EBM NOTEBOOK

Down with odds ratios!

As described in an earlier EBM note (1) and in our glossary, this journal reports the results of individual randomised trials in terms of relative risk reductions (RRRs), calculated by dividing the absolute difference in event rates between the control (control event rate [CER]) and experimental (experimental event rate [EER]) patients by the event rate for the controls: (CER – EER)/CER = RRR. From these same values, we also report the number of patients that would need to be treated (NNT) to prevent 1 additional event—1/(CER – EER)—or by its alternative calculation—1/(RRR × CER). Thus, in the example shown in Table 1, the RRR is 89% and the NNT is 4 (2).

However, we also report the results of overviews of several randomised trials, and these results appear not as RRRs but as relative odds, or odds ratios (ORs). There are reasons for this variation (although, as it happens, arguably no longer very good ones!). We will explain ORs, point out their properties (many of which interfere with their clinical application), and provide you with some practical help in applying them to individual patients.

When used to summarise an overview, an OR describes the odds of an experimental patient having an adverse event relative to a control patient. We can calculate the odds of a patient having an event by dividing the number of patients who have the event by the number of patients who do not. Hence, for the control group in Table 1, the odds of a patient having the event were $c/d = 9/21 = 0.43$, which compares to a risk of $c/(c + d) = 9/30 = 0.30$. If we mistakenly interpret odds as if they were risks, we will exaggerate the latter, especially with events that are more common.

The OR is calculated by dividing the odds in the experimental group by the odds in the control group—(a/b)/(c/d)—or equivalently through the "cross-products" calculation shown below Table 1—ad/bc. From this definition, it follows that efficacious treatments generate ORs < 1, which is analogous to the relative risk (RR) for the adverse event (EER/CER) being < 1. (We usually prefer to think in terms of RRRs, which are equivalent to 1 – RR, but for ease of comparison with ORs, please bear with us and think in terms of RRs.)

How did we get into this confusing situation of using ORs in the first place? The OR had its origins in case-control studies of drug side effects and of harmful agents and exposures, such as cigarette smoking. In these case-control studies, it is not possible to estimate RRs directly because the prevalence of the adverse outcome (required for calculating the RR) is not usually known. You can, however, calculate the OR in these situations, either by comparing the odds of incurring an adverse event in the exposed group and the control group (i.e., [a/b]/[c/d] = ad/bc) or by comparing the odds of exposure in the event and nonevent groups (i.e., [a/c]/[b/d] = ad/bc); both routes lead to the same answer, which will be > 1 when the exposure is harmful. Hence, the OR can be estimated when the prevalence of the events is unknown, as in most case-control studies. Moreover, because case-control studies typically are used for the study of rare events, the distortion of risk produced by interpreting ORs as if they were RRs is negligible (if necessary, refresh your memory by rereading the 3rd paragraph in this note).

When ORs came into use, several powerful and informative statistical methods were developed (by persons such as Nathan Mantel and William Haenszel) for use in analysing subgroups of patients and combining them (even when the latter were unbalanced for confounding factors) into a single overall estimate (3). Later, when scientists began to do overviews of multiple randomised trials and were seeking a statistical method for combining their results, the analogy with combining subgroups in case-control studies was recognised,

Table 1. Trimethoprim-sulfamethoxazole prophylaxis in cirrhosis*

Treatment group	Adverse event (infectious complications)		Totals
	Occurs	Does not occur	
Experimental (prophylaxis)	1 a	29 b	30 a+b
Control (no prophylaxis)	9 c	21 d	30 c+d
Totals	10 a+c	50 b+d	60 a+b+c+d

*From reference 2.
Control event rate = CER = c/(c+d) = 0.30; experimental event rate = EER = a/(a+b) = 0.033
Control event odds = c/d = 0.43; experimental event odds = a/b = 0.034
Relative risk reduction = RRR = (CER – EER)/CER = 89%
Number needed to treat = NNT = 1/(CER – EER) = 4, also = 1/(RRR × CER) = 4
Relative risk = EER/CER = 0.11, also = 1 – RRR = 0.11
Relative odds = odds ratio = OR = (a/b)/(c/d) = ad/bc = 0.08

and the Mantel-Haenszel method was adapted to this new use (soon joined by a computationally simpler method developed by Richard Peto [4] that provides good approximations to the OR when treatment effects are small and the trials being combined are large and balanced [5]). For these reasons, ORs are now commonly used in the analysis and reporting of overviews of randomised trials that have binary outcomes.

ORs, however, have 5 properties that interfere with their clinical application. First, because very few clinicians are facile at dealing with odds and relative odds, ORs are not useful in their original form at the bedside or in the examining room. Second, in many trials, ORs are not even similar to RRs: In many fields, controlled trials tend to study common adverse events, and it is in these situations that the approximation of the OR to the RR breaks down. Treating an OR as if it were an accurate estimate of the RR will overestimate both the likely benefits and harms of treatment (6), and this distortion becomes greater as the disease being treated becomes more severe and CERs increase.

Third, and as a result of the foregoing, ORs cannot be used in the same simple way as RRRs to calculate the corresponding NNTs for the treatments of interest. To extrapolate results from trials that have different patient expected event rates (PEERs), clinicians need to do separate and complicated calculations of the NNT for each PEER. Although we expect the NNT to decrease as the PEER rises for a treatment with a fixed relative effect, even this is not true for ORs! Looking down a column in Table 2 will show you that, for a fixed OR, the NNT initially decreases as the PEER rises (as expected), but it increases again when the PEER is above 0.5. This counterintuitive result occurs because the difference between the RR and the OR accelerates as event rates rise.

Fourth, when treatments generate a constant RRR for different CERs (e.g., antihypertensive drugs generate the same RRR for stroke among patients with both severe and mild hypertension), their ORs cannot be constant across these CERs (and vice versa). Finally, when clinicians draw up "league tables" of therapeutic efficacy, the order of treatments based on their RRRs may be different from the order based on ORs when the diseases and disorders in the table are of different severity and have different CERs.

Help is on the way (but not quite here yet). Prospects are very good that meta-analyses do not have to be done by using ORs. Statistical meth-

Table 2. Translating odds ratios to numbers needed to treat*

Patient's expected event rate	Odds ratios								
	0.9	0.85	0.8	0.75	0.7	0.65	0.6	0.55	0.5
0.05	209†	139	104	83	69	59	52	46	41‡
0.10	110	73	54	43	36	31	27	24	21
0.20	61	40	30	24	20	17	14	13	11
0.30	46	30	22	18	14	12	10	9	8
0.40	40	26	19	15	12	10	9	8	7
0.50§	38	25	18	14	11	9	8	7	6
0.70	44	28	20	16	13	10	9	7	6
0.90	101¶	64	46	34	27	22	18	15	12‖

* The numbers in the body of the table are the numbers needed to treat (NNTs) for the corresponding odds ratios (ORs) at that particular patient's expected event rate (PEER). To calculate the NNT for any OR and PEER:

$$NNT = \frac{1 - [PEER \times (1 - OR)]}{PEER \times (1 - OR) \times (1 - PEER)}$$

† The relative risk reduction (RRR) is 10%.
‡ The RRR is 49%.
§ For any OR, the NNT is lowest when the PEER = 0.50.
¶ The RRR is 1%.
‖ The RRR is 9%.

ods for combining both relative risks and absolute risk differences across trials are available (7), although some concern exists that they are not appropriate for all circumstances. Validation work is under way to outline the situations where they can be widely adopted for combining randomised trials into systematic reviews.

As soon as these clinically friendlier alternatives are used in reporting the results of overviews, they will appear in *Evidence-Based Medicine*. In the meantime, we will be adding Table 2 to our glossary, which permits our readers to identify NNTs for a range of ORs and PEERs. The intersection of the OR closest to that reported in the overview with the PEER that best represents the reader's patient will identify the corresponding NNT. For readers who want to do the full calculations, the formula appears below Table 2.

David L. Sackett, MD
NHS Centre for Evidence-Based Medicine
Oxford, England, UK
Jonathan J. Deeks, MSc
Douglas G. Altman, BSc
Centre for Statistics in Medicine
Oxford, England, UK

References

1. **Sackett DL.** On some clinically useful measures of the effects of treatment [EBM note]. *Evidence-Based Medicine.* 1996 Jan-Feb;1:37-8.
2. **Singh N, Gayowski T, Yu VL, Wagener MM.** Trimethoprim-sulfamethoxazole for the prevention of spontaneous bacterial peritonitis in cirrhosis: a randomized trial. Ann Intern Med. 1995;122:595-8.
3. **Rothman KJ.** Modern Epidemiology. Boston: Little, Brown and Co.; 1986:177-236.
4. **Yusuf S, Peto R, Lewis J, Collins R, Sleight P.** Beta blockade during and after myocardial infarction: overview of the randomized trials. Prog Cardiovasc Dis. 1985; 27:335-71.
5. **Greenland S, Salvan A.** Bias in the one-step method for pooling study results. Stat Med. 1990;9:247-52.
6. **Sinclair JC, Bracken MB.** Clinically useful measures of effect in binary analyses of randomized trials. J Clin Epidemiol. 1994; 47:881-90.
7. **Greenland S, Robins JM.** Estimation of a common effect parameter from sparse follow-up data. Biometrics. 1985;41:55-68.

Journals Reviewed for This Issue

Core Journals

Am J Med	Arch Gen Psychiatry	Circulation	J Intern Med
Am J Obstet Gynecol	Arch Intern Med	Clin Pediatr	J Neurol Neurosurg Psychiatry
Am J Psychiatry	Arch Pediatr Adolesc Med	Cochrane DB Sys Rev	J Pediatr
Am J Surg	Arch Surg	Diabetes Care	J Vasc Surg
Anaesthesia	Arthritis Rheum	Hypertension	Lancet
Anaesth Analg	BMJ	JAMA	N Engl J Med
Anesthesiology	Br J Gen Pract	J Am Board Fam Pract	Obstet Gynecol
Ann Intern Med	Br J Obstet Gynaecol	J Am Coll Surg	Pediatrics
Ann Surg	Br J Surg	J Gen Intern Med	Surgery
Arch Dis Child			

Journals for Continuing Review

Acta Obstet Gynecol Scand	Arch Neurol	Eff Health Care	J Fam Pract
Age Ageing	Br J Dermatol	Fertil Steril	J Infect Dis
Am J Cardiol	Br J Psychiatry	Gastroenterology	J Reprod Med
Am J Gastroenterol	Br J Rheumatol	Gut	Med Care
Am J Public Health	Can Fam Phys	Heart (formerly Br Heart J)	Med J Aust
Am J Respir Crit Care Med	Can Med Assoc J	J Am Acad Dermatol	Neurology
Ann Emerg Med	Chest	J Am Coll Cardiol	Spine
Ann Med	Clin Invest Med	J Am Geriatr Soc	Stroke
Arch Fam Med	Crit Care Med	J Clin Epidemiol	Thorax

BMJ Evidence-Based Medicine

Down with odds ratios!

David L. Sackett, Jonathan J. Deeks and Doughs G. Altman

Evid Based Med 1996 1: 164-166
doi: 10.1136/ebm.1996.1.164

Notes