Explaining Odds Ratios

Magdalena Szumilas, MSc[1]

What is an odds ratio?

An odds ratio (OR) is a measure of association between an exposure and an outcome. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure. Odds ratios are most commonly used in case-control studies, however they can also be used in cross-sectional and cohort study designs as well (with some modifications and/or assumptions).

Odds ratios and logistic regression

When a logistic regression is calculated, the regression coefficient (b1) is the estimated increase in the log odds of the *outcome per unit increase* in the value of the *exposure.* In other words, the exponential function of the regression coefficient (e^{b1}) is the odds ratio associated with a one-unit increase in the exposure.

When is it used?

Odds ratios are used to compare the relative odds of the occurrence of the outcome of interest (e.g. disease or disorder), given exposure to the variable of interest (e.g. health characteristic, aspect of medical history). The odds ratio can also be used to determine whether a particular exposure is a risk factor for a particular outcome, and to compare the magnitude of various risk factors for that outcome.

OR=1   Exposure does not affect odds of outcome

OR>1   Exposure associated with higher odds of outcome

OR<1   Exposure associated with lower odds of outcome

What about confidence intervals?

The 95% confidence interval (CI) is used to estimate the precision of the OR. A large CI indicates a low level of precision of the OR, whereas a small CI indicates a higher precision of the OR. It is important to note however, that unlike the p value, the 95% CI does not report a measure's statistical significance. In practice, the 95% CI is often used as a proxy for the presence of statistical significance if it does not overlap the null value (e.g. OR=1). Nevertheless, it would be inappropriate to interpret an OR with 95% CI that spans the null value as indicating evidence for lack of association between the exposure and outcome.

Confounding

When a non-casual association is observed between a given exposure and outcome is as a result of the influence of a third variable, it is termed confounding, with the third variable termed a confounding variable. A confounding variable is causally associated with the outcome of interest, and non-causally or causally associated with the exposure, but is not an intermediate variable in the causal pathway between exposure and outcome (Szklo & Nieto, 2007). Stratification and multiple regression techniques are two methods used to address confounding, and produce "adjusted" ORs.

Example

Data from an article published in the Journal in November 2008 will be used to illustrate how ORs (A) and 95% CIs (B) are calculated

In their article, Greenfield and colleagues looked at previously suicidal adolescents (n=263) and used logistic regression to analyze the associations between baseline variables such as age, sex, presence of psychiatric disorder, previous hospitalizations, and drug and alcohol use, with suicidal behaviour at six-month follow-up (Greenfield et al., 2008).

A) Calculating Odds Ratios

We will calculate odds ratios (OR) using a two-by-two frequency table

[1] Research Associate, Sun Life Financial Chair in Adolescent Mental Health, IWK Health Centre & Dalhousie University, Maritime Outpatient Psychiatry, Halifax, Nova Scotia

Corresponding Email: magdaszumilas@gmail.com

Outcome Status

		+	−
Exposure status	+	a	b
	−	c	d

Where

a = Number of exposed cases
b = Number of exposed non-cases
c = Number of unexposed cases
d = Number of unexposed non-cases

$$OR = \frac{a/c}{b/d} = \frac{ad}{bc}$$

$$OR = \frac{(n) \text{ exposed cases} / (n) \text{ unexposed cases}}{(n) \text{ exposed non-cases} / (n) \text{ unexposed non-cases}} = \frac{(n) \text{ exposed cases} \times (n) \text{ unexposed non-cases}}{(n) \text{ exposed non-cases} \times (n) \text{ unexposed cases}}$$

In the study, 186 of the 263 adolescents previously judged as having experienced a suicidal behaviour requiring immediate psychiatric consultation did not exhibit suicidal behaviour (non-suicidal, NS) at six months follow-up. Of this group, 86 young people had been assessed as having depression at baseline. Of the 77 young people with persistent suicidal behaviour at follow-up (suicidal behaviour, SB), 45 had been assessed as having depression at baseline.

What is the OR of suicidal behaviour at six months follow-up given presence of depression at baseline?

First we determine the numbers to use for (a), (b), (c), (d)

a: Number of exposed cases (+ +) = ?
b: Number of exposed non-cases (+ –) = ?
c: Number of unexposed cases (– +) = ?
d: Number of unexposed non-cases (– –) = ?

Q1: Who are the exposed cases (++ = a)?
A1: Youth with persistent SB assessed as having depression at baseline
a=45

Q2: Who are the exposed non-cases (+ – = b)?
A2: Youth with no SB at follow-up assessed as having depression at baseline
b=86

Q3: Who are the unexposed cases (– + = c)?
A3: Youth with persistent SB not assessed as having depression at baseline
c: 77(SB) –45(depression) = 32

Q4: Who are the unexposed non-cases (– – = d)?
A4: Youth with no SB at follow-up not assessed as having depression at baseline
d: 186(NS) –86(depression) = 100

Then we plug the values into the formula

a: Number of exposed cases (++) = 45
b: Number of exposed non-cases (+ –) = 86
c: Number of unexposed cases (– +) = 32
d: Number of unexposed non-cases (– –) = 100

$$OR = \frac{a/c}{b/d} = \frac{ad}{bc} = \frac{45/32}{86/100} = 1.63$$

Thus, the odds of persistent suicidal behaviour is 1.63 higher given baseline depression diagnosis compared to no baseline depression.

B) Calculating 95% confidence intervals

What are the confidence intervals for the OR calculated above?

Confidence intervals are calculated using the formula shown below

Upper 95% CI = e ^ [1n (OR) + 1.96 √(1/a + 1/b + 1/c + 1/d)]

Lower 95% CI = e ^ [1n (OR) – 1.96√ (1/a + 1/b + 1/c + 1/d)]

Plugging in the numbers from the table above, we get:

Upper 95% CI = e^ [1n (OR) + 1.96√ (1/45 + 1/86 + 1/32 + 1/100)] = 2.80

Lower 95% CI = e^ [1n (OR) – 1.96√ (1/45 + 1/86 + 1/32 + 1/100)] = 0.96

Since the 95% CI of 0.96 to 2.80 spans 1.0, the increased odds (OR 1.63) of persistent suicidal behaviour among adolescents with depression at baseline does not reach statistical significance. In fact, this is indicated in Table 1 of the reference article, which shows a p value of 0.07.

Interestingly, the odds of persistent suicidal behaviour in this group given presence of borderline personality disorder at baseline was twice that of depression (OR 3.8, 95% CI:1.6-8.7), and was statistically significant (p 0.002)

_____

This example illustrates a few important points. First, presence of a positive OR for an outcome given a particular exposure does not necessarily indicate that this association is statistically significant. One must consider the confidence intervals and p value (where provided) to determine significance. Second, while the psychiatric literature shows that overall, depression is strongly linked to suicide and suicide attempt (Kutcher & Szumilas, 2009), in a particular sample, with a particular size and composition, and in the presence of other variables, the association may not be significant.

Understanding odds ratios, how they are calculated, what they mean, and how to compare them is an important part of understanding scientific research. This article has covered the basics of odds ratios. For more information about odds ratios and other statistics used in medicine, the following website provides a link to the Centre for Statistics in Medicine at Oxford University, and a series of "Statistics Notes" published in *BMJ* by Doug Altman, Martin Bland, and others (http://www.csm-oxford.org.uk/index.aspx?o=1292).

References

Centre for Statistics in Medicine (Undated). *BMJ Statistics Notes*. Retrieved June 15, 2010. URL: http://www.csm-oxford.org.uk/index.aspx?o=1292

Greenfield, B., Henry, M., Weiss, M., Tse, S. M., Guile, J. M., Dougherty, G., Zhang, X., Fombonne, E., Lis, E., Lapalme-Remis, & Harnden, B. (2008). Previously suicidal adolescents: Predictors of six-month outcome. *Journal of the Canadian Association of Child and Adolescent Psychiatry, 17*(4), 197-201.

Kutcher, S., & Szumilas, M. (2009). *Suicide Risk Management. BMJ Point of Care* [Internet]. London: BMJ Publishing Group.

Oleckno, W. A. (2002). Essential Epidemiology: Principles and Applications. Long Grove, IL: Waveland Press.

Szklo, M., & Nieto, F. J. (2007). *Epidemiology: Beyond the basics. 2nd edition*. Sudbury, MA: Jones and Bartlett Publishers.