Evidence-based medicine

Michael D Rawlins

Abstract

Good evidence should form the basis for the practice of medicine. Evidence-based medicine encompasses not only pharmaceuticals and devices, but also the appropriate use of diagnostic tests, screening and clinical guideline development. There are, however, a number of different methods for establishing evidence. All have both advantages and disadvantages, with each posing potential problems, particularly from bias and confounding.

Keywords Case-control studies; clinical trial endpoints; hierarchies of evidence; historical controlled trials; randomized controlled trials; systematic reviews

Key points

The evidence underpinning evidence-based medicine can come from different designs:

- systematic reviews
- randomized controlled trials
- historical controlled trials
- case-control studies
- case-series, or very rarely
- case reports

There is no place, however, for 'hierarchies' of evidence.

Background

Evidence-based medicine has been defined as the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients.

Types of evidence

A variety of techniques are used in determining 'best evidence'.

Randomized controlled trials (RCTs)

In a typical RCT, two (or sometimes more) groups of patients are randomly allocated to different treatments. At the end of the study, the outcomes are compared, and if one intervention is 'significantly different' from the other(s), it may be concluded that it is better.

The comparator depends on the objectives of the study but can be an inert placebo, two (or more) different of doses of the investigative drug, or an 'active' control. Although RCTs typically investigate whether one treatment is better than the comparator, there are designs that assess whether one treatment is 'as good as', 'similar to' or 'worse than' alternatives (Table 1).

Pragmatic RCTs attempt to study whether a particular intervention is effective in circumstances approximating to normal conditions of use. Such studies — at least in theory — overcome the fact that many conventional RCTs are conducted in highly selected patient populations with, for example, no co-morbidities.

In *cross-over* RCTs, patients are randomized to one of two interventions of interest and then, after an appropriate interval, are switched to the alternative treatment. Cross-over trials are limited to the treatment of long-term conditions such as chronic pain or hypertension, and particular care has to be taken to ensure that there is no 'carry-over' effect from the first treatment to the second.

RCTs are unquestionably powerful for investigating the effectiveness of interventions. Most importantly, they minimize bias and confounding. *Bias* occurs when there are systematic errors associated with the design, conduct, analysis and reporting of the results of a clinical trial (Table 2). *Confounding* occurs when the relationship between the use of an intervention and the outcome is influenced by another factor, such as the underlying severity of the condition, which is unevenly distributed between the groups. In RCTs, it is often assumed that such confounders are equally divided between the groups, but the play of chance means that this is not always be the case. In analysing an RCT, the possibility of confounding can sometimes be taken in account using statistical techniques.

Despite their advantages, RCTs also have limitations:

- Although adjustments can be made to ensure that known *confounders* are accounted for, the problem of unknown confounders can remain.
- The results of an RCT in a defined group of patients may not be *generalizable to wider populations.* This particularly applies to elderly participants (if the study has been

Designs used in RCTs

Type of design	Explanation
Superiority trial	Tests whether one treatment is better than another
Equivalence trial	Tests whether two treatments have similar effects
Non-inferiority trial	Tests whether one treatment is less effective than another
Futility trial	A form of non-inferiority trial, used particularly during drug development, to test whether a new product appears promising for testing in larger patient populations

Table 1

Michael D Rawlins *GBE Kt BSc MD FRCP FMedSci is Honorary Professor, London School of Hygiene and Tropical Medicine, London, and Chair, Medicines and Healthcare products Regulatory Agency, London, UK. Competing interests: none declared.*

Please cite this article in press as: Rawlins MD, Evidence-based medicine, Medicine (2018), https://doi.org/10.1016/j.mpmed.2018.04.009

Potential biases in clinical trials

Type of bias	Explanation	Measures to control biases
Selection (or allocative) bias	Systematic differences between the comparison groups	Randomized treatment allocation Concealment of treatment assignment
Performance bias	Systematic differences in other aspects of care apart from the intervention under evaluation	Standardized study procedures Standard equipment
Exclusion bias	Systematic differences in withdrawals from the trial	Blinding or masking
Observational (or ascertainment) bias	Systematic differences in outcome assessments	Intention-to-treat analysis
Publication bias	Failure to publish trial results	Prospective registration of all trials Publication of both 'positive' and 'negative' trials

Table 2

confined to younger patients) and to patients with co-morbidities.

- RCTs can be of *insufficient size* to recognize less common adverse reactions.
- During the analysis of a trial, investigators may identify *subgroups* responding better or worse than others. If such potential subgroups have been identified in advance, subgroup analyses can be reasonable. If they have not been pre-selected, they should be disregarded because of the possibility that the findings are the result of chance.
- RCTs have become increasingly *expensive*: costs of $100 million per trial are commonplace, and in some instances costs can amount to > $500 million per trial.

Historical controlled trials (HCTs)

In an HCT, the effects of a particular treatment are compared with so-called historical controls. These can be implicit and based on what is reliable about the natural history of the particular condition, or they can be explicit and based on the experience of a cohort of patients previously studied.

Hip arthroplasty, for example, was introduced using historical controls. Without surgery, it was known that patients with osteoarthritis of the hip would continue to suffer pain and impaired mobility; however, after successful surgery, most patients became pain free and fully mobile.

HCTs suffer, at least potentially, from various disadvantages, particularly bias and confounding. In order for the results of an HCT to be considered reliable, four conditions should be satisfied:

- There should be a plausible basis for the beneficial effects of the intervention.
- The condition should have a known and predictable natural history.
- The intervention should not be expected to have adverse effects that would compromise its benefits.
- The intervention's effectiveness should be sufficiently great as to minimize the possibility of bias and confounding.

Case-control studies

In a case-control study, exposure to an intervention is compared among patients with or without the outcome of interest. The results are usually expressed as an '*odds ratio*' (Table 3) together with *95% confidence intervals*. An odds ratio of 1 suggests that the outcome of interest is similar in the two groups. An odds ratio significantly greater than 1 suggests that the outcome of interest is greater in the exposed group, while an odds ratio of significantly less than 1 indicates that the outcome is less in the exposed group.

An unmatched case-control study

Exposure status	Cases (with the condition)	Controls (without the condition)
Exposed group	a	b
Unexposed group	c	d

Odds ratio = (a/c) ÷ (b/d) = (ad) ÷ (bc).

Table 3

Case-control studies have been used to study the benefits of interventions as well as to assess their harms. In assessing the effectiveness of interventions, the approach is limited by two factors. First, the intervention must already be generally available; and, second, the possibility of confounding is substantial.

For example, case-control studies in the early 1990s suggested that women taking hormone replacement treatment had a reduced rate of ischaemic heart disease. Later large-scale RCTs showed, however, that this was not the case. It became clear that the case-control studies were confounded by an excess of wealthier women, taking hormone replacement therapy, who had fewer risk factors for ischaemic heart disease.

Case-control studies have been extensively — and successfully — used in assessing possible adverse effects from marketed interventions. They are capable of showing not only an association with the use of a particular intervention, but also the absence of such an association. Again, the design and analysis of case-control studies requires investigators to ensure that bias and confounding are minimized but, in the assessment of harms, this is less of a problem.

Please cite this article in press as: Rawlins MD, Evidence-based medicine, Medicine (2018), https://doi.org/10.1016/j.mpmed.2018.04.009