How to appraise clinical trials

Y K Loke

Abstract

Treatment decisions should be based on data from randomized controlled trials. However, trials are susceptible to bias, and there may be important limitations to the data. Critical appraisal of such studies will need to consider both the methodological rigour of the study and the applicability of the results to real-life clinical practice. Instead of simply glancing at the abstract, readers should work systematically through the trial report. The first port of call should be to establish the aims of the study, and then to consider whether the methods used are able to provide an unbiased answer. Particular attention should be directed towards patient allocation ensuring that the study groups are well balanced. Readers should check that the follow-up is adequate, and that there is sufficient blinding of investigator and participants so that preconceived notions do not influence recording of outcomes. The Results section should be reviewed in the light of the trial's objectives to confirm that the researchers have reported all data (positive or negative) that are relevant to the study question. Trials are scientific experiments with specific aspects that can markedly differ from real-life clinical practice. The results in a trial may not be achievable in a different treatment environment, or with less highly selected patients. Sophisticated treatment regimens in the trial may be difficult to deliver in daily practice. All of these potential differences should be considered when evaluating clinical trial results.

Keywords clinical trials; critical appraisal; evidence-based medicine; intention-to-treat; randomization; reporting bias; selection bias

What's new?

- Researchers may try to analyse trial results selectively and report them positively, downplaying less favourable findings in the trial report
- If there was no significant beneficial effect on pre-specified endpoints, there may be attempts to change the focus to more favourable outcome measures, or to engage in multiple subgroup analyses and data dredging to identify specific instances of benefit
- Researchers may use surrogate outcomes and prefer to test their interventions against placebo (rather than best-available treatment), so that positive findings will show up more quickly and easily

In evidence-based medicine, treatment decisions are made after careful assessment of the available clinical trial data. The ability to critically appraise trial reports is essential for doctors who rely on high-quality data to guide their practice. This means that the evaluation of trial data should go beyond a cursory look at the Abstract and Results sections (as busy clinicians are prone to take). At the outset, readers should consider the following questions:

- What was the primary research objective?
- Were these objectives refined into a specific research question?
- Is the question relevant to clinicians and patients?
- Were the design and conduct of the trial adequate to provide a reliable answer?

Y K Loke ***MBBS MRCP MD*** *is Senior Lecturer in Clinical Pharmacology at the University of East Anglia, UK. He qualified from the University of London and trained in clinical pharmacology in Oxford. His research interests include the evidence-based assessment of adverse effects, and he is co-convenor of the Cochrane Adverse Effects Methods Group.*
Competing interests: none declared.

These issues are of particular interest to pharmaceutical physicians, research ethics committees, funding bodies and regulatory authorities.

If the trial does provide a reliable answer to a specific and relevant question then the next considerations for the clinician are:

- What do these results mean for me, and for my patients?
- What influence, if any, should these findings have on medical practice?

Individual clinicians need to judge whether the trial data are applicable to the patients they treat in real-life practice. At a different level, medicines advisory committees, including the National Institute for Health and Clinical Excellence (NICE), use trial results as essential supporting evidence when deciding on the usefulness and cost-effectiveness of a drug across many groups of patients in the wider population.

With the above in mind, readers should work systematically through the trial report, from beginning to end.

Introduction and aim

Here it is important to check that the researchers had a clinically relevant, well-defined study question (hypothesis). Is the trial aimed at comparing the new treatment against the existing options, or simply comparing against placebo or no treatment? While demonstrating superiority over placebo may be an easier option, the data are much less useful to clinicians who need to decide if the new treatment offers an advantage over drugs in current use. If the trial aims to compare two active agents, then it is worth checking whether the control group is actually being given the best existing treatment option in routine clinical use.

Methods

Design and conduct of controlled clinical trials

Almost all pharmacological treatments are tested in controlled clinical trials before being licensed for sale. The essence of a controlled trial is that, in order to compare the effects of therapy, two or more patient groups of similar characteristics are exposed to differing treatments. These trials are scientific experiments in human beings, and should be conducted to rigorous methodological standards. However, as in any

experiment, the scientific integrity of the study, and the reliability of the results can be undermined by the presence of bias.

In simple terms, bias is any process (conscious or unconscious) that causes results to deviate systematically from the true values. There are a number of important areas where bias may crop up in a controlled clinical trial[1]:

- allocation of patients
- delivery of the treatments that are being evaluated
- assessment and reporting of treatment outcomes
- loss of patients to follow-up.

Allocation of patients to intervention arms

Ideally, the patient groups under comparison in a controlled clinical trial should have identical characteristics, and differ only with respect to the treatment arms to which they have been allocated. To achieve such a balance, all trial participants must share the same likelihood of ending up in any particular treatment arm. This is achieved through the process of random allocation — neither the doctor nor the patient knows, or has any influence on, the treatment group to which the patient will be allocated. This can only be achieved if allocation is on the basis of a truly random sequence that cannot be influenced by either the investigators or subjects. There are two important steps in the randomization process.

- Generating a truly random sequence — often using a computer, or random number tables. Failing that, drawing numbers out of a hat, or flipping a coin will have to suffice.
- Making sure that the trialists or the patients cannot work out the sequence, so that they cannot influence the treatment allocation process. This can be achieved through the use of a remote telephone randomization centre, or simply by using sealed opaque envelopes. Inadequate concealment of allocation may result in differences between groups of participants, as well as lack of blinding during the conduct of the trial.

Bias can be introduced inadvertently if these steps are not followed. In one study, patients were openly enrolled into treatment groups depending on the day of admission. This looks, on the surface, like a randomized process but those admitted on a Sunday may be different from those referred on a weekday, and this could lead to an imbalance in the groups. Furthermore, the triallist or the patient could choose their preferred treatment by arranging hospital admission for a specific day. For example, frail patients may prefer being allocated to what appears to be the 'gentler' treatment arm.

Equal delivery of the treatments under comparison

Ideally, patients in each group should be managed in exactly the same way, with the only difference being the specific therapeutic agents under evaluation. This may not always be the case, as illustrated by the following examples.

- The experimental drug was administered in the coronary care unit, while patients in the conventional therapy arm were looked after in general medical wards. Improved outcomes in the experimental group may simply have been due to closer supervision in coronary care, rather than to the drug itself.
- Patients in a study of a new endoscopic device were treated by the consultant who had undergone a special training course in the new technique. Meanwhile, other patients in the trial had their conventional procedure performed by the registrar. Readers should check the trial report to ensure that the treatment groups are indeed receiving the same standard of care.

Measuring treatment outcomes

Bias may not be a problem when measuring hard outcomes such as death or survival, but may creep in when dealing with outcomes that are subject to human interpretation (e.g. deciding the cause of death, reading an echocardiogram, assessing symptomatic change). For example, in a trial demonstrating the benefit of compression stockings in preventing travel-related thrombosis, calf vein clots were monitored by ultrasonographers who were aware of which patients had been using stockings.[2] These technicians may have believed that patients without stockings were at higher thrombotic risk, leading to more thorough scanning and highlighting of borderline abnormalities.

Blinding or masking of treatments has been introduced to get round this type of bias. In double-blind studies, neither the trialist nor the participant knows which treatment regimen is being given. It is worth checking that blinding of treatment is feasible — for example, one would be sceptical about adequate blinding in a trial comparing botulinum toxin to placebo for migraine, where the cosmetic benefit of botulinum is rapidly discernible to patients and investigators.[3]

Methods of data analysis and follow-up

There are numerous reasons why patients may drop out of trials. Some may develop adverse effects, while others may give up because they feel no better on the trial treatment. If these drop-outs are not accounted for, the results of a trial may be misleading (Figure 1) because the remaining patients are not representative of those who originally started on treatment.

In order to get round this type of bias, 'intention to treat' analysis is carried out. All randomized patients are included in the analysis according to the assigned treatment group, irrespective of whether or not they completed the trial. If such analysis is not possible, the trial should report on the numbers who dropped out, and the reasons why.

Reporting of results

Using the CONSORT flowchart (Figure 2)

Many medical journals have now signed up to a standardized reporting structure for randomized controlled trials.[4] A flowchart in the trial report provides a quick and simple way of assessing the flow of patients through the recruitment and randomization process.

Another key feature in the Results section is a table listing the baseline characteristics of the patients in their respective treatment groups. It is important to check that the treatment groups are similar, particularly when differences in patient risk factors (such as age, disease severity, socioeconomic status) may have skewed the treatment response.

Outcome reporting bias

The results should be assessed in light of information gleaned from the study Objectives and Methods sections. Is there a full report on the pre-defined primary outcome of interest? In the absence of