International Review of Pure and Applied Mathematics
Volume 15 • Number 1 • Jan-June 2019

ON COMPARATIVE EVALUATION OF ODDS RATIO, RELATIVE RISK AND ATTRIBUTABLE RISK

Shishir Kumar Jha

***Abstract*:** Relative risk (RR) is normally used in analysing the connection of exposure (causing disease) and the consequence of exposure resulting in mortality (or some serious disease e.g. stroke, heart attack etc.). However, in many situations odds ratio (OR) and attributable risk (AR) are either equally useful or sometimes give better understanding of the association between exposure and the outcome than the relative risk. Using a secondary data relating to eating habit of a group of people (resulting in prevention of deadly disease), a critical analysis has been performed to distinguish the findings of relative risk, attributable risk and odds ratio.

***Keywords*:** Attributable Risk, Hazard Rate, Odds Ratio, Relative Risk.

1. INTRODUCTION

In epidemiological and clinical trial studies, comparing risks of occurring of an event between groups (exposed and unexposed to a particular factor), the statistic under consideration is generally taken as relative risk. Probability of happening of some events in a group of people suffering from some medical complications are generally termed as risk, however epidemiologists also call it as incidence (vide: Savitz, 1992). Risk is generally termed as chance of an individual (without a disease) developing the disease during a span of time. The relative risk (RR) is defined as ratio of two incidence rates and so is relevant in prospective studies, however, the comparable idea of odds ratio (OR) is the other statistic considered to compare the risks between two groups of population with respect to happening of some events (vide: Indrayan, 2008). Odds ratio is pertinent to retrospective and cross-sectional studies. Both relative risk (RR) and odds ratio (OR) are used to measure the association between a chronic disease and possible hazard components. Mantel and Haenszel (1959) study is considered as beginning of the development of relative risk regression model. Cox (1972) developed a regression model with a set of explanatory variables to study relative risk by defining the hazard function. Further, the relative risk is also defined as ratio of two hazard rates which is fundamental to medical research, especially in clinical and epidemiological studies. Odds ratio is the ratio of proportion of happening of an event to not happening of that event as described by Bland and Altman (2000). Andrade (2015) also recommended to compute odds ratio in case- control studies and logistic regression analysis. In a case-control study, OR determines the link of an exposed population with some hazard and an outcome. Greenland (1987), Holland (1989) and Greenland and Holland (1991) discussed odds ratio estimation and showed that how conceptually it is different from relative risk. Similarly, difference in the incidence of happening of the event in the exposed group of population and incidence of happening of that in the unexposed population is termed as attributable risk (AR) (vide: Kirch, 2008). Davies, et.al. (1998) critically distinguished odds ratio and relative risk and showed that odds ratio generally, overstate the risk in comparison to relative risk. Andrade (2015) explains with examples the utility of attributable risk, relative risk and odds ratio in understanding the findings of clinical experiments.

2. METHODOLOGY

RR is the ratio of the possibility of the outcome (such as disease, death etc.) in those with exposure to risk compared with those without exposure. The structure of the study is summarized in the following 2 x 2 contingency **table 1**. This study is restricted to dichotomous variables only. Suppose, there are two variables under our study, exposure variable (A) and outcome variable (B).

Table 1

Variable B (Outcome)	Variable A (Exposed to risk)		
	Yes	No	Total
Yes (Disease)	n_{11}	n_{21}	$n_{.1}$
No (No disease)	n_{12}	n_{22}	$n_{.2}$
Total	$n_{1.}$	$n_{2.}$	n

The RR is given by the ratio of development of disease for the people unprotected against certain risk factor to the ratio of development of disease for the people protected against the risk factor. The OR is given by the proportion of odds endorsing development of disease in the unprotected group to the odds endorsing development of disease in the protected group.

i.e. $$\text{RR} = \frac{\frac{n_{11}}{n_{1.}}}{\frac{n_{21}}{n_{2.}}} = \frac{n_{11}}{n_{21}}\frac{n_{2.}}{n_{1.}} \tag{1}$$

and $$\text{OR} = \frac{\frac{\frac{n_{11}}{n_{1.}}}{\frac{n_{12}}{n_{1.}}}}{\frac{\frac{n_{21}}{n_{2.}}}{\frac{n_{22}}{n_{2.}}}} = \frac{n_{11}n_{22}}{n_{12}n_{21}} \tag{2}$$

Suppose, we multiply any column of the observed values by m (m > 0), e.g. if in the column one the intensity of the exposure changes such that the n_{11} becomes m$n_{11}$and n_{12} *becomes* mn_{12} in table 1 then

the revised relative risk $= \frac{\mathrm{m}n_{11}}{n_{21}}\frac{n_{2.}}{mn_{1.}} = \frac{n_{11}}{n_{21}}\frac{n_{2.}}{n_{1.}}$ [No change in RR value of eq. (1)] (3)

and the revised odds ratio $= \frac{mn_{11}n_{22}}{mn_{12}n_{21}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$ [No change in OR value of eq. (2)] (4)

If the outcome (disease cases) n_{11} and n_{21} in table 1 are multiplied by m (m > 0) then the revised relative risk $= \frac{n_{11}}{n_{21}} \frac{mn_{21}+n_{22}}{mn_{11}+n_{12}}$ (5)

If the outcome (no disease cases) n_{12} and n_{22} in table 1 are multiplied by m (m > 0) then

the revised relative risk $= \frac{n_{11}}{n_{21}} \frac{n_{21}+mn_{22}}{n_{11}+mn_{12}}$ (6)

and the revised odds ratio $= \frac{mn_{11}n_{22}}{mn_{12}n_{21}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$ = OR [No change in OR value of eq. (2)] (6i)

We observe that changes in exposure intensity does not result in any changes in relative risk or odds ratio. However, Changes in one level of outcome keeping the second level of outcome unchanged affects the value of RR but has no effect on the value of OR.

$$\text{RR} = \frac{\frac{n_{11}}{n_{1.}}}{\frac{n_{21}}{n_{2.}}} = \frac{\textit{probability of the occurance of an event in exposed group}}{\textit{probability of the occurance of an event in unexposed group}}$$ and

$$\text{OR} = \frac{\frac{\frac{n_{11}}{n_{1.}}}{\frac{n_{12}}{n_{1.}}}}{\frac{\frac{n_{21}}{n_{2.}}}{\frac{n_{22}}{n_{2.}}}} = \text{RR} \cdot \left(\frac{n_{1.}n_{22}}{n_{2.}n_{12}}\right) = \text{RR} \cdot \left(\frac{1-\frac{n_{21}}{n_{2.}}}{1-\frac{n_{11}}{n_{1.}}}\right) \text{ [using equations (1) and (2)]} \quad (7)$$

From equation (7), we see that OR is approximately equal to RR if the outcome probabilities are small.

Similarly, if $\frac{n_{21}}{n_{2.}} \geq \frac{n_{11}}{n_{1.}}$ then OR $\leq$ RR and

if $\frac{n_{21}}{n_{2.}} \leq \frac{n_{11}}{n_{1.}}$ then OR $\geq$ RR.

Now the attributable risk (AR) $= \frac{n_{11}}{n_{1.}} - \frac{n_{21}}{n_{2.}}$ (8)

and the AR fraction $= \frac{\frac{n_{11}}{n_{1.}} - \frac{n_{21}}{n_{2.}}}{\frac{n_{11}}{n_{1.}}} = 1 - \frac{\frac{n_{21}}{n_{2.}}}{\frac{n_{11}}{n_{1.}}} = 1 - \frac{1}{RR}$ [using (1)] (9)

Therefore, AR percentage $= (1 - \frac{\frac{n_{21}}{n_{2.}}}{\frac{n_{11}}{n_{1.}}})\text{x}100 = (1 - \frac{1}{RR})\text{x}100$ (10)

If we multiply any column of the observed values by m (m > 0) in table 1 then there is no change in the attributable risk, i.e. changes in exposure intensity does not result in any changes in AR like RR and OR.

If we multiply row one of the observed values in table1 by m (m > 0), then we have

$$\text{AR} = \frac{mn_{11}}{mn_{11}+n_{12}} - \frac{mn_{21}}{mn_{21}+n_{22}} \quad (11)$$

$\Rightarrow$ Changes in one level of outcome keeping the second level of outcome unchanged affects the value of AR.

3. NUMERICAL ILLUSTRATION

Kaelin and Bayona (2004) data (table 2) on the study of advantages of eating fish with respect to risk of stroke is used to analyse the comparative studies of attributable risk, relative risk and odds ratio. In earlier studies, it has been established that eating fish contributes in preventing the stroke. Therefore, we assume that the population who does not eat fish are exposed to the risk of stroke. Further, we ignore the fact that some members of the population eat fish occasionally. Such members of the population are taken in the category "never".

Table 2 (Eating fish and stroke)

Outcome	Eating Fish		
	Never (Exposed)	Frequently(unexposed)	Total
Stroke cases	n_{11} (82)	n_{21}(23)	$n_{.1}$(105)
No stroke cases	n_{12} (1549)	n_{22} (779)	$n_{.2}$ (2328)
Total	$n_{1.}$ (1631)	$n_{2.}$ (802)	n (2433)

RR of stroke for the group of persons with no fish in their menu in comparison to the group of people who eat fish frequently is

$$\text{RR} = \frac{65764}{37513} = 1.753 \text{ [using (1)]} \tag{12}$$

$$\text{OR} = \frac{63878}{35627} = 1.793 \text{ [using (2)]} \tag{13}$$

$$\text{AR} = \frac{82}{1631} - \frac{23}{802} = 0.0215 \text{ [using (8)]} \tag{14}$$

AR fraction = 0.4295 [using (9)]

and AR percentage = 0.4295 x 100 = 42.95

If we multiply any column of the observed values in the table 2 by m (m > 0) then the values of RR, OR and AR are the same as the values obtained in equations (12), (13) and (14) respectively, which proves that the change in the intensity of the exposure does not make any impact on the RR, OR and AR.

If we multiply row 1 of observed values (i.e. stroke cases) by m = 0.0001, 0.001, 0.01, 0.1, 0.5, 1, 2, 10, 100 and 1000, then the corresponding relative risks (RR) are obtained as 1.79296, 1.79292, 1.79255, 1.78879, 1.773, 1.753, 1.717, 1.518, 1.126 and 1.015 respectively [using equation (5)].

$$\Rightarrow \text{As } m \to \infty, \text{RR} \to 1 \tag{15}$$

Similarly, if we multiply row 2 of observed values (i.e. no stroke cases) by m = 0.0001, 0.001, 0.01, 0.1, 0.5, 1, 2, 10, 100 and 1000, then the corresponding relative risks (RR) are obtained as 1.0014, 1.0147, 1.1260, 1.5185, 1.7170, 1.7531, 1.7725,1.7888, 1.7925 and 1.7929 respectively [using equation (6)].

$$\Rightarrow \text{As } m \to 0, \text{RR} \to 1 \tag{16}$$

From equations (15) and (16) it is quite clear that exposure does not make any significant impact on the outcome if stroke cases are very large in comparison to non-stroke cases.

Now, if we multiply row 1 of observed values (i.e. stroke cases) by m = 0.1, 0.5, 1, 2, 10, 50, 75, 100 and 1000, then the corresponding attributable risks (AR) are obtained as 0.0023, 0.01124, 0.0215, 0.03998, 0.1182, 0.1296, 0.1099, 0.094 and 0.0142 respectively.

$$\Rightarrow \text{As } m \to 0, \text{AR} \to 0 \tag{17}$$

AR takes the maximum value when m lies between 50 and 75 for the given data.

$$\text{Also, as } m \to \infty, \text{AR} \to 0 \tag{18}$$

⇒ risk of stroke in exposed population (those who do not eat fish) is highest when m lies between 50 and 75 for the given data. Risk of stroke due to exposure is very low if $m \to 0$ or $m \to \infty$.

4. CONCLUSION

Intake of fish may be considered to be protective against the stroke, which is evident from the findings of (12) and (13). Increasing the intensity of exposure does not have any impact on odds ratio, attributable risk and relative risk. From (15) and (16) it is safe to conclude that if stroke cases increase proportionately in both the exposed and unexposed groups of population keeping non-stroke cases unchanged then the relative risk tends to 1 (RR taking value 1 means exposures have no association with the disease). However, odds ratio remains unchanged as proved in (6i) even if stroke cases increase proportionately in both the exposed and unexposed groups of population. Findings of (17) and (18) with respect to attributable risk indicates that the risk of disease tends to zero in both the extreme situations (i.e. when $m \to \infty$ or $m \to 0$) and peaks in between.

REFERENCES

[1] **Andrade, C. (2015):** Understanding relative risk, odds ratio and attributable risk: As simple as it can get. J. Clin. Psychiatry, 76(7), e857-e861.

[2] **Bland, J. M. and Altman, D. G. (2000):** The odds ratio. BMJ, 320: 1468.

[3] **Cox, D. R. (1972):** Regression models and life tables, Journal of the Royal Statistical Society, Series B, 34, 187-220.

[4] **Davies, H. T. O., Crombie, I. K. and Tavakoli, M. (1998):** When can odds ratios mislead? BMJ, 316(7136), 989-991.

[5] **Greenland, S. (1987):** Interpretation and choice of effect measures in epidemiologic analyses. American Journal of Epidemiology, 125, 761-768.

[6] **Greenland, S and Holland, P. (1991):** Estimating standardized risk differences from odds ratios, Biometrics, 47(1), 319-322.

[7] **Holland, P. W. (1989):** A note on the covariance of Mantel-Haenszel log-odds-ratio estimator and the sample marginal rates. Biometrics, 45, 1009-16.

[8] **Intrayan, A. (2008):** Medical Biostatistics. Chapman & Hall/CRC Biostatistics series, Boca Raton, London, New York.

[9] **Kaelin, M. A. and Bayona, M. (2004):** Attributable Risk Applications in Epidemiology. The Young Epidemiology Scholars Program (YES), www.collegeboard.com

[10] **Kirch, W. (2008):** Attributable risk proportion. Encyclopedia of Public Health. Springer, Dordrecht.

[11] **Mantel, N. and Haenszel, W. (1959):** Statistical aspects of the analysis of data from retrospective studies of disease. J. Nat. Cancer Inst. 22, 719-748.

[12] **Savitz, D. A. (1992):** Measurements, estimates and inferences in reporting epidemiologic study results. American Journal of Epidemiology, 135, 223-224.

Shishir Kumar Jha
Department of Statistics
Ramjas College
University of Delhi
Email: skjha_statistics@ramjas.du.ac.in