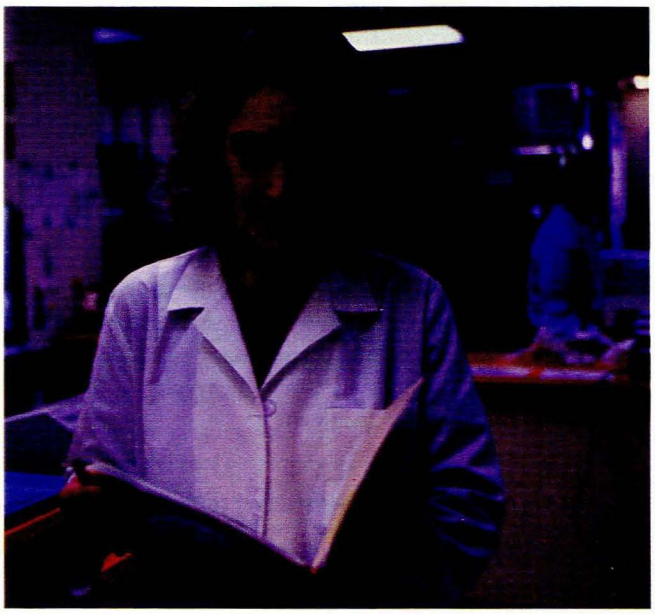
CLINICAL PRACTICE

'IS THIS TREATMENT WORTH WHILE?' HOW TO READ THE MEDICAL JOURNALS

P Rheeder, J A Ker



Doctors are confronted daily with the results of new trials evaluating drugs. Many practitioners are unfamiliar with the terminology and science underlying intervention studies and may find it difficult to interpret such studies.

PLACEBO COMPARATIVE STUDIES VERSUS EQUIVALENCE STUDIES

Until recently most trials have compared placebo with active treatment, the null hypothesis being that the two forms of therapy are equal. A P-value < 0.05 is interpreted as rejection of the null hypothesis, and the conclusion is that active therapy is superior to placebo.

More frequently, however, active therapy is compared with active therapy. In this setting, failure to reject the null hypothesis cannot be interpreted as equivalence. For equivalence to be demonstrated, a trial designed to show equivalence is needed. An equivalence study should state the aim of demonstrating equivalence, much larger numbers of participants are needed (often four times as many as in a placebo comparative study), and the analysis should be handled differently to that of a placebo-controlled trial.[1]

RANDOMISATION AND ITS CONCEALMENT

The natural history of a disease may lead to a favourable outcome in some patients but not others. This necessitates randomisation in order to spread possible extraneous prognostic factors evenly among the groups being compared, provided that the group is large enough.

Adequate concealment of random allocation (e.g. by ensuring that randomisation cannot be manipulated by seeing through an envelope) is, however, essential. Studies with doubtful concealment tend to show a greater treatment effect (therefore biased) compared with studies with adequate concealment.[2]

BLINDING

Blinding is needed to exclude bias. Randomised controlled trials are often done in a double-blind manner — the patient is blinded via the addition of placebo and the doctor is blinded because treatment and placebo are coded in such a manner that it is impossible to discern between the two. It is not always possible to blind with therapy as regimens involving more than one agent may be used. The PROBE design (prospective, randomised, open with blinded end-point evaluation) is often used in this instance.[3] The blinded evaluation of outcome is therefore a minimum requirement.

LOSS OF FOLLOW-UP

The number of patients lost to follow-up should be minimal and should preferably not be related to either form of treatment. Exact figures beyond which the validity of the study is doubtful are not clear. In general, loss of follow-up greater than 20% makes the results questionable; some journals will not publish results if follow-up was less than 80%.[4]

Paul Rheeder, a Specialist Physician, trained at the University of Pretoria and completed his training in Clinical Epidemiology at Erasmus University in the Netherlands in 1995. He currently occupies the Medihelp Chair in Clinical Epidemiology at the University of Pretoria, aiming to enhance patient-oriented research. His main research interests are in diabetes, cardiovascular disease and evidence-based care.

James Ker, a Specialist Physician trained at the University of Pretoria, has been actively involved for more than two decades in clinical teaching at the bedside, with special interest in using evidence-based methodology. He has also been in private practice as a Specialist Physician for 6 years.

COMPARABILITY OF CARE

Care between the two groups should be comparable, e.g. same time schedule and intensity of care. This safeguards patients benefiting from any intervention other than the one being studied.

INTENTION TO TREAT ANALYSIS

Patients should be analysed in the groups that they were initially randomised to. Failure to do so results in loss of the randomisation effect, and bias occurs. Many studies report both an intention to treat analysis and the results for those only complying with the given treatment.

PRECISION AND INTERPRETATION OF RESULTS

In an intervention study the outcome of interest is the reduction of one or more events (outcomes) and all important outcomes should be considered. The reader must ascertain if the outcome represents a hard end-point such as death or stroke rather than a surrogate end-point such as improvement in a certain physiological parameter, e.g. lipid level. Risk reduction can be defined in relative and absolute terms.[4]

Relative risk reduction (RRR) = risk in control group minus risk in the active treatment group divided by risk in control group.

HYPOTHETICAL EXAMPLE

Myocardial infarction (MI) rate on placebo 40% and MI rate on aspirin 20%:

$$(\text{RRR}) = \frac{40 - 20}{40} = 0.5 \times 100 = 50\%.$$

This risk reduction should be accompanied by a confidence interval (CI), which is an estimate of the precision of the risk reduction. The width of this CI reflects the sample size and the degree of uncertainty around this estimate. The RRR of 50% may seem very impressive; usually a RRR of ≥ 25% is clinically relevant.

The true effect, however, is very much dependent on the baseline risk. For example, if the MI rate on placebo was 0.4% and on aspirin 0.2%, then the RRR would still be 50%. The number of events in this last example is so low that implementing this therapy on large numbers of people in order to prevent a rare event is questionable.

The concept of absolute risk reduction (ARR) is easier to interpret from a clinical point of view: ARR = risk in the control group minus risk in the active treatment group = 40% − 20% = 20%. The RRR is 50%, yet the ARR is only 20%. ARR can then also be translated into a very useful concept; namely numbers needed to treat (NNT):[4]

$$\text{NNT} = 1/\text{ARR} = \frac{1}{0.2} = \frac{10}{2} = 5.$$

In this instance it means that 5 patients have to be treated in order to prevent one event.

There is no magical NNT figure above which treatment is worthwhile or cost effective. This is still left to the judgement of the physician and will depend on resources available.

EVALUATING AN ARTICLE ON THERAPY (FIG. 1)

Useful guidelines from McMaster University can be found on the Internet (**http://hiru.hirunet.mcmaster.co/ebm/userguid/default.htm**).

When reading a journal one approach would be:

Question 1. Does the particular disease and its therapy interest me?

Question 2. Are the results of the study valid (equivalence v. placebo, randomisation, blinding, comparative care, follow-up, intention to treat analysis).

Question 3. What is the treatment effect (and the uncertainty surrounding it) in relative and absolute terms as well as NNT?

EVALUATING AN ARTICLE ON THERAPY

1. Is the study valid?

a) Study design issues

- Randomised ☐
- Concealment of treatment allocation ☐
- Blinded
  - Single ☐
  - Double ☐
  - PROBE ☐
- Baseline comparability ☐
- Equivalence of care ☐

b) Statistical issues

- Adequate power and sample size ☐
- Equivalence or superiority study ☐
- Intention to treat analysis ☐

2. Are the results statistically and clinically significant?

Out-come	RRR (95% CI)	ARR (95% CI)	NNT (95% CI)

3. Can I apply these results in my practice?

- All relevant end-points assessed? ☐
- Are my patient(s) similar? ☐
- Is their risk similar or greater? ☐
- Do I need special expertise or resources? ☐
- Would benefit exceed possible harm? ☐
- What will my patient(s) prefer? ☐

Fig. 1. Critical appraisal worksheet.

Question 4. Given these results in a valid study, how applicable are they to my clinical setting (the profile of my patients, and their preferences, resources).

The last question is very important. Many practice- and patient-related factors will determine the ultimate use or non-use of a certain therapy. Patient preference is a vital component of evidence-based medicine.[5] In a study of anti-thrombotic treatment for atrial fibrillation, 20 out of 100 patients at risk for thrombo-embolism declined warfarin. Patients declining warfarin were inclined to seek a higher level of benefit than those taking it, as measured by the minimal clinically important difference.[6]

Keeping up with the medical literature demands critical appraisal skills of the practising physician. This brief article and the evaluation worksheet constructed by the authors (Fig. 1) could aid in developing appraisal of the validity of interventional studies.

1. Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous methods. *BMJ* 1996; **313:** 36-39.
2. Schultz K, Chalmers I, Hayes R, Altman D. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273:** 408-412.
3. Hansson L, Zanchetti A, Carruthers SG, *et al.* for the HOT study group. Effects of intensive blood-pressure lowering and low-dose aspirin in patients with hypertension: principal results of the Hypertension Optimal Treatment (HOT) randomised trial. *Lancet* 1998; **351:** 1755-1762.
4. Sacket DL, Scott Richardson W, Rosenberg W, Haynes RB. Evidence-based medicine. How to practice and teach EBM. Churchill Livingstone, 1997.
5. Haynes RB, Sackett DL, Gray JMA, Cook DJ. Guyatt GH. Transferring evidence from research into practice (Editorial). *ACP J Club* 1996; **125:** A-14.
6. Howitt A, Armstrong D. Implementing evidence based medicine in general practice: audit and qualitative study of antithrombotic treatment for atrial fibrillation. *BMJ* 1999; **318:** 1324-1327.

CASE REPORT

UNRECOGNISED ACUTE RENAL FAILURE FOLLOWING THE COMRADES MARATHON

Roal van Zyl-Smit, Phillip Mills, Louis Vogelpoel

A 40-year-old experienced and well-trained marathon runner attempted his second Comrades Marathon (a 90 km ultramarathon). He felt perfectly fit and well before the race and, as was his custom, took one 2 mg loperamide (Imodium) tablet before the race as prophylaxis against diarrhoea. During the course of the race he took 6 paracetamol 500 mg/chlormezanone 100 mg (Bessenol) tablets for the expected aches and pains. He completed the race over a period of 9 hours, having taken what appears to have been adequate volumes of fluids. The discomfort experienced during the race was similar to that he had experienced during previous races.

On the day following the race he experienced bilateral loin pain for which he took another 2 Bessenol tablets, and during the following 5 days he used a total of 11 Stopayne tablets (paracetamol/caffeine/codeine phosphate/meprobamate). During the entire period following the race he had mild muscle aches and pains and did not notice any diminution in urine output.

Nine days after the marathon the patient attended a social function and was noted to have an elevated jugular venous pressure by an astute general physician also present at the function. A medical consultation was arranged for the following day.

The patient was noted to be generally well but with a blood pressure of 200/140 mmHg, pulse 60/min, heart clinically normal, bilateral basal pulmonary crackles and mild tenderness over both renal angles. His urine contained 2+ protein, 1+ blood and moderate numbers of granular casts. Plasma creatinine was 713 μmol/l, urea 32.7 mmol/l, potassium 5.5 mmol/l, lactic dehydrogenase (LDH) 568 U/l (normal range 170 - 350 U/l), alanine aminotransferase (ALT) 159 U/l (normal range 1 - 25 U/l), gamma-glutamyltransferase (GGT) 175 U/l (normal range 0 - 65 U/l), serum calcium 2.3 mmol/l, inorganic phosphate 2.08 mmol/l and serum urate 0.68 mmol/l (normal

Roal van Zyl-Smit is a consultant nephrologist at Groote Schuur Hospital, and has completed two Comrades Marathons. Louis Vogelpoel is the astute physician who noted the raised venous pressure in the case described. Philip Mills, the medical registrar involved in handling the case, is now a consultant cardiologist.