Submit to: *Journal of Advanced Nursing*

Type: Editorial

Title: Using risk and odds ratios to assess effect size for meta-analysis outcome measures

Authors: Mousa **ALAVI,** Glenn E **HUNT,** Denis C **VISENTIN,** Roger **WATSON,** Deependra K **THAPA,** Michelle **CLEARY**

Mousa ALAVI, PhD, Department of Psychiatric Nursing, School of Nursing and Midwifery, Isfahan University of Medical Sciences, Hezarjarib Avenue, Isfahan, Iran.
***corresponding author, Email: m_alavi@nm.mui.ac.ir ORCID: https://orcid.org/0000-0003-4847-2915

Glenn E. HUNT, PhD, Discipline of Psychiatry, Concord Clinical School, The University of Sydney, NSW, Australia. Email: glenn.hunt@sydney.edu.au ORCID: http://orchid.org/0000-0002-8088-9406

Denis C. VISENTIN, PhD, College of Health and Medicine, University of Tasmania, Sydney, NSW, Australia. Email: denis.visentin@utas.edu.au ORCID: https://orcid.org/0000-0001-9961-4384

Roger WATSON, RN, PhD, FAAN, Faculty of Health Sciences, University of Hull, Hull, UK. Email: r.watson@hull.ac.uk ORCID: https://orcid.org/0000-0001-8040-7625

Deependra K. **THAPA,** MPH, MSc, College of Health and Medicine, University of Tasmania, Sydney, NSW, Australia. Email: deependrakaji.thapa@utas.edu.au ORCID: https://orcid.org/0000-0002-5689-0837

Michelle CLEARY, RN, PhD, College of Health and Medicine, University of Tasmania, Sydney, NSW, Australia. Email: michelle.cleary@utas.edu.au ORCID: http://orcid.org/0000-0002-1453-4850

Conflict of Interest statement: Nil

Author contributions: All authors have agreed on the final version and meet at least one of the following criteria recommended by the ICMJE (http://www.icmje.org/recommendations/)]:

- Substantial contributions to conception and design, acquisition of data or analysis and interpretation of data; drafting the article or revising it critically for important intellectual content.

Using risk and odds ratios to assess effect size for meta-analysis outcome measures

1. INTRODUCTION

Best practice is built on the principle of aggregating all available evidence on a topic to make a clinical decision on the most appropriate intervention for the situation at hand. Systematic reviews and meta-analyses are powerful tools that summarize the evidence for current best practice guidelines for the available interventions for a particular problem (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009). Meta-analysis combines the results of multiple studies to produce an aggregated and more precise estimates of the benefits of the interventions. Meta-analysis of high-quality randomized trials are considered the highest level of evidence to inform practice.

When reading the healthcare literature, several measures of the effect of an intervention on an outcome are available to judge whether the evidence presented can be applied to clinical practice. It is important to be able to understand, correctly interpret and honestly communicate these reported measures (Thapa, Visentin, Hunt, Watson, & Cleary, 2020). However, it is not uncommon for clinicians and researchers to be confused about the differences between the various effect measures available (Tufanaru, Munn, Stephenson, & Aromataris, 2015). These will be outlined further in this editorial, with a focus on the odds ratio and risk ratio.

2. MEASURES OF EFFECT SIZE IN CLINICAL STUDIES

The three most common designs used to assess the effectiveness of a given intervention are case-control studies, cohort studies and randomized controlled trials (RCTs; Knol, Algra, & Groenwold, 2012). These studies estimate the measures of association, and hence the effectiveness of the intervention using various measures of effect (Schäfer & Schwarz, 2019). Quantitative indicators of the magnitude and direction of the effect of any intervention on a respective outcome is called the *effect size* (Tufanaru et al., 2015).

Key statistics applied in a meta-analysis allow researchers to draw conclusions through comparing standardized effect sizes across several studies (Feingold, 2017, Lakens, 2013). By undertaking a meta-analysis, we aim to arrive at a weighted combination of the effect sizes reported by several studies as a pooled estimation of the outcome of interest. In addition, meta-analysis often consists of the test for effect (i.e. risk factor or treatment effect), which can be represented either as a p-value and/or a confidence interval expressing the range of likely effect sizes (Visentin & Hunt, 2017). Moreover, it consists of a test for heterogeneity, whether the effect varies across the studies included.

Identification of the outcomes reported in the individual studies to be extracted for the meta-analysis is a crucial step in meta-analysis. The choice of effect measure reported depends on the type of outcome variable used in a particular study, which can be dichotomous, continuous or ordinal based on the outcome measures used. For dichotomous outcomes, risk ratio or relative risk (RR), odds ratio (OR) and risk difference (RD) can be used to assess differences between two groups. A mean difference or a standardized mean difference can be used to assess between-group comparisons of continuous outcomes (Higgins, Li, & Deeks, 2019). The interventions may have no effect, decrease the risk of an adverse outcome, or increase the chance of a desired outcome. In situations where the interventions are shown to reduce the occurrence of adverse outcomes, the OR and RR will be less than 1, and the RD negative. Where the intervention increases the occurrence of a desired outcome, the OR and RR will be greater than 1, with a positive RD (Deeks, Higgins, & Altman, 2008). Alternatively, when the intervention is ineffective the OR and RR will be near one, and the RD close to zero.

Two examples will be used to illustrate similarities and differences between RR and OR in this editorial. The first describes an RCT that assesses the impact of implementing a psychological support program on the status of depression as a dichotomous outcome event among 200 nurses caring for cancer patients. The second example describes a case-control study to assess

Substance Use Disorder among 620 subjects, of which 200 patients (cases) have a Generalized Anxiety Disorder and 420 people without Generalized Anxiety Disorder (controls) are recruited to assess Substance Use Disorder differences between the groups. The studies are summarized in Table 1.

3. DIFFERENCES BETWEEN RISK, ODDS, OR AND RR

The terms 'risk' and 'odds' are often used interchangeably. However, they have specific meanings in statistics and are calculated in different ways (Tufanaru et al., 2015). It is important to understand the differences between RR and OR when interpreting the results of a meta-analysis as there are subtle differences that need to be considered when interpreting the effect sizes (Deeks et al., 2008).

The risk describes the probability of an event, usually an adverse health outcome. It is a decimal number between 0 and 1, sometimes converted to a percentage or presented as number of events per 1000 people. For example, in the RCT example of 100 nurses who did not receive psychological support, 70 had depression – a risk of 0.7, 70 per 100 or 700 per 1000 nurses.

The RR is the ratio of risk of the outcome event in one group (e.g., intervention group or treated subjects) to the risk of the outcome event in another group (e.g., control group or untreated subjects. The risks in the treatment and control groups in the RCT example are 0.35 and 0.7 respectively, giving an RR of (35/100)/(70/100) = 0.5 (Table 1). This means that the risk of being depressed in the treatment group is half the risk in the group that did not receive the treatment, indicating better mental health for those receiving the psychological support program.

The RR provides a relative measure of association. However, we also have an absolute measure of association, the RD, commonly expressed as a percentage, calculated from the difference of the risks of the outcome event occurring between the two groups (i.e. the risk in one group

minus the risk for another group). For the study RCT example, this is 35/100 – 70/100 = –0.35 = –35% (Table 1b). The RD helps to put the RR into context, as it takes into account the incidence of the outcome. For example, in our illustrated RCT, the RR of 0.5 indicates reducing the risk of depression from 70% to 35% following participation in the program that corresponds to RD of 35%. The same RR statistics of 0.5 could also indicate reducing the risk of depression from 0.7% in the control group to 0.35% in the intervention group, which corresponds to RD of 0.35% which is much less (i.e. one hundredth) than former RD of 35%. Reporting the absolute RD effect size along with the RR may assist to avoid misinterpretation (Tufanaru et al., 2015).

Odds refer to the ratio of the probability of an event occurring to the probability of it not occurring within a group. Odds can also be defined as the risk (or probability) of an event occurring over the risk of the event being absent (Scott, 2008). Odds and risks can sometimes be computed through the following formulae: risk = odds/(1+odds), odds = risk/(1-risk) (Deeks et al., 2008) but these relationships depend on the study design and other factors (see below). Odds are also expressed as log-odds in some studies.

The interpretation of odds is more difficult than that of risk as researchers commonly think in terms of probability (risk) rather than odds even in studies that report OR (O'Connor, 2013). The confusion between risk and odds can lead to an incorrect interpretation of the OR as a multiplier for risk of the outcome (Martinez et al., 2017). One way to ensure the correct interpretation is to convert the odds to risks. When the outcome event is rare (i.e. less than 10%), the difference between the odds and risk is small, but when the outcome events are higher (e.g., clinical trials), the difference between the odds and risks would be large (Deeks et al., 2008). Since many clinical conditions have low incidence, the *rare disease assumption* approximation is often valid and may contribute to the common interpretation of OR as a risk multiplier.

The OR is preferred and the most popular measure of effect used in meta-analysis of dichotomous data (Bakbergenuly, Hoaglin, & Kulinskaya, 2019; Tufanaru et al., 2015). One reason for its popularity is that it is the main output of the logistic regression, the statistical method widely used in epidemiological studies (Martinez et al., 2017). Another reason is that OR can also be used in cross-sectional analytical studies, in addition to case-control studies. RR however requires longitudinal studies (cohort or RCT) which assess the incidence of the outcome in each group. In the absence of an assessment of incidence, the risk cannot be assessed. In cross-sectional and case-cohort studies only odds can be assessed, not risk. The OR is the ratio of odds in one group (i.e. the cases in a case control study or intervention group participants of an RCT) divided by the odds of the event in another group (i.e. the controls). In our case-control study example, the OR is (105/95)/(180/240) = 1.47 (Table 1b).

In case-control studies the OR is often a good approximation of RR since the outcome event is usually rare (i.e. less than 10%; Zlowodzki et al., 2007). The OR could also be calculated in RCTs or cohort studies (Knol et al., 2012); however, it can overestimate the magnitude of the effect or response in RCTs or cohort studies when the frequency of outcome event is large (Ospina, Nydam, & DiCiccio, 2012). When the outcome event rate increases or as the treatment effect becomes large, the OR will progressively diverge from the RR (Scott, 2008).

In our RCT illustration, the OR is (35/65)/(70/30) = 0.23 compared with the RR of 0.5 (Table 1a). As can be seen, the difference between the two is substantial because in our example the incidence of outcome event was high (54%). If the OR is incorrectly interpreted as a RR, it may lead to obtaining an overestimate of the risk, especially when the outcome is frequent.

It is preferable to avoid reporting odds ratios in RCTs and cohort studies to avoid such misinterpretations. If ORs are reported in these types of studies, the research team should be cautious about misinterpretation of it as a RR; particularly where the outcome is frequent or when the OR is not close to 1. For both OR and RR a value of 1 means the same estimated

effects for both interventions (i.e. intervention and control). In studies without any outcome event in the control group, neither OR not RR can be calculated and in studies where all subjects receiving the intervention experience the outcome event, the OR cannot be calculated (Knol et al., 2012).

5. SUMMARY AND CONCLUSION

Although both OR and RR can be used to compare relative likelihood of desired outcomes between groups (Simon, 2001); they are different, particularly when the outcome event is frequent. However, the difference between the RR and OR is usually small for large studies and should not be a concern in terms of the accuracy of results. The issue is where the OR is misinterpreted as a RR. Unfortunately, this kind of misinterpretation of effect measure is often seen in primary studies as well as in systematic reviews and meta-analyses (Deeks et al., 2008). It is worth remembering that the OR is a good approximation of RR only under certain circumstances (Hosmer, Lemeshow, & Sturdivant, 2013). Other subtle differences between the two ratios are apparent when assessing heterogeneity between studies, Egger's test and funnel plot asymmetry (Papageorgiou, Tsiranidou, Antonoglou, Deschner, & Jäger, 2015).

An appropriate use of OR is in case-control studies where, usually, a dichotomous outcome variable is considered, and logistic regression is often adopted for the data analysis (Lee, Tan, & Chia, 2009). It is suggested to report the results of RCTs and systematic reviews in terms of RR by default (J. Deeks, 1998) and to avoid use of "risk of X" when the odds are the measure of an event (O'Connor, 2013). Finally, as the RR is easier to interpret and researchers often use it by default, it is preferable where it can be calculated (Simon, 2001).

To avoid misinterpretation, researchers using cohort studies and randomized clinical trials should report RRs where possible, and other studies that use ORs should take care in interpreting this measure of effect (Knol et al., 2012). If the RR effect measure is used, it is

important to correctly interpret its magnitude, and the point estimate alone cannot be the basis for judging and interpreting the effectiveness of an intervention. The range of likely values for the RR should inform the interpretation of the effect, where statistical significance can be assumed if the 95% confidence interval (CI) around the RR does not include 1 (Scott, 2008). Where the intervention intends to prevent an undesirable outcome, an RR less than 1 indicates efficacy and in trials where the intervention aims to promote a positive event, a RR of more than 1 indicates intervention efficacy.

References

Bakbergenuly, I., Hoaglin, D. C., & Kulinskaya, E. (2019). Pitfalls of using the risk ratio in meta-analysis. *Research Synthesis Methods, 10*(3), 398-419. doi: 10.1002/jrsm.134

Deeks, J. (1998). When can odds ratios mislead? Odds ratios should be used only in case-control studies and logistic regression analyses. *British Medical Journal, 317*(7166), 1155-1156. doi: 10.1136/bmj.317.7166.1155a

Deeks, J. J., Higgins, J. P., & Altman, D. G. (2008). Chapter 9: Analysing data and undertaking meta-analyses In J. P. Higgins & S. Green (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions*. West Sussex: John Wiley & Sons, Ltd.

Feingold, A. (2017). Meta-analysis with standardized effect sizes from multilevel and latent growth models. *Journal of Consulting and Clinical Psychology, 85*(3), 262-266. doi:10.1037/ccp0000162

Higgins JPT, Li T, Deeks JJ (editors). Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions version 6.0* (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.

Hosmer, D.W., Lemeshow, S. & Sturdivant, R.X. (2013) Applied Logistic Regression. 3rd Edition, John Wiley & Sons, Hoboken, NJ. doi.org/10.1002/9781118548387

Knol, M. J., Algra, A., & Groenwold, R. H. H. (2012). How to Deal with Measures of Association: A Short Guide for the Clinician. *Cerebrovascular Diseases, 33*(2), 98-103. doi: 10.1159/000334180

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*(863). doi: 10.3389/fpsyg.2013.00863

Lee, J., Tan, C. S., & Chia, K. S. (2009). A practical guide for multivariate analysis of dichotomous outcomes. *Annals of the Academy of Medicine, 38*(8), 714-719.

Martinez, B. A. F., Leotti, V. B., Silva, G. S.E., Nunes, L. N., Machado, G., & Corbellini, L. G. (2017). Odds Ratio or Prevalence Ratio? An Overview of Reported Statistical Methods and Appropriateness of Interpretations in Cross-sectional Studies with Dichotomous Outcomes in Veterinary Medicine. *Frontiers in Veterinary Science, 4*(193). doi: 10.3389/fvets.2017.00193

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS med, 6*(7), e1000097. doi: 10.1371/journal.pmed.1000097

O'Connor, A.M. (2013). Interpretation of odds and risk ratios. *Journal of Veterinary Internal Medicine, 27*(3), 600-603. doi: 10.1111/jvim.12057.

Ospina, P. A., Nydam, D. V., & DiCiccio, T. J. (2012). Technical note: The risk ratio, an alternative to the odds ratio for estimating the association between multiple risk factors and a dichotomous outcome. *Journal of Dairy Science, 95*(5), 2576-2584. doi: 10.3168/jds.2011-4515

Papageorgiou, S. N., Tsiranidou, E., Antonoglou, G. N., Deschner, J., & Jäger, A. (2015). Choice of effect measure for meta-analyses of dichotomous outcomes influenced the identified heterogeneity and direction of small-study effects. *Journal of Clinical Epidemiology, 68*(5), 534-541. doi: 10.1016/j.jclinepi.2015.01.004

Schäfer, T., & Schwarz, M. A. (2019). The Meaningfulness of Effect Sizes in Psychological Research: Differences Between Sub-Disciplines and the Impact of Potential Biases. *Frontiers in Psychology, 10*(813). doi: 10.3389/fpsyg.2019.00813. eCollection 2019.

Scott, I. (2008). Interpreting risks and ratios in therapy trials. *Australian Prescriber,* 31:12–16. doi:10.18773/austprescr.2008.008

Simon, S. D. (2001). Understanding the odds ratio and the relative risk. *Journal of Andrology, 22*(4), 533-536. doi:10.1002/j.1939-4640.2001.tb02212.x

Thapa, D.K., Visentin, D.C., Hunt, G.E., Watson, R. and Cleary, M. (2020), Being honest with causal language in writing for publication. *Journal of Advanced Nursing, 76,* 1285-1288. doi:10.1111/jan.14311

Tufanaru, C., Munn, Z., Stephenson, M., & Aromataris, E. (2015). Fixed or random effects meta-analysis? Common methodological issues in systematic reviews of effectiveness. *International Journal of Evidence-Based Healthcare, 13*(3), 196-207. doi: 10.1097/XEB.0000000000000065.

Visentin, D. C., & Hunt, G. E. (2017). What do the stats mean? Improving reporting of quantitative nursing research. *International Journal of Mental Health Nursing, 26*(4), 311-313. doi: 10.1111/inm.12352

Zlowodzki, M., Poolman, R. W., Kerkhoffs, G. M., Tornetta III, P., Bhandari, M., & International Evidence-Based Orthopedic Surgery Working Group. (2007). How to interpret a meta-analysis and judge its value as a guide for clinical practice. *Acta Orthopaedica, 78*(5), 598-609. doi: 10.1080/17453670710014284

Table 1.

a) Number of subjects assigned to interventions and outcome events (100 each in treatment and control groups) in the RCT example

Intervention	Depression status		
	Yes	No	Total
Treatment group Psychological Support	35	65	100
Control group No Psychological Support	70	30	100

b) Number of subjects with and without current substance use disorder (200 patients and 420 controls) in the case-control study example

Substance Use Disorder status	Generalized Anxiety Disorder status	
	Yes	No
Yes	105	180
No	95	240
Total	200	420