Interpretation of dichotomous outcomes: risk, odds, risk ratios, odds ratios and number needed to treat

Introduction

Clinical research often investigates whether a patient’s diagnosis, prognostic characteristics or treatment are related to his or her clinical outcomes. The outcomes may be continuous measures, which can occupy any point on a scale, for example: pain intensity on a Visual Analogue Scale (0 to 100) or body temperature in degrees. Other outcomes are measured on ordinal scales, where the differences between adjacent categories may vary along the scale, for example: manual muscle testing grades of none, trace, poor, fair, good and normal. Some continuous outcomes may not be strictly linear (because each step on the scale may not have exactly the same magnitude), but are often treated as such, for example: the Berg Balance Scale (0 to 56 points). In contrast, it is generally clear whether an outcome is dichotomous because it only has two states (such as yes/no, better/not better, or dead/alive). Dichotomous outcomes can also be derived from continuous scales (eg, categorising temperatures as febrile/not febrile) or from ordinal scales (eg, categorising muscle contraction as present/absent). Such dichotomisation is generally not recommended, primarily due to loss of information. However, if the threshold used is nominated a priori and based on a clinically relevant point on the scale, then dichotomisation of a continuous outcome may be appropriate.

The reporting and interpretation of studies with dichotomous outcomes can be challenging. There are several ways to report the findings about dichotomous outcomes, including: proportions, percentages, risk, odds, risk ratios, odds ratios, number needed to treat, likelihood ratios, sensitivity, specificity, and pre-test and post-test probability – each of which has a different meaning. The purpose of this two-part series is to describe the correct interpretation of commonly used methods of reporting dichotomous outcomes. This first paper focuses on risk, odds, risk ratio, odds ratio and number needed to treat. The second paper will focus on sensitivity, specificity and likelihood ratios.

Risk and odds

The chance of a dichotomous outcome occurring can be described in terms of risk or odds. While these terms may sound similar, they have different meanings and methods of calculation.

Risk

Risk is defined as the probability of an event during a specified period of time. In clinical studies, risk is often calculated as the number of people in whom an outcome occurs divided by the total number of people assessed. There are various ways to express risk. The raw count of a dichotomous outcome can be summarised as a proportion or a percentage. For example, in an observational study of 96 independent ambulatory older people, 59 had a fall during a 1-year period. This could be reported as a proportion (59/96 = 0.61) or as a percentage (61%). If the outcome was recorded as having occurred during a specific period of observation, as in this example, this measure of risk is called an incidence (Strictly speaking, this is an incidence proportion, ie, the proportion of people in whom the outcome occurs during a specified period of time. An incidence rate measures the occurrence of the outcome per unit of person-time). Although the outcome may be positive (eg, recovery) or negative (eg, falling), we still talk about the ‘risk’ of that outcome. If the outcome of interest was recorded at a single time point, such as the number of people in a community sample who have neck pain, this measure is called a prevalence.

Odds

Odds are calculated differently from risk and are defined as the number of people in whom the outcome (eg, a fall) occurred divided by the number of people in whom the outcome did not occur. The odds of falling (using the same data from Maki et al above) would be 59/37 or 1.59. Note that the odds and the risk for the same data are not the same (Box 1), because their denominators are different.

Comparing risk or odds between groups

The example above from Maki et al describes the risk or odds of an event (a fall) in one population (independent ambulatory older people). However, many clinical studies investigate the difference in risk or odds for a dichotomous outcome based on certain patient characteristics or exposures. For example, we may want to know if older people fall more often than younger people, or if trial participants who were randomised to receive a Tai Chi intervention fell less often than those who were randomised to no treatment. This is commonly performed by calculating the ratio of either the risk (risk ratio) or the odds (odds ratio) between the groups of interest.

A risk ratio (also known as relative risk) is calculated by dividing the risk of the outcome (eg, a fall) in people with a characteristic or exposure (eg, received a Tai Chi intervention) by the risk of the same outcome in the people who do not have that characteristic or exposure (eg, control group). In a similar way, odds ratios are calculated by dividing the odds of the outcome in people who have a particular characteristic by the odds of the same outcome in people who do not have that characteristic. A worked example for calculating both risk ratios and odds ratios can be seen in Box 1. It is critical to note that the value of the relative risk and the odds ratio are different, despite being based on the same data, and therefore the interpretation is also different.

It is useful to recognise that risk ratios and odds ratios can be greater than or less than 1. When greater than 1, this indicates that people with the clinical characteristic (positive test result, prognostic characteristic or treatment exposure) are more likely to have the outcome of interest (diagnosis, prognostic outcome or treatment outcome) than people without that clinical characteristic. When risk ratios or odds ratios are less than 1, this indicates...
that these people with the clinical characteristic are less likely to have the outcome of interest.

The possible range of odds ratios is from 0 to infinity. In contrast, the range of risk ratios starts at 0 but has an upper limit that depends on the risk in the reference group (the people who do not have the clinical characteristic). For example, if the risk in the reference group is 50%, the maximum value that a relative risk can have is 2.0, because this would indicate that 100% of people with the clinical characteristic have the outcome of interest.

### Interpretation of risk ratios

Risk ratios are relatively easy and intuitive to interpret. For example, a study investigating risk factors for non-contact anterior cruciate injury reported a risk ratio of 2.8 for the presence of generalised joint laxity. A correct interpretation would be to say that the people with generalised joint laxity had 2.8 times the risk of having an anterior cruciate injury compared with those without generalised joint laxity. As mentioned previously, risk ratios of less than 1 mean the risk of the outcome is less in those with the characteristic than in those without. For example, Kiely et al investigated the risk of stroke among people with moderate-to-high physical activity and compared it with the risk among people with low physical activity. They reported an risk ratio of 0.8. This was correctly interpreted as: moderate-to-high physical activity was associated with a relative reduction in the risk of stroke by 20% compared with low physical activity.

### Absolute risk reduction and relative risk reduction

To correctly interpret the clinical importance of risk ratios, it is critical to understand that the absolute risk reduction (also called...
absolute risk difference, because the change in risk is not always a reduction) and relative risk reduction mean different things. In the stroke example above, the relative risk was 0.8, indicating a 20% relative reduction in risk associated with moderate-to-high physical activity. How much this reduces the absolute risk of stroke is dependent on the baseline risk. If the risk of stroke in the low physical activity participants were 10%, an absolute risk of 0.8 would mean that the risk of a stroke in those who performed moderate-to-high physical activity would be 8% (0.10 x 0.8). This means the absolute risk reduction was 2% (ie, from 10% down to 8%).

If, however, the risk of stroke in the low physical activity participants were 1%, then the same relative risk would only result in an absolute risk reduction of 0.2%. This demonstrates the importance of distinguishing the relative risk reduction and the absolute risk reduction. The baseline risk is very important to the absolute risk reduction but not considered in relative risk reduction. The same risk ratio (eg, 0.8) will have a greater impact on absolute risk reduction, the more common the outcome is.

Number needed to treat

Another statistic that is used to help in the interpretation of risk ratios is the number needed to treat (NNT), which is a simple way of understanding how many patients need to be treated for one patient, on average, to benefit. For example, let us imagine that we conducted a clinical trial to determine if participants who were randomised to receive a Tai Chi intervention fell less often than those who were randomised to no treatment. If the incidence of falling was 30% in the Tai Chi group and 60% in the no-treatment group, the absolute risk reduction would be 30% and the ratio would be 0.5. The formula for the NNT is 100%/absolute risk reduction. So, in our example, this would be 100%/30% = 3.3. This means that 3.3 patients would need to be treated with the Tai Chi intervention to prevent one fall. Because it is not possible to treat a fraction of a patient, the NNT may be reported with the decimal places rounded off (in this example, to 3) or conservatively rounded up to the next highest whole number of participants (in this example, to 4).

It is worth noting that, while the treatment was very effective (it halved the rate of falls), the NNT was affected by both the effectiveness of the treatment and the risk in the reference group (the no-treatment group). So, if the risk in the no-treatment group had been only 10%, a treatment with the same relative risk would have an NNT of 20 (100%/absolute risk reduction = 100/5 = 20). The reason for this is that most people (nine out of 10) had no fall during the follow-up period, so even a very effective treatment would need to be given to many people in that population before it prevented a fall.

Interpretation of odds ratios

Odds ratios are somewhat more difficult and less intuitive to interpret than risk ratios. An odds ratio of 2 from a randomised, controlled trial means that the intervention doubled the odds of the outcome occurring compared with the control group. However, because we tend not to think in terms of odds, this is not easy to interpret. Importantly, an odds ratio of 2 is not the same as an risk ratio of 2. For the same study data, the odds ratio will usually be further from 1 than the risk ratio. Therefore, when greater than 1, odds ratios are almost always substantially higher than risk ratios and when less than 1, odds ratios are almost always substantially lower than risk ratios. They only become similar when the incidence or prevalence is very low.17 Holcomb et al18 found that in nearly half of the studies they reviewed, the odds ratio was more than 20% larger than the risk ratio when using the same data. They also found that 26% of the studies they reviewed incorrectly interpreted odds ratios as risk ratios. It could be asked why odds ratios are used if they are often misinterpreted. One explanation is that odds ratios have mathematical properties that make them more easily managed within some statistical procedures. More thorough explanations are available for interested readers.10

Confidence intervals

All of the statistics mentioned above, from basic risk and odds through to NNT, can be calculated with 95% CIs. Briefly, with each statistic, the 95% CI indicates the range of uncertainty around the estimate. An excellent and more detailed explanation of the interpretation of 95% CI for dichotomous measures has previously been presented in this journal.13 With risk ratios or odds ratios, a 95% CI that crosses 1 indicates that the result is not statistically significant, whereas a 95% CI that does not include 1 indicates that the difference between groups can be attributed to the distinguishing treatment, exposure or characteristic between groups. The 95% CI around the NNT is easy to interpret when the result is statistically significant, but it is particularly unintuitive when the result is not statistically significant, so readers are referred to further explanation elsewhere.12

Summary

Statistics that summarise dichotomous outcome measures, including risk ratios, odds ratios, absolute risk reduction and relative risk reduction, are commonly used, but have different meanings. A good understanding of these terms will enable readers of clinical studies to ensure that they correctly interpret the clinical importance of the findings reported.

Acknowledgements: Nil.
Competing interests: Nil.
Provenance: Not commissioned. Peer reviewed.

Mark Hancocka and Peter Kenta,b,c
aFaculty of Medicine and Health Sciences, Macquarie University, Sydney, Australia
bSchool of Physiotherapy and Exercise Science, Curtin University, Perth, Australia
cDepartment of Sports Science and Clinical Biomechanics, University of Southern Denmark, Denmark

References