Debate & Analysis

Against all odds?

Improving the understanding of risk reporting

GPs are increasingly called upon to make or guide patients with choices about medical interventions, but there is gathering evidence that clinicians' understanding of risk is poor and, correspondingly, that their ability to communicate risk is deficient.[1] In this article we aim to improve health professionals' understanding of risk reporting and clarify common misunderstandings in interpreting risk, odds, relative risk (RR), and odds ratios (ORs).

RISK

The risk of an event (or disease) happening is simply the number of times it occurs divided by the number of occasions on which it could potentially occur. It is usually expressed as a proportion or as a percentage. For example, to give the risk of breast cancer is to say that for every eight women, one will develop breast cancer, which is a lifetime risk of 1/8, or 0.13.[2] We might also express this by saying 12% of women develop breast cancer.

This lifetime risk is not the same as the prevalence of breast cancer, which is the total number of cases in a given population at a specific time. Prevalence is dependent on factors including survivability and associated comorbidity, and is closer to 2% in the UK.[2]

Using simple frequency statements — such as 'one in eight women will develop breast cancer' — avoids a potential source of miscommunication with patients. Gigerenzer's work highlights this issue in the description of a psychiatrist who was in the habit of advising patients that, on taking fluoxetine, there was 'a 30–50% chance of developing a sexual problem'.[3] He learned that patients interpreted this as meaning they would encounter difficulties during half their sexual encounters. He had intended to communicate that of every 10 people taking fluoxetine, three to five of them would experience difficulties. This exemplifies an important principle: physicians tend to consider risk in terms of the group of patients that they treat, whereas patients interpret risk as applying to their own individual case.

ODDS

Odds are another way of representing probability; they are used in the betting industry but have been adopted by healthcare researchers with increasing frequency since the 1990s.[4] The odds are the probability of an event occurring, divided by the probability of that event not occurring. For

example, the odds that a single throw of a dice will produce a six are one to five, or 1/5 — in multiple throws of a dice, there will be one six for every five other throws.

In health, odds are the number of people who experience an event, divided by the number of those who do not; the odds of a woman developing breast cancer are 1/7, which is 0.14. However, saying that 0.14 women develop breast cancer for every one that does not, makes little intuitive sense. Indeed, odds are much easier to understand if they are >1 than if they are <1; for example, for every one woman who develops breast cancer, seven will not — the 'odds against' developing breast cancer are seven to one. Unfortunately, as we will see below, ORs are often based on odds of <1.

Whether risk or odds are easier to interpret is subjective, but many authors consider the former to be more intuitive.[5] Health researchers often imply that having an understanding of odds is dependent on having an insight into the betting industry;[4,6] however, the analogy between health statistics and gambling is a perilous one. Odds in health are an accurate representation of frequencies in a specific setting whereas, in sports betting, the bookmakers' odds are a statement of the return on your stake; this is an indication of not only how likely something is to happen, but also how much you will win.[7]

THE RELATIONSHIP BETWEEN RISK AND ODDS

When a condition is rare, risk and odds are approximately equal. For example, in our breast cancer example, the lifetime risk is 1/8 (or 0.13) and the lifetime odds 1/7 (or 0.14). As conditions become rarer, the approximation becomes more exact. The lifetime risk of developing pancreatic cancer is 1/86 (or 0.0116) and its lifetime odds 1/85 (or 0.0117).[2] However, for more common conditions, the approximation is poor: the lifetime risk of developing any cancer is estimated to be 1/3 (or 0.33), but the lifetime odds are 1/2 (or 0.50).[2]

Table 1 not only illustrates how, at low values, the odds and the risk of an event are equivalent — a proportion of 0.01 is the equivalent of odds of 0.01 — but also demonstrates that, as the proportion rises, the odds diverge, such that, above values of 20%, the odds start to bear no relation to the proportions. In addition, risk ranges from

zero to one, whereas odds range from zero (the event will never happen) to infinity (the event will always happen).

COMPARING OUTCOMES IN PROSPECTIVE STUDIES USING RELATIVE RISKS AND ODDS RATIOS

In prospective studies, such as clinical trials and observational studies, groups of participants with different characteristics are followed up to determine whether one group is more prone to a particular outcome than another. The difference in outcomes between an 'exposed group' and an 'unexposed group' (in a clinical trial, the intervention and control groups) can be measured on the risk or the odds scales.

The RR is calculated as the risk in the exposed group divided by the risk in the unexposed group. It tells us how much risk has increased or decreased from an initial level; most authors consider it to be readily understood.[5,8,9] The OR is calculated as the odds in the exposed group divided by the odds in the unexposed group. As many authors have recognised, the OR is more difficult to visualise.[6,9,10] How would we interpret an OR of, say, 0.5 or 3? One suspects that many readers will have no intuitive feel for the size of the difference when expressed in this way.

Both RRs and ORs are >1 when the outcome of interest is more common in the exposed group (an example being a prospective trial showing harm); both are <1 when the outcome is more common in the unexposed group (typically, a prospective trial showing benefit). Confidence intervals (CIs) around each ratio, as well as P-values, can be used to assess whether differences in the

Table 1. Comparison of risk and odds

Risk as percentage (%)	Risk as decimal	Odds as decimal	Odds as fraction
1	0.01	0.01	1/99
5	0.05	0.05	1/19
10	0.10	0.11	1/9
20	0.20	0.25	1/4
30	0.30	0.5	1/2
50	0.50	1	1/1
60	0.60	1.5	6/4
70	0.70	2.3	7/3
80	0.80	4	8/2
90	0.90	9	9/1
100	1.00	Infinity	10/0

two groups in a particular sample reflect an underlying phenomenon or may have arisen by chance.

RR and ORs both depend on timescale — in a prospective study of mortality, no matter how different the two groups, the RR and OR will approach 1 if the study continues until most participants have died. A form of RR that is independent of the study length is the hazard ratio; discussion of this measure is beyond the scope of this article.[11]

SHOULD RELATIVE RISK OR THE ODDS RATIO BE USED?

Use of odds ratio as an approximation to relative risk

Just as risk and odds are close in value in prospective studies when the baseline risk is low, so too are the values for RR and OR — the so-called 'rare disease assumption'[11] (Box 1).[12] This has led to the common practice of allowing readers to assume that the OR is an RR in order to gain an intuitive feel for the size of the effect demonstrated in a study.[6,11] The problem is that, although the two measures will always move in the same direction, their values diverge when baseline risk (or odds) is high in either of the two groups being compared. Moreover, as the baseline prevalence increases, divergence is more pronounced the further the OR is from unity (1) this is illustrated in Figure 1. If the OR is <1, it is always smaller than the RR. Conversely, if the OR is >1, it is always bigger than the RR; it is unsafe to merely interpret an OR as though it were an RR, especially when the rare disease assumption fails, as this tends to exaggerate the apparent benefit or impact being studied.

Do odds ratio and relative risk get confused?

Many columns of print have been devoted to whether, in prospective studies, effects should be described using RR or ORs. The debate stems from the premise that, if the RR is more intuitively understandable than an OR, then, even when baseline risks are high, readers may interpret ORs as risk ratios and, hence, reach an incorrect conclusion.[5,13]

How commonly the two are confused is illustrated by a study of papers published in two journals reporting that 26% of the articles that used an OR interpreted it as a risk ratio.[9] It has even been argued that the effect may be exploited by healthcare researchers hoping to make their findings more dramatic and publishable.[14] For example in a study comparing National Psoriasis Foundation members with non-members, an OR of 24 for having heard of calcipotriene was reported as 'members [being] more than 20-fold more likely to have heard of' the treatment, although the actual RR was 3.5.[15]

When should odds ratio not be interpreted as relative risk?

Altman and colleagues advise that the OR should not be interpreted as an approximate RR unless the events are rare in both groups, 'say, less than 20–30%'.[16] Misinterpretation of the OR as an approximation to the RR is particularly ill-advised when the event rate is high in only one group.[16] A more conservative benchmark of 10% has been suggested as the threshold above which, 'authors, reviewers, and editors should consider presenting and interpreting odds ratios with caution.'[5]

The impact of an intervention (effect size) also affects the discrepancy between RR and OR (Figure 1). Taking both factors into account, an article in *Bandolier* concluded that, 'as both the prevalence (initial risk) and the OR increase, the error in the approximation quickly becomes unacceptable.'[17] As ORs are often reported without the corresponding RR, various authors have offered guidance on when to expect a problematic degree of divergence. Davies asserted that:

'So long as the event rate in both the intervention and control groups is less than 30%, and the effect size no more than moderate (say, a halving or doubling of risk), then interpreting an OR as a RR will overestimate the size of effect by less than one-fifth.'[10,18]

Holcomb *et al* suggested a rule of five:

'If the OR is greater than 1 but no greater than 5, and the occurrence of the outcome in the unexposed group is no greater than 5%, then the OR exceeds the risk ratio by less than 20%.'[9]

Does it matter?

In defence of ORs, Davies has also argued that misinterpretation of them does not materially affect qualitative judgements.[10] To illustrate this point, it was argued that reporting specialist stroke units as reducing adverse outcomes by an OR of 0.66 (instead of by an RR of 0.81), at worst, suggested a reduction in poor outcomes of one-third

Figure 1. Divergence between RR and OR with increase in baseline risk or effect size (risk ratio).
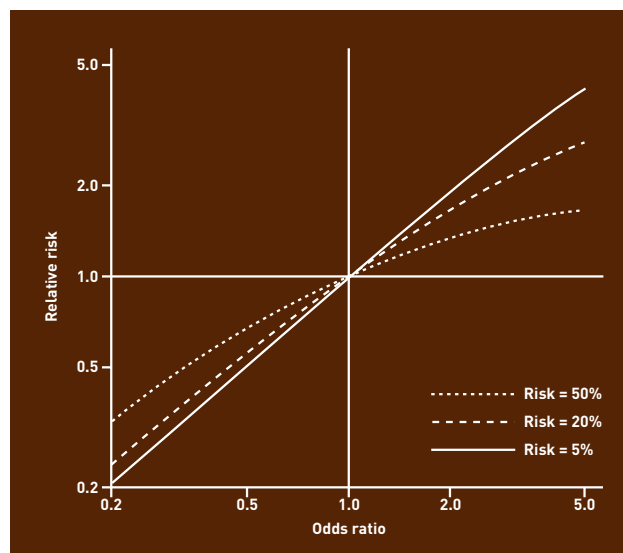
Box 2. How odds ratios are the ideal statistical tool for case–control studies[a]

The following data from a cross-sectional study show the prevalence of hayfever and eczema in 11-year-old children.[23] Should we address the risk (or odds) of hayfever in children with eczema, or the risk (or odds) of eczema in children with hayfever?

	Hayfever	No hayfever	Total
Eczema	141	420	561
No eczema	928	13 525	14 453
Total	1069	13 945	15 014

What is probability that a child with eczema will also have hayfever?
For 11-year olds with eczema, risk of hayfever is 141/561 (25%).
For 11-year olds without eczema, risk of hayfever is 928/14 453 (6.4%).
RR = 25.1/6.4 = 3.9, so a child with eczema is 3.9 times as likely to have hayfever.
OR = 141/420/928/13 525 = 4.9, demonstrating 4.9 greater odds of hayfever if a child has eczema.
What is probability that a child with hayfever will also have eczema?
For 11-year olds with hayfever, risk of eczema is 141/1069 (13.2%).
For 11-year olds without hayfever, risk of eczema is 420/139 459 (3%).
RR = 4.28, so a child with hayfever is 4.38 times as likely to have eczema.
OR = 141/928/420/13 525 = 4.9, demonstrating 4.9 greater odds of eczema if a child has hayfever.
Note that the RR depends on the choice of analysis, but the OR is always 4.9.

For more detail see Bland and Altman.[4] OR = odds ratio. RR = relative risk.[a]

instead of the true reduction of one-fifth, and that this was unlikely to affect the conclusion that stroke units improve outcomes and should be provided. Interestingly, this argument looks less secure now, in current cash-constrained times when cost efficacy and opportunity costs are seemingly at the forefront of most commissioning decisions. When the baseline event rate is high, interpretation of the OR as the RR results in a systematic and important underestimate of the number needed to treat.[8] Moreover, for clinicians to balance the good and bad outcomes of an intervention requires full understanding of the probability of each.[13] Many authorities seem to agree that authors reporting on prospective studies should use the RR or risk reduction in preference to ORs.[5,8,10,13]

Can use of odds ratios in prospective studies be defended?
Some authors defend the OR over the RR.[18,19] Sometimes, it is an arbitrary decision whether or not to treat an outcome, or its converse, as the 'event'. For instance, trials can choose to report the RR of dying, or of surviving. Any such choice affects the RR and OR very differently. For example, in Schulman and colleagues' study of cardiac catheterisation rates and ethnic group, the OR for referral in black versus white ethnic groups is 0.6, and the OR for non-referral is 1.7 (the reciprocal of 0.6).[20] Had the authors reported the RR, they could have quoted an RR of 0.9 for referral or 1.6 for non-referral, which could be translated into a 10% lower referral rate, or 60% higher non-referral rate. Each is correct, yet gives a different impression of the magnitude of any association between ethnic group and referral. ORs circumvent this effect by accounting for 'success' or 'failure' symmetrically (demonstrated by the derivation in Box 2).[21] However, we are not persuaded that this pleasing mathematical property is worth the risk of ORs being misinterpreted as an RR; in particular, in this case, the use of ORs did not prevent misreporting in the popular press.[22]

WHEN ARE ODDS THE STATISTIC OF CHOICE?
There are two settings when it makes sense to use ORs: case–control studies and logistic regression.

In a cross-sectional study, it may be arbitrary which variable is the outcome and which is the exposure. The same OR is obtained either way, for reasons explained by Bland and Altman (Box 2);[4] however, this does not apply to the RR.

A special type of cross-sectional study is the case–control study, in which the percentage of cases does not reflect the risk (baseline prevalence) in the population, but the design chosen by the investigators. Simply varying the number of cases and controls included in the study would vary the value obtained for the calculated risks in the exposed and unexposed groups; these are, therefore, meaningless, as would be the RR. The odds in the exposed and unexposed groups are also meaningless; however, the OR is invariant to the number of subjects recruited, as well as to how we order the categories in the rows and columns.[4]

It has also been argued for case–control studies that, as the outcome of interest is usually rare, the OR also gives a good approximation to the RR.[11] The advantage is that, in this situation, the researcher benefits from the intuitive understanding offered by RR, but gets a value that is not unduly swayed by the study design. For instance, a GP with prior Guillain-Barre syndrome (GBS), trying to decide whether to be vaccinated against H1N1, identified a case–control study reporting an OR of 18.6 (95% CI = 7.5 to 46.4) for the development of GBS in the 2 months following influenza-like illness, and the contrasting finding of an OR of GBS developing after H1N1 vaccination of 0.16 (95% CI = 0.02 to 1.25).[23] An approximate 19-fold risk increase after native infection, compared with no increase or possibly reduced risk after vaccination, enabled a decision to be made.

The second reason for the widespread use of ORs is that, unlike risks, odds can take any value between zero and infinity. This mathematical property is not of interest in itself, but it makes possible logistic regression analyses. These allow researchers to examine the effect of multiple exposure variables on a dichotomous outcome, or the effect of one exposure variable while adjusting for potential confounders.[11] ORs are the natural output of logistic regression software and are, therefore, widely quoted in the scientific literature.

For these reasons the OR, as well as the RR, will continue to be used in clinical research, and users of the medical literature will continue to need to be familiar with both.

ADDRESS FOR CORRESPONDENCE

Christine A'Court
Department of Primary Care Health Sciences, Oxford University, 2nd Floor, 23–38 Hythe Bridge Street, Oxford, OX1 2ET.

E-mail: christine.acourt@btinternet.com

Christine A'Court,
Clinical Researcher, Department of Primary Care Health Sciences, University of Oxford, Oxford.

Richard Stevens,
University Research Lecturer in Statistics, Department of Primary Care Health Sciences, University of Oxford, Oxford.

Carl Heneghan,
Reader in Evidence Based Medicine, Department of Primary Care Health Sciences, University of Oxford, Oxford.

REFERENCES

1. Spiegelhalter DJ. Understanding uncertainty. *Ann Fam Med* 2008; **6(3):** 196–197.

2. Cancer Research UK. http://info.cancerresearchuk.org/ (accessed 17 Jan 2012).

3. Gigerenzer G (ed.). *Reckoning with risk: learning to live with uncertainty.* London: Penguin, 2002.

4. Bland JM, Altman DG. The odds ratio. Statistics notes. *BMJ* 2000; **320(7247):** 1468.

5. Katz KA. The (relative) risks of using odds ratios. *Arch Dermatol* 2006; **142(6):** 761–764.

6. Westergren A, Karlsson S, Andersson P, *et al.* Eating difficulties, need for assisted eating, nutritional status and pressure ulcers in patients admitted for stroke rehabilitation: information point: odds ratio. *J Clin Nurs* 2001; **10(2):** 268–269.

7. Fortune Palace. *Bookies odds explained — fractions, decimals, percentages.* http://www.fortunepalace.co.uk/odds.html (accessed 17 Jan 2012).

8. Bracken MB, Sinclair JC. When can odds ratios mislead? Avoidable systematic error in estimating treatment effects must not be tolerated. *BMJ* 1998; **317(7166):** 1156; author reply 1156–1157.

9. Holcomb WL, Chaiworapongsa T, Luke DA, Burgdorf KD. An odd measure of risk: use and misuse of the odds ratio. *Obstet Gynecol* 2001; **98(4):** 685–688.

10. Davies HT, Crombie IK, Tavakoli M. When can odds ratios mislead? *BMJ* 1998; **316(7136):** 989–991.

11. Altman DG. *Practical statistics for medical research.* London: Chapman & Hall, 1991.

12. A'Court C, Stevens R, Sanders S, *et al.* Type and accuracy of sphygmomanometers in primary care: a cross-sectional observational study. *Br J Gen Pract* 2011 **61(590):** DOI: 10.3399/bjgp11X593884.

13. Deeks J. When can odds ratios mislead? Odds ratios should be used only in case-control studies and logistic regression analyses [letter]. *BMJ* 1998; **317(7166):** 1155–1157.

14. Sibanda T. The trouble with odds ratios. *BMJ* 2003: http://www.bmj.com/content/316/7136/989?tab=responses (accessed 25 Jan 2012).

15. Nijsten T, Rolstad T, Feldman SR, Stern RS. Members of the national psoriasis foundation: more extensive disease and better informed about treatment options. *Arch Dermatol* 2005; **141(1):** 19–26.

16. Altman DG, Deeks J , Sackett DL. Odds ratios should be avoided when events are common [letter]. *BMJ* 1998; **317(7168):** 1318. Erratum in: *BMJ* 1998; **317(7172):** 1505.

17. Deeks J. Swots corner: what is an odds ratio? *Bandolier* 1996; **3(3):** 6–7.

18. Senn S. Rare distinction and common fallacy (rapid response). *BMJ* 1999: http://www.bmj.com/content/317/7168/1318.1?tab=responses (accessed 25 Jan 2012).

19. Cook TD. Advanced statistics: up with odds ratios! A case for odds ratios when outcomes are common. *Acad Emerg Med* 2002; **9(12):** 1430–1434.

20. Schulman KA, Berlin JA, Harless W, *et al.* The effect of race and sex on physicians' recommendations for cardiac catheterisation. *N Engl J Med* 1999; **340(8):** 618–626.

21. Strachen DP, Butlan BK, Anderson HR. Incidence and prognosis of asthma and wheezing illness from early childhood to age 33 in a national British cohort. *BMJ* 1996; **312(7040):** 1195–1199.

22. Verghese A. *Showing doctors their biases.* New York US: New York Times, 1 March 1999. http://www.nytimes.com/1999/03/01/opinion/showing-doctors-their-biases.html (accessed 17 Jan 2012).

23. Price LC. Should I have H1N1 flu vaccination after Guillain-Barre syndrome. *BMJ* 2009; **339:** b3577.