Int J Public Health 53 (2008) 165–167
1661-8556/08/030165-3
DOI 10.1007/s00038-008-7068-3

When to use the odds ratio or the relative risk?

Carsten Oliver Schmidt, Thomas Kohlmann

Institute for Community Medicine, Methods in Community Medicine, University of Greifswald, Greifswald

Submitted: 12 December 2007; Revised: 16 January 2008; Accepted: 21 January 2008

Relative Risk and odds ratio in epidemiology

The relative risk (RR) and the odds ratio (OR) are the two most widely used measures of association in epidemiology. The direct computation of relative risks is feasible if meaningful prevalences or incidences are available. Cross-sectional data may serve to calculate relative risks from prevalences. Cohort study designs allow for the direct calculation of relative risks from incidences. The situation is more complicated for case-control studies. If meaningful prevalences or incidences are not available, the OR provides a valid effect measure: It describes the ratio of disease odds given exposure status, or alternatively the ratio of exposure odds given the disease status. Computationally, both approaches lead to the same result.

The OR for a given exposure is routinely obtained within logistic models while controlling for confounders. The availability of this approach in standard statistical software largely explains the popularity of this measure. However, it does not have as intuitive an interpretation as the relative risk. This is where problems start: OR's are often interpreted as if they were equivalent to relative risks while ignoring their meaning as a ratio of odds. It is for instance common to describe an OR of "2" in terms of a "twofold risk" of developing a disease given exposure. This inaccuracy entails potentially serious problems because the OR always overestimates the RR. This can easily be deduced from the mathematical formulas as depicted in Table 1 because of the way the denominators differ.

Discrepancies between relative risk and odds ratio

Two main factors influence the discrepancies between RR and OR: The initial risk of an event, and the strength of the association between exposure and the event. Figure 1 illustrates a simplified relation between OR and RR without consideration of confounding. In case of an incidence of 20 %, an OR of 10 corresponds to a RR of less than 4. If the incidence is 50 %, an OR of 5 corresponds to a RR of less than 2. Generally speaking, two factors contribute to high discrepancies between OR and RR: a high initial incidence of an outcome of interest, and a high OR. Under the so called "rare disease assumption" the OR may provide an acceptable approximation of the RR. A percentage of 10 % is a frequently mentioned figure for this purpose but still this cut-off point cannot ensure a close correspondence between both measures as can be inferred from Figure 1.

Yet it should be kept in mind, that there is nothing wrong with the OR as long as it is interpreted as such. There is no need at all for any transformation per se. In some applied situations,

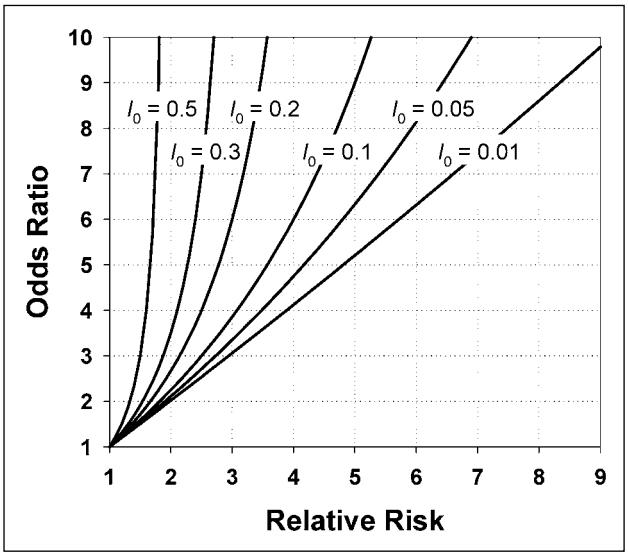

Figure 1 Relationship between Odds Ratio and Relative Risk for various incidence rates

however, it may be useful to report relative risks for better communication even though the rare disease assumption does not hold. Several approaches may be chosen to achieve this goal:

1. Conversion formulas for OR's have been proposed (see Table 1)[1, 2]. Confidence intervals for the RR can be obtained by applying the same correction to the confidence interval bounds of the OR. Despite the popularity of this approach results are susceptible to confounding and may produce biased estimates and confidence intervals[3, 4].
2. Mantel-Haenszel estimates may be used to infer relative risks in simple situations with a categorical exposure variable and a categorical confounder. However, if adjustment for several confounders and/or continuous covariates is needed this approach may be difficult to apply.
3. Relative risks may directly be calculated within the generalized linear regression model framework. Instead of the logistic regression, log-binomial, and Cox / poisson regression, among others, have been proposed for this purpose[4, 5, 6]. Frequent problems encountered with the log-binomial model include failures to converge, in particular in case of continuous covariates, and out of bound estimates for the predicted probabilities[3, 7]. The latter is also the case for poisson regression. Furthermore, robust variance estimates should be used with poisson models to obtain valid confidence intervals because the poisson distribution overestimates the binomial distribution for frequent events. The generalized linear model is implemented in major statistical software packages (SAS, STATA, SPSS).

Comparing relative risk and odds ratio, an example

For illustrative purposes, these approaches have been applied on data from a German health survey, a population based

Formal display

		Disease	
		yes	no
Exposed	yes	a	b
	no	c	d

Prevalence[1a)] / Incidence[1b)] among exposed, $I_1 = \frac{a}{a+b}$ and unexposed $I_0 = \frac{c}{c+d}$

Relative Risk (RR) und aprocimation to Odds Ratio (OR) under the rare disease assumption

$$RR = \frac{I_1}{I_0} = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} \xrightarrow[\text{if a, c small}]{} \approx \frac{\frac{a}{b}}{\frac{c}{d}} = OR$$

Conversion of Odds Ratio into Risk Ratio (Holland, 1989; Zhang und Yu 1998)

$RR = \frac{OR}{1 - I_0 + I_0 * OR}$ alternatively RR = OR * $\frac{1 - I_1}{1 - I_0}$

[1a)] using cross-sectional data; [1b)] using longitudinal data

Table 1 Calculation of relative risks, odds ratio, and conversion

	target variables at 10 month follow up			
	any back pain during 3 months prior to follow-up (incidence: 62 %)		disabling back pain during 3 months prior to follow-up (incidence: 4 %)	
	point estimate	95 % CI	point estimate	95 % CI
Logistic regression (OR)	3.75	3.15–4.46	3.47	2.44–4.94
Zhang Yu (RR)	1.40	1.36–1.43	3.16	2.31–4.27
Mantel Haenszel (RR)	1.34	1.29–1.39	3.07	2.23–4.23
Log-Binomial (RR)	1.34	1.29–1.40	3.09	2.25–4.26
Poisson (RR)	1.35	1.25–1.45	3.10	2.24–4.28

Analysis conducted among subjects without back pain at baseline.

Table 2 Comparing odds ratios and different approaches to calculate relative risks

prospective cohort study of 9267 adults[8]. We predicted "any" back pain experience during the 3 months previous to follow-up (62 %), and disabling back pain (4 %) at 10 month follow-up among subjects without back pain at baseline. The predictor variable was low vs. high somatization. Subjective health (poor-moderate vs. high) was included in the model as a confounding variable. Table 2 shows that the OR substantially overestimates the RR for incident "any" back pain but there is a comparatively small overestimation for "severe" back pain. This is due to the different incidences of both events. The other approaches lead to similar results in this "real data" example. However, the correction formula slightly overestimates the relative risk. The latter observation is an indicator of present confounding between both predictor variables. Differences between statistical approaches can be more pronounced with other data. Therefore it may be useful to cross check results, for instance by comparing results from different statistical approaches when computing risk ratios, and to examine the predicted probabilities when using regression models. Normally, the regression approaches are preferable over the correction formula but the researcher needs to be aware of their shortcomings. Valid estimates for risk ratios may also be based on logistic regression even when encountering problems with the before mentioned models[3], and some other approaches for cohort as well as for case-control designs are available[9]. Their description is beyond the scope of this short overview.

References

1. Holland PW. A note on the covariance of the Mantel-Haenszel log-odds-ratio estimator and the sample marginal rates. Biometrics 1989;45:1009–16.

2. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998;280:1690–1.

3. Localio AR, Margolis DJ, Berlin JA. Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression. J Clin Epidemiol 2007;60:874–82.

4. McNutt LA, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am J Epidemiol 2003;157:940–3.

5. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 2003;3:21.

6. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;15:702–6.

7. Blizzard L, Hosmer DW. Parameter estimation and goodness-of-fit in log binomial regression. Biom J 2006;48:5–22.

8. Schmidt CO, Raspe H, Pfingsten M, Hasenbring M, Basler HD, Kohlmann T. Back pain in the German adult population. Prevalence, severity, and sociodemographic correlates in a multi-regional survey. Spine 2007;32:2005–11.

9. Greenland S. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. Am J Epidemiol 2004;160:301–5.

Address for correspondence

Carsten Oliver Schmidt
Institute for Community Medicine
Methods in Community Medicine
University of Greifswald
Walther Rathenau Str. 48
17487 Greifswald
Germany
e-mail:
carsten.schmidt@uni-greifswald.de