

Effect Measures in Meta-Analysis

Md. Belal Hossain

Department of Statistics, Biostatistics and Informatics, Dhaka University, Dhaka – 1000, Bangladesh.

(Received: 9 July 2012; Accepted: 3 November 2013)

Abstract

When conducting a meta-analysis of randomised controlled trials outcomes, appropriate choice of the effect measure is important. This article demonstrates on various types of effect measures in meta-analysis, for example, binary, continuous and ordinal outcomes. A general fixed effects model and a random effects model are employed for combining these outcomes in meta-analysis. Six trials totaling 1876 patients from a meta-analysis of randomised controlled trials evaluating the efficacy and drawbacks of limited (D1) versus extended lymphadenectomy (D2) for proven gastric adenocarcinoma are analysed for binary and continuous outcomes. An individual patient data consisting of five randomised trials of anti-cholinesterase drug tacrine in patients with Alzheimer's disease is also discussed for ordinal outcomes.

Keywords: Meta-analysis; effect measure; binary, continuous and ordinal outcomes; fixed effects model; random effects model.

I. Introduction

There has been an increasing interest in the development of appropriate measures to ensure that public policy and decision making are based on results of reliable research. The evidence based scientific research has been helping decision makers to determine which interventions are doing good and which are actually harmful, particularly in the health care area. Meta-analysis is a statistical technique which concerns with the analysis of the data extracted from independent studies. It also estimates overall measures of association or effect size and assesses the sensitivity of the results.

Measures of outcomes need to be calculated for each of the studies in a meta-analysis before they can be quantitatively combined. In a meta-analysis of randomised controlled trials (RCTs) a comparative estimate of treatment effect such as the odds ratio is often chosen, while cohort studies it is common to employ relative risk as a measure of the risk between groups. Outcomes have been classified as one of the three groups, depending on the type of data from which they are derived. Those based on binary data, such as whether patients are alive or dead, diseased or nondiseased. Continuous data are based on continuous outcomes such as blood pressure, length of surgery, length of hospital stay etc. Finally, outcomes that are based on ordered categorical (ordinal) data, for example, pain relief, dementia, head injury, tonsil size etc.

The effect measures used for binary outcomes are odds ratio (OR), relative risk (RR), risk difference (RD), and arcsine difference (AS) etc. For continuous outcomes mean difference (MD) and standardised mean difference (SMD) are widely used in meta-analysis. For ordinal data, log OR and generalised odds ratio (GOR) are used as effect measures (Agresti¹; Whitehead *et al*²).

Next we discuss briefly the commonly used effect measures for binary, continuous and ordinal outcomes in meta-analysis. Then a general fixed effects model (FEM) and a random effects model (REM) are discussed for combining these effect measures in meta-analysis (Sutton *et al*³; DerSimonian and Laird⁴). Finally, meta-analyses of six RCTs totaling 1876 patients evaluating the efficacy and

drawbacks of limited (D1) versus extended lymphadenectomy (D2) for proven gastric adenocarcinoma are shown for binary and continuous outcomes (Memon *et al*⁵). A meta-analysis with individual patient data (IPD) consisting of five randomised trials of anti-cholinesterase drug tacrine in patients with Alzheimer's disease is also discussed for ordinal outcomes.

II. Effect Measures

Binary outcomes
Odds ratio

The odds ratio (OR) is defined as the ratio of two odds of interest and is calculated from Table 1 as

$$OR = \frac{(a/n)/(b/n)}{(c/n)(d/n)} = \frac{ad}{bc},$$

where, a, b, c, and d are the cell frequencies of the four cells in a RCT setting for a 2x2 table. For desirable outcomes OR greater than one indicates improvement by the new treatment while an OR less than one means the new treatment is less effective. For undesirable outcomes the converse is true.

The following large sample variance of the log OR is commonly used for the purpose of pooling in meta-analysis and for computing the CI: $Var(\ln OR) = 1/a + 1/b + 1/c + 1/d$.

Table 1. Data of a single RCT

Intervention	Success/ Alive	Failure/ Dead	Total
New treatment	a	b	a+b
Control	c	d	c+d
Total	a+c	b+d	n

For zero cell count problem, it is recommended adding 0.5 to each cell of all the studies of 2x2 tables. This also reduces the bias caused by one or more small cells in the study table. Under the normality assumption of the ln OR, a 95% CI for ln OR can be found as: $\exp[\ln OR \pm 1.96\sqrt{\text{var}(\ln OR)}]$.

Relative risk

The relative risk or risk ratio (RR) is defined as the probability of an event in the treatment group divided by the probability of an event in the control group (Table 1) as $RR = (a/(a+b))/(c/(c+d))$. Useful estimated variance

expression of the log RR can be found as $var(\ln RR) = 1/a - 1/(a+b) + 1/c - 1/(c+d)$.

There are debates on the choice of binary effect measures between OR and RR. Some researchers prefer OR as an effect measure than RR because OR can be estimated and interpreted for RCTs and case-control study reasonably for rare outcomes. Moreover, some study designs select subjects on the basis of outcome rather than the treatment type (case-control study). RR can not be used for this study design.

Risk difference

The risk difference (RD) provides an indication of the impact of the treatment or exposure. The RD is defined for a 2x2 table simply as risk in the experimental group minus risk in the control group and is calculated as $RD = a/(a+d) - c/(c+d)$. The variance estimate of RD can be found as $var(RD) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$,

where $p_1 = a/(a+b)$, $p_2 = c/(c+d)$, $n_1 = a+b$, and $n_2 = c+d$. A 95% CI for RD can be computed as $RD \pm 1.96\sqrt{var(RD)}$.

Arcsine difference

The arcsine difference (AS) is defined as

$AS = \arcsin\sqrt{\frac{a}{a+b}} - \arcsin\sqrt{\frac{c}{c+d}}$. The AS is a measure rarely used in medical science which handles zero frequencies naturally. The asymptotic variance of AS which does not depend on the event of probability is given as

$$var(AS) = \frac{1}{4(a+b)} + \frac{1}{4(c+d)}$$

Continuous outcomes

Mean difference

The measure of treatment effect (MD) for continuous data is given by

$MD = \mu_t - \mu_c$, where, μ_t and μ_c are the mean responses in the treatment and control groups, respectively.

The variance of this treatment difference is

$var(MD) = \sigma^2(1/n_t + 1/n_c)$, where n_t is the within-study sample size in the treatment group, n_c is the within-study sample size for the control group, and σ^2 is the variance, assumed common to both groups. An alternative sample variance for MD is given as

$$s^2 = \frac{(n_t - 1)(s^t)^2 + (n_c - 1)(s^c)^2}{n_t + n_c - 2}$$

(WMD) in meta-analysis under FEM can be found as

$$WMD = \bar{T} = \frac{\sum_{i=1}^k \omega_i T_i}{\sum_{i=1}^k \omega_i}$$

where $\omega_i = 1/var(T_i)$ is the weight and T_i is the effect size for the i th study. A 100(1- α)% CI for pooled estimate can be calculated as

$$\bar{T} \pm z_{\alpha/2} / \sqrt{\sum_{i=1}^k \omega_i}$$

Standardised mean difference

The standardised mean difference (SMD) is estimated by

$$d_i = \frac{\bar{X}_{ti} - \bar{X}_{ci}}{s^*}$$

where \bar{X}_{ti} and \bar{X}_{ci} are the sample mean responses in the treatment and control groups respectively, and s^* is the pooled estimate of the standard deviations. The estimate d_i has small sample bias and the bias can be removed using a simple correction that produces an unbiased estimate of the population SMD. The variance of d_i is difficult to compute exactly. However, a very good approximation of the variance of d_i is given by

$$V_{di} = \frac{n_{ti} + n_{ci}}{n_{ti}n_{ci}} + \frac{d_i^2}{2(n_{ti} + n_{ci})}$$

More simpler variance approximation is also available if the n_{ti} and n_{ci} are large and the population variances are equal as

$$V_{di} = \frac{n_{ti} + n_{ci}}{n_{ti}n_{ci}}$$

The use of SMD has been criticised by saying that studies with identical results may vaguely appear to yield different results. The transformation can even make a study whose original estimate was smaller in magnitude than another study appear greater and vice versa. There are other continuous outcome measures rarely used in medicine, such as correlation coefficient etc.

Ordinal outcomes

Generalised odds ratio

Let J be the number of comparison groups with L ordered outcome categories in each group.

Table 2. Contingency table for the i th study

Groups	Category 1	Category 2	...	Category L	Sample Size
Treatment	X_{i11}	X_{i12}		X_{i1L}	n_{i1}
Control	X_{i21}	X_{i22}		X_{i2L}	n_{i2}

In Table 2, X_{ijl} is the count of the l th category in the j th group for the i th study, n_{ij} is the total count of j th group for the i th study. For $J=2$, the i th study is modeled as two multinomial distributions with parameters (n_{i1}, π_{i1}) and (n_{i2}, π_{i2}) for a $2 \times L$ contingency table.

The GOR is defined as the ratio of the proportions of concordant and discordant pairs (Agresti¹) in a contingency table. A pair is said to be concordant if the subject ranked higher on groups also ranks higher on categories or vice versa. Without loss of generality we assume that the response in category 1 is more severe than the response in category l , where $l < l'$.

Mathematically, the GOR for the i th study is defined as

$$\Gamma_i = \Pi_{ci} / \Pi_{di}$$

$$\text{where } \Pi_{ci} = \sum_{r=1}^{L-1} \sum_{s=r+1}^L \Pi_{ir|1} \Pi_{is|2} \text{ and } \Pi_{di} = \sum_{r=2}^L \sum_{s=1}^{r-1} \Pi_{ir|1} \Pi_{is|2}$$

Here, Π_{ci} denotes the probability that the response of a randomly selected subject from group 2 (control) is more severe than the response of a randomly selected subject from group 1 (treatment). Similarly, Π_{di} denotes the probability that the response of a randomly selected subject from group 1 is more severe than the response of a randomly selected subject from group 2. The data with zero cell count is analysed by adding $1/L$ to each entry before calculation of the GOR. The value of $\Gamma_i = 1$, represents identical comparison groups as it is in the OR.

For an arbitrary number of outcome categories (L) in RCTs

in which each row is modeled as an independent multinomial distribution, the estimated variance of the i th study is

$$v_i = \hat{\omega}_i^{-1} = \sum_{l=1}^{L-1} \sum_{j=1}^2 \frac{1}{n_{ij} \hat{\pi}_{ijl} (1 - \hat{\pi}_{ijl})}, \text{ where } n_{ij} \text{ is the}$$

total count of the j th group for the i th study, $\hat{\pi}_{ijl} = X_{ijl} / n_{ij}$ is the MLE of π_{ijl} and X_{ijl} is the count of the l th category in the j th group for the i th study.

III. General Fixed and Random Effects Models

Generally two models are frequently used in meta-analysis. The fixed effects model (FEM) which assumes that studies being modeled are homogeneous. The random effects model (REM) assumes different effect sizes for the studies used in meta-analysis. The general FEM is given by $\hat{\theta}_i = \theta_i + e_i$, where e_i is the error with which $\hat{\theta}_i$ estimates θ_i . For the FEM, $\text{var}(\hat{\theta}_i) = v_i$ and for REM $\text{var}(\hat{\theta}_i) = \tau_\theta^2 + v_i$, where τ_θ^2 is the between study variance and v_i is the variance due to sampling error for the i th study. If $\tau_\theta^2 = 0$, the above REM would reduce to the FEM.

Fixed effects model

The inverse variance weighted method is the most widely used amongst all fixed effects models. For k independent studies if $\hat{\theta}_i$ represents logarithm of the effect measure and v_i represents the variance of the effect estimator, then assuming $\theta_1 = \theta_2 = \dots = \theta_k = \theta_0$, a pooled estimate of the treatment effect is given by

$$\hat{\theta}_0 = \sum_i \omega_i \hat{\theta}_i / \sum_i \omega_i. \text{ The estimated variances needed to}$$

form the confidence intervals of the effect measures are presented in previous sections.

Assuming $\hat{\theta}_i$ s are normally distributed, an approximate 100(1- α)% CI for the i th effect measure is given by the formula $\exp[\hat{\theta}_i \pm z_{\alpha/2} \omega_i^{-1/2}]$, where $z_{\alpha/2}$ is the 100 $\times\alpha/2$ percentage point of a standard normal distribution. An estimator of the variance of the pooled estimator of θ_0 is given by $\hat{\omega}^{-1} = \text{var}(\hat{\theta}_0) = 1 / \sum_{i=1}^k \omega_i$.

If $\hat{\theta}_0$ is assumed to be normally distributed, an approximate 100(1- α)% CI for the population effect, θ_0 , is given by $\exp[\hat{\theta}_0 \pm z_{\alpha/2} \hat{\omega}^{-1/2}]$ for the meta analysis.

Random effects model

For the standard random effects model (DerSimonian and Laird⁴), let $\hat{\tau}_\theta^2$ be the estimate of the between study

variance. Define $\bar{\omega}$ and s_W^2 to be the mean and variance of the weights from the k studies:

$$\bar{\omega} = \sum_{i=1}^k \omega_i / k \text{ and } s_W^2 = \frac{1}{k-1} \left(\sum_{i=1}^k \omega_i^2 - k \bar{\omega}^2 \right).$$

Further, define

$$U = (k-1) \left(\bar{\omega} - \frac{s_W^2}{k \bar{\omega}} \right) \text{ and } Q = \sum_{i=1}^k \omega_i (\hat{\theta}_i - \bar{\theta})^2,$$

where Q is the heterogeneity statistic, also known as

Cochran's χ^2 statistic (Cochran⁶) for testing the

$H_0 : \theta_1 = \theta_2 = \dots = \theta_k = \theta_0$. The estimated component of variance due to inter-study variation in effect size, $\hat{\tau}_\theta^2$, is calculated as

$$\hat{\tau}_\theta^2 = \begin{cases} 0 & \text{if } Q \leq k-1 \\ (Q - (k-1)) / U & \text{if } Q > k-1. \end{cases} \quad (1)$$

Then adjusted weights ω_i^* for each of the studies can be

calculated as $\omega_i^* = \frac{1}{[1/\omega_i] + \hat{\tau}_\theta^2}$. A 100(1- α)% CI

for θ_i is given by $\exp[\hat{\theta}_{iR} \pm z_{\alpha/2} / \sqrt{\omega_i^*}]$ under the assumption of normality of $\hat{\theta}_{iR}$.

The point estimate for the mean treatment effect of all studies, θ_0 , can be computed by

$$\hat{\theta}_{0R} = \sum_{i=1}^k \omega_i^* \hat{\theta}_i / \sum_{i=1}^k \omega_i^* \text{ with } \text{var}(\hat{\theta}_{0R}) = 1 / \sum_{i=1}^k \omega_i^*.$$

If normality of $\hat{\theta}_{0R}$ is assumed, a 100(1- α)% CI

for θ_0 is given by $\exp[\hat{\theta}_{0R} \pm z_{\alpha/2} / \sqrt{\sum_{i=1}^k \omega_i^*}]$.

IV. Results and Discussion

For binary outcomes (30-day mortality rate) meta-analyses are conducted using OR, RR, RD and AS. The estimates along with their 95% CIs are presented in Table 3.

To avoid the computation of reciprocal of zeros among observed values in the calculation of the original effect measure, say OR, we add 0.5 to each count in the 2 \times 2 contingency table. Mean difference (MD) and standardised mean difference (SMD) are used for continuous outcomes. For ordinal outcomes GOR is used as an effect measure. Both fixed and random effects models, developed by using the inverse variance weighted method are used to combine the data (Sutton *et al*³). Heterogeneity among studies is assessed using the Q statistic (Cochran⁶) and I^2 index (Higgins and Thompson⁷).

The published clinical trials that did not report the mean and standard deviation, but rather reported the size of the trial, the median and range, estimates of the mean and standard deviation are obtained according to Hozo *et al*⁸.

Table 3. Estimate of OR, RR, RD and AS along with their CIs of 30-day mortality with gastric adenocarcinoma

Study	OR[CI]	RR[CI]	RD[CI]	AS[CI]
Dent <i>et al</i>	0.96 [0.02; 50.33]	0.96 [0.012; 46.14]	-0.00 [-0.08; 0.08]	-0.00 [-0.29; 0.29]
Robertson <i>et al</i>	0.37 [0.01; 9.56]	0.38 [0.02; 9.04]	-0.03 [-0.12; 0.06]	-0.09 [-0.35; 0.18]
Bonenkamp <i>et al</i>	0.62 [0.39; 0.99]	0.65 [0.42; 0.99]	-0.03 [-0.07; -0.00]	-0.06 [-0.13; -0.00]
Cuschieri <i>et al</i>	0.46 [0.23; 0.93]	0.50 [0.26; 0.94]	-0.06 [-0.12; -0.00]	-0.11 [-0.21; -0.01]
Degiuli <i>et al</i>	3.44 [0.14; 85.63]	3.39 [0.14; 81.99]	0.01 [-0.02; 0.05]	0.06 [-0.09; 0.22]
Wu <i>et al</i>	1.01 [0.02; 51.30]	1.01 [0.02; 50.40]	0.00 [-0.02; 0.02]	0.00 [-0.13; 0.13]
MA[CI]				
FEM	0.59 [0.40; 0.85]	0.61 [0.43; 0.86]	-0.01 [-0.02; 0.00]	-0.05 [-0.10; -0.01]
REM	0.59 [0.40; 0.85]	0.61 [0.43; 0.86]	-0.01 [-0.03; 0.01]	-0.05 [-0.10; -0.01]

Table 4. Estimate of GOR along with CI for tacrine data

Study	Groups	C1	C2	C3	C4	C5	T	GOR(CI)	MA(CI)
1	Tacrine	4	23	45	22	2	96	1.28 (0.17, 9.71)	1.54
	Placebo	2	22	54	29	3	110		
2	Tacrine	14	119	180	54	6	373	1.22 (0.12, 11.99)	(1.15, 2.19)
	Placebo	1	22	35	11	3	72		
3	Tacrine	13	20	24	10	1	68	1.14(0.19, 6.68)	Q = 0.341, df=4
	Placebo	7	16	17	10	3	53		
4	Tacrine	21	106	175	62	17	381	1.93 (0.63, 5.91)	p=0.987
	Placebo	8	24	73	52	13	170		
5	Tacrine	3	14	19	3	0	39	1.54 (0.11, 22.13)	I ² = 0%[0%, 89%]
	Placebo	2	13	18	7	1	41		

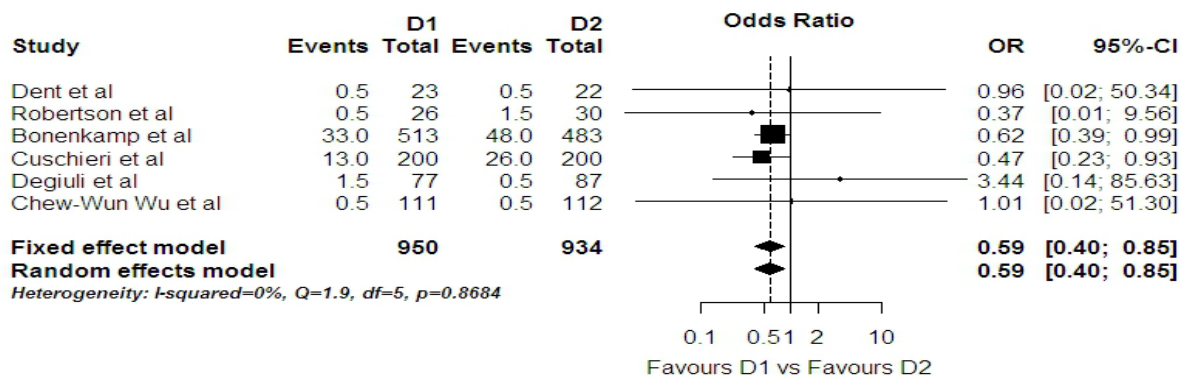


Fig. 1. Forest plot of mortality of the six studies with gastric carcinoma

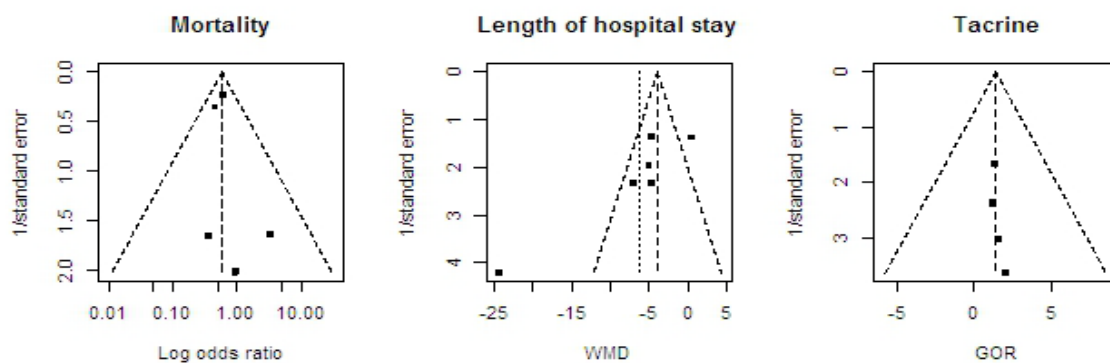


Fig. 2. Funnel plots of 30-day mortality rate, length of hospital stay and tacrine show absence of publication bias, presence of publication bias and absence of publication bias respectively.

Funnel plots are used to determine the presence of publication bias in the meta-analysis. All estimates and graphs are produced using the meta and rmeta packages in R (<http://cran.rproject.org/web/packages/rmeta/index.html>)⁹.

OR, RR, RD and AS for 30-day mortality rate: In Table 3, we present study specific study estimates using OR, RR, RD and AS along with their 95% CIs for the binary data of 30-day mortality rate. Results from meta-analyses under both the model approaches are shown using the four estimators. For OR, it shows 41% reduction in 30-day mortality rate (OR=0.59, CI= [0.40, 0.85], Z=-2.78, p= 0.0054) for D1 group than D2. For RR, it results 39% reduction in 30-day mortality rate (RR=0.61, CI=[0.43; 0.86], Z=-2.78, p=0.0054). Using RD, we find 1% reduction of 30-day mortality rate for D1 group than D2 (RD=-0.01, CI=[-0.03; 0.01], Z= -1.19, p= 0.23) and using AS, it is reduced by 5% (AS=-0.05, CI=[-0.10; -0.01], Z=-2.37, p=0.02). Weights for these effect measures are quite similar under FEM and REM (Not included due to space limitation of the journal). Due to the absence of significant heterogeneity (For OR, Q=1.86, df=5, p=0.87; for RR, Q=1.76, df=5, p=0.88; for RD, Q= 8.71, df=5, p=0.12; and for AS, Q=4.52, df=5, p=0.48), both models produce exactly the same result of Q. Funnel plot (precision vs. log OR) in Figure 2 demonstrates symmetry for 30-day mortality suggesting absence of publication bias.

MD and SMD for length of hospital stay

Under FEM it shows 3.97 days reduction in hospital stay (MD=-3.97, CI= [-5.46; -2.49], Z= -5.24, p<0.0001) and under REM it results 6.37 days reduction in hospital stay (MD= -6.37, CI= [-10.66, -2.08], Z=-2.91, p=0.0036). There exists significant heterogeneity among these trials (Q= 36.34, df=5, p<0.01; $\tau_{\theta}^2 = 23.47$; $I^2 = 86.2\%$ [72.2%; 93.2%]), hence results produced by the REM would be preferable as it takes into account the between study variation appropriately. Using SMD it also shows significant reduction in hospital stay under both FEM (SMD= -0.25, CI= [-0.34; -0.16], Z= -5.36, p <0.0001) and REM (SMD= -0.37, CI= [-0.61; -0.12], Z= -2.98, p= 0.0029). Funnel plot (figure not shown) for this outcome variable shows points falling outside the 95% CI limits suggesting presence of publication bias.

GOR for tacrine data

For ordinal measure an IPD of five RCTs (Table 4) of anti-cholinesterase drug tacrine in patients with Alzheimer's disease (Whitehead *et al*² is used. GOR is used to measure the effect sizes for these ordinal categorical outcomes. Using GOR one can estimate the effect sizes of studies with ordinal outcomes for any number of outcome categories as

well as for any number of comparison groups. Proportional odds model using log odds ratio arbitrarily splits the J×L table into 2×2 tables and incurs loss of information.

The categories of this trial are made by the Clinical Global Impression of Change scale (CGIC). The CGIC is based on a seven point scale where 1, 2 and 3 represent 'very much improved', 'much improved' and 'minimally improved' respectively, 4 indicates 'no change', and 5, 6 and 7 represent 'minimally worse', 'much worse' and 'very much worse' respectively. For the analysis purpose they combined categories 1 and 2, and 6 and 7 as there were very few patients in the two extreme categories. The second last column of Table 5 contains the GOR and their associated CIs. The last column contains the pooled GOR and its CI. The meta-analysis using GOR results 1.54 suggesting that there are 1.54 times as many tacrine-placebo pairs in the sample for which tacrine has improved the alzheimer's disease as there are pairs for which placebo has improved the alzheimer's disease.

The tests of heterogeneity are statistically insignificant Q = 0.341, df=4, p=0.987; and $I^2 = 0\%$ [0%, 89%]. Forest plot in Figure 1 for 30-day mortality rate is favouring D1. The text and values, on the left, are study identification, number of cases (n), sample sizes (N), odds ratio (OR), and lower (LL) and upper limits (UL) of 95% confidence interval (CI). In the graph, squares indicate point estimates of treatment effect (odds ratio for D1 over D2 groups) with the size of the squares representing the weight attributed to each study. The horizontal lines represent 95% CI for OR. The pooled estimate for 30-day mortality rate is the pooled OR obtained by combining all ORs of the 5 studies using the inverse variance weighted method, and is represented by the diamond. The horizontal length of the diamond depicts the 95% CI. Values to the left of the vertical line favour D1.

V. Conclusion

This article demonstrates on various types of effect measures for binary, continuous and ordinal outcomes in meta-analysis. The details of these effect measures along with their estimated variances are presented. While OR and RR are measuring how beneficial or harmful a treatment be as compared to the placebo, RD measures the impact of the treatment. AS can handle zero frequencies appropriately. For continuous data MD and SMD are frequently used. For ordinal outcomes GOR is used as an effect measure. The GOR is simple and it has straightforward interpretation. It is also free from model assumption. The procedures of conducting meta-analysis under general fixed and random effects model are shown on three different examples. For all these measures tests of heterogeneity are shown using Q

statistic and quantified by I^2 statistic. Funnel plot is used to identify presence of publication bias. As meta-analysis is quite new topic in Bangladesh, I hope this article will facilitate understanding and conducting different types of meta-analyses in Bangladesh.

References

1. Agresti, A., 1980. Generalised odds ratios for ordinal data. *Biometrics*, **36**, 59-67.
2. Whitehead, A., R. Z. Omar, J. P. T. Higgins, E. Savaluny, R. M. Turner, and S. G. Thompson, 2001. Meta-analysis of ordinal outcomes using individual patient data. *Stat Med*, **20**, 2243-2260.
3. Sutton, A. J., K. R. Abrams, D. R. Jones, T. A. Sheldon, and F. Song, 2000. *Methods for Meta-Analysis in Medical Research*, West Sussex, PO19 1UD, England: John Wiley & Sons, Ltd.
4. DerSimonian, R. and N. Laird, 1986. Meta-analysis in clinical trials. *Control Clin Trials*, **7**, 177-88.
5. Memon, M. A., M. S. Subramanya, S. Khan, M. B. Hossain, E. Osland, and B. Memon, 2011. Meta-analysis of D1 versus D2 gastrectomy for gastric adenocarcinoma. *Annals of Surgery*, **253(5)**, 900-911.
6. Cochran, W. G., 1954. The combination of estimates from different experiments. *Biometrics*, **10**, 101-29.
7. Higgins, J. P. and S. G. Thompson, 2002. Quantifying heterogeneity in a meta-analysis. *Stat Med*, **21(11)**, 1539-58.
8. Hozo, S. P., B. Djulbegovic, and I. Hozo, 2005. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*, **5**, 13.
9. Lumley T. The rmeta Package Version 2.14. <http://cran.rproject.org/web/packages/rmeta/index.html>