

What is an Odds Ratio? What does it mean?

Allison Shorten,¹ Brett Shorten²

10.1136/eb-2015-102206

¹Yale University School of Nursing, West Haven, Connecticut, USA
²Guilford, Connecticut, USA

Correspondence to:
Dr Allison Shorten,
 Yale University School of Nursing, PO Box 27399, West Haven, CT 06516-7399, USA;
 allison.shorten@yale.edu

When we read the results of healthcare research, we often read about relationships or associations between different treatments and their outcomes for patients. Health outcomes can be improved with the use of different types of medications, engagement in health education programmes and changes in lifestyle factors such as targeted improvements in nutrition or reduction in smoking. Depending on the research design, we can use statistical tools to assist us in identifying the most effective healthcare interventions, with the aim of achieving the best health outcome for patients. This can help us make decisions about which treatments we should provide or which public health programmes should be funded, in the hope that we can improve the overall health of our patients and the communities we serve. We often see odds ratios (ORs) being used in research to explain whether interventions contribute to improvements in health. They are used to help researchers demonstrate associations between interventions and outcomes in both positive and negative directions.

To explain how ORs are calculated and what they mean, we will use two examples. Let's first consider a hypothetical example of a randomised controlled trial (RCT). We are often interested in programmes that will reduce rates of smoking in the community. In particular, we often see studies that focus on reducing the rates of smoking in pregnant women, with the aim of achieving health improvements for both women and their newborn babies. Let's consider a study designed to test a smoking cessation programme for pregnant women. **Table 1** provides a summary of the hypothetical results.

From **table 1**, we see that 200 women were randomised to either the intervention group (where pregnant women received the smoking cessation programme), or a control group (pregnant women did not receive the smoking cessation programme). On first look, it appears that the smoking cessation intervention seems to have been successful. There were 30 of 100 women (30%) who quit smoking in the intervention group, yet only 15 of 100 women in the control group (15%) quit smoking. How certain can we be that this is a relationship we can be confident about?

The next step in exploring the veracity of this apparent relationship between the smoking cessation programme and the number of pregnant women who quit smoking, might be to use a method of statistical analysis involving ORs. Of course, the differences observed between the two groups of pregnant women may be due to other 'chance factors' such as the characteristics of the particular sample of women selected for this study.

Table 1 Hypothetical data on cessation of smoking among pregnant women

	Quit smoking	Did not quit	Total
Intervention group	a=30	b=70	100
Control group	c=15	d=85	100

To examine this possibility, we may calculate and analyse what is known as an 'unadjusted' OR. Calculation of an 'unadjusted' OR is actually quite simple. Note that when we first calculated the percentage of intervention and control group women who quit smoking, we did that by dividing the number of women who quit by the *total* for each group (eg, for the treatment group the cessation *rate* was 30/100=0.30, or 30%). OR analysis modifies this calculation by dividing the number of women quitting by the number *not quitting* to obtain a smoking cessation *ratio* for each group (ie, 30/70 for the intervention group and 15/85 for the control group). From **table 1**, the 'unadjusted' OR for smoking cessation can be simply calculated as the ratio for the intervention group divided by the ratio for the control group, or (a/b)/(c/d). From the hypothetical data in **table 1**, this gives (30/70)/(15/85), or 2.43. Try this yourself to verify that you can obtain this result, or enter the **table 1** data into an online OR calculator.¹

So, from our calculated OR of 2.43, the odds of a pregnant woman quitting smoking are almost two and a half times that of mothers who did not receive the intervention. However, as with virtually all inferential statistics analyses, we must investigate further to determine how likely it is that this result could have occurred simply due to the characteristics of the particular sample selected for the study. In this case, a CI would be calculated, within which we are confident (usually 95% confidence is specified) that the true OR lies, based on the sample data. This calculation can also be easily obtained using an online CI calculator.¹ In this case, the interval calculated is approximately 1.21 to 4.87, with a p value of 0.01. The lower bound implies that women from the intervention group would be approximately 1.2 times as likely to cease smoking as women in the control group. The upper bound of 4.87 implies that women in the intervention group would be almost five times more likely to quit smoking than women in the control group. Note that an OR of exactly one implies that there is no difference in outcome between the intervention and control groups. Since the 95% CI does not contain the value of one within the interval, we are more than 95% confident that a difference in smoking cessation rate actually exists between the two groups. The reported p value of 0.01 suggests that there is actually a 99% chance that a difference really exists.

OR analyses can also be used in efforts to control for potential confounding factors in research, as we will see in the next example. This time the example contains data from a published study designed to examine the potential benefits of using corticosteroids in late preterm pregnancy.² This particular study was based on a retrospective cohort of 167 pregnant women who underwent amniocentesis at 34–37 weeks gestation, to determine fetal lung maturity and received a negative result.² Of these women, 83 received antenatal corticosteroids (the treatment group) and 84 did not (the control group).²



Table 2 NICU Admission rates by use of antenatal corticosteroids

	Admitted to NICU	Not admitted	Total
Steroids administered	a=14	b=69	83
No steroids	c=24	d=60	84

NICU, neonatal intensive care unit.

Table 2 provides results according to whether the newborn baby was admitted to a neonatal intensive care unit (NICU) after delivery. In the treatment group, only 14 of the 83 infants were admitted to the NICU (16.9%), yet in the control group, 24 of 84 babies (28.6%) were admitted to the NICU.² Again, at first look, it appears as though the treatment may have been effective in reducing the rate of admission of newborn babies to the NICU.

For this example, we can calculate an OR for NICU admission of $(14/69)/(24/60)=0.51$, implying that babies whose mothers received the treatment were only half as likely to require NICU admission, with a 95% CI of 0.24 to 1.07 and p value of 0.07. Note that the CI implies that, although the true odds of babies in the treatment group requiring NICU admission may be as little as one quarter those of the control group, the interval does include the value of one. Therefore, there is a chance that the true NICU admission rates could be the same for both groups. This is reinforced by the p value of 0.07, which suggests that based on this study, we are less than 95% confident that a true difference exists. In statistical terms, the difference found is not 'statistically significant'.

What does this mean? We can conclude that this result deserves further study, perhaps using a RCT, to further examine whether the relationship exists or not. It is important to note that the study has *definitely not* proven that the treatment is ineffective, based solely on

the p value being less than the conventional 0.05 level or that the upper bound of the CI (1.07) was greater than 1. After all, the weight of evidence in this study points to the likelihood that the treatment was effective.

When studies use retrospective cohort data, rather than a prospective RCT design, it is more likely that unmeasured factors will be present to confound the results. For example, there are many other factors affecting the need to admit babies to a NICU, such as mode of delivery and gestational age. To account for this, studies will usually try to identify and control for as many of these possible 'confounders' as they can, often via the use of some form of regression analysis, particularly 'logistic' regression. In this case, the paper would be expected to report 'adjusted' ORs, which are calculated from regression coefficients after possible confounding variables have been taken into account. In fact the published study from which this sample is drawn used regression analysis to examine important factors associated with a range of outcomes for babies in the sample studied.² When it comes to the interpretation of both the 'adjusted' OR and associated CI and p values, they are essentially the same as what has been outlined for the simple 'unadjusted' OR reported earlier. Of course, ORs are only one of many statistical tests that can be used to help us form conclusions about research, and to help us make decisions about the strength of evidence on which healthcare is based.

Competing interests None declared.

References

1. MedCalc Easy to Use Statistical Software. https://www.medcalc.org/calc/odds_ratio.php
2. Yinon Y, Haas J, Mazaki-Tovi S, *et al.* Should patients with documented fetal lung immaturity after 34 weeks of gestation be treated with steroids? *Am J Obstet Gynecol* 2012;207:222. e1-4.