Interpreting risks and ratios in therapy trials

Ian Scott, Director of Internal Medicine and Clinical Epidemiology, Princess Alexandra Hospital, and Associate Professor of Medicine, University of Queensland, Brisbane

Summary

To appreciate the significance of clinical trial results, clinicians need to understand the mathematical language used to describe treatment effects. When comparing intervention and control groups in a trial, results may be reported in terms of relative or absolute risk (or probability), or as more statistically sophisticated entities based on odds and hazard ratios. When events in the intervention group are significantly less frequent than in the control group, then relative risk, odds ratio and hazard ratio (and their confidence intervals) will be less than 1.0. If the converse holds true, these values will be greater than 1.0.

Key words: clinical trials, number needed to treat, odds, statistics.

Introduction

In randomised trials and systematic reviews of trials, the effects of new treatments on dichotomous outcomes (such as death vs survival) can be expressed in several ways including relative risk, absolute risk, odds ratio and hazard ratio. These figures help to determine if the new treatment has an advantage over other treatments or placebo.

Ways of expressing treatment effects

The absolute risk, number needed to treat, relative risk and odds ratio can be calculated by compiling a 2x2 table of study data. Values can then be derived using the equations shown in the box.

Absolute risk

Absolute risk reduction, also termed risk difference, is the difference between the absolute risk of an event in the intervention group and the absolute risk in the control group.

In a trial of 441 patients at risk of developing pressure ulcers, patients were randomised to receive a sheepskin mattress overlay (intervention group) or usual treatment (control group) during their hospital stay.[1] The data from the trial can be represented in a 2x2 table (see Table 1).

Box
Calculations

	Bad outcome	Good outcome	Total patients
Intervention group	a	c	$a+c$
Control group	b	d	$b+d$

Measure	Equation
Absolute risk	$\left(\dfrac{b}{b+d}\right) - \left(\dfrac{a}{a+c}\right)$
Number needed to treat	$\dfrac{1}{\left(\dfrac{b}{b+d}\right) - \left(\dfrac{a}{a+c}\right)}$
Relative risk	$\left(\dfrac{a}{a+c}\right) \div \left(\dfrac{b}{b+d}\right)$
Odds ratio	$\left(\dfrac{\frac{a}{a+c}}{\frac{c}{a+c}}\right) \div \left(\dfrac{\frac{b}{b+d}}{\frac{d}{b+d}}\right) = \left(\dfrac{a}{c}\right) \div \left(\dfrac{b}{d}\right)$

Table 1
Trial data

	Patients with ulcer(s)	Patients with no ulcers	Total patients
Sheepskin group	21	197	218
Control group	37	186	223

The absolute risk reduction can then be calculated by subtracting the proportion of patients with ulcers in the sheepskin group from that in the control group.

$$\left(\frac{37}{223}\right) - \left(\frac{21}{218}\right) = 0.07 \text{ (or 7.0\%)}$$

Almost 17% of patients in the control group developed ulcers compared to 10% in the sheepskin group after 20 days of observation. This means that the absolute risk of developing ulcers in the sheepskin group was 7% less than in the control group.

If a treatment is effective and reduces the risk of an unwanted event, we see an absolute risk reduction. Conversely, if the treatment does not work and in fact increases the risk of the event, then we see an absolute risk increase.

It may be difficult to conceptualise the clinical relevance of the absolute risk reduction. The reciprocal of this value (1/absolute risk reduction) gives the number of patients who need to be treated for a certain period of time to prevent one event.

This is termed the number needed to treat and can be useful for comparing the effectiveness of a number of different interventions. So in the ulcer trial, 14 patients need to have a sheepskin overlay for 20 days to prevent one of them from getting an ulcer.

It is important to appreciate that absolute risk will vary according to the event rates in both patient groups, whereas the relative risk usually remains unchanged across the spectrum of disease severity (see Table 2). Putting this another way, in 'low risk' patients (those with mild hypertension in Table 2) the absolute risk reduction will be small whereas in 'high risk' patients (those with moderate hypertension) absolute risk reduction will be larger. For both groups the relative risk (and relative risk reduction) is the same.[2]

Relative risk

Relative risk, also known as risk ratio, is the risk of an event in the experimental group divided by that in the control group. For the sheepskin trial, this can be calculated from the data in Table 1.

$$\left(\frac{21}{218}\right) \div \left(\frac{37}{223}\right) = 0.58$$

In the trial, 10% of patients in the sheepskin group developed ulcers compared to 17% in the control group. So the risk of getting ulcers with a sheepskin overlay was 0.58 of that in the control group.

In most trials where the treatment intends to prevent an undesirable outcome such as death or complication (prevention trials), efficacy will be denoted by a relative risk of less than 1.0. Treatment harm, reflecting an increased risk of an event (including adverse effect), will be denoted by a relative risk of more than 1.0. However, in trials where the treatment intends to reduce active disease (treatment trials) and promote a positive event, such as disease remission or symptom abatement, a relative risk of more than 1.0 confirms treatment efficacy. A relative risk of 1.0 indicates no difference between comparison groups. In all cases, statistical significance is assumed if the 95% confidence interval (CI) around the relative risk does not include 1.0.

The relative risk reduction equals the amount by which the relative risk has been reduced by treatment and is calculated as 1 – relative risk. For example in the sheepskin trial, sheepskin overlays reduced the risk of patients getting ulcers by 0.42 (1 – 0.58) or 42%.

Odds ratio

Odds are the number of times an event happens divided by the number of times it does not within a group. Odds can also be expressed as the risk (or probability) of an event occurring over the risk of an event not occurring. To provide a numerical example: if 1/5 of the patients in a study suffer a stroke, the odds of their having a stroke is (1/5) ÷ (4/5) or 0.20/0.80, or 0.25. As the denominator is the same in both top and bottom expressions, it cancels out, leaving the number of patients with the event (1) divided by the number of patients without the event (4).

The odds ratio is the odds of an event occurring in one group divided by the odds of the same event in another group. In the sheepskin trial, the odds ratio can be calculated by dividing the odds of getting an ulcer in the sheepskin group by the odds in the control group.

$$\left(\frac{21}{197}\right) \div \left(\frac{37}{186}\right) = 0.54$$

The odds were about 0.11 in the sheepskin group and 0.20 in the control group. This means that the odds of developing an ulcer in the sheepskin group were 0.54 of that in the control group. Put another way, patients with a sheepskin overlay were half as likely to develop ulcers as patients given usual treatment.

Odds ratio is similar to relative risk. In the sheepskin trial the relative risk was 0.58 and the odds ratio was 0.54. For most clinical trials where the event rate is low, that is less than 10%

Table 2

Relation between relative risk, absolute risk and odds ratio [2]

In an overview of randomised controlled trials of hypertension management, rates of stroke were measured in patients randomised to receive the experimental treatment or control. Results were analysed according to the severity of hypertension.

Disease severity	Event rate in control group (or AR)	Event rate in experimental group (or AR)	RR (RRR)	ARR	NNT	OR
Moderate hypertension	20%	12%	0.60 (0.40)	8%	13	0.54
Mild hypertension	1.5%	0.9%	0.60 (0.40)	0.6%	167	0.60

AR	absolute risk		ARR	absolute risk reduction
RR	relative risk		NNT	number needed to treat to prevent one stroke
RRR	relative risk reduction		OR	odds ratio

Hazard ratio and time-to-event analysis [1]

In a randomised controlled trial, 441 patients assessed on admission as having low to moderate risk of developing pressure ulcers were randomised to receive a sheepskin mattress overlay for the duration of hospital stay or usual treatment (control group) as determined by ward staff. Patients were followed for up to 20 days after randomisation and assessed daily for the onset of pressure ulcers. The results were reported as follows:

	Sheepskin group	Control group
a. Total number of patients	218	223
b. Total number of bed days observed	1728	1561
c. Total number of ulcers	27	58
d. Number of patients with ulcer(s)	21	37
e. Mean bed days per patient	7.9	7.0
f. Cumulative incidence risk (95% CI)	9.6% (6.1%–14.3%)	16.6% (12.0%–22.1%)
g. Relative risk	0.58 (0.35–0.96)	1.0 (referent group)
h. Incidence rate per 100 bed days (95% CI)	1.6 (1.0–2.3)	3.7 (2.8–4.8)
i. Incidence rate ratio (95% CI)	0.42 (0.26–0.67)	1 (referent group)
j. Hazard ratio	0.39 (0.22–0.69)	1 (referent group)

CI   confidence interval

Cumulative incidence risk (f) is the total number of patients who developed one or more ulcers (d)/number of patients for each group (a).

Relative risk (or risk ratio) (g) is the ratio of cumulative incidence risk (f) in sheepskin vs control group (9.6%/16.6% = 0.58).

Incidence rate (h) per 100 bed days is the total number of ulcers (c)/total number of bed days observed (b).

Incidence rate ratio (i) is the ratio of incidence rate per 100 bed days (h) in sheepskin vs control group (1.6/3.7 = 0.42).

Hazard ratio (j) is estimated using Cox proportional hazards regression applied to Kaplan-Meier time-to-event curves for ulcer-free survival (Fig. 1).

of all participants have an event, the odds ratio and relative risk can be considered interchangeable. The relative risk and odds ratio will also be closer together when the treatment effect is small (that is, odds ratio and relative risk are close to 1) than when treatment effect is large. However, as the event rate increases above 15% or as the treatment effect becomes huge, the odds ratio will progressively diverge from the relative risk.

Fortunately, this is rarely a problem. Consider a meta-analysis of ligation versus sclerotherapy for oesophageal varices, which demonstrated a re-bleeding rate of 47% with sclerotherapy, as high an event rate as one is likely to find in most trials.[3] The odds ratio associated with treatment with ligation was 0.52, a large effect. Despite the high event rate and large effect, the relative risk was 0.60, not very different from the odds ratio. Thus choosing one measure or the other is unlikely to have an important influence on most treatment decisions.

The odds ratio is gradually losing favour as a measure of treatment effect[4], particularly as data from which relative risk is derived can also be used to calculate absolute risk reduction and number needed to treat, which are more clinically useful.

Hazard ratio

Hazard ratio is a measure of relative risk over time in circumstances where we are interested not only in the total number of events, but in their timing as well. The event of interest may be death or it may be a non-fatal event such as readmission or symptom change.

Table 3 shows results of the study on pressure ulcers in hospitalised patients.[1] Results were expressed in several ways including:

■ relative risk (row g), which is based on comparing the proportions of patients between groups who developed ulcers by study end (which the authors of the study termed cumulative incidence risk)

■ incidence rate ratio (row i), which is a time-dependent relative risk comparing the rates of ulcers over time (in this case, per 100 bed days) between groups.

Note that the relative risk and the incidence rate ratio were different, 0.58 versus 0.42, with the time-dependent relative risk suggesting a greater benefit from intervention than the overall relative risk, and which is also fairly close to the estimated hazard ratio of 0.39 (row j).

In contrast to the overall relative risk, both the time-dependent relative risk and hazard ratio take into account the timing of events which may not be evenly distributed throughout the study period.

The hazard ratio equals a weighted relative risk over the entire

duration of a study and is derived from a time-to-event curve or Kaplan-Meier curve. This curve describes the status of both patient groups at different time points after a defined starting point. In the sheepskin study, events in the intervention group are not only less frequent overall than in the control group but they are delayed in time (Fig. 1). As some patients will be followed for a longer period of time than others (because they were recruited or randomised into the trial at an earlier time or because they remained in the study while others dropped out), the time-to-event curve usually extends beyond the mean follow-up duration.

As the trial progresses, at some point prediction of treatment effect becomes very imprecise (in our example at 20 days) because there are few patients available to estimate the probability of the outcome of interest. Confidence intervals around the survival curves would capture the precision of the estimate. Ideally then, we would estimate relative risk by applying an average, weighted for the number of patients available, over the entire study duration. Statistical methods allow just such an estimate which is the hazard ratio.

This derived (or 'crude') hazard ratio then needs to be 'adjusted' or corrected for differences in the two groups at baseline that might influence the outcome of interest. This issue is less of a concern if randomisation has rendered both groups similar in terms of their baseline characteristics. In our example, patients in the intervention group compared to control were older (mean age 63.2 years vs 61.1 years), more acutely ill (51% were emergency admissions vs 43%), and had greater prevalence of medical, as opposed to surgical, diagnoses (35% vs 27%). Applying the Cox proportional hazards regression model produces an adjusted hazard ratio which takes account of such imbalances.

In every other way the hazard ratio is similar to odds ratio and relative risk wherein treatment efficacy is denoted by a hazard ratio of less than 1.0 in prevention trials and a hazard ratio of more than 1.0 in treatment trials.

Statistical significance

If there is a statistically significant difference in outcomes between treatment and control groups, the observed difference is very unlikely to have occurred due to the play of chance, even after accounting for imprecision in the difference related to the total number of events in both groups.

P values

Statistical significance is defined arbitrarily in terms of a p value of less than 0.05. The p value however does not directly indicate the chance of an effect being present or not being present. Instead it tells us how often chance alone would give apparently favourable results. A p value of less than 0.05 tells us that there is less than 5% probability that chance alone would lead to such

Fig. 1

Kaplan-Meier curve for time to onset of first pressure ulcer *

| Sheepskin: | 218 | 141 | 49 | 18 | 6 |
| Referent: | 223 | 147 | 26 | 4 | 2 |
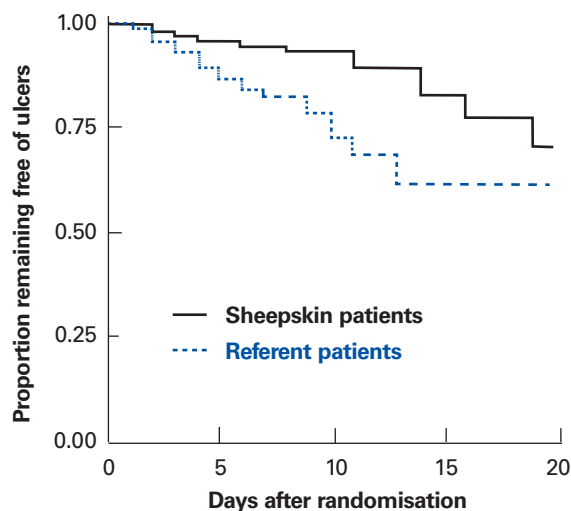
*   Jolley DJ et al. Preventing pressure ulcers with the Australian medical sheepskin: an open-label randomised controlled trial. MJA 2004;180:324–327. ©Copyright 2004. The Medical Journal of Australia – reproduced with permission.

Kaplan-Meier estimates show the time to onset of first pressure ulcer in 441 hospitalised patients at risk of developing pressure ulcers. Patients were randomised to receive either a sheepskin mattress overlay or usual treatment (referent group).

Predicting the effect of the sheepskin intervention becomes very imprecise as the number of patients in each group decreases with time.

favourable results, but it says nothing directly about whether chance is the best explanation for the results.

Confidence intervals

Confidence intervals give us an estimate of the precision of the results. Conventionally 95% confidence intervals are used which, if the same trial were to be repeated many times over, define the range of values within which the true estimate would be found in 95% of occasions. The confidence interval represents the range of values within which we are 95% confident that the true population estimate lies. If the number of events such as death occurring over time is fairly small (as occurs with small samples and/or low case fatality rate), then the precision with which the true probability of the event can be estimated is relatively low, as reflected in wider confidence intervals. Narrower confidence intervals indicate more precise results. The 95% confidence intervals represent almost two standard deviations around the mean.

It is important to remember that the result is statistically significant if the confidence intervals do not cross the null value, such as 1.0 for relative risk and 0 for absolute risk reduction.

Conclusion

An understanding of the commonly used statistical measures of benefit is necessary if clinicians are to gain an appreciation of the efficacy of different therapies. For the majority of clinical trials, relative risk and odds ratio can be considered interchangeable as a measure of the relative change in the risk of a preventable event. The hazard ratio is a related measure that weights the risk change according to when events occur over time. Absolute risk reduction represents the absolute change in risk (expressed in percentage points) and its reciprocal represents the number of patients who would need to be treated over a given period of time to prevent one event.

References

1. Jolley DJ, Wright R, McGowan S, Hickey MB, Campbell DA, Sinclair RD, et al. Preventing pressure ulcers with the Australian medical sheepskin: an open-label randomised controlled trial. Med J Aust 2004;180:324-7.
2. Collins R, Peto R, MacMahon S, Hebert P, Fiebach NH, Eberlein KA, et al. Blood pressure, stroke, and coronary heart disease. Part 2, Short-term reductions in blood pressure: overview of randomised drug trials in their epidemiological context. Lancet 1990;335:827-38.
3. Laine L, Cook D. Endoscopic ligation compared with sclerotherapy for treatment of esophageal variceal bleeding. A meta-analysis. Ann Intern Med 1995;123:280-7.
4. Sackett DL, Deeks JJ, Altman DG. Down with odds ratios! Evid Based Med 1996;1:164-6.

Conflict of interest: none declared

On the correct use of eye drops

Michael Steiner, Eye Surgeon, Sydney

Summary

Drops are a common vehicle for administering drugs to the eye, but they must be instilled correctly. To limit wastage and systemic absorption a single drop should usually be prescribed. If the patient needs to use two types of drop their instillation should be separated by at least three minutes. Most eye drops contain a preservative, but they should not be kept beyond the expiry date on the label.

Key words: expiry dates, instillation, ophthalmic solutions.

(Aust Prescr 2008;31:16–17)

Introduction

Patients should be instructed on how to use their eye drops. They need to know about the frequency and the method of administration, and how the drops should be stored.

One drop or two?

Only one drop should be used at a time. A second drop may wash out the first or increase the possibility of systemic absorption and toxicity. A second drop can often end up on the skin of the eyelids and the patient is then more likely to develop a contact allergy. Using two drops also doubles the cost of the medication.

How often?

The type of drug and the patient's condition determine the frequency of instillation. In some serious infective or inflammatory conditions the drops may need to be used as frequently as half hourly (although generally only while the patient is awake). In contrast, the most commonly used treatments for glaucoma only need to be instilled once a day.

How to use eye drops

The method of instilling the drops is important. If it is not done properly, the drops have almost as much chance of landing on the cheek as in the eye.

It is important that patients wash their hands and remove any contact lenses before using the drops. Many eye drops contain the drug in suspension rather than in solution. These drops should always be shaken before use.

The cap should be removed from the bottle but never put down on the table in such a way that it may become contaminated. It should either be put on its side or held carefully in the other hand.

During instillation it is very important that patients do not touch their eye with the tip of the bottle. This could both abrade the cornea and contaminate the remaining drops.

In the traditional method of instilling drops (see Fig. 1) the bottle is held upside down in one hand between the thumb and index finger and with the other hand the lower eyelid is gently pulled