

Risks in Biomedical Science – Absolute, Relative, and Other Measures

RISK DIFFERENCE, RISK RATIO, AND ODDS RATIO

Pneumonia is a life-threatening disease for which a variety of treatments have been developed. In the 1800s, the practice of blood-letting was used as treatment for a number of diseases. Pierre Louis (1787–1872) is credited for conducting an experiment assigning 77 pneumonia patients to ‘bleed early’ ($n=41$) or ‘bleed late’ ($n=36$) groups. He analyzed the numbers of patients in each group who survived: 23/41 (survival rate or probability of 56%) in the ‘bleed early’ group versus 27/36 (75%) in the ‘bleed late’ group. In this example, the ‘bled late’ group survived more [Table 1].

Next, how do we determine how much more or better a survival statistic is? We can compare survival rate in terms of difference or ratio, that is, $75\%-56\%=19\%$ (risk difference [RD]) or $75/56=1.34$ (risk ratio [RR]). [Strictly speaking, survival ‘rate’ is incorrect.^[1] Also, rate ratio can be computed similarly, and hazard ratio (HR) is instantaneous rate/risk ratio; see below].

In contrast, when we compare sexes of babies, we may use a somewhat different ratio. As an example, if there were 20 boys and 40 girls, there is a ratio of 20:40 or $20/40=0.5$, such that there are twice as many girls as boys, or half as many boys as girls (of course, this is an extreme example for illustration; real ratio may be like 105:100). Using this measure, namely ‘odds’, the likelihood of survival is 1.3 (23/18) times the likelihood of death for the ‘bleed early’ group, and it is 3 (27/9) times that for the ‘bled late’ group (more rigorously, in probability). This measure is often used in gambling, e.g., the odds of winning vs. losing. In terms of odds, the ‘bled late’ group shows higher survival – but again how much higher? Odds difference (OD) is $3-1.3=1.7$ and odds ratio (OR) is $3/1.3=2.3$. Then we arrive at four values: RD, RR, OD, and OR. When two treatments are equal in performance (in population parameter or expectation, not necessarily in actual

sample), RD and OD will be 0, and RR and OR will be 1, which often serves as the null hypothesis or finding. We often refer to difference as ‘absolute measure’ and to ratio as ‘relative measure’.^[2] In practice, OR is widely used, partly because of computational advantages and convenience (e.g., logistic regression with covariates, case-control studies), whereas OD is rarely used.

ABSOLUTE RISK REDUCTION (ARR) AND NUMBER NEEDED TO TREAT (NNT)

There are alternative (and possibly less familiar) ways to express risk difference (RD) and its function: $RD=ARR$ (absolute risk reduction) and $1/RD=NNT$ (number needed to treat).^[3] For a graphic example:



Arm A:  and Arm B: , where red denotes ‘cured’ and blue ‘uncured’. Arm A yielded 70% (7/10) cure and Arm B yielded 50% (5/10) cure; hence, RD is 20% ($=70\%-50\%$). If 10 persons received Arm A (instead of Arm B), 2 more patients can be cured, or equivalently, 20% of extra benefit is expected. Thus, if we treat 10 patients, we can expect 2 additional cures and if we treat 20 patients, 4 additional cures are expected, etc. The reciprocal of the number, RD, can be interpreted as the number needed to treat (NNT) to get 1 expected *additional* cure. In this example, $NNT=5$. A treatment showing a lower NNT could be clinically better, assuming that all other conditions are comparable (as in a well-conducted controlled trial or with causal RD estimated from an observational study). The most ideal NNT (over a given period) is 1, where everyone improves with treatment and no one improves with control. [Strictly speaking, the NNT to reduce the adverse outcome *count* by 1. After all, if we waited longer it might be that everyone improved and treatment just sped up the process, making $RD=0$ and

Table 1: Pierre Louis’s bloodletting experiment on pneumonia

Table	Dead	Alive	Total	Dead/total ratio = Risk	Dead/alive ratio = Odds
Bleed early	18	23	41	18/41=44%	18/23=78%
Bleed late	9	27	36	9/36=25%	9/27=33%
Summary measures				Risk difference =44–25=19% Risk ratio=44/25=1.8	Odds difference=78–33=45% Odds ratio=78/33=2.4

Risk difference ($=ARR =1/NNT$). ARR, absolute risk reduction; NNT, number needed to treat

NNT = ∞.] Number needed to harm (NNH) or number needed to screen (NNS) can be similarly calculated.^[4]

CONFIDENCE INTERVALS, COMPUTATIONS AND MATHEMATICAL PROPERTIES

After identifying a point estimate (RD/NNT/RR/OR), we may want to obtain a confidence interval (CI) for statistical inference about the population. Note that communication of the meaning and interpretation of CIs is challenging. CIs mean only one thing: in repeated sampling, a 95% CI contains the true parameter 95% of the time.^[5,6]

Despite their popularity, most software packages do not provide all four measures together. Thus, we provide a tool to calculate RD/NNT/RR/OR and corresponding 95% CIs in easily downloaded and used Excel format: <https://tinyurl.com/OR-RR-RD-NNT>. Here, we should note some mathematical annoyance: when we use a reciprocal, 0 becomes +/−∞. Thus, ∞ may be inside the CI of NNT!^[7] Users should not panic.

Also, when assessing and interpreting risk, RR may be more intuitive than OR as we generally want to know the probability of success (=number of successes/number of attempts) rather than odds (=number of successes/number of failures). If we want RR but only OR is available, it is useful to understand the mathematical relationship, say, for conversion: With D=disease and E=exposure,

$$\begin{aligned} \text{OR} &= \frac{P(D|E)/P(D|\text{not } E)}{P(\text{not } D|E)/P(\text{not } D|\text{not } E)} \\ &= \text{RR} * \frac{P(\text{not } D|\text{not } E)}{P(\text{not } D|E)}. \end{aligned}$$

If $\text{RR} > 1$, then $\text{OR} > \text{RR}$; and if $\text{RR} < 1$, then $\text{OR} < \text{RR}$.

For more on definitions and relationships between OR vs. RR vs. HR, see references.^[8-10]

CONSORT

The CONSORT guidelines are invaluable for current best practices although there is room for disagreement and further clarification or specialization. Let us review their recommendations: (<http://www.consort-statement.org/checklists/view/657-harms/1015-binary-outcomes>).

17a. Outcomes and estimation

For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)

For each outcome, study results should be reported as a summary of the outcome in each group (for example, the number of participants with or without the event

and the denominators, or the mean and standard deviation of measurements), together with the contrast between the groups, known as the effect size. For binary outcomes, the effect size could be the risk ratio (relative risk), odds ratio or risk difference; for survival time data, it could be the hazard ratio or difference in median survival time; and for continuous data, it is usually the difference in means. Confidence intervals should be presented for the contrast between groups. A common error is the presentation of separate confidence intervals for the outcome in each group rather than for the treatment effect. Trial results are often more clearly displayed in a table rather than in the text.

For all outcomes, authors should provide a confidence interval to indicate the precision (uncertainty) of the estimate. A 95% confidence interval is conventional, but occasionally other levels are used. Many journals require or strongly encourage the use of confidence intervals. They are especially valuable in relation to differences that do not meet conventional statistical significance, for which they often indicate that the result does not rule out an important clinical difference. The use of confidence intervals has increased markedly in recent years, although not in all medical specialties. Although *P* values may be provided in addition to confidence intervals, results should not be reported solely as *P* values

For both binary and survival time data, expressing the results also as the number needed to treat for benefit or harm can be helpful (see item 21).

17b. Binary outcomes

For binary outcomes, presentation of both absolute and relative effect sizes is recommended

When the primary outcome is binary, both the relative effect (risk ratio (relative risk) or odds ratio) and the absolute effect (risk difference) should be reported (with confidence intervals), as neither the relative measure nor the absolute measure alone gives a complete picture of the effect and its implications. Different audiences may prefer either relative or absolute risk, but both doctors and lay people tend to overestimate the effect when it is presented in terms of relative risk. The size of the risk difference is less generalisable to other populations than the relative risk since it depends on the baseline risk in the unexposed group, which tends to vary across populations. For diseases where the outcome is common, a relative risk near unity might indicate clinically important differences in public health terms. In contrast, a large relative risk when the outcome is rare may not be so important for public health (although it may be important to an individual in a high-risk category).

Acknowledgement

The author thanks Ms. Caron Modeas for English editing service, and Drs. Tancredi, Elston, Greenland, Shuster, Zhao, Kaufman and Boos for useful advice and comments.

Financial support and sponsorship

H. Bang is partly supported by the National Institutes of Health through grant UL1 TR001860 and R01 AR076088.

Conflicts of interest

There are no conflicts of interest.

Jeehyoung Kim¹, Heejung Bang²

¹Department of Orthopedic Surgery, Seoul Sacred Heart General Hospital, Seoul, Korea²Division of Biostatistics, Department of Public Health Sciences, School of Medicine, University of California, Davis, CA, USA

Address for correspondence: Jeehyoung Kim, Department of Orthopedic Surgery, Seoul Sacred Heart General Hospital, Seoul, Korea.
E-mail: jeehyoung.kim@gmail.com

Received: 18 April 2020 **Revised:** 18 April 2020
Accepted: 19 April 2020 **Published:** 22 July 2020

REFERENCES

1. Elandt-Johnson RC. Definition of rates: some remarks on their use and misuse. *Am J Epidemiol* 1975;102:267-71.
2. Rose G. *Rose's Strategy of Preventive Medicine* 2008. Oxford University Press.
3. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Eng J Med* 1988;318:1728-33.
4. Fenton JJ, Weyrich MS, Durbin S, Liu Y, Bang H, Melnikow J. Prostate-specific antigen-based screening for prostate cancer:

- Evidence report and systematic review for the US Preventive Services Task Force. *J Am Med Assoc* 2018;319:1914-31.
5. Greenland S, Poole C. Problems in common interpretations of statistics in scientific articles, expert reports, and testimony. *Jurimetrics* 2011;113-29.
6. Kim J, Bang H. Three common misuses of P values. *Dent Hypotheses* 2016;7:73-80.
7. Altman DG. Confidence intervals for the number needed to treat. *Br Med J* 1998;317:1309-12.
8. Symons MJ, Moore DT. Hazard rate ratio and prospective epidemiological studies. *J Clin Epidemiol* 2002;55:93-9.
9. Jewell NP. *Statistics for Epidemiology* 2003. Chapman and Hall/CRC.
10. Rothman KJ, Greenland S. *Modern Epidemiology* 1998. Philadelphia: Lippincott-Raven.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

Access this article online

Quick Response Code:



Website:

www.dentalthypotheses.com

DOI:

10.4103/denthyp.denthyp_41_20

How to cite this article: Kim J, Bang H. Risks in Biomedical Science – Absolute, Relative, and Other Measures. *Dent Hypotheses* 2020;11:69-71.

© 2020 Dental Hypotheses | Published by Wolters Kluwer - Medknow