

# The use of odds ratio in the large population-based studies: Warning to readers

Luigi Di Lorenzo  
Valeria Coco  
Francesco Forte  
Giovanni Felice Trinchese  
Alfonso Maria Forte  
Marco Pappagallo

Rehabilitation Unit, Gait Analysis Lab, Neuroscience Department. "RUMMO" Hospital, Benevento, Italy  
Biomedical Research Centre, Gruppo Forte Salerno  
The New Medical Home for Pain Management and Medical Mentoring, New York, USA

## Corresponding author:

Luigi Di Lorenzo  
Rehabilitation & Gait Analysis Lab Unit, AO RUMMO  
Benevento, Italy  
E-mail: luigidilorenzo2005@libero.it

---

## ALL AUTHORS STATE THAT:

*"The undersigned Authors transfer the ownership of copyright to your Journal should their work be published in this journal. They state that the article is original, has not been submitted for publication in other journals and has not already been published. They state that they are responsible for the research that they have designed and carried out; that they have participated in drafting and revising the manuscript submitted, which they approve in its contents. They also state that the research reported in the paper was undertaken in compliance with the Helsinki Declaration and the International Principles governing research on animals".*

---

## Summary

**When researchers conduct large prospective studies, they provide results generating statistical analysis; therefore readers need considerable familiarity with descriptive and inferential statistics. If quantitative judgments are based on interpreting odds ratios as though they were relative risks, they are unlikely to be seriously in error. Because of the calculating method, the OR is often less precise than the RR in estimating the strength of an association, and this should definitely be kept in mind by anyone who reads and interprets the results of a large population based-study.**

*KEY WORDS: statistic analysis, relative risk, odds ratio, risk ratio, multivariate analysis.*

## Introduction

When researchers conduct large prospective studies, they provide results generating statistical analysis<sup>1</sup> and therefore readers need considerable familiarity with descriptive and inferential statistics<sup>2-4</sup>. When preparing a research project, it is also important to make realistic, well-researched and supported assumptions<sup>1</sup>. For these reasons, distinguished journals request "authors reporting clinical trials to indicate who carried out the analysis" and how to generate analysis to determine if an exposure to a treatment is associated with a specific outcome<sup>5</sup>. In recent years, odds ratio (OR) has become widely used in medical reports and Bland<sup>6</sup> clearly explained why: firstly, OR provides an estimate (with confidential interval) for the relationship between two binary ("yes or no") variables. Secondly, it enables us to examine the effects of other variables on that relationship, using logistic regression. Thirdly, it has a special and very convenient interpretation in case-control studies<sup>6</sup>. Despite its usefulness – he concluded – OR can cause difficulties in interpretation<sup>6</sup>. Unfortunately, after more than a decade, after reading recent large cohort studies we think that it is still useful "to review this debate and also discuss the choice of odds ratios in logistic regression and case-control studies"<sup>6-8</sup>.

## Relative risks as measures of outcome changes

The most direct way to determine if an exposure to a treatment is associated with an outcome is to prospectively follow two groups and observe the frequency with which each group develops the outcome. When researchers carry out a study, after describing the main measures of the frequency of a symptom, they usually study the correlations between variables through the analysis of measures of association. The most common measures of association are the relative risks and odds ratio. However, we should keep on mind that there are other similar measures, with the same meaning but different calculation methods (Hazard ratio, rate ratio, etc.) and others with equal calculation method but with different meanings (attributable risk or risk differential)<sup>7</sup>. The relative risk (RR) compares the frequency of an outcome between groups and represents the ratio of the probability of the out-

come occurring in the treatment group *versus* a control group. Although it is well-known that the two measures evaluate different quantities in general, the odds ratio has been misinterpreted as the relative risk in some studies, and thus contributed to incorrect conclusions. RR is a key measure and can be calculated in cohort studies but not in cross-sectional or case-control studies<sup>9,7,8</sup>. Indeed, the OR is an estimate of the strength of an association between an exposure and a disease. It can be easily shown that the OR is a good approximation to the RR when the incidence or risk rate is low – for instance, in rare diseases – and can largely overestimate the RR when the outcome is common in the study population<sup>9</sup>. Just as an example, we cite OR based conclusions of two recent large population-based studies<sup>10,11</sup>. In the first neurological prospective study<sup>10</sup>, authors studied 10,528 consecutive patients admitted with acute ischemic stroke, analyzing the association of atrial fibrillation (AF) with mortality and disability. On the base of the analysis, they concluded that AF is associated with an increased risk of death and severe disability, suggesting that increased uptake of Oral Anticoagulant therapy in suitable candidates could substantially reduce risks. In the second study cited<sup>11</sup>, authors generated a large population-based case-control study including 67,162 patients, in which they estimated OR for pneumonia hospitalization among persons who were divorced. In these studies, both research groups used OR as key measure of outcome changes. A group of authors<sup>11</sup> used OR as it is probably the best available measure of association used in case-control studies where it is not possible to calculate the “incidence”, and the others did it because it can provide greater precision in some complex statistical tests<sup>11</sup>. OR is therefore a measure widely used in prospective large population-based study<sup>2-4,7</sup> in which the researchers use to calculate it in order to express the results of the most complex statistical tests.

### The odds ratio: misconceptions and doubts

Odds ratio (OR) represents the approximate estimate of the relative risk of the strength of association between risk factor and disease. It is calculated by the formula  $OR = A \cdot C / B \cdot D = AD / B \cdot C$ . Behind misconceptions<sup>5-8</sup>, OR is not considered a perfect and true measure of risk as it relates to the probability of having a disease already, while the term “risk” is implicit in the idea of an event that will occur in the future. With an extreme simplification we could say that, from a conceptual point of view, OR and RR are relatively similar, and the OR approximates RR in cohort studies and it is well known that such an approximation is poor and can generate misleading conclusions<sup>5-8</sup>. Odds ratios are a way of representing probability and OR is used in the majority of systematic reviews and large cohort studies, but if we interpret it as a relative risk, we could overstate the effects. The extent of overstatement increases as both the initial risks increases and the OR departs from units. Standard tables of interpretation of OR and RR

report as “modest” the association when OR is > zero up to 3, 0 (up to 1, 3 “faint”, up to 1, 7 “modest” and up to 3, 0 “moderate”). OR calculations in the context of a large number of cases is not easy and OR can be often less precise when estimating the strength of the RR in an association<sup>12</sup>. In particular, OR tends to increase the result, to make it look greater than it is: OR is always higher than RR if this is greater than 1, it is always lower than RR if this is lower. This approximation is negligible in studies where the incidence or prevalence of the disease is relatively low. The difference between OR and RR becomes greater, instead, the more the incidence or prevalence are elevated<sup>4,5</sup>. Early last year, Wang<sup>13</sup> presented a method extending a previous popular method<sup>14</sup> reporting it to be able to convert the odds ratio to the relative risk, if an odds ratio and/or a confidence interval as well as the sample sizes for the treatment and control group are available. Namely, the method developed is reported to be useful to approximate the RR based on the adjusted OR from logistic regression or other multiple regression model<sup>13,14</sup>.

### Conclusion

As stated by Wang<sup>13</sup>, epidemiologists are often interested in “*comparing a risk of a binary outcome between a treatment and control group, or between exposed and unexposed*”. Such an outcome can be an onset of a symptom or condition and the OR and RR is the important measures. In a case-control study, OR is often used as a surrogate for RR. OR evaluates whether the probability of a study outcome is the same for two groups and can approximate RR or risk ratio, which is a more direct measure than the odds ratio. The approximation between the two values is negligible in studies where the incidence or prevalence of the disease is relatively low, but the difference is more significant as most are high incidence or prevalence. In other words, as pointed out by several eminent<sup>7?</sup> “The overestimation of the strength of the association on the part of the RR is much more pronounced as there is a high prevalence or incidence. So far, if quantitative judgments are based on interpreting odds ratios as though they were relative risks, they are unlikely to be seriously in error. Because of the calculation method, OR is often less precise than RR in estimating the strength of an association, and this should definitely be kept in mind by anyone who reads and interprets the results of a large population based-study.

### Financial disclosure statement

NA

### References

1. Wang D, Bakhai A, Del Buono A, Maffulli N. Sample size determination for clinical research. *Muscle Ligament Tendon Journal*. 2013;3(3)116-117.

2. Geenhalgh T. How to read a paper: the basics of EBM BMJ Books, 2006.
3. Bophal R. Concepts of epidemiology: An integrated introduction to the ideas, theories, principles and methods of epidemiology. Oxford University Press, USA. 2002.
4. Senn S. Statistical Issues in Clinical Trials in Neurology Clinical Trials in Neurology. 2001;95-103.
5. Padulo J, Oliva F, Frizziero A, Maffulli N. Muscle, Ligaments and Tendons Journal. Basic principles and recommendations in clinical and field science research. MLTJ. 2013;3(4), 250-252.
6. Bland MJ, Altman DG. The odds Ratio. Statistics note. BMJ. 2000; vol.320:1468-1469.
7. Manzoli L, Villari P, Boccia A. Measures of association: relative risks and similar. (3)27-37. In textbook "Epidemiology and Management in Health (Italian).Edi-ermes. 2008. Milano. ISBN 978-88-7051-320-2
8. Davies H. When can Odds Ratio mislead? BMJ. 1998;316: 989.
9. Schmidt AO, Kohlmann T. When to use the odds ratio or the relative risk? Int J Public Health. 2008;53:165-167.
10. McGrath ER, Kapral MK, Fang J, et al., Association of atrial fibrillation with mortality and disability after ischemic stroke Neurology. 2013;vol. 81, no. 9:825-832.
11. Mor A, Ulrichsen SP, Svensson E, Berencsi K, Thomsen RW. Does marriage protect against hospitalization with pneumonia? A population-based case-control study Clinical Epidemiology. 2013;5(1):397-405.
12. Kabaila P, Tissera D. Effect of a preliminary test of homogeneity of stratum-specific odds ratios on their confidence intervals. Electronic Journal of Statistics. 2012;6:672-685.
13. Wang Z. Converting Odds ratio to relative risk in cohort studies with partial data information.
14. Zhang J, Yu KF. "What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. Journal of the American association. 1998;280 (19): 1690-1691.