

# Measures of association: an overview with examples from Canadian emergency medicine research

Andrew Worster, MD;\* Brian H. Rowe, MD†

• SEE RELATED EDITORIAL, PAGE 175.

## ABSTRACT

The authors of studies often report their results using abbreviated terms such as RR, OR, ARR, RRR and NNT. These terms are quantities that express the strength of association between the dependent and independent variables and are collectively referred to as *measures of association*. The similarity between these measures and the multiple terms by which each is referred can be confusing. The purpose of this article is to explain in a straightforward manner the purpose, derivation, and limitations of some of the more commonly used categorical measures of association, including relative risk, odds ratio, absolute and relative risk reduction and number needed to treat, using results from recent emergency medicine studies published by Canadian researchers.

## RÉSUMÉ

Les auteurs d'études présentent souvent leurs résultats en utilisant des abréviations comme RR, RC (OR en anglais), ARR, RRR et NNT. Ces termes sont des quantités qui expriment le degré d'association entre les variables dépendantes et indépendantes et qu'on appelle collectivement les mesures d'association. La similitude entre ces mesures et les multiples termes utilisés pour les identifier peuvent prêter à confusion. Le présent article a comme objectif d'expliquer clairement le but, la dérivation et les limites de certaines des mesures les plus couramment utilisées, notamment le risque relatif, le rapport des cotes (odds ratio), la réduction absolue et relative du risque et le nombre nécessaire pour traiter, en ayant recours aux résultats d'études récentes en médecine d'urgence publiés par des chercheurs canadiens.

**Key words:** Evidence-based medicine, emergency medicine, decision-making, therapeutics

## Clinical scenario

A 32-year-old woman presents to the emergency department (ED) with acute asthma. After 3 hours of treatment she feels much better, and her pulmonary function measurement is now 70% of predicted. A recent article suggests that she would have less risk of relapse if you added inhaled corticosteroids to her oral prednisone treatment.<sup>1</sup> The study reported a *p* value of 0.049, a relative risk reduction of 48%, an absolute risk reduction of 12% and a num-

ber needed to treat of 9. The patient asks why you think she should take this and whether it's any better than the echinacea she currently takes. How do you summarize these different results for yourself, for junior learners and for the patient?

## The basics

Familiarity with 2×2 tables is central to understanding each of the measures of association.<sup>2</sup> Table 1, a standard

\*Research Coordinator, Research Division of Emergency Medicine, Hamilton Health Sciences Corporation, and McMaster University, Hamilton, Ont.; †Division of Emergency Medicine and Department of Public Health Sciences, University of Alberta, and Capital Health Authority, Edmonton, Alta.

Received: Feb. 2, 2001; final submission: Apr. 28, 2001; accepted: May 15, 2001

This article has been peer reviewed.

$2 \times 2$  table, presents the dependent variable (outcome event) across the top and the independent variables along the side, with the experimental treatment listed above the control treatment. This format is important because deviation from it can alter future calculations.

When we speak of “outcomes” in clinical trials, we are typically referring to things that happen to patients (e.g., death, surgery, stroke or disease recurrence). As these examples reveal, most outcomes in clinical trials are negative events — especially death! An exception is in cardiac arrest trials, where a commonly reported outcome is survival-to-hospital discharge. In addition, many outcomes are dichotomous; that is, the patient either experienced the event or did not. However, in order to express outcomes in a  $2 \times 2$  table format, outcomes need not be dichotomous (nor must they be negative). For instance, we can take a continuous variable such as peak expiratory flow rate (PEFR) and set a cut-off point, such as 200 L/min, to differentiate a severe from a not-severe patient group. In this case, the outcome is binary, but not necessarily a negative event.

## Risk vs. odds

Next we need to define and differentiate two terms: *risk* and *odds*. *Risk* is simply the probability of an event occurring.<sup>2</sup> The term “risk” generally implies a negative event; we don’t describe people as being at risk of winning a lottery. However, the clinical event need not necessarily be unfavourable to the patient even though the results of the trial are described in terms of risk. *Odds* differ from *risk* in that odds involves 2 probabilities instead of 1, and these 2 probabilities are expressed in the form of a ratio. Specifically, *odds* is the ratio of the probability of an event occurring to the probability of the event not occurring.<sup>2</sup> This order in the ratio — occurrence over non-occurrence — is important if one is to interpret the results accurately. For example, if of every 100 ED chest pain patients with a normal initial ECG, 60 are found to have an acute coronary syndrome (ACS) and 40 are not, the *risk* or probability of these patients developing ACS is 60/100 or 0.6.<sup>2</sup> However,

**Table 1. A typical  $2 \times 2$  table**

	Outcome events		Total no. of patients per group
	Yes	No	
Treatment (exposure)	A	B	(A + B)
Control (no exposure)	C	D	(C + D)
Total	(A + C)	(B + D)	(A + B + C + D)

the *odds* among these 100 patients of developing an ACS are 60:40 or 1.5.<sup>2</sup>

## Relative risk (RR)

One way to measure the results in a cohort study or randomized controlled trial (RCT) is to compare the number of outcome events in the treatment group vs. the number in the control group.<sup>3</sup> The relative risk (RR), or risk ratio, does just that. It is a measure of the probability of occurrence of an event in those who received the experimental treatment (exposed group) compared (relative) to those who received the standard or control treatment (not exposed).<sup>4</sup>

The relative risk is defined as (see Table 1 for definitions):

$$\begin{aligned} \text{RR} &= \text{Incidence in treatment group} / \text{Incidence in control group} \\ &= [A / (A + B)] / [C / (C + D)] \\ &= A(C + D) / C(A + B) \end{aligned}$$

An RR of 1 indicates that the outcome rate is the same in both groups, therefore, the experimental group is no better or worse off than the control group. An  $\text{RR} < 1$  indicates that the outcome rate is lower in the exposed (treatment) group than the unexposed (control) group. This signifies an effective treatment. If the outcome is a negative event such as death, an  $\text{RR} < 1$  would indicate that there are fewer deaths in the treatment group than in the control group. In this case, the experimental treatment is superior to the standard treatment. Also, the further the RR is from 1, the greater the effect size or difference between the 2 groups. It follows that an  $\text{RR} > 1$  indicates the reverse, suggesting that the experimental treatment is inferior to the control treatment.

Table 2 was taken from an RCT of acute asthma treatment, wherein ED patients were discharged with oral corticosteroids plus either budesonide (an inhaled corticosteroid) or placebo.<sup>1</sup> The main outcome event was relapse, defined as an unscheduled visit for worsening asthma symptoms within 21 days of discharge. Of the 94 patients randomized to budesonide, 12 (12.8%) relapsed. Conversely, 23 (24.5%) of the 94 patients randomized to placebo relapsed.

**Table 2. Asthma relapse with budesonide vs. placebo**

	Relapse within 21 days		Total
	Yes	No	
Treatment (budesonide)	12	82	94
Control (placebo)	23	71	94
Total	35	153	188

The incidence of the primary outcome (relapse) was 12/94 (0.128) in the inhaled corticosteroid (ICS) group and 23/94 (= 0.245) in the control (placebo) group. The RR is, therefore,  $0.128/0.245 = 0.52$ . We can conclude that the relative risk of asthma relapse with ICS is 0.52. When the outcome is an adverse event, an  $RR < 1$  suggests the treatment is superior to the control. An RR of 0.52, therefore, shows ICS to be superior to placebo. Expressed as a percent (52%), we can say that the risk of relapse with ICS is approximately 52% (or one half) of that with placebo.

Unfortunately, RR does not tell us the baseline event rate. In this example, the baseline event rate would be the rate of asthma relapse in patients discharged without ICS.<sup>5</sup> RR only reveals whether the outcome rate in the treatment group is more ( $RR > 1$ ), the same ( $RR = 1$ ), or less ( $RR < 1$ ) than it is in the control group. Another limitation is that the RR cannot be calculated from non-cohort or RCT studies. This is because in case-control studies the investigator selects the number of case and control subjects according to whether they have experienced the outcome event and, subsequently, predetermines the proportion of individuals with the outcome event.<sup>4,5</sup> In such studies, the odds ratio (OR) is a more correct measure of association to apply.

## Relative risk reduction

The relative risk reduction (RRR) is a quantification of how much the treatment reduces the risk of the outcome relative to the baseline outcome event rate. Wherein the baseline risk of an outcome event is the event rate in the control group defined by  $C/(C + D)$ , the RRR is the difference in outcome event rates between the control and treatment groups divided by the outcome event rate in the control group.

$$\begin{aligned} RRR &= ([C / (C + D)] - [A / (A + B)]) / [C / (C + D)] \\ &= (0.245 - 0.128) / 0.245 = 0.48 \end{aligned}$$

Since the RRR is expressed as a percent:  $0.48 \times 100\% = 48\%$ . The shorter method of performing this calculation is from the RR:<sup>5</sup>

$$\begin{aligned} RRR &= (1 - RR) \times 100\% \\ &= (1 - 0.52) \times 100\% \\ &= 48\% \end{aligned}$$

From the budesonide example we see that ICS reduced the baseline risk of asthma relapse at 21 days by 48% relative to placebo.

Although the RRR is the most commonly reported measure of dichotomous treatment effect, it gives the decision-maker no indication of the absolute number of people who

might benefit from the intervention and, therefore, no indication of the cost-effectiveness of the intervention.<sup>6</sup> In our ICS example, an RRR of 48% appears very impressive, and for those patients who benefit, it is very impressive.<sup>6,7</sup> However, it doesn't tell us what the risk of relapse is without ICS. If the risk of asthma relapse was rare (e.g., 1:1000), an RRR = 48% means that only about 1:2000 patients would benefit. In this case, the relative risk reduction is high, but the absolute benefit to the population is low.<sup>6</sup> Alternatively, small relative risk reductions may provide significant benefit to the population if the condition is common. An example of a common disorder with significant impact on society is stroke. A treatment that produced a 5% relative reduction in strokes would confer a huge benefit to the population as a whole.<sup>6,7</sup>

## Absolute risk reduction (ARR)

The ARR has several synonyms, including *attributable relative risk*, *attributable risk*, *attributable risk reduction*, and *risk difference*.<sup>5,6</sup> Where RR gives a ratio of risk (an indication of the relative risk of exposure) and RRR gives the reduction in baseline risk relative to control, the ARR combines both the baseline outcome event rate and the risk reduction in the treatment group to give the actual risk difference between the treatment and control groups.<sup>5</sup>

$$\begin{aligned} ARR &= \text{Incidence in control group} - \text{Incidence in treatment group} \\ &= [C / (C + D)] - [A / (A + B)] \end{aligned}$$

Using the results from the budesonide example in Table 2:

$$ARR = 0.245 - 0.128 = 0.117$$

we convert these numbers to percent by simply multiplying by 100 to arrive at:

$$ARR = 24.5\% - 12.8\% = 11.7\%$$

This means that the absolute reduction in 21-day relapse from asthma in the ICS patients is 11.7%.<sup>2</sup> Hence, 11.7% fewer patients relapsed in the ICS group than in the placebo group. This is a much lower number than the RRR, making the treatment effect appear less impressive. Expressing results in relative terms (*the risk of relapse in the treatment group is 52% of that in the placebo group*) can make them seem very different than expressing them in absolute terms (*11.7% fewer patients relapsed*). The ARR provides more information about treatment effectiveness, in that it tells us what the consequences are of not treating the patient (i.e., the percentage of control patients who experience the outcome event).<sup>5,6</sup> The main limitation of the ARR is the usefulness of this number to the clinician

who has to decide whether or not to start his or her patients on this therapy.

## Number needed to treat (NNT)

NNT, the number of patients who need to be given the treatment in order to prevent one outcome event, is a useful measure for explaining the likelihood of benefit from treatment.<sup>6</sup> Moreover, NNT is intuitive to both physicians and patients. Like the ARR, NNT provides a treatment effect size result that incorporates both the baseline outcome event rate and the risk reduction in the treatment group. It can be argued, the ARR is more useful than the NNT because it expresses the results in more “concrete” terms and allows for treatment effect comparisons to be made with other disorders.<sup>6</sup> The NNT is simply the reciprocal of the ARR:

$$\text{NNT} = 1 / \text{ARR}$$

$$= 1 / 0.117 = 8.54$$

Hence, by inverting 11.7% i.e., we arrive at 8.54.<sup>2</sup> Therefore, the number of asthma patients we need to treat with ICS to prevent one relapse is 9.<sup>1</sup> NNT changes inversely with both the baseline outcome event rate and the RR; therefore, an increase in event rate or relative risk mean that fewer patients need to be treated in order to prevent one outcome event. Conversely, a decrease in the risk of the event will result in a higher NNT.

## Odds ratio (OR)

As with the ARR, the OR has a number of synonymous terms, including *relative odds* and *cross-product ratio*.<sup>1</sup> Other terms often incorrectly used synonymously with OR include *relative risk (RR)* and *risk ratio*. The OR is the ratio of the probability of an outcome event in one group (treatment or exposure) to the probability of the outcome event the other group (control or non-exposure).<sup>2,3</sup> The OR is a useful measure of treatment effect in RCTs and, as with the RR, the OR indicates that the outcome rate in the treatment group is more ( $\text{OR} > 1$ ), the same ( $\text{OR} = 1$ ), or less ( $\text{OR} < 1$ ) than in the control group.<sup>6</sup> In case-control studies on etiology, the OR indicates the probability of exposure to causative agent given the outcome.

$$\text{OR} = \text{AD}/\text{CB}$$

Many confuse RR and OR. Unlike RR, the OR is not an accurate measure of risk.<sup>4,6</sup> OR tends to slightly overestimate RR, but as the frequency of outcome events or dis-

ease incidence becomes smaller and approaches zero, the OR and RR will become increasingly equal.<sup>4</sup> Thus for rare events, the OR is a good approximation of RR. We can also see from the OR formula where the term *cross-product ratio* comes from.<sup>8</sup>

$$\text{OR} = \text{AD}/\text{CB}$$

$$\text{RR} = \text{A}(\text{C} + \text{D}) / \text{C}(\text{A} + \text{B})$$

From the example above we can calculate the OR:

$$\begin{aligned}\text{OR} &= (12 \times 71) / (23 \times 82) \\ &= 852 / 1886 \\ &= 0.452\end{aligned}$$

The OR (0.452) in this example is quite different from the RR (0.522).<sup>2</sup> However, if we reduce the incidence (A and C) by a factor of 10 (Table 3), then re-calculate the OR, we will see that it more closely approximates the RR.

$$\begin{aligned}\text{OR} &= (12 \times 917) / (23 \times 928) & \text{RR} &= 12(23 + 917) / 23(12 + 928) \\ \text{OR} &= 11,004 / 21,344 & \text{RR} &= 11,280 / 21,620 \\ \text{OR} &= 0.516 & \text{RR} &= 0.522\end{aligned}$$

As predicted, the OR approximates the RR as the incidence becomes very small.

Odds ratios are also easily generated from logistic regression models, where the purpose is to quantify the predictive relationship between the independent and dependent variables — using the former to predict the latter.<sup>8</sup> An excellent example comes from the Ontario Prehospital Advanced Life Support (OPALS) Study Group, who analyzed 6331 cardiac arrest calls to determine how strongly each of 5 independent factors contributed to the survival of prehospital cardiac arrest patients.<sup>9</sup> The strongest association came from the variables, “witnessed arrest” and “bystander CPR,” which had ORs of 3.9 and 3.7 respectively. Knowing this, we can say that a cardiac arrest victim who suffered a *witnessed* arrest is approximately 3.9 times more likely to survive than one whose arrest was *not witnessed*. Similarly, patients who receive bystander CPR are approximately 3.7 times more likely to survive to hospital discharge than those who do not. The odds ratio is also the measure of choice in meta-analysis (systematic reviews) where it refers to the probability that the outcome will occur in the treatment group.<sup>6</sup> In another

**Table 3. Asthma relapse with lower incidence rates**

	Relapse of symptoms within 21 days		Total
	Yes	No	
Treatment (budesonide)	12	928	940
Control (placebo)	23	917	940
Total	35	1845	1880

Canadian study that pooled the results of 6 RCTs comparing mortality rates of patients treated with prehospital thrombolysis to those treated with in-hospital thrombolysis, the OR was 0.83.<sup>10</sup> Since the OR < 1, we can conclude that the outcome (all cause mortality) is less common in the experimental group (prehospital thrombolysis recipients) than in the control group (in-hospital thrombolysis recipients).

## Precision

Measures of association, like all measures, are merely estimates of the truth. Therefore, it is often helpful to calculate 95% confidence intervals (CIs) around the derived estimate. Confidence intervals describe the precision of the estimate and help determine whether the result is statistically significant. The 95% CI represents the value range that the *true* value will fall within 95% of the time. In other words, although we can rarely know the exact true value (e.g., the proportion of asthmatic patients who relapse — in the whole world), if we know the 95% CI, we can be 95% certain that the *true* population value lies somewhere within this interval. Since an OR or RR value of 1.0 indicates no difference between the treatment and control groups, a 95% CI that includes the value 1.0 is not statistically significant. In other words, regardless of the value of our measure of association, if it has a 95% CI that includes 1.0, it is very possible that there is no treatment effect. Similarly, CIs for the ARR or RRR that include a value of zero are also not statistically significant.

## Summary

This article has attempted to explain and clarify the advantages and disadvantages of common measures of association. With respect to RR and OR, readers should remember that if these parameters are > 1, the outcome event is more common in the treatment group. If they are = 1, the outcome rate is the same in the treatment and control groups, and if they are < 1, the outcome rate is lower in the treatment group than in the control group.

In the clinical scenario described at the beginning of this article, you would explain to the patient that, while there are no guarantees she will remain relapse-free, her chance of asthma relapse within 3 weeks of ED discharge is about twice as great without ICS than with ICS (RRR = 48%; RR = 0.52). You might also explain that approximately 12% (ARR = 11.7%) fewer patients would be expected to relapse when treated with ICS than with placebo, and that of every 9 patients treated with ICS, at least 1 will avoid relapse. Furthermore, you're unaware of any studies on the

efficacy of echinacea in the treatment of acute asthma in the emergency setting and are unable to comment on its possible effectiveness. However, you are aware that high quality evidence fails to support the use of a variety of homeopathic treatments in chronic asthma.<sup>11</sup>

**Acknowledgments:** We wish to thank Dr. David Sackett for his guidance in writing this article. We would also like to thank Drs. Laurie J. Morrison and Ian Stiell for so promptly responding to our request for studies published by Canadian emergency physicians.

This paper was funded by the Hamilton Health Sciences Corporation, Hamilton, Ont.

## References

- Rowe BH, Bota GW, Fabris L, Therrien SA, Milner RA, Jacono J. Inhaled budesonide in addition to oral steroids to prevent asthma relapse following discharge from the emergency department: a randomized controlled trial. *JAMA* 1999;281:2119-26.
- Last JM. A dictionary of epidemiology. 3rd ed. New York: Oxford University Press; 1995.
- McKibbon A, Eady A, Marks S. PDQ: Evidence-based principles and practice. Hamilton (ON): BC Decker Inc.; 1999.
- Streiner DL, Norman GR. PDQ: Epidemiology. 2nd ed. St. Louis: Mosby; 1996.
- Jaeschke R, Guyatt G, Shannon H, Walter S, Cook D, Heddle N. Basic statistics for clinicians. 3. Assessing the effects of treatment: measures of association. *CMAJ* 1995;152(3):351-7.
- Laupacis A, Sackett DL, Roberts RR. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;318(26):1728-33.
- Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 1994;271(1):59-63.
- Lang TA, Secic M. How to report statistics in medicine: annotated guidelines for authors, editors, and reviewers. Philadelphia: American College of Physicians; 1997.
- Stiell IG, Wells GA, Field BJ, Spaite DW, De Maio VJ, Ward R, et al, for the OPALS Study Group. Improved out-of-hospital cardiac arrest survival through the inexpensive optimization of an existing defibrillation program: OPALS study phase II. Ontario Prehospital Advanced Life Support. *JAMA* 1999;281:1175-81.
- Morrison LJ, Verbeek PR, McDonald A, Sawadsky B: A meta-analysis of prehospital thrombolysis and the effect on mortality in acute myocardial infarction. *JAMA* 2000;283:2686-92.
- Linde K, Jobst KA. Homeopathy for chronic asthma [Cochrane review]. In: The Cochrane Library, Issue 1, 2001. Oxford: Update Software.

**Correspondence to:** Dr. Andrew Worster, Department of Emergency Medicine, Hamilton Health Sciences Corporation, 237 Barton St. E, Hamilton ON L8N 3Z5; 905 521-2100 x73136, fax 905 527-7051, worster@hhsc.ca