Improved Meta Analysis Using Predicted Relative Risk

Shahjahan Khan and Md Masud Hasan
Department of Mathematics & Computing
Australian Centre for Sustainable Catchments
University of Southern Queensland
Toowoomba, Queensland, AUSTRALIA
Emails: khans@usq.edu.au and hasan@usq.edu.au

Abstract

This paper proposes a new method of improved meta analysis to combine relative risk for both homogeneous and heterogeneous set of studies. The standard meta analyses don't give any conclusive result when the effects of heterogenous studies are combined. The proposed improved meta analysis uses the predicted relative risk, and chi-square test to check the heterogeneity of the effects. Confidence intervals for the relative risks obtained via improved method concentrate more towards the value of the pooled estimate than that of the standard meta analysis. Exclusion of identified studies with outliers from the analysis brings the results of the remaining studies closer to the pooled estimate. An illustration shows that the new method improves the results and provide conclusive estimate of the relative risk.

Key Words: Relative risk, predicted relative risk, odds ratio, chi-square test, standard meta analysis.

1 Introduction

Meta analysis is a statistical procedure that integrates the results of several independent studies to be combinable. It usually integrates and summarizes the findings from many clinical studies of treatments because an individual study of a particular treatment may not be as conclusive as several studies put together. A meta analysis also helps to gain greater objectivity and precision by including all the available evidence from randomized trials that pertain to the issue [1]. Since the 1980's there has been an upsurge in the application of meta analysis in medical research. Over the same period there have been great strides in the development and refinement of the associated statistical methodology. These developments have mainly been due to

greater emphasis on evidence-based medicine and the need for reliable summaries of the vast and expanding volume of clinical research.

If there exists homogeneity in the treatment effects among the trials considered, standard meta analysis provides conclusive results. However the method is not useful for integrating treatment effects which are heterogenous in character. The most obvious cause of the heterogeneity in the treatment effects relates to the ages of the participants, or more particularly the average age of experiencing the event during follow up, since it is well known that the relative risk associated with a particular disease with a given treatment declines with advancing age [2]. Other factors causing non-homogeneous treatment effects are environment in which the trials are conducted. Despite the laudable attempts to achieve objectivity in reviewing scientific data, considerable subjective judgment is necessary in carrying out meta analyses. These judgments include those about which studies are "relevant" and which studies are methodologically sound enough to be included, as well as the issue of whether and how to investigate sources of heterogeneity [3]. Saleh et al.[4] introduced an improved meta analysis for handling studies having statistically heterogeneous treatment effects and achieved reliable (minimum mean square sense) quantitative conclusions.

Most of the meta analyses are performed for combining odds ratios. Odds ratio is the ratio of odds of treatment group to control group. Such studies consider the odds rather than the actual risks for the treatment and placebo groups, and hence it is not necessarily a better estimate of the treatment effects. In mathematical epidemiology, relative risk is a ratio of the probability of the event occurring in the exposed group versus the control (non-exposed) group. Relative risk is used frequently in the statistical analysis of binary outcomes where the outcomes of interest has relatively low probability. It is thus often more suited to analyze clinical trial data, that are used to compare the risk of developing a disease in people not receiving the new medical treatment versus people who are receiving an established treatment. The relative risk is easily interpretable than other parameters used in medical science. While comparing a new treatment to the control, if a relative risk is less than one, then it would indicate an improvement on the new treatment; whereas a ratio greater than one would imply the new treatment was less effective than any treatment received by the control group.

In this paper we consider both homogeneous and heterogenous sets of treatments and perform standard meta analyses for relative risks. We estimate relative risks, weights and construct confidence intervals for individual studies as well as combining them. We test the heterogeneity of trials for both the cases and check the validity of the meta analyses. We then perform improved meta analysis of relative risk for

heterogeneous set of studies where usual meta analysis fails to provide conclusive result. Improve meta analysis combines the estimated relative risk with the value of the test statistic to get the treatment effect. This method it is expected to yield better estimate of the treatment effect than conventional meta analyses.

2 Standard meta analysis

Standard meta analysis is performed for the data produced by randomised control trails (RCT). The idea is to integrate or combine studies or groups which are more or less homogenous with respect to the parameter to be estimated to make conclusion on the effectiveness any new drug or medical procedure.

2.1 Methodology

Outcome variables in terms of odds ratio or relative risk or weighted mean difference are commonly used in meta analysis. Here we consider meta analysis based on the relative risk. From the data we first estimate the relative risks and weights of individual studies. The weight of a study is the inverse of the variance of the respective study, and the combined effect is calculated as a weighted average of the relative risk from individual studies. Weighted geometric mean is the appropriate estimate here.

Let for the ith study, the number of respondents for treatment and control groups with success and failure criteria be given as **Table 1**.

Table 1 Number of individuals for different subgroups.

	Success	Failure	Total
Treatment	x_{1i}	n_{1i}-x_{1i}	n_{1i}
Control	x_{2i}	n_{2i}- x_{2i}	n_{2i}

The different steps for conducting meta analysis of relative risk is as follows:

- Calculate the relative risk of the treatment and log relative risk. For the ith study, the estimated risk for the treatment and control groups are $R_{ti} = \frac{x_{1i}}{n_{1i}}$ and $R_{ci} = \frac{x_{2i}}{n_{2i}}$ respectively. Then the estimated relative risk for ith study is

$$\hat{\theta}_i = RR = \frac{R_{ti}}{R_{ci}}.$$

  The corresponding log relative risk is defined as the natural logarithm of the estimated relative risk. Thus for the ith study the log RR becomes

$$\hat{l}_i = ln(\hat{\theta}_i).$$

- Calculate the variance of log relative risk of the treatment. For the ith study the variance of the log RR is defined as

$$\hat{\sigma^2}_i = \text{Var}[\ln(\hat{\theta}_\text{i})] = \frac{1}{\text{x}_{1\text{i}}} - \frac{1}{\text{n}_{1\text{i}}} + \frac{1}{\text{x}_{2\text{i}}} - \frac{1}{\text{n}_{2\text{i}}}.$$

- Calculate the weights of the ith study by taking the inverse of the variance of the log RR as

$$w_i = \frac{1}{\text{Var}[\ln(\hat{\theta})_\text{i}]}.$$

- Calculate the estimated pooled log relative risk of the treatment by taking weighted geometric mean of the individual log relative risks from all the studies from all the studies . Thus the estimated pooled log RR becomes

$$\hat{\theta} = \exp\left[\frac{\sum w_i \times ln(\hat{\theta}_i)}{\sum w_i}\right].$$

- Calculate the variance of the estimated Pooled log RR as the weighted average of variances of individual studies as

$$\sigma^2(\hat{\theta}) = \frac{\sum w_i^2 \times \hat{\sigma}_i^2}{n^2}.$$

- Calculation of confidence intervals for individual studies: The $(1-\alpha)100\%$ confidence interval for the log RR [5] is obtained as

$$\hat{li} \pm z_{\alpha/2} \times \sqrt{w_i^{-1}}.$$

  Accordingly the $(1-\alpha)100\%$ confidence interval for the RR is obtained as

$$\exp[\hat{li} \pm z_{\alpha/2} \times \sqrt{w_i^{-1}}].$$

Usually meta analyses are done with large samples so that the use of $Z_{\alpha/2}$, the $\alpha/2$ level of critical value from N(0,1) is justified. In standard meta analysis, the combined estimate of the log RR, $\hat{\theta}$ is a represented value of the individual RR's and so in plot of the CI's the vertical line through $\hat{\theta}$ as a representative value of the individual RR's and so the vertical line through $\hat{\theta}$ is expected to go through all the trials and their concomitant error bars.

If there exists heterogeneity among the treatment effects, meta analysis fails to give any conclusive result of the pooled treatment effect [6]. So it is necessary to test the heterogeneity whether the treatment effects for the population from which the samples of individual studies come differ significantly. Here the null hypothesis is

$H_0 : \theta_1 = \theta_2 = \cdots = \theta_k = \theta$ against the alternative, H_A: at least one of the pairs (θ_i, θ_j) differ $for\ i, j = 1, 2, \cdots, k$. An appropriate test statistic is

$$\chi^2 = \sum_{i=1}^{n} [\hat{l}_i - ln(\hat{\theta})] \times w_i$$

Under the null hypothesis χ^2 follows approximately chi-square distribution with (k-1) degrees of freedom. The value of the test statistic measures the departure of the individual values of the parameter from its common value in log scale. The cut off point $\chi^2_{k-1}(\alpha)$ at α -level based on the central chi-squared distribution with (k-1) d.f. If $\chi^2_{obs} > \chi^2_{k-1}(\alpha)$, we don't accept the null hypothesis, whereas for $\chi^2_{obs} < \chi^2_{k-1}(\alpha)$ we don't reject the null hypothesis.

2.2 Analytical Results

Here we perform two standard meta analyses, one for studies with homogenous treatment effects and other for heterogenous treatment effects. For the first study we consider the data from the results of treatments of infectious disease mononucleosis (IDM) with a history of tonsillectomy (T) among seven age groups of students aged 18-24. The data are taken from Schork and Remington (2000, see Example 5; p. 218) [7]. Students diagnosed as having IDM were compared with the "control" (C) group students who are free from the disease. The different age groups are treated as separate studies. Result of meta analysis from the set of homogenous studies are shown in Table 2. It presents the relative risks, weights and 95% confidence intervals of RR for all the studies as well as the pooled estimate of the RR. The confidence interval for the Pooled RR is constructed using the standard deviation combined by Mantel-Haenszel [8] method.

Table 2 Standard meta analysis for homogenous studies.

age	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**Meta Analysis**
IDM	6/23	3/42	12/41	8/46	5/15	2/9	4/9	
Control	17/49	26/96	34/112	48/139	45/118	29/66	36/75	
RR	0.75	0.26	0.96	0.50	0.87	0.51	0.93	**0.66**
Weight	6.19	2.96	12.59	8.55	6.80	2.45	6.52	
95% CI	.34-1.65	.08-.82	.55-1.68	.26-.98	.41-1.85	.14-1.77	.43-1.99	**.53-.81**

The result from standard meta analysis is also represented in Figure 1. Here the values of relative risks for the studies (age groups) varies from 0.26 for students aged 19 years to 0.96 for those aged 20. If we consider the age groups as individual studies, we may get some misleading conclusion from the study. For the age group 19, the result indicates a $(1-0.26)\times 100\% = 74\%$ reduction in tonsillectomy for the treatment

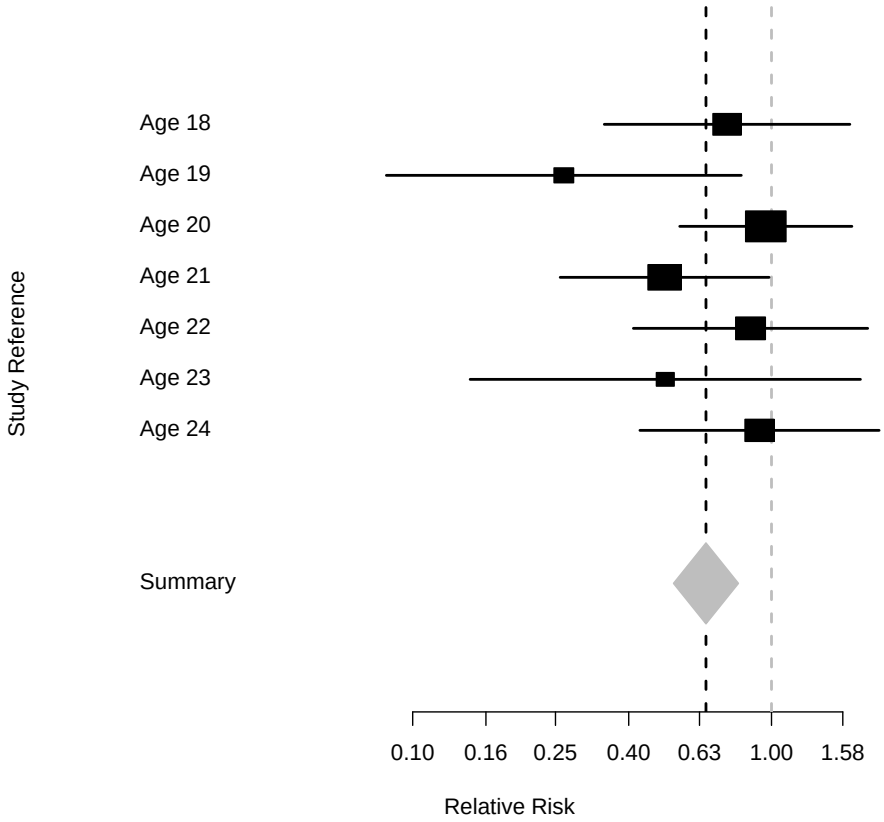

Figure 1: 95% Confidence intervals for estimated relative risk

group with respect to control group. However, for the age group 20, it shows only 4% reduction rate. For the age groups, 18, 21, 22, 23, and 24, the percentages of reduction rates in the history of tonsillectomy for the treatment groups with respect to control groups are 25, 50, 13, 49 and 7 respectively. This variation in the relative risks of individual studies may be due to sampling fluctuations of the studies due to relatively smaller sample sizes.

From the meta analysis, we get combined relative risk as 0.66. We see that the value of the pooled relative risk lies between the 95% confidence limits of the individual CI's for all the studies. The dotted vertical line of the graph from the estimated value of the pooled effect passes through all the strata and their concomitant error bars. We then carry out a test to check the heterogeneity of the trials. The observed value of the chi-square statistic is found to be $\chi^2 = 6.53$. The critical at the 5% significance level value for the test with 6 degrees of freedom is $\chi^2_6(.05)$=12.59. The observed value of the chi square statistic is less than its critical value at the 5% sig-

nificance level indicating that there is not enough sample evidence to reject the null hypothesis. From the result we can conclude that there exists no significant difference among the population relative risks for different age groups. In such a case we get a conclusive result by combining the individual studies. We accept the common value of the relative risk as 0.66, which corresponds to an estimated 34% reduction of the relative risk of the IDM attributed due to tonsillectomy.

Table 3 Standard meta analysis for heterogenous studies.

Trials	**Diuretics**	**Control**	**RR**	**Weight**	**CI**
Wesley	14/131	14/136	1.04	7.82	0.52-2.09
Flowers	21/385	17/134	0.43	10.38	0.23-0.79
Menzies	14/57	24/48	0.49	13.38	0.29-0.84
Fallis	6/38	18/40	0.35	5.85	0.16-0.79
Cuadros	12/1011	35/760	0.26	9.12	0.13-0.49
Landerman	138/1370	175/1336	0.77	87.09	0.62-0.95
Krans	15/506	20/524	0.78	8.87	0.40-1.05
Tervila	6/108	2/103	2.86	1.54	0.59-13.85
Campbill	65/153	40/102	1.08	41.59	0.80-1.47
Meta-analysis			**0.71**		**0.62-0.82**

Heterogeneous Studies:

Table 3 represents the results for another meta analysis. The data shown here come from a meta analysis of nine randomized controlled trials investigating the use of diuretics to prevent preeclampsia. We use the data from Thompson and Pocock (1987, Lancet 338) [9]. For each treatment group, the proportion of patients developing preeclampsia is shown. For all the nine studies the relative risks, weights and confidence intervals are presented in the table. The relative risk for the Cuadros trial is 0.26 which implies a 74% reduction in the risk of preeclampsia to Diuretics imposed group with respect to control group. Whereas, for the Tervila trial treatment group shows 1.86 times more risk of the disease with respect to the control group. Thus some of the treatments considered here shows completely opposite results. However the pooled relative risk for the meta analysis is 0.71 which indicates a 29% reduction in relative risk in diuretics group with respect to control group. The value of the pooled relative risk is not contained in the 95% confidence limits for the RR of the trials Cuadrons and Campbill. These two may be the outliers for the study. From Figure 2 we see that the vertical line through the point 0.71 does not pass through the concomitant error bars for the studies Cuadros and Campbell. To test $H_0 : \theta_1 = \theta_2 = \cdots = \theta_k = \theta$ vs H_A: at least one of the pairs (θ_i, θ_j) differ for

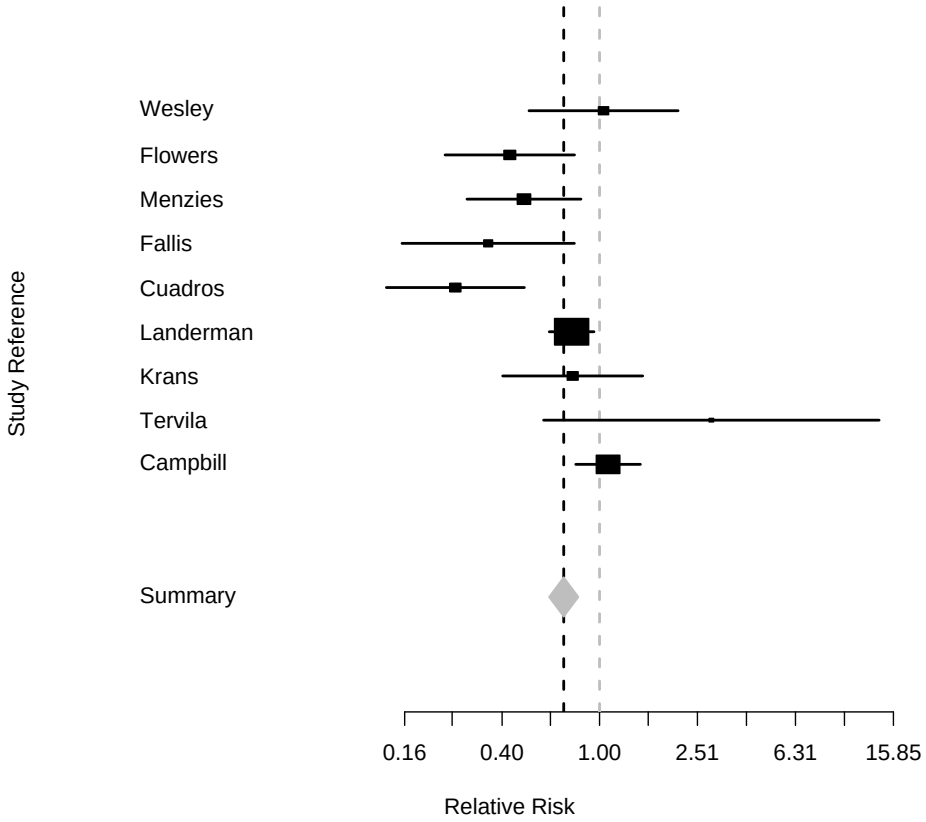

Figure 2: 95% Confidence intervals for relative risk

$i, j = 1, 2, \ldots, k$, the observed χ^2 value for the study is 28.81 with 8 degrees of freedom. The critical value is $\chi^2_8(.05) = 15.07$, which is less than the observed value. So the null hypothesis is rejected and we conclude that the population relative risks for the studies differs significantly. In this case the meta analysis to combine the individual studies in estimating the parameters is not statistically valid and fail to provide any valid conclusive result. An alternative methodology is pursued to conduct meta analysis when the the studies are heterogeneous with respect to the treatment effects.

3 Improved Meta Analysis

The proposed improved meta analysis uses the predicted relative risk, rather than the estimated relative risk, in the computation of the confidence intervals and pooled results.

3.1 Methodology

The standard meta analysis of relative risk in Table 3 is not conclusive as there exists heterogeneity among the studies. The vertical line through the estimated value of combined relative risk does not pass through all the trials and their concomitant error bars. To overcome the problem of combining treatment effects for heterogeneity set of studies, Saleh et al. [4] introduced an improved estimate of meta analysis. They carried out a test of homogeneity of individual treatment effects and combined the value of test statistic with the treatment effects to get improved meta analysis. Combining the test statistic and usual estimated values, there exists improved estimators in the theory of inference. Some of such estimators are preliminary test estimator (PTE) [10, 11], Stein type estimator (SE) and positive-rule Stein type estimator (PRSE) [12]. On the basis of these theories, Saleh et al. [4] proposed a new estimator for odds ratios namely predicted odds ratio (POR's). Here the method is adopted for meta analysis of the RR.

Steps for conducting improved meta analysis are as follows.

Performing a standard meta analysis of the set of study effects. Testing the homogeneity of the studies and check weather the study effects differ significantly. Get the observed value of the test statistic and combine it with treatment effects to get the log predictive relative risks as

$$\ln(\text{PRR}) = \text{common } \ln(\text{RR}) + \text{c (observed } \ln(\text{RR}) - \text{common } \ln(\text{RR})) \tag{3.1}$$

where $c = 1 - [\nu - 2]/[\chi^2 - \text{value}]$ in which $\nu = k - 1$, the degrees of freedom of the χ^2 statistic.

We estimate 95% confidence intervals for the study effects and check whether the value of the estimated pooled effect passes through the confidence limits of all the trials and their concomitant error bars.

If there exists any outlier in the set of studies, we perform another meta analysis excluding the trial from our analysis.

If there are more than one outliers then we consider the amount of diversity of the respective study effect and the weight of the study before excluding any of them from the final study. Finally we perform another improved meta analysis with the rest of trials and check the results.

3.2 Analytical Results

In the second meta analysis, the observed value of χ^2 is 28.81 with 8 degrees of freedom. The value of the constant for calculating the predicted relative risk is

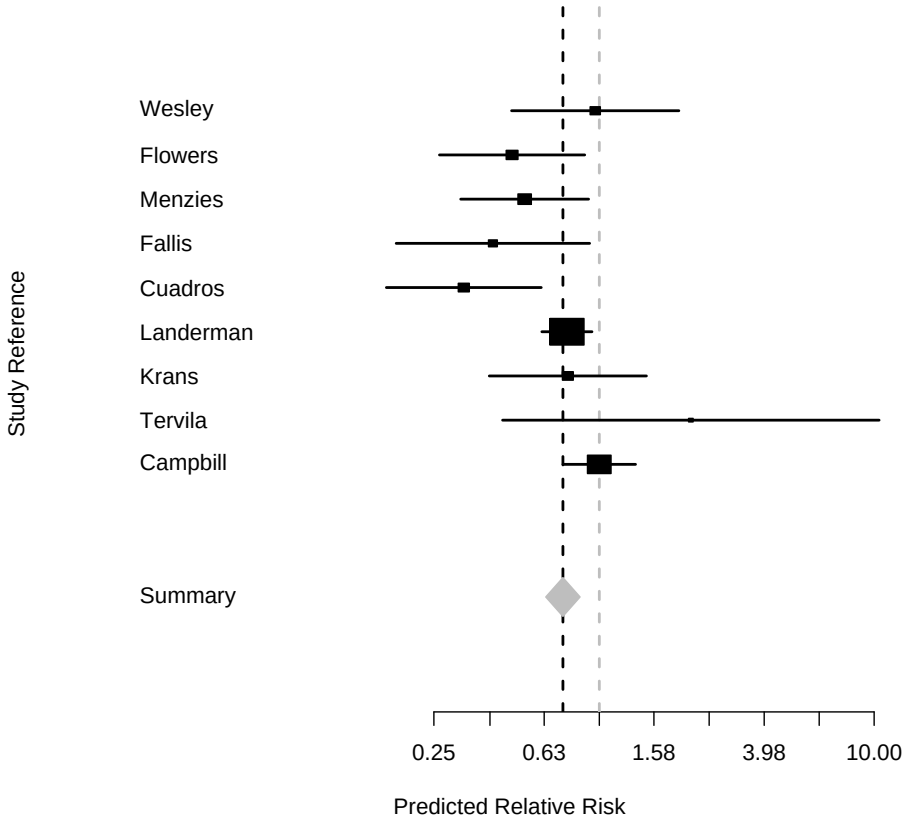

Figure 3: 95% Confidence Intervals for Predicted Relative Risk for all trials

$c = (1 - 6/28.81) = 0.79$. Using the value we get the estimated log predicted relative risk for Wesley trial as
$ln(PRR) = -0.34249 + 0.79(0.039221 - (-0.34249)) = -0.049$ and hence the PRR for the trial becomes $exp(-0.049) = 0.960$ Predicted relative risks and 95% confidence intervals for all the nine studies are shown in Table 4 and presented in Figure 3.

Table 4 Improved meta analysis for nine studies.

Trials	RR	CI lower limit	CI upper limit
Wesley	0.97	0.48	1.94
Flowers	0.48	0.26	0.88
Menzies	0.53	0.31	0.94
Fallis	0.41	0.18	0.94
Cuadros	0.32	0.17	0.61
Landerman	0.76	0.62	0.94
Krans	0.77	0.40	1.48
Tervila	2.15	0.44	10.34
Campbill	1.00	0.74	1.38
Meta-analysis	**0.71**	**0.62**	**0.82**

From Tables 4 and 5 we see that the predicted relative risks are much closer to the pooled relative risk than the relative risks estimated by standard meta analysis. We get a better set of estimates. However there exists two extreme values here. The PRR for Cuadros trial is 0.32 which implies 68% reduction in the risk of preeclampsia attributed to the use of diuretics. Whereas, the Tervila trial shows 1.13 times increase of the risk for the diuretics user group. For rest the of the trials, the PRR's take values between 0.41 and 1.00.

From Figure 3 we see that the confidence intervals for the PRR's of the trials, Wesley, Landermad, Tervila amd Campbell, are shifted to the left and that for the other trials shifted to the right. All the CI's are closer to the value of the pooled estimate of the PRR than relative risk for standard meta analysis. Unlike the standard meta analysis the vertical line through the pooled PRR passes through the concomitant error bar of the PRR of Campbill trial. However the confidence limits of the PRR of Cuadros trial don't contain value of the combined estimate. In all aspects Cuadros trial is found to be a outlier of the set of studies. It may be due to the clinical difference of the study from the others but the actual reason is not known to us. We then perform the meta analysis excluding the Cuadros trial.

Table 5 Improved Meta Analysis for Eight Studies (except the study of Cuadros).

	Standard Meta Analysis			**Improved Meta Analysis**		
Trials	**RR**	**CI lower**	**CI upper**	**PRR**	**CI lower**	**CI upper**
Wesley	1.04	0.52	2.09	0.94	0.47	1.90
Flowers	0.43	0.23	0.79	0.52	0.29	0.96
Menzies	0.49	0.29	0.84	0.57	0.34	0.98
Fallis	0.35	0.16	0.79	0.46	0.20	1.03
Landerman	0.77	0.62	0.95	0.77	0.63	0.95
Krans	0.78	0.40	1.50	0.78	0.40	1.50
Tervila	2.86	0.59	13.85	1.86	0.38	8.98
Campbill	1.08	0.80	1.47	0.97	0.71	1.31
Meta-analysis	**0.78**	**0.67**	**0.90**			

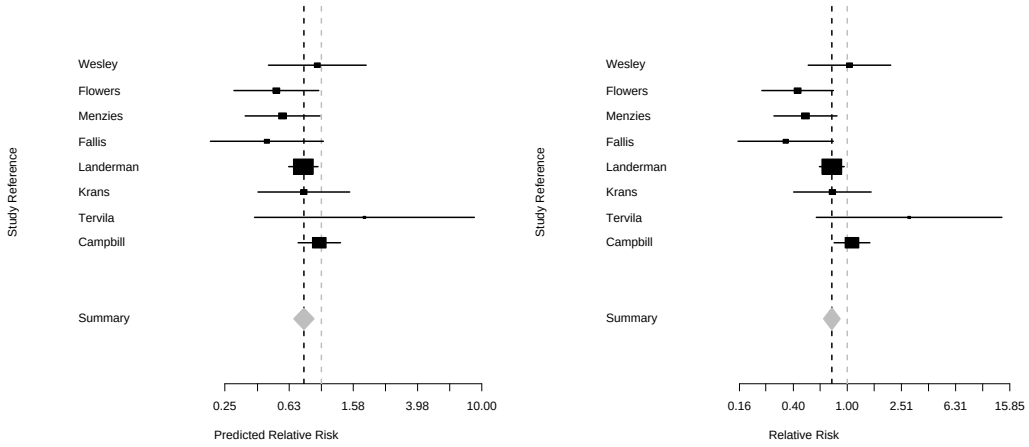

Figure 4: 95% Confidence intervals for predicted relative risk and relative risk of eight studies

Revised Analysis:

We again test the heterogeneity of the eight studies considered for final study. The observed value of chi-square for the studies is 18.05 with 7 degrees of freedom. The observed value is larger than the critical value (14.07) of chi-square at .05 level with the same degrees freedom. There still exist heterogeneity among the studies. The conventional meta analysis is not able to give a conclusive result and the pooled value of the relative risk does not pass through the confidence limits of all the trials and their concomitant error bars. We perform improved meta analysis once again. The value of c in this case is $\frac{7-2}{18.05} = 0.72$.

The results of both standard and improved meta analyses from the remaining eight studies are presented in Table 5. Figure 4 shows the graphical representation of the two sets of study effects. Standard meta analysis of the 8 studies shows the estimated relative risks of individual studies ranges from 0.35 to 2.86. It also reveals the fact that there are three studies with estimated relative risks more than one. Moreover there exists very high variability among the treatment effects. Based on the available results, if we consider Tervila's trial as an outlier, there still exists heterogeneity. Fallis' trial indicates 65% reduction in the estimated relative risk of preeclampsia for the treatment group with respect to the control group. Whereas for the Campbill's trial, treatment group possesses 8% more risk than the control group. Excluding the study of Cuadros from the analysis, we see that the predicted relative risks of the other trials are getting closer to the value of the pooled PRR with the exception of Tervila (with 1.86 PRR) all other PRRs lie between 0.46 and

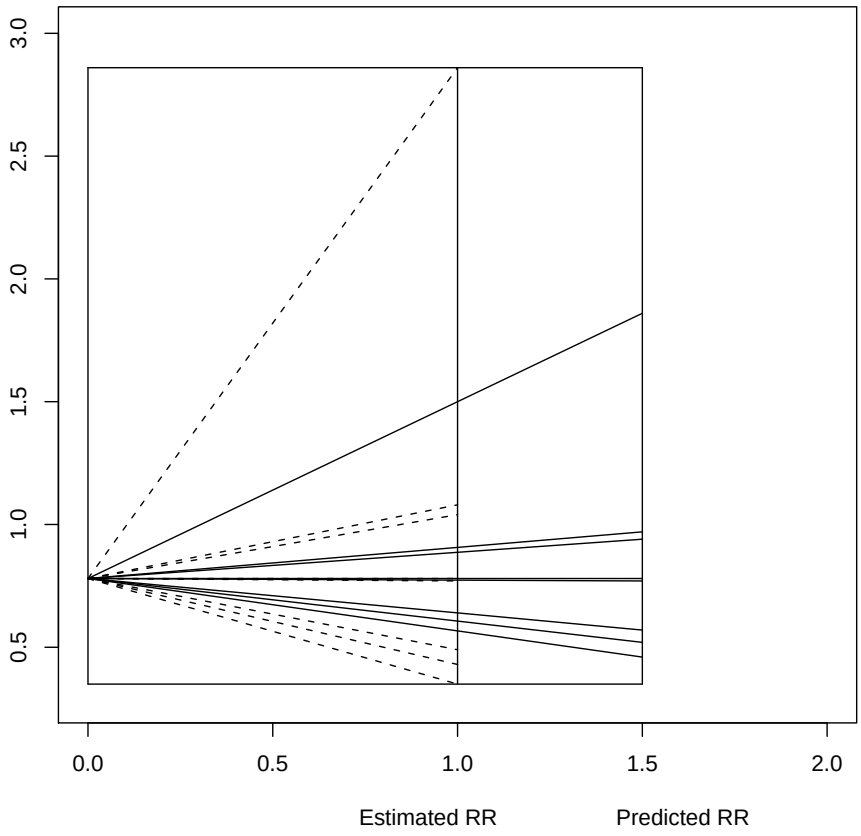

Figure 5: Display of estimated relative risks and predicted relative risks of eight studies

0.97. The values of PRRs indicate a reduction in the risk of preeclampsia for the treatment group with respect to the control group. The percentages of reduction in the relative risks due to the treatment varies from 3 to 54. The study of Tervila possesses relatively lower effect, and hence has a little impact on the RR of combined study. The estimated pooled value of the predicted relative risk is 0.78. The vertical line from this point passes through the 95% confidence limits of all the study effects and their concomitant error bars. So it may be considered that the pooled estimate of the effects is a representative value of all the studies.

The ray-plot in Figure 5 demonstrates how the predicted relative risks shrink towards the pooled estimate of the relative risks. The vertical lines, here labeled as the estimated relative risk and predicted relative risk respectively, hold the rays emitting from the pooled value of the RR from all the studies. It is clear from the

figure that the rays to the line of predicted RR are concentrating more towards the pooled RR (at the centre) than those to the line of estimated RR. Thus the improved estimates of RR's based on the predicted RR are providing better results with respect to the spread of the estimates.

4 Concluding remarks

The foregoing analyses show that the improved meta analysis based on the predicted relative risk provide shorter confidence interval compared to the standard meta analysis based on estimated relative risk. Also, the new method helps shrink the limits of the confidence intervals closer to the pooled estimate than the standard method. Particularly the benefit of the use of the new method is evident when the effects of the studies are heterogenous.

Acknowledgements

The authors are thankful to Professor A K Md E Saleh of Carleton University, Canada for his constructive criticism and thoughtful suggestions to improve the final version of the paper.

REFERENCES

1. Dickersen K, Berlin J. Meta-analysis: State-of-the-Science. *Epidemiologic Reviews* 1992; **14**: 154-176.
2. Manolio TA, Pearson TA, Wenger NK, Barrett-Connor E, Payne GH, Harlan WR. Cholesterol and heart disease in older persons and women. *Review of an NHLBI workshop. A. Epidemiol.* 1992; **2**: 167-176.
3. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ* 1994; **309**: 1351-1355.
4. Saleh AKMdE, Hassanein KM, Hassanein RS, Kim HM. Quasi-empirical bayes methodology for improving meta analysis *Journal of Biopharmaceutical Statistics*. 2006; **16**: 77-90.
5. Katz D, Dpatista J, Azen SP, Pike MC. Obtaining confidence intervals for the risk ratio in cohort studies. *Biometrics*, 1978; **34**: 469-474.
6. Egger M, Smith GD, Phillips AN. Meta analysis: Principles and procedures. *BMJ* 1997; **315**: 1533-1537.
7. Schork MA, Remington RD. Statistics with Applications to the Biological and Health Sciences. Prentice Hall. New Jersey: 2000.
8. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 1959; **22**: 719-748.
9. Thompson SG, and Pocock SJ. Can meta analysis be trusted? *Lancet* 1987; **338**: 1127-1130.
10. Bancroft TT. On biases in estimation due to the use of preliminary test of significance, *Ann. Math. Stat* 1944; **15**: 190-194.
11. Han CP, Bancroft TA. On pooling means when variance is unknown, *J. Am. Stat. Assoc.* 1968; **63**: 1333-1342.
12. Saleh AKMdE. Theory of Preliminary test and Stein-Type Estimation with Applications. Wiley, New Jersey. 2006; 175.