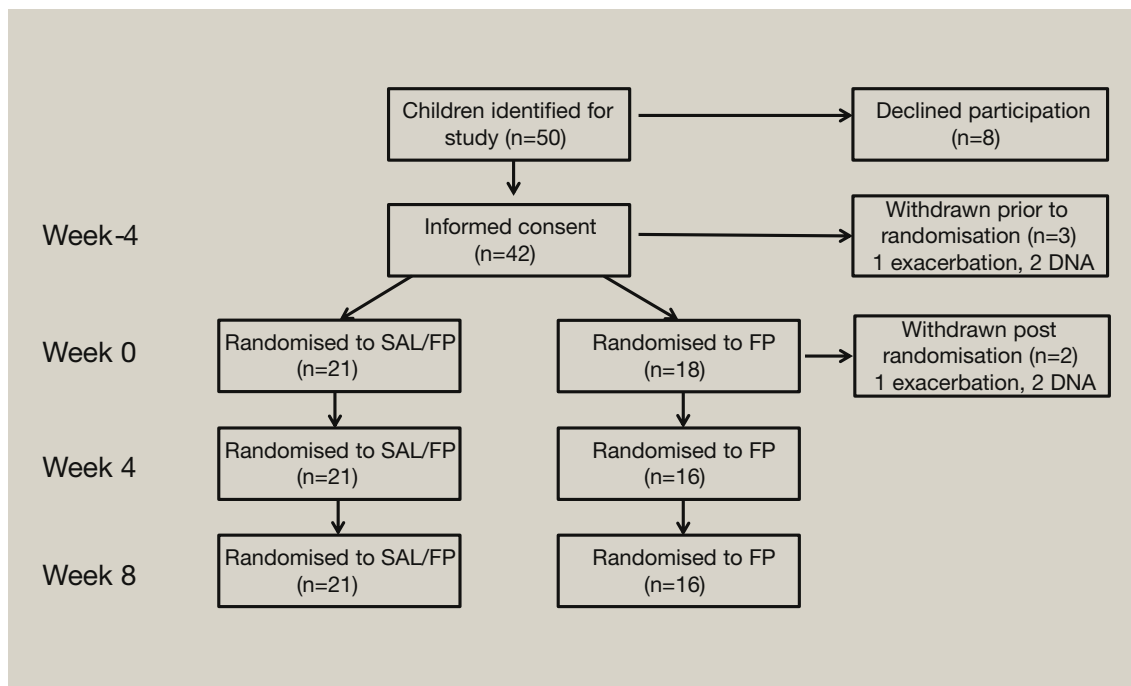


Figure 1 Example of a flow diagram. *Data taken from Carroll et al. 2010. See references and further reading. FP = fluticasone propionate, SAL = salmeterol.

Sometimes a common sense approach is required in assessing this and some knowledge of the treatments being offered is helpful. If a treatment is known to have very predictable effects or side effects then blinding might be more difficult. For instance, a medication with a very particular taste, appearance or smell can be difficult to blind, particularly if individuals in the study have taken it before.

An example of a treatment that is difficult to blind for is nebulized adrenaline (see references and further reading). When this was compared against nebulized 0.9% saline in a research study that I helped to conduct, the researchers became aware quite quickly that babies receiving adrenaline would often look pale and become tachycardic. It was fairly easy for parents and clinicians to guess which group a child had been randomized to.

Controls

Where were controls recruited from? Was this the same source as participants or were they recruited from elsewhere. This may well affect the validity of a study.

Demographics

One way to check quickly to see if controls are from a different source or likely to behave differently from the intervention group is to check through the demographics for intervention and control groups. These should be reported and all the characteristics that might influence outcome should be listed e.g. age and sex. Be careful with small studies which assume that because there is not a statistically significant difference between measured demographics that these are not important. If there is a difference then ask yourself, 'could this account for part or all of the reported effect'.

Equality of treatment

It is important to check the protocol description and ensure that the control groups is treated equally to the intervention group *in all aspects aside from the experimental intervention*. Many studies purport to study a single treatment but in reality are studying a multitude of effects.

If treatment groups spend longer with healthcare professionals then outcomes might be influenced by this extra contact time. Therefore, to be valid then this must be accounted for in experimental design. If it seems likely to have occurred due to extra visits or interventions in the intervention group then this is an important limitation.

Is the study important?

Deciding whether a study addresses an important topic is somewhat subjective. Small improvements in outcomes for important or common diseases can be very important. All improvements in care and outcome are of course very important to families and clinicians. However, it is helpful to ask some questions of any study at the outset.

It is useful to see if the researchers have undertaken a power calculation. This gives an indication of whether the study was adequately powered to look for any observed effect. I then perform a 'sense check' on whether this effect is clinically important. For instance, if the effect is on a clinical score, has this been validated and do the individual components of the score all have clinical importance.

Some studies measure outcomes that are removed from clinical outcomes that matter. A reduction in C-reactive protein levels without reduction in hospital stays or mortality is probably unimportant. A reduction in wheeze (audibility) is often reported as part of a composite measure of respiratory distress, but this

may not equate to reduced requirement for oxygen or hospitalization. Even these may be dependent upon the doctors and nurses judgement and are less helpful than 'hard' end points such as death and/or disability.

Is the study applicable (to my patients)?

Who, where and when?

This is not always easy to answer. However, it is helpful to consider the Who, where and how of any study.

Who is particularly important. Studies in very different population groups to your own may significantly limit the applicability of a study to your own practice. Paediatricians are very aware that many adult studies are erroneously applied to paediatric populations so age is a particularly important consideration. However, there are other important considerations. For instance, some conditions have different responses to treatment that are dependent upon ethnicity.

Where is also important. Some of the settings for reported studies are so different in context from your own that research might not be easily transferred. Children in developing countries are likely to present to hospital (and researchers) much later and therefore results may be misleading for those working in the developed world and vice versa.

When is important when considering older studies. In some fields the clinical treatment has evolved so much that reported interventions may no longer be applicable. For instance studies of asthma treatments in the era prior to widespread use of inhaled corticosteroids.

Assessing the Greatness of benefit

Be careful with small or very large studies

Alongside a subjective judgement about the importance of a particular treatment, it is useful to consider in real terms *how great the observed benefit is*. Many recent studies will be powered to detect a 'minimally important clinical difference'. It is helpful to see whether this has been determined by anyone other than the researchers themselves. As patients and families become more involved with research, it is likely that end-points will begin to become more relevant.

Small studies which report a very large effect must be viewed with caution. It helps to look carefully at any reported confidence intervals. When these are wide and particularly if the p value for any observed effect is close to 0.05 be very careful. Remember

that a study with a p value of 0.04 will be reporting a chance finding 1 in 25 occasions.

Some very large studies report a very high level of statistical significance, but a small clinical difference. A judgement here is required in the reviewer about whether any such observations are important. For instance an improvement of 2% in lung function of a child with asthma is unlikely to be of clinical significance to the individual but if a study is large enough it might be able to detect it.

Two useful quick checks of benefit are to calculate the Number Needed to Treat (NNT) and the Number Needed to Harm (NNH) for any proposed intervention.

Acceptability

A final check is to see if the proposed intervention is likely to be acceptable to your own patients. This may depend upon cultural, religious and societal expectations. It is useful to see whether ethical approval has been sought and given – and if so from where. It is very unlikely that a recent randomized controlled trial will be published without ethical approval.

As a final check, I routinely check to see who has funded the research and make a judgement about whether this may have introduced bias either into the protocol design or the presentation of the results. Funders, such as drug companies may have a significant financial interest in some studies being published. Some sources of funding might not be acceptable or at the very least questionable e.g. the tobacco industry. ◆

FURTHER READING

- Carroll WD, Boit P, Cliff I, Clayton S, Jones PW, Lenney W. The childhood evaluation of salmeterol tachyphylaxis study (CHEST). *Paed All Immunol* 2010; **21**: 336–44.
- CONSORT website. <http://www.consort-statement.org/> (accessed 3 November 2019).
- Critical Appraisal Skills Programme. CASP (randomised controlled trial) checklist. 2018 [online]. Available at: https://casp-uk.net/wp-content/uploads/2018/03/CASP-Randomised-Controlled-Trial-Checklist-2018_fillable_form.pdf (accessed 3 November 2019).
- Hariprakash S, Alexander J, Carroll WD, et al. Randomized controlled trial of nebulized adrenaline in acute bronchiolitis. *Ped All Immunol* 2003; **14**: 134–9.