How to appraise clinical trials

YK Loke
K Mattishent

Abstract

Treatment decisions should be based on reliable data. However, randomized controlled trials are susceptible to bias and other important limitations. Critical appraisal of such studies must consider both the methodological rigour of the study and the applicability of the results to routine clinical practice. Readers should work systematically through trial reports. They should establish the aims of the study and 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. There should be adequate follow-up and sufficient blinding of the 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. Critical appraisers should also consider how closely the conditions of the trial (e.g. selection of patients, follow-up arrangements) reflect real-world medicine, allowing the results to be generalizable to routine clinical practice.

Keywords Clinical trials; critical appraisal; evidence-based medicine; intention-to-treat; publication bias; randomization; reporting bias

Introduction and aims

Here it is important to check that the researchers had a clinically relevant, well-defined study question (hypothesis and pre-specified primary outcome). Is the trial aiming to compare the new treatment against the existing options, or simply to compare it against placebo or no treatment? Researchers should specify whether the trial aims to demonstrate superiority of a new intervention, or simply non-inferiority or equivalence between two existing treatments. While demonstrating superiority over placebo may be an initial option, the data are much less useful to clinicians who should decide if the new treatment offers an

YK Loke ***MBBS MRCP MD*** *is Professor of Medicine and Pharmacology at the University of East Anglia, Norwich, UK. 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.*

K Mattishent ***MBBS MRCP*** *is Academic Clinical Fellow in Older People's Medicine at the University of East Anglia, Norwich, UK. Her research interests are in pharmacoepidemiology, particularly with regard to the prevention and management of adverse events in older patients with multiple morbidities. Competing interests: none declared.*

Key points

- Treatment decisions should be guided by evidence obtained from clinical trials
- The reliability of trial results should be evaluated by looking for potential sources of the bias in the design and conduct of the trial
- Researchers may try to selectively analyse and report significant or striking results, and downplay less favourable findings, in the trial report

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. 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?
- Did the investigators report all the outcomes (positive and negative) that they set out to measure?

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, 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 must 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 Care Excellence, 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.

advantage over drugs in current use. If the trial aims to compare two active agents, it is worth checking whether the control group is actually being given the best existing treatment option available in routine clinical use.

The sample size and power calculation of a clinical trial are of particular importance. Estimates of sample size are based on the number of participants needed to reliably measure a clinically meaningful difference in effect. An inadequate sample size can result in failure to detect or rule out a (genuine) difference between the treatments.

Methods

Design and conduct of controlled clinical trials

Almost all pharmacological treatments are tested in controlled clinical trials before being given a marketing authorization (license). 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 on 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 the results to deviate systematically from the true values. There are a number of important areas where bias can 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 by 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 the 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 patients cannot work out the sequence, so that they cannot influence the treatment allocation process. This can be achieved by using a remote telephone randomization centre or simply by using sealed opaque envelopes. Inadequate concealment of allocation can result in differences between groups of participants, as well as lack of blinding during the conduct of the trial.

Bias can be inadvertently introduced if these steps are not followed. In one study, patients were openly enrolled into treatment groups depending on the day of admission. This might appear to be a randomized process but those admitted on a Sunday might be different from those admitted on a weekday, and this could lead to an imbalance in the groups. Furthermore, the trialist (or 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 except for 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 might simply have been the result of closer supervision in coronary care rather than of the drug itself.
- Patients in a study of a new endoscopic device were treated by a specialist who had undergone a dedicated training course in the new technique. Meanwhile, other patients in the trial had their conventional procedure performed by a trainee doctor. 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. These technicians may have believed that patients without stockings were at higher thrombotic risk, leading to more rigorous 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.

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 dropouts 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.

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

