Scientific Research Publishing

Open Journal of Statistics, 2022, 12, 49-69
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

# Minimum MSE Weighted Estimator to Make Inferences for a Common Risk Ratio across Sparse Meta-Analysis Data

**Chukiat Viwatwongkasem[1*], Sutthisak Srisawad[2], Pichitpong Soontornpipit[1], Jutatip Sillabutra[1], Pratana Satitvipawee[1], Prasong Kitidamrongsuk[1], Hathaikan Chootrakool[1]**

[1]Department of Biostatistics, Faculty of Public Health, Mahidol University, Bangkok, Thailand
[2]Nursing Research Center, Faculty of Nursing, Mahidol University, Nakhon Pathom, Thailand
Email: *chukiat.viw@mahidol.ac.th, sutthisak.sri@mahidol.ac.th, pichitpong.soo@mahidol.ac.th, jutatip.sil@mahidol.ac.th, pratana.sat@mahidol.ac.th, prasong.kit@mahidol.ac.th, hathaikan.cho@mahidol.ac.th

**How to cite this paper:** Viwatwongkasem, C., Srisawad, S., Soontornpipit, P., Sillabutra, J., Satitvipawee, P., Kitidamrongsuk, P. and Chootrakool, H. (2022) Minimum MSE Weighted Estimator to Make Inferences for a Common Risk Ratio across Sparse Meta-Analysis Data. *Open Journal of Statistics*, **12**, 49-69.
https://doi.org/10.4236/ojs.2022.121004

**Received:** December 31, 2021
**Accepted:** February 11, 2022
**Published:** February 14, 2022



## Abstract

The paper aims to discuss three interesting issues of statistical inferences for a common risk ratio (*RR*) in sparse meta-analysis data. Firstly, the conventional log-risk ratio estimator encounters a number of problems when the number of events in the experimental or control group is zero in sparse data of a 2 × 2 table. The adjusted log-risk ratio estimator with the continuity correction points $(c_1, c_2) = \left(\frac{1}{6}, \frac{1}{6}\right)$ based upon the minimum Bayes risk with respect to the uniform prior density over (0, 1) and the Euclidean loss function is proposed. Secondly, the interest is to find the optimal weights $\hat{f}_j$ of the pooled estimate $\hat{\theta}_w = \sum_{j=1}^{k} \hat{\theta}_{cj} \hat{f}_j$ that minimize the mean square error (MSE) of $\hat{\theta}_w$ subject to the constraint on $\sum_{j=1}^{k} \hat{f}_j = 1$ where $\hat{\theta}_{cj} = \log\left(\widehat{RR}_{cj}\right) = \log\left(\hat{p}_{c1j}\right) - \log\left(\hat{p}_{c2j}\right)$, $\hat{p}_{c1j} = \left(X_{1j} + c_1\right) / \left(n_{1j} + 2c_1\right)$, $\hat{p}_{c2j} = \left(X_{2j} + c_2\right) / \left(n_{2j} + 2c_2\right)$. Finally, the performance of this minimum MSE weighted estimator adjusted with various values of points $c = c_1 = c_2$ is investigated to compare with other popular estimators, such as the Mantel-Haenszel (MH) estimator and the weighted least squares (WLS) estimator (also equivalently known as the inverse-variance weighted estimator) in senses of point estimation and hypothesis testing via simulation studies. The results of estimation illustrate that regardless of the true values of *RR*, the MH estimator achieves the best performance with the smallest MSE when the study size is rather large ($k \geq 16$) and the sample sizes within each study are small. The MSE of WLS estimator and the proposed-weight estimator adjusted by $c = 1/6$, or $c = 1/3$, or $c = 1/2$ are close together and they are the best when the sample sizes are moderate to large ( $n_{1j} \geq 16$ and $n_{2j} \geq 16$ ) while the study size is rather small.

**Keywords**

Minimum MSE Weights, Adjusted Log-Risk Ratio Estimator, Sparse Meta-Analysis Data, Continuity Correction

## 1. Introduction

Frequently, biostatisticians would like to evaluate the effects of treatments or risk factors in terms of risk difference, relative risk (risk ratio), and/or odds ratio between two independent sample groups (e.g., treatment or control, presence or absence of a risk factor) and binary outcomes (e.g., disease or non-disease, success or failure, dead or alive) in a 2 × 2 table. Let $p_1$ be the probability of outcome in the treatment or exposed group and $p_2$ be the probability of outcome in the control or unexposed group. On comparing two independent groups, the popular effect parameters are defined by the risk difference $RD = p_1 - p_2$, the relative risk $RR = p_1 / p_2$, or the odds ratio $OR = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$. Obviously, all three effect sizes are related to the estimation of proportion $p$ on each arm. It has widely been known that the conventional proportion estimator $\hat{p} = X/n$ is a good choice for estimating $p$ in general, but may not be in sparse data. For example, in sparse data coping with the small number of events $X$ and the small sample size $n$, the variance of $\hat{p}$, estimated by $\frac{\hat{p}(1-\hat{p})}{n}$, can cause a problem with a value of 0 when $X = 0$ or $X = n$. To solve the problem, a continuity correction term $c$ is often added to each cell of each group in the 2 × 2 table, yielding to $\hat{p}_c = (X + c)/(n + 2c)$ in each arm. Yate [1] first used the continuity correction of 0.5 in the approximation of a discrete distribution to a continuous one in 1934. And it seems that the correction value of 0.5 has been used extensively until now, for examples: Lane [2], Stijnen *et al.* [3], White *et al.* [4], Lui and Lin [5], Sankey *et al.* [6], Gart and Zwefel [7], Walter [8], and Cox [9] used value 0.5 adjustment for zero observations in each cell of the 2 × 2 table. Another choice of $c$ for this class such as 0.25, 0.5 and 1 had been suggested by Li and Wang [10]; 1/6 by Turkey [11] and Sánchez-Meca and Marin-Martinez [12]; Böhning and Viwatwongkasem [13] showed that the simple adjusted estimate $\hat{p}_c = (X + c)/(n + 2c)$ with $c = 1$ performed surprisingly well with the smallest average MSE. In addition, Agresti and Caffo [14] also suggested a value of 1 to solve the zero observations; however, a correction value of 2 was recommended by MaClave and Sincich [15].

Our focus of interest is not on the simple adjusted estimate $\hat{p}_c = (X + c)/(n + 2c)$, but on the logarithm of $\hat{p}_c$ instead, leading to the final interest to the logarithm of the risk ratio estimate, $\hat{\theta}_c = \log \widehat{RR}_c = \log \hat{p}_{c1} - \log \hat{p}_{c2} = \log\left(\frac{X_1 + c_1}{n_1 + 2c_1}\right) - \log\left(\frac{X_2 + c_2}{n_2 + 2c_2}\right)$. The reason is that we are interested in the logarithm function, because it is widely known that the risk ratio has a non-symmetric and rather right-skewed distribution about the null value of 1; consequently, the natural logarithm of the risk ratio is needed to transform to be more reliable for the normal distribution. For estimating $\log\left(\frac{p + c/n}{1 + 2c/n}\right)$, Pettigrew, Gart and Thomas [16] proposed the moment estimator $\log\left(\frac{X + c}{n + 2c}\right)$ with the bias and the first four cumulants by means of asymptotic Taylor's series expansion. Likewise, we re-derive and expand on the details of the solution properties for the adjusted estimator of the log-risk ratio, $\hat{\theta}_c = \log \widehat{RR}_c = \log \hat{p}_{c1} - \log \hat{p}_{c2} = \log\left(\frac{X_1 + c_1}{n_1 + 2c_1}\right) - \log\left(\frac{X_2 + c_2}{n_2 + 2c_2}\right)$.

Therefore, the first objective of the study is to find the optimal points $(c_1, c_2)$ based upon continuity correction of this adjusted log-risk ratio estimator under various settings, such as the smallest bias, the smallest average bias, the smallest MSE, and the smallest average MSE.

Secondly, after obtaining the optimal value of $(c_1, c_2)$ which is the best choice when considering the smallest bias and/or the smallest MSE of $\hat{\theta}_c$. The next focus of interest is concerned with sparse data in meta-analysis studies that combine the various risk ratios from $k$ studies to produce a single summary risk ratio. Under a common risk ratio overall studies or homogeneity of risk ratios across $k$ studies, the second aim of the study is to find the optimal weights $\hat{f}_j$ of the pooled estimate $\hat{\theta}_w = \sum_{j=1}^{k} \hat{\theta}_{cj} \hat{f}_j$ that minimize the MSE of $\hat{\theta}_w$ subject to the constraint on $\sum_{j=1}^{k} \hat{f}_j = 1$ where $\hat{\theta}_{cj} = \log\left(\widehat{RR}_{cj}\right) = \log\left(\hat{p}_{c1j}\right) - \log\left(\hat{p}_{c2j}\right)$, $\hat{p}_{c1j} = \left(X_{1j} + c_1\right) / \left(n_{1j} + 2c_1\right)$, $\hat{p}_{c2j} = \left(X_{2j} + c_2\right) / \left(n_{2j} + 2c_2\right)$. Indeed, these optimal weights $\hat{f}_j$ actually come from the roughly similar formula to the work of Viwatwongkasem *et al.* [19] that was used for the risk difference in multi-center studies.

Finally, the final objective of the study of interest is about comparing the performance of well-known summary effect estimators, such as the Mantel-Haenszel (MH) estimator stated with its variance estimate by Greenland and Robin [20] and the weighted least square (WLS) estimator or equivalently known as the inverse-variance weighted estimator to the new adjusted summary effect estimator $\hat{\theta}_w$ with various adjustment points $(c_1, c_2)$ in terms of point estimation and hypothesis testing via simulation studies.

The rest of the paper is organized as follows. Section 2 contains the derivative methods and the results of the adjusted log-risk ratio estimator. Section 3 discusses on the methodology and the outcomes of the minimum MSE weights of the adjusted summary relative risk estimator. Section 4 states on the other well-known estimators and tests as the comparative candidates to compare the performances of making inference for a common relative risk across meta-analysis studies. Section 5 consists of the simulation plans for studying the performances in senses of the estimation, the type I errors and the power of the tests. Section 6 contains the results of Section 5. Section 7 is about the discussion and the recommendation.

## 2. Adjusted Estimator of the Logarithm of the Risk Ratio

In a single study ( $k=1$ ), since the conventionally popular estimator of the logarithm of risk ratio, obtained by $\hat{\theta}=\log\widehat{RR}=\log\hat{p}_1-\log\hat{p}_2=\log\left(\frac{X_1}{n_1}\right)-\log\left(\frac{X_2}{n_2}\right)$, may have a problem when data are sparse. To solve the problem, a continuity correction term $(c_1,c_2)$ is often added to each cell of the 2 × 2 table, leading to the adjusted estimate as $\hat{\theta}_c=\log\widehat{RR}_c=\log\hat{p}_{c1}-\log\hat{p}_{c2}=\log\left(\frac{X_1+c_1}{n_1+2c_1}\right)-\log\left(\frac{X_2+c_2}{n_2+2c_2}\right)$. In addition, the adjusted risk ratio estimator can be further found as $\widehat{RR}_c=\frac{\hat{p}_{c1}}{\hat{p}_{c2}}=\frac{(X_1+c_1)/(n_1+2c_1)}{(X_2+c_2)/(n_2+2c_2)}$. Due to the work of Pettigrew, Gart and Thomas [16], we re-derive and extend the results to get the expectation, bias, variance, and MSE of $\hat{\theta}_c$ in the following:

$$E\left(\hat{\theta}_c\right)=\left(\log p_1-\log p_2\right)+B\left(\hat{\theta}_c\right) \tag{1}$$

$$\begin{aligned}B\left(\hat{\theta}_c\right)=&\frac{2c_1-1+p_1\left\{1-2\left(2c_1\right)\right\}}{2n_1p_1}\\&+\frac{-6c_1^2+6\left(2c_1p_1\right)^2+12c_1q_1+4q_1\left(q_1-p_1\right)-9q_1^2}{12\left(n_1p_1\right)^2}\\&-\left[\frac{2c_2-1+p_2\left\{1-2\left(2c_2\right)\right\}}{2n_2p_2}\right.\\&\left.+\frac{-6c_2^2+6\left(2c_2p_2\right)^2+12c_2q_2+4q_2\left(q_2-p_2\right)-9q_2^2}{12\left(n_2p_2\right)^2}\right]+O\left(n^{-3}\right)\end{aligned} \tag{2}$$

$$V\left(\hat{\theta}_c\right)=\frac{q_1}{n_1p_1}+\frac{q_1\left(1-2c_1\right)+\frac{1}{2}q_1^2}{\left(n_1p_1\right)^2}+\frac{q_2}{n_2p_2}+\frac{q_2\left(1-2c_2\right)+\frac{1}{2}q_2^2}{\left(n_2p_2\right)^2}+O\left(n^{-3}\right) \tag{3}$$

$$\begin{aligned}MSE\left(\hat{\theta}_c\right)=&\frac{q_1}{n_1p_1}+\frac{\left(c_1-2c_1p_1\right)\left(c_1-2c_1p_1-q_1\right)+q_1\left(1-2c_1\right)+\frac{3}{4}q_1^2}{\left(n_1p_1\right)^2}+\frac{q_2}{n_2p_2}\\&+\frac{\left(c_2-2c_2p_2\right)\left(c_2-2c_2p_2-q_2\right)+q_2\left(1-2c_2\right)+\frac{3}{4}q_2^2}{\left(n_2p_2\right)^2}+O\left(n^{-3}\right)\end{aligned} \tag{4}$$

when $q_1=1-p_1$ and $q_2=1-p_2$.

The first setting to find the optimal point $(c_1,c_2)$ that minimizes the smallest bias of $\hat{\theta}_c$ is investigated. The first derivatives of the bias of $\hat{\theta}_c$ with respect to $c_1$ and $c_2$ are given by

$$\frac{\partial B\left(\hat{\theta}_c\right)}{\partial c_1}=\frac{1+\left(-1+n_1\right)p_1-2n_1p_1^2+c_1\left(-1+4p_1^2\right)}{n_1^2p_1^2} \tag{5}$$

$$\frac{\partial B\left(\hat{\theta}_c\right)}{\partial c_2}=-\left[\frac{1+\left(-1+n_2\right)p_2-2n_2p_2^2+c_2\left(-1+4p_2^2\right)}{n_2^2p_2^2}\right] \tag{6}$$

Setting $\frac{\partial B\left(\hat{\theta}_c\right)}{\partial c_1}=0$ and $\frac{\partial B\left(\hat{\theta}_c\right)}{\partial c_2}=0$, the critical roots $\left(c_1,c_2\right)$ are obtained as follows:

$$c_1=\frac{-1+p_1-n_1p_1+2n_1p_1^2}{-1+4p_1^2},\quad p_1\neq 0.5 \tag{7}$$

$$c_2=\frac{-1+p_2-n_2p_2+2n_2p_2^2}{-1+4p_2^2},\quad p_2\neq 0.5 \tag{8}$$

The sufficient conditions of the second-order derivatives of the bias of $\hat{\theta}_c$ to guarantee the solutions of (7) and (8) being minimum points are that the determinants $D_1$ and $D_2$, evaluated at critical points, are all positive where $D_1=\frac{\partial^2 B\left(\hat{\theta}_c\right)}{\partial c_1^2}$ and $D_2=\begin{vmatrix}\frac{\partial^2 B\left(\hat{\theta}_c\right)}{\partial c_1^2} & \frac{\partial^2 B\left(\hat{\theta}_c\right)}{\partial c_1\partial c_2}\\ \frac{\partial^2 B\left(\hat{\theta}_c\right)}{\partial c_1\partial c_2} & \frac{\partial^2 B\left(\hat{\theta}_c\right)}{\partial c_2^2}\end{vmatrix}$. Unfortunately, it is impossible to find the minimum point $\left(c_1,c_2\right)$ such that $\hat{\theta}_c=\log\left(\frac{X_1+c_1}{n_1+2c_1}\right)-\log\left(\frac{X_2+c_2}{n_2+2c_2}\right)$ has the smallest bias for all $\left(p_1,p_2\right)$. In addition, the solution to (7) and (8) is practically inflexible because it cannot be applied when $p_1$ and $p_2$ equal 0.5. Therefore, we further consider the other settings, such as the smallest average bias and the smallest MSE, to find the optimal points $\left(c_1,c_2\right)$. However, these settings still fail, they cannot provide the optimal points $\left(c_1,c_2\right)$.

Another successful setting for finding the optimal values $\left(c_1,c_2\right)$ is coming from an average MSE or equivalently known as Bayes risk. Suppose that the squared error loss function is given by $\text{Loss}=\left(\hat{\theta}_c-\theta\right)^2$. The risk as the expectation of loss or the MSE of $\hat{\theta}_c$ in such this case, is given by $\text{Risk}=MSE\left(\hat{\theta}_c\right)=E\left(\hat{\theta}_c-\theta\right)^2$. The prior distribution is usually constructed based on the previously scientific knowledge or the prior observed data; especially, in this paper we set the prior distribution based on **Table 1**. For given the prior uniform density $g\left(p_1,p_2\right)=1_{[0,1]}\times 1_{[0,1]}$ over [0, 1] × [0, 1], the Bayes risk of $\hat{\theta}_c$ (or the average MSE of $\hat{\theta}_c$) with respect to the Euclidean loss function is obtained as

$$\text{Bayes risk}=m\left(c_1,c_2\right)=\int_0^1\int_0^1 MSE\left(\hat{\theta}_c\right)g\left(p_1,p_2\right)\mathrm{d}p_1\mathrm{d}p_2$$

**Table 1.** High sparse data of a multi-center clinical trial in CALGB study (Cancer and Leukemia Group B) from Lipsitz *et al.* [17] and Cooper *et al.* [18].

| Center $j$ | Treatment | | Control | | $\log\left(\frac{X_{1j}}{n_{1j}}\right)-\log\left(\frac{X_{2j}}{n_{2j}}\right)$ |
|---|---|---|---|---|---|
| | $X_{1j}$ | $n_{1j}$ | $X_{2j}$ | $n_{2j}$ | |
| 1 | 3 | 4 | 1 | 3 | 0.811 |
| 2 | 3 | 4 | 8 | 11 | 0.031 |
| 3 | 2 | 2 | 2 | 3 | 0.405 |
| 4 | 2 | 2 | 2 | 2 | 0.000 |
| 5 | 2 | 2 | 0 | 3 | * |
| 6 | 1 | 3 | 2 | 3 | −0.693 |
| 7 | 2 | 2 | 2 | 3 | 0.405 |
| 8 | 1 | 5 | 4 | 4 | −1.609 |
| 9 | 2 | 2 | 2 | 3 | 0.405 |
| 10 | 0 | 2 | 2 | 3 | * |
| 11 | 3 | 3 | 3 | 3 | 0.000 |
| 12 | 2 | 2 | 0 | 2 | * |
| 13 | 1 | 4 | 1 | 5 | 0.223 |
| 14 | 2 | 3 | 2 | 4 | 0.288 |
| 15 | 2 | 4 | 4 | 6 | −0.288 |
| 16 | 4 | 12 | 3 | 9 | 0.000 |
| 17 | 1 | 2 | 2 | 3 | −0.288 |
| 18 | 3 | 3 | 1 | 4 | 1.386 |
| 19 | 1 | 4 | 2 | 3 | −0.981 |
| 20 | 0 | 3 | 0 | 2 | * |
| 21 | 2 | 4 | 1 | 5 | 0.916 |

*Either $X_{1j}$ or $X_{2j}$ is 0, leading to the infinity value problem of the natural logarithm.

$$m(c_1,c_2)=\int_0^1\int_0^1 \frac{q_1}{n_1p_1}+\frac{(c_1-2c_1p_1)(c_1-2c_1p_1-q_1)+q_1(1-2c_1)+\frac{3}{4}q_1^2}{(n_1p_1)^2}+\frac{q_2}{n_2p_2}+\frac{(c_2-2c_2p_2)(c_2-2c_2p_2-q_2)+q_2(1-2c_2)+\frac{3}{4}q_2^2}{(n_2p_2)^2}\mathrm{d}p_1\mathrm{d}p_2$$

$$m(c_1,c_2)=\frac{(-1+c_1+3c_1^2)n_2^2-n_1n_2^2+n_1^2(3.25-c_2-3c_2^2+n_2)}{n_1^2n_2^2} \tag{9}$$

The first and second order of partial derivatives of $m(c_1,c_2)$ with respect to

$c_1$ and $c_2$ are as follows: $\frac{\partial m(c_1,c_2)}{\partial c_1}=\frac{1+6c_1}{n_1^2}$, $\frac{\partial m(c_1,c_2)}{\partial c_2}=\frac{-1-6c_2}{n_2^2}$, $\frac{\partial^2 m(c_1,c_2)}{\partial c_1^2}=\frac{6}{n_1^2}$, $\frac{\partial^2 m(c_1,c_2)}{\partial c_2^2}=\frac{-6}{n_2^2}$, $\frac{\partial^2 m(c_1,c_2)}{\partial c_1 \partial c_2}=0$. Unfortunately, the result of Bayes risk shows that the critical point $(c_1,c_2)=\left(-\frac{1}{6},-\frac{1}{6}\right)$ is not a minimum point. With the conditions of $D_1=\frac{\partial^2 m(c_1,c_2)}{\partial c_1^2}$ and $D_2=\begin{vmatrix}\frac{\partial^2 m(c_1,c_2)}{\partial c_1^2} & \frac{\partial^2 m(c_1,c_2)}{\partial c_1 \partial c_2}\\ \frac{\partial^2 m(c_1,c_2)}{\partial c_1 \partial c_2} & \frac{\partial^2 m(c_1,c_2)}{\partial c_2^2}\end{vmatrix}$, the critical point $(c_1,c_2)$ is a saddle point since $D_2=\frac{-36}{n_1^2 n_2^2}<0$. However, in particular case, we let $c=c_1=c_2$ and try again to find the minimum point. Fortunately, with the condition of $n_2>n_1$, the optimal point $c$ with the smallest average MSE is obtained by $c=-\frac{1}{6}$; equivalently, with the condition of $n_1>n_2$, the minimum point is $c=\frac{1}{6}$. The solution of $c=\frac{-1}{6}$ or $c=\frac{1}{6}$ looks well and can be considered as an appropriate one in practice.

## 3. Minimum MSE Weights of the Adjusted Summary Relative Risk Estimator

Under a common risk ratio overall $k$ studies or homogeneity of risk ratios across $k$ studies, the optimal weights $f_j$ are investigated to minimize the MSE of $\hat{\theta}_w$ of the form $\hat{\theta}_w=\sum_{j=1}^{k}\hat{\theta}_{cj} f_j$, subject to the constraint on $\sum_{j=1}^{k} f_j=1$, where $\hat{\theta}_{cj}=\log\left(\widehat{RR}_{cj}\right)=\log\left(\hat{p}_{c1j}\right)-\log\left(\hat{p}_{c2j}\right)$, $\hat{p}_{c1j}=\left(X_{1j}+c_1\right)/\left(n_{1j}+2c_1\right)$ and $\hat{p}_{c2j}=\left(X_{2j}+c_2\right)/\left(n_{2j}+2c_2\right)$. The MSE of $\hat{\theta}_w=\sum_{j=1}^{k}\hat{\theta}_{cj} f_j$, used under a true common risk ratio $\theta$ across all $k$ studies, is given by

$$MSE\left(\hat{\theta}_w\right)=E\left(\hat{\theta}_w-\theta\right)^2=E\left(\sum_{j=1}^{k}\hat{\theta}_{cj} f_j-\theta\right)^2$$

To find the optimal weights $f_j$ with the constraint on $\sum_{j=1}^{k} f_j=1$, we form the auxiliary function $\phi$ in extending $\text{MSE}\left(\hat{\theta}_w\right)$ under the Lagrange's method where $\lambda$ is a Lagrange multiplier as follows:

$$\phi=E\left(\sum_{j=1}^{k}\hat{\theta}_{cj} f_j-\theta\right)^2+\lambda\left(\sum_{j=1}^{k} f_j-1\right)$$

$$\phi=E\left(\sum_{j=1}^{k}\hat{\theta}_{cj} f_j\right)^2-2\theta E\left(\sum_{j=1}^{k}\hat{\theta}_{cj} f_j\right)+\theta^2+\lambda\left(\sum_{j=1}^{k} f_j-1\right)$$

$$\phi=Var\left(\sum_{j=1}^{k}\hat{\theta}_{cj} f_j\right)+\left(E\left(\sum_{j=1}^{k}\hat{\theta}_{cj} f_j\right)\right)^2-2\theta E\left(\sum_{j=1}^{k}\hat{\theta}_{cj} f_j\right)$$
$$+\theta^2+\lambda\left(\sum_{j=1}^{k} f_j-1\right)$$

$$\phi = \sum_{j=1}^{k} f_j^2 Var\left(\hat{\theta}_{cj}\right) + \left(\sum_{j=1}^{k} f_j E\left(\hat{\theta}_{cj}\right)\right)^2 - 2\theta\left(\sum_{j=1}^{k} f_j E\left(\hat{\theta}_{cj}\right)\right) + \theta^2 + \lambda\left(\sum_{j=1}^{k} f_j - 1\right)$$

$$\phi = \sum_{j=1}^{k} f_j^2 Var\left(\hat{\theta}_{cj}\right) + \left(\sum_{j=1}^{k} f_j E\left(\hat{\theta}_{cj}\right) - \theta\right)^2 + \lambda\left(\sum_{j=1}^{k} f_j - 1\right)$$

From the previous section, we have the following results:

$$V_j = Var\left(\hat{\theta}_{cj}\right) = \frac{q_{1j}}{n_{1j} p_{1j}} + \frac{q_{1j}\left(1-2c_1\right) + \frac{1}{2} q_{1j}^2}{\left(n_{1j} p_{1j}\right)^2} + \frac{q_{2j}}{n_{2j} p_{2j}} + \frac{q_{2j}\left(1-2c_2\right) + \frac{1}{2} q_{2j}^2}{\left(n_{2j} p_{2j}\right)^2} + O\left(n^{-3}\right)$$

and

$$E_j = E\left(\hat{\theta}_{cj}\right) = \left(\log p_{1j} - \log p_{2j}\right) + \left(\frac{2c_1 - 1 + p_{1j}\left\{1 - 2\left(2c_1\right)\right\}}{2n_{1j} p_{1j}} - \frac{2c_2 - 1 + p_{2j}\left\{1 - 2\left(2c_2\right)\right\}}{2n_{2j} p_{2j}}\right) + O\left(n^{-2}\right)$$

The partial derivatives with respect to $\lambda$ and $f_j$ yield

$$\frac{\partial \phi}{\partial \lambda} = \sum_{j=1}^{k} f_j - 1$$

$$\frac{\partial \phi}{\partial f_j} = 2 f_j V_j + 2\left(\sum_{j=1}^{k} f_j E_j - \theta\right) E_j + \lambda$$

Setting $\frac{\partial \phi}{\partial \lambda} = 0$, $\frac{\partial \phi}{\partial f_j} = 0$ and solving the equations, it yields

$$\lambda = -\frac{2}{a} - \frac{2b\left(\sum_{j=1}^{k} f_j E_j - \theta\right)}{a}$$

$$f_j = \left(\frac{V_j^{-1}\left(1 + \tau_j \theta\right)}{a}\right) - \left(\frac{V_j^{-1} \tau_j}{a + \sum_{j=1}^{k} \tau_j E_j V_j^{-1}}\right)\left(\frac{\sum_{j=1}^{k} V_j^{-1} E_j \left(1 + \tau_j \theta\right)}{a}\right)$$

where $a = \sum_{j=1}^{k} \frac{1}{V_j} = \sum_{j=1}^{k} V_j^{-1}$, $b = \sum_{j=1}^{k} \frac{E_j}{V_j}$, $\tau_j = aE_j - b$.

More details of derivation can be found from the work of Viwatwongkasem *et al.* [19] that used the minimum MSE weights for the risk difference in multi-center studies. After replacing the unknown parameter quantities with their sample estimates, the minimum MSE weighted estimate is obtained in the following:

$$\hat{\theta}_w = \sum_{j=1}^{k} \hat{\theta}_{cj} \hat{f}_j \tag{10}$$

where

$$\hat{\theta}_{cj} = \log\left(\widehat{RR}_{cj}\right) = \log\left(\hat{p}_{c1j}\right) - \log\left(\hat{p}_{c2j}\right), \quad \hat{p}_{c1j} = \left(X_{1j} + c_1\right)/\left(n_{1j} + 2c_1\right),$$

$$\hat{p}_{c2j} = \left(X_{2j} + c_2\right)/\left(n_{2j} + 2c_2\right), \quad \hat{q}_{c1j} = 1 - \hat{p}_{c1j}, \quad \hat{q}_{c2j} = 1 - \hat{p}_{c2j}$$

$$\hat{f}_j = \left(\frac{\hat{V}_j^{-1}\left(1 - \hat{\tau}_j \hat{\theta}_{pool}\right)}{\hat{a}}\right) - \left(\frac{\hat{V}_j^{-1}\hat{\tau}_j}{\hat{a} + \sum_{j=1}^{k} \hat{\tau}_j \hat{E}_j \hat{V}_j^{-1}}\right) + \left(\frac{\sum_{j=1}^{k} \hat{E}_j \hat{V}_j^{-1}\left(1 + \hat{\tau}_j \hat{\theta}_{pool}\right)}{\hat{a}}\right)$$

$$\hat{E}_j = \left(\log \hat{p}_{c1j} - \log \hat{p}_{c2j}\right) + \left(\frac{\left\{c_1 - 0.5 + \left(0.5 - 2c_1\right)\hat{p}_{c1j}\right\}}{n_{1j}\hat{p}_{c1j}} - \frac{\left\{c_2 - 0.5 + \left(0.5 - 2c_2\right)\hat{p}_{c2j}\right\}}{n_{2j}\hat{p}_{c2j}}\right) + O\left(n^{-2}\right)$$

$$\hat{V}_j = \frac{\hat{q}_{c1j}}{n_{1j}\hat{p}_{c1j}} + \frac{\hat{q}_{c1j}\left(1 - 2c_1\right) + 0.5\hat{q}_{c1j}^2}{\left(n_{1j}\hat{p}_{c1j}\right)^2} + \frac{\hat{q}_{c2j}}{n_{2j}\hat{p}_{c2j}} + \frac{\hat{q}_{c2j}\left(1 - 2c_2\right) + 0.5\hat{q}_{c2j}^2}{\left(n_{2j}\hat{p}_{c2j}\right)^2} + O\left(n^{-3}\right)$$

$$\hat{\theta}_{pool} = \log\left(\hat{p}_1\right) - \log\left(\hat{p}_2\right)$$

where

$$\hat{p}_1 = \frac{\sum_{j=1}^{k} n_{1j}\hat{p}_{1j}}{\sum_{j=1}^{k} n_{1j}} = \frac{\sum_{j=1}^{k} X_{1j}}{\sum_{j=1}^{k} n_{1j}}, \quad \hat{p}_2 = \frac{\sum_{j=1}^{k} n_{2j}\hat{p}_{2j}}{\sum_{j=1}^{k} n_{2j}} = \frac{\sum_{j=1}^{k} X_{2j}}{\sum_{j=1}^{k} n_{2j}}$$

$$\hat{a} = \sum_{j=1}^{k} \frac{1}{\hat{V}_j} = \sum_{j=1}^{k} \hat{V}_j^{-1}, \quad \hat{b} = \sum_{j=1}^{k} \frac{\hat{E}_j}{\hat{V}_j}, \quad \hat{\tau}_j = \hat{a}\hat{E}_j - \hat{b}$$

Assuming that a normal approximation is reliable, the minimum MSE weights Z-test for testing $H_0: \theta = \theta_0$, ($\theta_0 = \log RR_0$) is

$$Z_{cw} = \frac{\sum_{j=1}^{k} \hat{f}_j \hat{\theta}_{cj} - \theta_0}{\sqrt{\sum_{j=1}^{k} \hat{f}_j^2 \hat{V}ar\left(\hat{\theta}_{cj} \mid H_0\right)}} = \frac{\sum_{j=1}^{k} \hat{f}_j \hat{\theta}_{cj} - \theta_0}{\sqrt{\sum_{j=1}^{k} \hat{f}_j^2 \hat{V}_j}}$$

We will reject $H_0$ at $\alpha$ level for two-sided test if $\left|Z_{cw}\right| > Z_{\alpha/2}$ where $Z_{\alpha/2}$ is the upper $100\left(\alpha^{th}\right)$ percentile of the standard normal distribution. Alternatively, reject $H_0$ when the p-value (*p*) is less than or equal to $\alpha$ where $p = 2\left(1 - \Phi\left(\left|Z_{cw}\right|\right)\right)$ and $\Phi\left(Z\right)$ is the cumulative standard normal distribution.

## 4. Other Well-Known Estimators and Tests for Making Inference for a Common Relative Risk

Under a common risk ratio or homogeneity of risk ratios across *k* studies, we wish to compare the performance of the minimum MSE weighted estimator adjusted by various points $c = 1/6$, $c = 1/3$, $c = 1/2$ with the other well-known summary risk ratio estimators, such as the Mantel-Haenszel (MH) estimator and the weighted least square (WLS) estimator or equivalently known as the inverse-variance weighted estimator via a simulation study. According to these well-known estimators, we will present briefly both estimators.

**Mantel-Haenszel Weights (MH)**

For Mantel-Haenszel (MH) relative risk estimator overall centers/studies from binomial data, the estimator has been proposed by

$$\widehat{RR}_{MH} = \frac{\sum_{j=1}^{k} n_{2j} X_{1j} / N_j}{\sum_{j=1}^{k} n_{1j} X_{2j} / N_j} \quad (11)$$

where $N_j = n_{1j} + n_{2j}$.

The variance estimator of the log-relative risk of Mantel-Haenszel was proposed by Green and Robin [20] based on unconditional binomial distribution is given by

$$\hat{V}\left[\log \widehat{RR}_{MH}\right] = \frac{\sum_{j=1}^{k} D_j}{\left(\sum_{j=1}^{k} R_j\right)\left(\sum_{j=1}^{k} S_j\right)} \quad (12)$$

where:

$D_j = \left(n_{2j} n_{1j} t_j - x_{1j} x_{2j} N_j\right) / N_j^2$, $R_j = x_{1j} n_{2j} / N_j$, $S_j = x_{2j} n_{1j} / N_j$, $t_j = x_{1j} + x_{2j}$ and $N_j = n_{1j} + n_{2j}$. Note that under a binomial sparse-data model, the Mantel-Haenszel relative risk is consistent in sparse stratification.

Assuming that a normal approximation is valid, the Mantel-Haenszel's Z-test for testing $H_0 : \theta = \theta_0$, ($\theta_0 = \log RR_0$) is

$$Z_{MH} = \frac{\log \widehat{RR}_{MH} - \theta_0}{\sqrt{\hat{V}\left(\log \widehat{RR}_{MH} \mid H_0\right)}}$$

We will reject $H_0$ at $\alpha$ level for two-sided test if $\left|Z_{MH}\right| > Z_{\alpha/2}$ where $Z_{\alpha/2}$ is the upper $100\left(\alpha^{th}\right)$ percentile of the standard normal distribution. Alternatively, reject $H_0$ when the p-value ($p$) is less than or equal to $\alpha$ where $p = 2\left(1 - \Phi\left(\left|Z_{MH}\right|\right)\right)$ and $\Phi(Z)$ is the cumulative standard normal distribution.

**Weighted least square (WLS) estimator**

The weighted least square (WLS) estimator or equivalently known as the inverse-variance weighted estimator of the log-relative risk overall centers/studies is

$$\hat{\theta}_{WLS} = \log \widehat{RR}_{WLS} = \sum_{j=1}^{k} w_j \log \widehat{RR}_j \Big/ \sum_{j=1}^{k} w_j \quad (13)$$

where

$$\log \widehat{RR}_j = \log \hat{p}_{1j} - \log \hat{p}_{2j} = \log\left(\frac{X_{1j}}{n_{1j}}\right) - \log\left(\frac{X_{2j}}{n_{2j}}\right)$$

$$w_j = \frac{1}{V\left[\log \widehat{RR}_j\right]} = \left(\frac{1 - p_{1j}}{n_{1j} p_{1j}} + \frac{1 - p_{2j}}{n_{2j} p_{2j}}\right)^{-1} \quad (14)$$

Practically, the weights $w_j$ are often replaced by their estimates.

$$\hat{w}_j = \frac{1}{\hat{V}\left[\log \widehat{RR}_j\right]} = \left(\frac{1-\hat{p}_{1j}}{n_{1j}\hat{p}_{1j}} + \frac{1-\hat{p}_{2j}}{n_{2j}\hat{p}_{2j}}\right)^{-1} = \left(\frac{1}{X_{1j}} - \frac{1}{n_{1j}} + \frac{1}{X_{2j}} - \frac{1}{n_{2j}}\right)$$

The variance of the summary estimator $\hat{\theta}_{WLS}$ is given by

$$V\left(\hat{\theta}_{WLS}\right) = \frac{\sum_{j=1}^{k} w_j^2 V\left[\log \widehat{RR}_j\right]}{\left(\sum_{j=1}^{k} w_j\right)^2} = \frac{\sum_{j=1}^{k} w_j^2 \left(\frac{1}{w_j}\right)}{\left(\sum_{j=1}^{k} w_j\right)^2} = \frac{1}{\sum_{j=1}^{k} w_j} \quad (15)$$

Assuming that a normal approximation is valid, the weighted least square Z-test for testing $H_0 : \theta = \theta_0$, ($\theta_0 = \log RR_0$) is

$$Z_{WLS} = \frac{\sum_{j=1}^{k} \hat{w}_j \log \widehat{RR}_j \Big/ \left(\sum_{j=1}^{k} \hat{w}_j\right) - \theta_0}{\sqrt{1 \Big/ \left(\sum_{j=1}^{k} \hat{w}_j\right)}}$$

We will reject $H_0$ at $\alpha$ level for two-sided test if $\left|Z_{WLS}\right| > Z_{\alpha/2}$ where $Z_{\alpha/2}$ is the upper $100\left(\alpha^{th}\right)$ percentile of the standard normal distribution. Alternatively, reject $H_0$ when the p-value ($p$) is less than or equal to $\alpha$ where $p = 2\left(1-\Phi\left(\left|Z_{WLS}\right|\right)\right)$ and $\Phi(Z)$ is the cumulative standard normal distribution.

## 5. Simulation Plan for the Estimation, Studying the Type I Error and the Power of the Test

We present here a simulation study using the following designs:

**Parameters:** Let the common relative risk be some constants ($RR$ = 1, 2 and 4) and generate the baseline proportion $p_{2j}$ in the control arm for the $j^{th}$ study from a uniform distribution in which the range corresponds to the values of $RR$. If $RR = 1$, then $p_{2j} \sim U(0, 0.9)$; if $RR = 2$, then $p_{2j} \sim U(0, 0.45)$; and if $RR = 4$, then $p_{2j} \sim U(0, 0.23)$. The correspondent proportion risks in the exposure arm is $p_{1j} = p_{2j} \times RR$. The sample size $n_{1j}$ and $n_{2j}$ in each study are fixed and varied as 4, 8, 16, 32 and 100. The number of studies is 1, 2, 4, 8, 16, 32 and 100.

**Statistic 1** (estimation of $RR$): Generate $x_{1j}$ and $x_{2j}$ from the binomial distribution with parameter $\left(p_{1j}, n_{1j}\right)$ and $\left(p_{2j}, n_{2j}\right)$ for each study $j$ ($j = 1, 2, \cdots, k$). All summary estimates are calculated. The procedure is replicated 5000 times. From these replicates, we compute bias, variance, and mean square error (MSE) for the adjusted relative risk estimator with the proposed weight to compare the performance with Mantel-Haenszel (MH) estimator and the weighted least square estimator.

**Statistic 2** (studying the type I error under $H_0 : RR = RR_0$): Generate $x_{1j}$ and $x_{2j}$ from the binomial distribution with parameter $\left(p_{1j}, n_{1j}\right)$ and $\left(p_{2j}, n_{2j}\right)$ and replicate these generations 5000 times for every procedure. From these replicates, the number of null hypothesis rejections when $H_0$ is true under three Z-tests is counted for the actual type I error.

$$\text{The actual type I error} = \frac{\text{Number of rejections of } H_0 \text{ when } H_0 \text{ is true}}{\text{Number of replicates}\left(5000 \text{ times}\right)}$$

**Statistic 3** (studying the power of the test under $H_1$): Before comparing the power of tests, all test statistics should be calibrated to handle the same type I error rate under the null hypothesis. Under the alternative hypothesis with the random effects model, the powers of three candidate tests are compared. We need to revise the parameter setting for studying the power of the test. Let $p_{2j}$ be a uniform distribution over [0, 0.25] and we assume that $\theta_j = \log RR_j$ follows as $\theta_j = 0.1 + U_m = 0.1 + mm(2U - 1)$ where $U_m$ is a uniform $[-mm, mm]$ random variable for a given $mm = 0.2, 0.4, 0.6$, or equivalently, $U$ is a uniform over (0, 1). Note that these parameter settings provide: $E(\theta_j) = 0.1$ and $Var(\theta_j) = (2 \times mm)^2 / 12$. Consequently, we still have $p_{1j} = p_{2j} \times RR$. Binomial variates $x_{1j}$ and $x_{2j}$ are also generated with parameter $(p_{1j}, n_{1j})$ and $(p_{2j}, n_{2j})$, respectively. All proposed test statistics under this alternative hypothesis are computed and replicated 5000 times. From these replicates, the number of null hypothesis rejection is counted for the power of the test.

$$\text{The actual power of the test} = \frac{\text{Number of rejections of } H_0 \text{ when } H_1 \text{ is true}}{\text{Number of replicates}\left(5000 \text{ times}\right)}$$

## 6. Results from Simulation Studies

### 6.1. Comparative Performance for Point Estimation

Under a constant of relative risk ($RR$ = 1, 2 and 4), the performance in terms of bias, variance, and mean square error (MSE) of several summary relative risk estimators are compared. Results show that increasing $k$ can decrease the variance and the MSE of all estimators and the increase of both $n_{ij}$ can also decrease the variance of all estimators while fixing $k$. The unbalance cases of $n_{ij}$ ($i = 1, 2$ and $j = 1, \cdots, k$) have less affected on the order performance of MSE estimators. The summary adjusted relative risk estimator in meta-analysis study of size $k$ has shrinkage estimator to be a simple adjusted relative risk estimator in one single study case. The optimal point ($c = c_1 = c_2$) providing the bias, variance, and MSE of $\hat{\theta}_w$, adjusted by $c = -1/6$ is identical to $c = 1/6$. By these, the numerical evidence has confirmed the derivation process of finding the root $c$ and it is very useful in practice.

For a single center study ($k = 1$), regardless of a true value of RR, the proposed estimator adjusted by $c = 1/3$ performs the best with smallest MSE.

For a multi-center study of size $k$, when $RR = 1$, the WLS estimator is the best in sense of the smallest MSE ignoring the sample size $n_{ij}$. In case $RR = 2$, $k = 4$ the proposed estimator adjusted by $c$ performs the best with the smallest MSE. Another issue, when $RR = 2, k = 16$, the MH estimator achieves the smallest MSE when sample sizes are small. The MSE of the WLS estimator and the proposed estimator adjusted by $c$ are close together when the sample sizes are moderate and large ( $n_{ij} \geq 16$ ). For the case $RR = 4$ and $k = 4$ or 16, the MH estimator well performs generally with the smallest MSE. Some comparisons of the performances of all summary estimators when $n_{1j} = 4,8$ and $n_{2j} = 4,8,16,32$ are depicted in **Figures 1(a)-(f)** for bias; **Figures 2(a)-(f)** for variance; and **Figures 3(a)-(f)** for MSE.

In summary for the performance of estimators, regardless of the true values of *RR*, the MH estimator achieves the best performance with the smallest MSE when the study size is rather large ( $k \geq 16$ ) and the sample sizes within each study are small. The MSE of WLS estimator and the proposed-weight estimator adjusted by $c = 1/6$, $c = 1/3$, $c = 1/2$ are close together and they are the best when the sample sizes are moderate to large ( $n_{ij} \geq 16$ ) while the study size is rather small.

### 6.2. Studying the Type I Error

A type I error comparison between the tests is considered by comparing the actual (empirical) type I error ( $\hat{\alpha}$ ) with the nominal level of significance. In this study, the evaluation of the ability to control type I error probability for two-sided tests is based on Cochran limits as follows.

At $\alpha = 0.01$ significant level, the actual $\hat{\alpha}$ value is between [0.005, 0.015].

At $\alpha = 0.05$ significant level, the actual $\hat{\alpha}$ value is between [0.04, 0.06].

At $\alpha = 0.10$ significant level, the actual $\hat{\alpha}$ value is between [0.08, 0.12].

If the actual type I error or the empirical alpha lies within those Cochran limits, then the statistical test can control type I error rate.

For a single study ( $k = 1$ ), regardless of the true values of *RR*, it is unfortunate that almost all tests shown in **Table 2** cannot control type I error rate. There are some type I error rates lying in the Cochran limits.

For a meta-analysis study of size *k*, also displayed in **Table 2**, regardless of the true values of RR, the Mantel-Haenszel's Z-Test can control type I error rate when the sample size either in treatment or control group is moderate to large ( $n_{1j} \geq 16$ or $n_{2j} \geq 16$ ). In addition, the weighted least square Z-test and the proposed weight Z-test adjusted by $c = 1/6$, $c = 1/2$, $c = 1$ can handle type I error rate when both $n_{ij}$ are large. But the proposed Z-test adjusted by $c = 2$ cannot control type I error rate into Cochran's limit at almost all situations.

### 6.3. Studying the Power of the Test

Usually, the empirical power of the tests will be compared under the same type I error value. In summary, the Mantel-Haenszel Z-test performs best when $n_{ij}$ is moderate to large with satisfying the type I error value within Cochran's range limit, regardless of the study size *k*. The inverse variance weighted Z-test is good when $n_{ij}$ is large. For the proposed Z-test adjusted by $c = 1$, $c = 2$, the Z-tests perform well under the same type I error value when *k* is small ( $k \leq 4$ ) and $n_{ij}$ is large ( $n_{ij} \geq 32$ ). The results in **Table 3** illustrate the comparison of the power of tests under the same rate of the type I error.

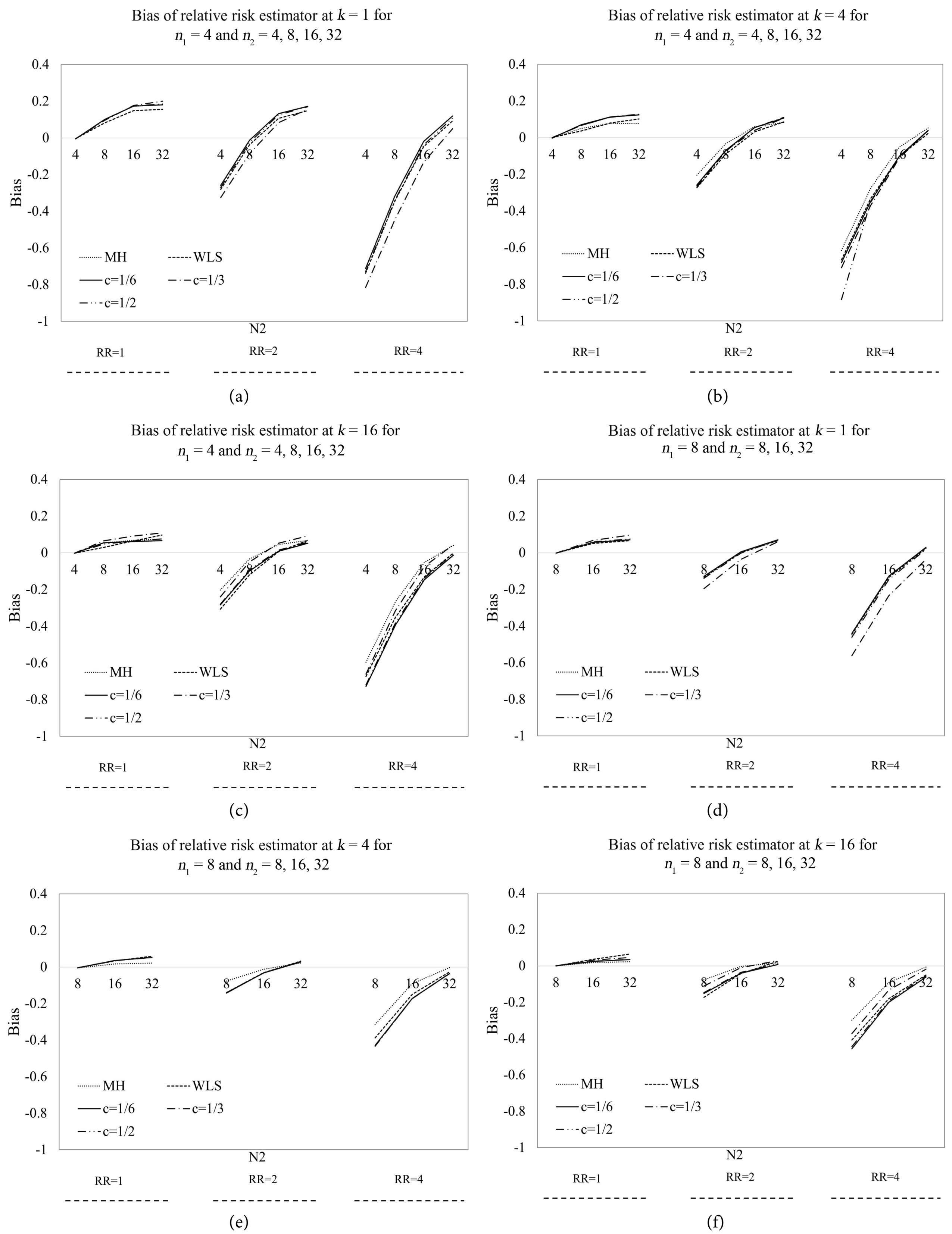


**Figure 1.** Bias comparison of relative risk between well-known estimators and adjusted relative risk estimators at $k = 1$, $k = 4$ and $k = 16$ for $n_1 = 4$ and $n_2 = 4, 8, 16$ and 32 ((a)-(c)) and at $k = 1$, $k = 4$ and $k = 16$ for $n_1 = 8$ and $n_2 = 8, 16$ and 32 ((d)-(f)).

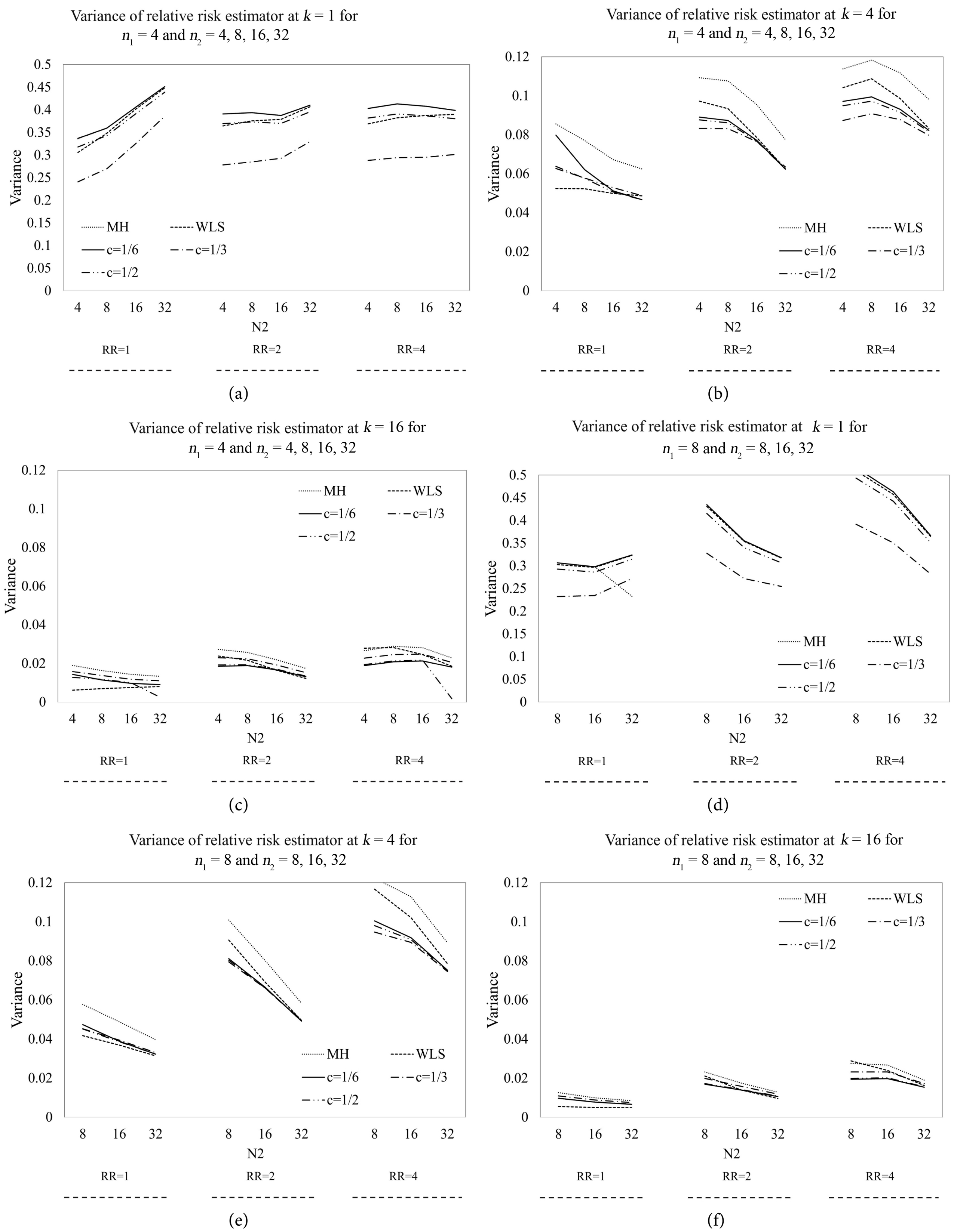


**Figure 2.** Variance comparison of relative risk between well-known estimators and adjusted relative risk estimators at $k = 1$, $k = 4$ and $k = 16$ for $n_1 = 4$ and $n_2 = 4$, 8, 16 and 32 ((a)-(c)) and at $k = 1$, $k = 4$ and $k = 16$ for $n_1 = 8$ and $n_2 = 8$, 16 and 32 ((d)-(f)).

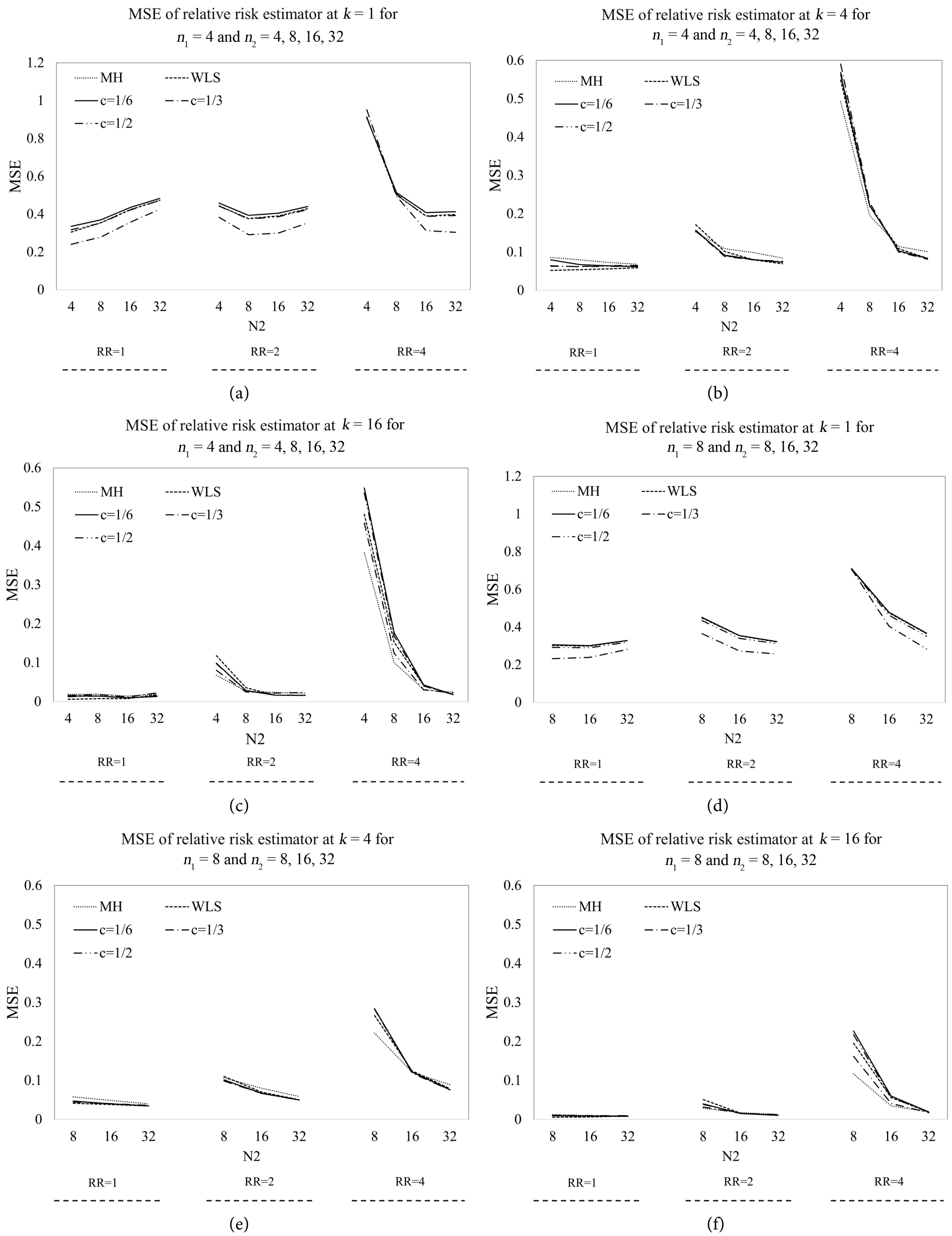


**Figure 3.** MSE comparison of relative risk between well-known estimators and adjusted relative risk estimators at $k = 1$, $k = 4$ and $k = 16$ for $n_1 = 4$ and $n_2 = 4$, 8, 16 and 32 ((a)-(c)) and at $k = 1$, $k = 4$ and $k = 16$ for $n_1 = 8$ and $n_2 = 8$, 16 and 32 ((d)-(f)).

**Table 2.** Comparison of the empirical type I error for testing $H_0: RR = RR_0$ in a center study ( $k = 1$ ) and multi-center study ( $k = 4$ , 16 and 32) at 5% significant level.

| $RR$ | $k$ | $n_{1j}$ | $n_{2j}$ | $Z_{MH}$ | $Z_{WLS}$ | $Z_{cw}$ $(c = 0.05)$ | $Z_{cw}$ $(c = 1)$ | $Z_{cw}$ $(c = 2)$ | $Z_{cw}$ $(c = 1/6)$ |
|---|---|---|---|---|---|---|---|---|---|
| | | 16 | 16 | 2.34 | 1.90 | 0.74 | 1.36 | 13.72 | 0.74 |
| | | 16 | 32 | 3.88 | 3.22 | 2.64 | 2.46 | 10.06 | 2.96 |
| 1 | 1 | 32 | 32 | 3.12 | 3.06 | 2.00 | 2.54 | **4.44** | 2.14 |
| | | 32 | 100 | **5.08** | **4.96** | 3.82 | **4.46** | 9.44 | 3.78 |
| | | 100 | 100 | **4.28** | **4.28** | 3.78 | 3.96 | **4.52** | 3.80 |
| | | 16 | 16 | 2.94 | 2.96 | 1.68 | 3.52 | 18.62 | 1.68 |
| | | 16 | 32 | 2.46 | 0.46 | 1.06 | 2.62 | 14.00 | 1.06 |
| 2 | 1 | 32 | 32 | 2.90 | 2.90 | 1.96 | 3.64 | 14.28 | 1.96 |
| | | 32 | 100 | 3.90 | 3.90 | 2.70 | **4.02** | 6.58 | 2.78 |
| | | 100 | 100 | **4.14** | **4.14** | 3.44 | **4.38** | 8.16 | 3.40 |
| | | 16 | 16 | 3.56 | 2.06 | 2.20 | **4.28** | 20.20 | 2.14 |
| | | 16 | 32 | **4.64** | **4.36** | **4.32** | **5.86** | 15.32 | **4.34** |
| 1 | 4 | 32 | 32 | **5.00** | **4.46** | **4.34** | **5.84** | 10.06 | **4.32** |
| | | 32 | 100 | **4.60** | 6.36 | **5.34** | 6.40 | 8.96 | **5.30** |
| | | 100 | 100 | **4.80** | **4.44** | **5.16** | **5.80** | 6.56 | **5.12** |
| | | 16 | 16 | 3.78 | **4.82** | 2.66 | **5.78** | 21.24 | 2.52 |
| | | 16 | 32 | **4.52** | 3.48 | 2.14 | **5.04** | 18.28 | 2.06 |
| 2 | 4 | 32 | 32 | **4.12** | **5.04** | 2.90 | **5.04** | 10.86 | 2.86 |
| | | 32 | 100 | **5.18** | **4.56** | 3.88 | **5.36** | 8.50 | 3.80 |
| | | 100 | 100 | **5.12** | **5.24** | **4.38** | **5.24** | 6.40 | **4.30** |
| | | 16 | 16 | **4.50** | 2.54 | 6.20 | 9.32 | 29.64 | **5.52** |
| | | 16 | 32 | **4.76** | 6.20 | 7.02 | 10.88 | 24.38 | 6.88 |
| 1 | 16 | 32 | 32 | **5.06** | 3.90 | 7.08 | 9.38 | 15.74 | 7.04 |
| | | 32 | 100 | **5.22** | 10.66 | 7.90 | 10.64 | 14.26 | 7.74 |
| | | 100 | 100 | **4.88** | **4.94** | 7.74 | 8.60 | 9.64 | 7.70 |
| | | 16 | 16 | **4.56** | 1.82 | 7.24 | 10.70 | 30.96 | 6.20 |
| | | 16 | 32 | **4.94** | 9.16 | 9.06 | 14.18 | 27.12 | 8.64 |
| 1 | 32 | 32 | 32 | **4.54** | 3.86 | 7.18 | 9.50 | 15.80 | 7.10 |
| | | 32 | 100 | **4.64** | 15.92 | 9.06 | 13.16 | 17.02 | 8.76 |
| | | 100 | 100 | **5.48** | 5.36 | 8.05 | 8.87 | 9.24 | 7.89 |

**Bold Values** denote that the statistical tests can control the type I error.

**Table 3.** Comparison of the empirical power of test (percent).

| *mm* | *k* | $n_{1j}$ | $n_{2j}$ | $Z_{MH}$ | $Z_{WLS}$ | $Z_{cw}$ $(c = 0.05)$ | $Z_{cw}$ $(c = 1)$ | $Z_{cw}$ $(c = 2)$ | $Z_{cw}$ $(c = 1/6)$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.2 | 1 | 32 | 32 | 2.08# | 2.08# | 0.44# | 2.42# | 9.36 | 0.40# |
| | | 32 | 100 | 4.80 | 4.80 | 3.24# | **6.38** | 24.02# | 3.24# |
| | 4 | 32 | 32 | 5.10 | 3.34 | 1.02 | **7.16** | 26.94# | 0.86 |
| | | 32 | 100 | **9.74** | 11.68# | 8.16 | 13.30# | 22.84# | 8.00 |
| | | 100 | 100 | 10.68 | 9.38 | 7.92 | **11.70** | 18.18# | 7.70 |
| | 16 | 32 | 32 | 10.08# | 6.28# | 2.60# | 18.12# | 53.30# | 2.22# |
| | | 32 | 100 | 17.80# | 30.98# | 19.70# | 17.28# | 21.16# | 19.72# |
| | | 100 | 100 | **23.44** | 21.46 | 18.70# | 27.02# | 38.56# | 18.24# |
| | 32 | 32 | 32 | **15.06** | 9.60 | 5.70# | 25.10# | 62.38# | 4.92# |
| | | 32 | 100 | 31.08# | 57.40# | 39.24# | 24.68# | 20.74# | 39.84# |
| | | 100 | 100 | 40.02# | 37.30# | 33.50# | 45.56# | 58.48# | 32.70# |
| | 100 | 100 | 100 | 83.40 | 81.52 | 76.68 | 86.24 | 92.26# | 78.28 |
| 0.4 | 1 | 32 | 32 | 2.86# | 2.86# | 0.80# | 3.34# | **10.80** | 0.72# |
| | | 32 | 100 | 7.32 | 7.32 | 5.50# | **8.90** | 26.70 | 5.50# |
| | 4 | 32 | 32 | 6.78 | 4.60 | 1.80 | **9.64** | 32.10 | 1.68 |
| | | 32 | 100 | **13.12** | 15.94# | 11.40 | 17.20# | 27.42 | 11.14 |
| | | 100 | 100 | 16.24 | 14..78 | 12.94 | **17.92** | 25.90 | 12.56 |
| | 16 | 32 | 32 | 13.54# | 9.74# | 4.92# | 21.88# | 57.70 | 4.34# |
| | | 32 | 100 | 24.56# | 42.44# | 27.56# | 23.20# | 26.10 | 27.66# |
| | | 100 | 100 | **32.74** | 31.76 | 27.98# | 36.58# | 47.58 | 27.72# |
| | 32 | 32 | 32 | **20.70** | 15.72 | 8.84# | 31.88# | 67.96 | 8.22# |
| | | 32 | 100 | 39.74# | 68.88# | 47.60# | 32.58# | 27.36 | 48.12# |
| | | 100 | 100 | 53.04# | 52.20# | 47.32# | 56.42# | 66.48 | 46.82# |
| | 100 | 100 | 100 | 89.74 | 90.14 | 87.24 | 91.42 | 94.80 | 87.04 |
| 0.6 | 1 | 32 | 32 | 4.74# | 4.74# | 1.70# | 5.22# | **12.94** | 1.64# |
| | | 32 | 100 | 10.48 | 10.48 | 7.88# | **11.98** | 30.46 | 7.88# |
| | 4 | 32 | 32 | 10.58 | 10.54 | 5.02 | **16.94** | 58.58 | 4.72 |
| | | 32 | 100 | **19.36** | 23.52# | 17.22 | 23.34# | 34.18 | 16.94 |
| | | 100 | 100 | 24.86 | 23.70 | 20.44 | **26.56** | 34.90 | 20.26 |
| | 16 | 32 | 32 | 20.86# | 16.38# | 9.94# | 30.10# | 64.52 | 9.00# |
| | | 32 | 100 | 33.58# | 55.00# | 36.32# | 32.20# | 34.86 | 36.50# |
| | | 100 | 100 | 45.72 | **45.96** | 40.72# | 48.60# | 57.60 | 40.22# |

**Continued**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 32 | 32 | **32.26** | 27.66 | 17.90# | 42.94# | 75.32 | 16.84# |
| 32 | 32 | 100 | 52.74# | 82.28# | 60.64# | 45.68# | 36.78 | 61.36# |
| | 100 | 100 | 64.74# | 66.66# | 60.44# | 67.14# | 74.80 | 60.08# |
| 100 | 100 | 100 | 95.78 | **96.82** | 94.56 | 96.22 | 97.66 | 94.36 |

#denotes that the statistical tests cannot control type I error rate under $H_0 : RR = 1$.

## 7. Discussion and Recommendation

The main question rises which continuity correction values are the best choice for the adjusted relative risk estimator in a center study and multi-center study with sparse data. Due to the conventional continuity correction, most of investigators such as Yate [1], Lane [3], Stijnen *et al.*, [4], Lui and Lin [6], Sankey *et al.*, [7], Gart and Zwefel [8], Walter [9] and Cox [10], suggest to use $c = 0.5$. However, in this study with the smallest average MSE of $\hat{\theta}_c = \log \widehat{RR}_c$, the optimal point $(c_1, c_2) = (1/6, 1/6)$ can perform the best for the point estimation. The minimum point $(c_1, c_2) = (1/6, 1/6)$ of $\hat{\theta}_c$ agrees with the suggestion of Turkey [12], which is very useful and the most appropriate in a practice way.

For estimation of fixing $\theta$ ($\theta = \log RR$) in a center ($k = 1$), regardless of a true value of *RR*, the proposed estimator adjusted by $c = c_1 = c_2 = 1/3$ performs the best with the smallest MSE. For a meta-analysis study of size *k*, in general the MH estimator achieves the smallest MSE when the sample size $n_{ij}$ is small while the study size is rather large ($k \geq 16$). The MSE of the WLS estimator and the proposed estimator adjusted by the various values of *c* are closed together and they are the best when the sample sizes are moderate to large ($n_{1j} \geq 16$ and $n_{2j} \geq 16$) while the study size is rather small. This finding is consonant with the work of Viwatwongkasem *et al.* [19]. Since the true value of *RR* is usually not available in practice as mentioned earlier, we suggest to choose the proposed relative risk estimator adjusted by $c = 1/6$ that can minimize the Bayes risk with respect to uniform prior (0, 1) and Euclidean loss function.

For the empirical power of the test under $H_1 : \theta_j = 0.1 + U_m(-mm, mm)$, regardless of the study size *k*, the MH Z-test performs the best with the highest power when both $n_{ij}$ are moderate to large. The inverse-variance weighted Z-test is good when $n_{ij}$ is large. In accordance with Soulakova and Bright [21], the empirical power of the test will increase when the sample sizes increase. For the power of the proposed Z-test adjusted by $c = 1$, $c = 2$, the Z-test has the higher performance under the same type I error when *k* is small ($k \leq 4$) and $n_{ij}$ is large ($n_{ij} \geq 32$).

If we don't know information about parameter *RR*, we recommend to use the adjusted estimator $\hat{\theta}$ by using continuity correction defined by $c_1 = c_2 = 1/6$, or 1/3, or 1/2 in a center study. For a multi-center study of size *k*, we recommend to use adjusted $\hat{\theta}_w$ defined by $c_1 = c_2 = 1/6$, or 1/3, or 1/2, including optimal weights $\hat{f}_j$.

Generally, the effect of exposures or the effect of treatments with binary outcomes covers the risk difference, the relative risk, and the odds ratio. Obviously, all three conventional effect estimators have the same problem of the zero values in sparse data. The conventional proportion estimate of $\hat{p} = X/n$ is in need of replacement by $\hat{p}_c = (X+c)/(n+2c)$ to solve this problem. Therefore, the recommendation for a further study is to use these ideas, such as the smallest MSE, the smallest Bayes risk to fine the appropriate point $(c_1, c_2)$ in estimating the odds ratio parameter.

## Acknowledgements

The study was partially supported for publication by the Faculty of Public Health, Mahidol University, Bangkok, Thailand.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1] Yates, F. (1934) Contingency Tables Involving Small Numbers and the Chi-Squared Test. *Supplement to the Journal of the Royal Statistical Society*, **1**, 217-235. https://doi.org/10.2307/2983604

[2] Lane, P.W. (2013) Meta-Analysis of Incidence of Rare Events. *Statistical Methods in Medical Research*, **22**, 117-132. https://doi.org/10.1177/0962280211432218

[3] Stijnen, T., Hamza, T.H. and Özdemir, P. (2010) Random Effects Meta-Analysis of Event Outcome in the Framework of the Generalized Linear Mixed Model with Applications in Sparse Data. *Statistics in Medicine*, **29**, 3046-3067. https://doi.org/10.1002/sim.4040

[4] White, I.R., Daniel, R. and Royston, P. (2010) Avoiding Bias Due to Perfect Prediction in Multiple Imputation of Incomplete Categorical Variables. *Computational Statistics and Data Analysis*, **54**, 2267-2275. https://doi.org/10.1016/j.csda.2010.04.005

[5] Lui, K.J. and Lin, C.D. (2003) A Revisit on Comparing the Asymptotic Interval Estimators of Odds Ratio in a Single 2 × 2 Table. *Biometrical Journal*, **45**, 226-237. https://doi.org/10.1002/bimj.200390008

[6] Sankey, S.S., Weissfeld, L.A., Fine, M.J. and Kapoor, W. (1996) An Assessment of the Use of the Continuity Correction for Sparse Data in Meta-Analysis. *Communications in Statistics—Simulation and Computation*, **25**, 1031-1056. https://doi.org/10.1080/03610919608813357

[7] Gart, J.J. and Zweifel, J.R. (1967) On the Bias of Various Estimators of the Logit and Its Variance with Application to Quantal Bioassay. *Biometrika*, **54**, 181-187. https://doi.org/10.1093/biomet/54.1-2.181

[8] Walter, S.D. (1975) The Distribution of Levin's Measure of Attributable Risk. *Biometrika*, **62**, 371-372. https://doi.org/10.1093/biomet/62.2.371

[9] Cox, D.R. (1970) The Continuity Correction. *Biometrika*, **57**, 217-219. https://doi.org/10.1093/biomet/57.1.217

[10] Li, L. and Wang, X. (2017) Meta-Analysis of Rare Binary Events in Treatment

Groups with Unequal Variability. *Statistical Methods in Medical Research*, **28**, 263-274. https://doi.org/10.1177/0962280217721246

[11] Tukey, J.W. (1977) Exploratory Data Analysis. Addison-Wesley Publishing Company, Boston.

[12] Sánchez-Meca, J. and Marín-Martínez, F. (2000) Testing the Significance of a Common Risk Difference in Meta-Analysis. *Computational Statistics and Data Analysis*, **33**, 299-313. https://doi.org/10.1016/S0167-9473(99)00055-9

[13] Böhning, D. and Viwatwongkasem, C. (2005) Revisiting Proportion Estimators. *Statistical Methods in Medical Research*, **14**, 147-169. https://doi.org/10.1191/0962280205sm393oa

[14] Agresti, A. and Caffo, B. (2000) Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures. *The American Statistician*, **54**, 280-288. https://doi.org/10.1080/00031305.2000.10474560

[15] McClave, J.T. and Sincich, T.T. (2012) Statistics. Pearson Education, London.

[16] Pettigrew, H.M., Gart, J.J. and Thomas, D.G. (1986) The Bias and Higher Cumulants of the Logarithm of a Binomial Variate. *Biometrika*, **73**, 425-435. https://doi.org/10.1093/biomet/73.2.425

[17] Lipsitz, S.R., Dear, K.B.G., Laird, N.M. and Molenberghs, G. (1998) Tests for Homogeneity of the Risk Difference When Data Are Sparse. *Biometrics*, **54**, 148-160.

[18] Cooper, M.R., Dear, K.B.G., McIntyre, O.R., Ozer, H., Ellerton, J.Cannellos, G., Duggan, B. and Schiffer, C. (1993) A Randomized Clinical Trial Comparing Melphalan/Prednisone with and without $\alpha$-2b Interferon in Newly-Diagnosed Patients with Multiple Myeloma: A Cancer and Leukemia Group B Study. *Journal of Clinical Oncology*, **11**, 155-160. https://doi.org/10.1200/JCO.1993.11.1.155

[19] Viwatwongkasem, C., Jitthavech, J., Böhning, D. and Lorchirachoonkul, V. (2012) Minimum MSE Weights of Adjusted Summary Estimator of Risk Difference in Multi-Center Studies. *Open Journal of Statistics*, **2**, 48-59. https://doi.org/10.4236/ojs.2012.21006

[20] Greenland, S. and Robin, J.M. (1985) Estimation of a Common Effect Parameter from Sparse Follow-up Data. *Biometrics*, **45**, 55-68.

[21] Soulakova, J.N. and Bright, B.C. (2013) Applications of Asymptotic Confidence Intervals with Continuity Corrections for Asymmetric Comparisons in Noninferiority Trials. *Pharmaceutical Statistics*, **12**, 147-155. https://doi.org/10.1002/pst.1566