Meta-analysis With Zero-event Studies: A Comparative Study With Application to COVID-19 Data

Jiajin Wei

Hong Kong Baptist University

Enxuan Lin

Shenzhen Research Institute of Big Data

Jiandong Shi

Hong Kong Baptist University

Ke Yang

Hong Kong Baptist University

Zongliang Hu

Shenzhen University

Xian-Tao Zeng

Center for Evidence-Based and Translational Medicine Zhongnan Hospital of Wuhan University

Tiejun Tong (✉ tongt@hkbu.edu.hk)

Department of Mathematics, Hong Kong Baptist University, Hong Kong, China <https://orcid.org/0000-0003-0947-3990>

Methodology

Keywords: Continuity correction, COVID-19 data, Meta-analysis, Relative risk, Zero-event studies

Posted Date: March 31st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-365606/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

RESEARCH

Meta-analysis with zero-event studies: a comparative study with application to COVID-19 data

Jiajin Wei¹, Enxuan Lin², Jiandong Shi¹, Ke Yang¹, Zongliang Hu³, Xian-Tao Zeng⁴ and Tiejun Tong^{1*}

*Correspondence:

tongt@hkbu.edu.hk

¹Department of Mathematics,

Hong Kong Baptist University,

Hong Kong, China

Full list of author information is available at the end of the article

Abstract

Background: Meta-analysis is a statistical method to synthesize evidence from a number of independent studies, including those from clinical studies with binary outcomes. In practice, when there are zero events in one or both groups, it may cause statistical problems in the subsequent analysis.

Methods: In this paper, by considering the relative risk as the effect size, we conduct a comparative study that consists of four continuity correction methods and another state-of-the-art method without the continuity correction, namely the generalized linear mixed models. To further advance the literature, we also introduce a new method of the continuity correction for estimating the relative risk.

Results: From the simulation studies, the new method performs well in terms of mean squared error when there are few studies. In contrast, the generalized linear mixed model performs the best when the number of studies is large. In addition, by reanalyzing a recent COVID-19 data, it is evident that the double-zero-event studies impact on the estimate of the mean effect size.

Conclusion: We recommend the new method to handle the zero-event studies when there are only few studies in the meta-analysis, or instead use the GLMM when the number of studies is large. The double-zero-event study may be informative, and so we suggest not excluding them.

Keywords: Continuity correction; COVID-19 data; Meta-analysis; Relative risk; Zero-event studies

Background

Meta-analysis is a statistical method to synthesize evidence from a number of independent studies that addressed the same scientific questions [1, 2]. In clinical studies, experimental data are commonly composed of binary outcomes, and consequently, meta-analyses of binary data have attracted increasing attention in evidence-based medicine [3, 4]. For each study, an effect size is reported to quantify the treatment effect by comparing the event probabilities between the treatment group and the control group, including the odds ratio (OR), the relative risk (RR), and the risk difference (RD). In meta-analysis, when the study-specific effect size is estimated based on a two-by-two contingency table, the zero-event problem in one or both groups frequently occurs, which may cause an unexpected calculation complication in the statistical inference of the effect size. If the study involves a zero event in one group, we refer to it as a *single-zero-event study*; and if the study involves zero

events in both groups, we refer to it as a *double-zero-event study* [5]. Vandermeer et al. [6] and Kuss [7] applied random sampling techniques and found that 30% of meta-analyses from the 500 sampled Cochrane reviews included one or more single-zero-event studies, while 34% of the reviews involved at least one meta-analysis with a double-zero-event study.

As a recent example, Chu et al. [8] conducted several meta-analyses to evaluate the effectiveness of physical distancing, face masks, and eye protection on the spread of three coronaviruses, which caused severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS) or coronavirus disease 2019, also known as COVID-19. Specifically, they considered RR as the effect size and applied the random-effects model to pool the observed effect sizes with an inverse-variance weight assigned to each study [9, 10]. As a result, for their meta-analysis on physical distancing, they concluded that the risk of infection will be significantly decreased with a further physical distance. We note, however, that there are 8 single-zero-event studies and 7 double-zero-event studies among a total of 32 studies. In particular for the 7 studies on COVID-19 data, 4 of them are single-zero-event studies and 2 of them are double-zero-event studies. To escape the zero-event problem, Chu et al. [8] excluded the double-zero-event studies from their meta-analyses, which, however, may introduce an estimation bias to the overall effect size [7]. More recently, Xu et al. [11] revisited 442 meta-analyses with or without the double-zero-event studies, and then by a comparative study, they concluded that the double-zero-event studies do contain valuable information and should not be excluded from the meta-analysis.

Inspired by the aforementioned examples, we provide a selective review on the existing methods for meta-analysis that can handle the zero-event studies. For ease of presentation, we will mainly focus on the random-effects model with RR as the effect size, whereas the same comparison also applies to OR and RD. For a given study, we let n_1 be the number of samples in the treatment group with X_1 being the number of events, and n_2 be the number of samples in the control group with X_2 being the number of events. Let also X_1 follow a binomial distribution with parameters n_1 and $p_1 > 0$, and X_2 follow a binomial distribution with parameters n_2 and $p_2 > 0$. We further assume that X_1 and X_2 are independent of each other. Then to estimate $RR = p_1/p_2$, the maximum likelihood estimator is known as

$$\widehat{RR} = \frac{X_1/n_1}{X_2/n_2} = \frac{X_1 n_2}{X_2 n_1}. \quad (1)$$

Note that $\widehat{RR}$ is often right-skewed. To derive the statistical inference on RR, researchers often apply the log scale so that the resulting estimator is more normally distributed. Specifically by Agresti [12], the approximate variance of $\ln(\widehat{RR})$ is

$$\text{var}[\ln(\widehat{RR})] \approx \frac{1}{X_1} - \frac{1}{n_1} + \frac{1}{X_2} - \frac{1}{n_2}. \quad (2)$$

By (1) and (2), when there are zero events in one or both groups, the classic method for estimating RR suffers from the zero-event problem and will no longer be applicable.

To have a valid estimate of RR, originated from Haldane [13], one often recommends to add 0.5 to the counts of events and non-events if some count is zero

[14, 15]. This method is referred to a correction method and has been extensively used in meta-analysis to deal with the zero-event studies. For further developments on the continuity correction, one may refer to Sweeting et al. [16], Carter et al. [17], and the references therein. On the other side, there are also statistical models without the continuity correction to handle meta-analysis with zero-event studies, such as the generalized linear mixed models [4, 18, 19].

The remainder of this paper is organized as follows. In Section 2, we first review the random-effects model and the existing methods with the continuity correction, and then propose a new method of the continuity correction for estimating RR. In Section 3, we review the generalized linear mixed models for meta-analysis. In Section 4, we conduct simulation studies to evaluate the performance of the reviewed methods and our new method. In Section 5, we apply all the well performed methods to a recent meta-analysis on COVID-19 data for further evaluation of their performance. We then conclude the paper in Section 6 with some interesting findings, and provide the supplementary materials in the Appendix.

Methods

Methods with the continuity correction

Suppose that there are k studies in the meta-analysis, and y_i for $i = 1, \dots, k$ are the observed effect sizes for each study. By DerSimonian and Laird [20], the random-effects model can be expressed as

$$y_i = \theta + \zeta_i + \epsilon_i, \quad (3)$$

where θ is the mean effect size, ζ_i are the deviations of each study from θ , and ϵ_i are the sampling errors. We further assume that ζ_i are independent and identically distributed random variables from $N(0, \tau^2)$, ϵ_i are independent random errors from $N(0, \sigma_i^2)$, and that they are independent of each other. In addition, τ^2 is referred to as the between-study variance, and σ_i^2 are referred to as the within-study variances.

For the random-effects model in (3), by the inverse-variance method the mean effect size θ can be estimated by

$$\hat{\theta} = \frac{\sum_i w_i^* y_i}{\sum_i w_i^*}, \quad (4)$$

where $w_i^* = 1/(\sigma_i^2 + \tau^2)$ are the weights assigned to each individual study [21]. In meta-analysis, the within-study variances σ_i^2 are routinely estimated by the variances of the observed effect sizes, denoted by $\text{var}(y_i)$. While for the between-study variance, DerSimonian and Laird [20] proposed the method of moments estimator as

$$T^2 = \frac{Q - k + 1}{C}, \quad (5)$$

where $Q = \sum_i w_i (y_i - \sum_i w_i y_i / \sum_i w_i)^2$ is known as the Q statistic, and $C = \sum_i w_i - \sum_i w_i^2 / \sum_i w_i$ with $w_i = 1/\sigma_i^2$ for $i = 1, \dots, k$.

We note, however, that the random-effects model may suffer from the zero-event problem. Taking RR as an example, if we apply the random-effects model for meta-analysis, then the effect sizes y_i will be the observed $\ln(\text{RR})$ values. Now for estimating $\log(\text{RR})$, if we plug in $\widehat{\text{RR}}$ from formula (1) directly, then $\ln(\widehat{\text{RR}})$ will not be well defined when the studies involve the zero events, and so is for the variance estimate of $\ln(\widehat{\text{RR}})$ in formula (2). Consequently, without a valid estimate of the effect size and of its within-study variance, the random-effects model cannot be applied to estimate the mean effect size by the inverse-variance method. This shows that a correction on $\widehat{\text{RR}}$ is often desired in meta-analysis with some studies involving zero events.

Existing methods with the continuity correction

Let $c_1 > 0$ and $c_2 > 0$ be two values for the continuity correction. To overcome the zero-event problem, one common approach is to estimate p_1 by $(X_1 + c_1)/(n_1 + 2c_1)$ and estimate p_2 by $(X_2 + c_2)/(n_2 + 2c_2)$. Plugging them into (1) and (2), we have

$$\widetilde{\text{RR}}(c_1, c_2) = \frac{X_1 + c_1}{n_1 + 2c_1} \cdot \frac{n_2 + 2c_2}{X_2 + c_2}. \quad (6)$$

Accordingly, the 95% confidence interval (CI) of RR is

$$\exp \left\{ \ln(\widetilde{\text{RR}}(c_1, c_2)) \pm 1.96 \sqrt{\frac{1}{X_1 + c_1} - \frac{1}{n_1 + 2c_1} + \frac{1}{X_2 + c_2} - \frac{1}{n_2 + 2c_2}} \right\}. \quad (7)$$

For the values of c_1 and c_2 in (6), there are mainly three suggestions in the literature that widely used for the random-effects meta-analysis.

- (i) When $c_1 = c_2 = 0.5$, it yields the Haldane estimator [13] as

$$\widetilde{\text{RR}}_{\text{Haldane}} = \begin{cases} \frac{X_1 + 0.5}{n_1 + 1} \cdot \frac{n_2 + 1}{X_2 + 0.5} & X_1 = 0 \text{ or } n_1, X_2 = 0 \text{ or } n_2, \\ \frac{X_1 n_2}{n_1 X_2} & \text{otherwise.} \end{cases} \quad (8)$$

- (ii) When $c_1 = n_1/(n_1 + n_2)$ and $c_2 = n_2/(n_1 + n_2)$, it yields the TACC estimator [16] as

$$\widetilde{\text{RR}}_{\text{TACC}} = \begin{cases} \frac{X_1 + c_1}{n_1 + 2c_1} \cdot \frac{n_2 + 2c_2}{X_2 + c_2} & X_1 = 0 \text{ or } n_1, X_2 = 0 \text{ or } n_2, \\ \frac{X_1 n_2}{X_2 n_1} & \text{otherwise.} \end{cases} \quad (9)$$

For the balanced case when $n_1 = n_2$, the TACC estimator is equivalent to the Haldane estimator. Also to implement this estimator, one may apply *metabin* in the R package “meta” with the setting `incr = “TACC”` [22].

- (iii) When $c_1 = c_2 = 1$, it yields the Carter estimator [17] as

$$\widetilde{\text{RR}}_{\text{Carter}} = \frac{X_1 + 1}{n_1 + 2} \cdot \frac{n_2 + 2}{X_2 + 1}. \quad (10)$$

Besides the continuity correction methods in family (6), another alternative is to estimate p_1 by $(X_1 + c_1)/(n_1 + c_1)$ and estimate p_2 by $(X_2 + c_2)/(n_2 + c_2)$. Then with $c_1 = c_2 = 0.5$, it yields the Pettigrew estimator [23] as

$$\widetilde{\text{RR}}(0.5, 0.5) = \frac{X_1 + 0.5}{n_1 + 0.5} \cdot \frac{n_2 + 0.5}{X_2 + 0.5}, \quad (11)$$

and the 95% CI of RR as

$$\exp \left\{ \ln(\widetilde{\text{RR}}(0.5, 0.5)) \pm 1.96 \sqrt{\frac{1}{X_1 + 0.5} - \frac{1}{n_1 + 0.5} + \frac{1}{X_2 + 0.5} - \frac{1}{n_2 + 0.5}} \right\}. \quad (12)$$

Moreover, to avoid a zero standard error, Hartung and Knapp [24] suggested not to correct X_1 and X_2 when $X_1 = n_1$ and $X_2 = n_2$.

A hybrid method with the continuity correction

Note that the existing methods are all constructed to first estimate p_1 and p_2 , and then take their ratio as an estimate of $\text{RR} = p_1/p_2$. Nevertheless, noting that p_2 is in the denominator rather than in the numerator, inverting an optimal estimate for p_2 may not necessarily yield an optimal estimate for $1/p_2$. In this section, we propose a hybrid method that is to estimate p_1 and $1/p_2$ directly, and then take their product to estimate RR.

For the estimation of p_1 , we show in Appendix 1 that the mean squared error (MSE) of $(X_1 + c_1)/(n_1 + 2c_1)$ is smaller than the MSE of $(X_1 + c_1)/(n_1 + c_1)$ in most settings. We thus consider to apply $(X_1 + c_1)/(n_1 + 2c_1)$ to estimate p_1 in RR. While to estimate the reciprocal of p_2 , one may consider $(n_2 + 2c_2)/(X_2 + c_2)$ as in (6). Or instead, another option can be to consider $(n_2 + c_2)/(X_2 + c_2)$ as originated in (11), see also [25] and [26] for more discussion. And if we take the latter one, then a hybrid estimator of RR can be constructed as

$$\widehat{\text{RR}}(c_1, c_2) = \frac{X_1 + c_1}{n_1 + 2c_1} \cdot \frac{n_2 + c_2}{X_2 + c_2}. \quad (13)$$

For the optimal values of c_1 and c_2 in (13), our simulation studies in Appendices 2 and 3 show that $c_1 = 0.5$ and $c_2 = 0.5$ are among the best options. In view of this, our new hybrid estimator is taken as follows:

$$\widehat{\text{RR}}(0.5, 0.5) = \frac{X_1 + 0.5}{n_1 + 1} \cdot \frac{n_2 + 0.5}{X_2 + 0.5}, \quad (14)$$

whereas the 95% CI of RR is given as

$$\exp \left\{ \ln(\widehat{\text{RR}}(0.5, 0.5)) \pm 1.96 \sqrt{\frac{1}{X_1 + 0.5} - \frac{1}{n_1 + 1} + \frac{1}{X_2 + 0.5} - \frac{1}{n_2 + 0.5}} \right\}. \quad (15)$$

Comparison of the continuity correction methods

In this section, we conduct a numerical study to compare the finite sample performance of the existing and new methods in Sections 2.1 and 2.2. For ease of presentation, we refer to the confidence intervals associated with (8), (9), (10), (11) and

(14) as the Haldane interval, the TACC interval, the Carter interval, the Pettigrew interval, and the hybrid interval, respectively.

To generate the data, we let $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $p_1 = p_2 \times \text{RR}$ with RR ranging from 0.2 to $\min\{5, 1/p_2\}$. We also consider different combinations of the sample sizes, whereas for the sake of brevity, only the results for balanced samples with $n_1 = n_2 = 10$ or 50 are presented, whereas the results for the unbalanced samples are postponed to Appendix 4. Recall that the Haldane and TACC intervals are the same when $n_1 = n_2$, and we thus present the results for the Haldane interval only. With $N = 100,000$ repetitions for each setting, we generate random numbers from the binomial distributions with parameters (p_1, n_1) and (p_2, n_2) to yield the estimates of RR and their CIs. We then compute the frequencies of the true RR falling in the CIs as the coverage probability estimates. Moreover, the expected lengths of the CIs on the log scale are computed by $N^{-1} \sum_{s=1}^N (\ln(\text{UL})_s - \ln(\text{LL})_s)$, where UL_s and LL_s are the upper and lower limits of the s th CI.

[width=15.5cm]Figure1.eps

Figure 1 Comparison of the four CIs of RR with $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $n_1 = n_2 = 10$. The dot-dashed lines represent the simulation results of the Haldane interval, the dashed lines represent the simulation results of the Carter interval, the dotted lines represent the simulation results of the Pettigrew interval, and the solid lines represent the simulation results of the hybrid interval.

[width=15.5cm]Figure2.eps

Figure 2 Comparison of the four CIs of RR with $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $n_1 = n_2 = 50$. The dot-dashed lines represent the simulation results of the Haldane interval, the dashed lines represent the simulation results of the Carter interval, the dotted lines represent the simulation results of the Pettigrew interval, and the solid lines represent the simulation results of the hybrid interval.

For $p_2 = 0.05$ or 0.15 , the top four panels of Figures 1 and 2 show that the Haldane interval is the most conservative interval in most settings, and it provides the longest expected lengths compared to the other three intervals. The Carter interval may have downward spikes in the left or right tail, although it leads to the shortest expected lengths. We also note that the simulation results of the Pettigrew interval and the hybrid interval are comparable. Their coverage probabilities and expected lengths are intermediate between those of the other two intervals in most settings.

From the bottom four panels of Figures 1 and 2 with $p_2 = 0.85$ or 0.95 , it is evident that the Haldane interval has a satisfactory performance in most settings with the coverage probabilities around the nominal level. In contrast, the Carter interval fails to provide enough large coverage probabilities in most settings, so does the Pettigrew interval when n_1 and n_2 are small. Note also that the coverage probabilities of the hybrid interval are comparable to the Haldane interval as long as p_2 is not extremely large. Moreover, the hybrid interval yields shorter expected lengths than the Haldane interval.

To sum up, when p_2 is small, the Pettigrew interval and the hybrid interval are less conservative than the Haldane interval in most settings. While for large p_2 , the

Haldane interval and the hybrid interval perform better than the Pettgrew interval in terms of the coverage probabilities. In addition, the expected lengths of the hybrid interval are always shorter than the Haldane interval. This shows that the hybrid interval can serve as a good alternative for the interval estimation of RR.

The generalized linear mixed models

The generalized linear mixed models (GLMMs) are extensions of the generalized linear model, which include both the fixed and random effects as linear predictors [12]. Different types of the GLMMs have been proposed in the literature including a few reviews and comparison studies [4, 27]. Among the existing models, the bivariate GLMM has been well recognized and being recommended for estimating RR in meta-analysis [18].

Let p_{i1} and p_{i2} be the event probabilities in the treatment and control groups of the i th study, respectively. The bivariate GLMM is represented as

$$\begin{aligned} g(p_{i1}) &= \Omega_1 + \zeta_{i1}, \\ g(p_{i2}) &= \Omega_2 + \zeta_{i2}, \end{aligned} \quad (16)$$

where $g(\cdot)$ is the link function, Ω_1 and Ω_2 are the fixed effects, and the random effects are given by

$$\begin{pmatrix} \zeta_{i1} \\ \zeta_{i2} \end{pmatrix} \stackrel{\text{ind}}{\sim} N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_1^2 & \rho\tau_1\tau_2 \\ \rho\tau_1\tau_2 & \tau_2^2 \end{pmatrix} \right).$$

By [28], the mean effect size based on model (16) was defined as

$$\text{RR}_{\text{GLMM}} = \frac{E(p_1)}{E(p_2)} = \frac{\int_{-\infty}^{\infty} g^{-1}(\Omega_1 + t) \tau_1^{-1} \phi(t/\tau_1) dt}{\int_{-\infty}^{\infty} g^{-1}(\Omega_2 + t) \tau_2^{-1} \phi(t/\tau_2) dt}, \quad (17)$$

where $E(p_1)$ and $E(p_2)$ are the mean event probabilities in the control and treatment groups, $g^{-1}(\cdot)$ is the inverse function of the link, and $\phi(\cdot)$ is the probability density function of the standard normal distribution. For the logit link, Zeger et al. [29] proposed an approximate formula $E(p_j) \approx \text{expit}(\Omega_j / \sqrt{1 + C^2 \tau_j^2})$ with $C = 16\sqrt{3}/(15\pi)$. For the probit link, $E(p_j) = \Phi(\Omega_j / \sqrt{1 + \tau_j^2})$, where $j = 1$ or 2 , and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. While for the other links, there does not exist a closed form of formula (17) and so a numerical approximation is often needed [30].

For the parameter estimation in model (16), Jackson et al. [4] provided a detailed introduction for the implementation based on the R package “lme4” in their model 6. Alternatively, one may also apply the function *meta.biv* in the R package “altmeta” maintained by Lin and Chu [31], in which the 95% CI of RR can be derived by the bootstrap resampling method.

Results

Simulation studies

In this section, we compare the performance of the reviewed methods on handling meta-analysis with zero-event studies, including the continuity correction methods

and the generalized linear mixed models. Among the existing continuity correction methods, we note that the Haldane and TACC estimators are comparable and among the best when estimating the mean effect size, in contrast to the other two methods including the Carter and Pettigrew estimators. Hence, for the sake of brevity, we only present the results of the Haldane and TACC estimators in the main text but provide the simulation results for all four methods in Appendix 5. Besides the Haldane estimator, we also consider the newly introduced hybrid estimator and the GLMM with the logit link for further comparison.

To conduct the meta-analysis, we consider $k = 3, 6$ and 12 as three different numbers of studies. Also by (3), we let $\theta = \ln(\text{RR})$ be the mean effect size that ranges from $\ln(0.2)$ to $\ln(5)$, and then generate the random effects ζ_i from $N(0, \tau^2)$ with $\tau^2 = 0.5$ or 1 . For each study, the numbers of samples in the two groups (n_{i1}, n_{i2}) are assumed to be equal and are randomly sampled from the set $\{5, 6, \dots, 50\}$. In addition, we generate the event probabilities of the control group p_{i2} from the uniform distribution with values from 0.01 to $\min(0.99, 1/\exp(\theta))$. Then accordingly, the event probabilities of the treatment group are given by $p_{i1} = \exp(\theta + \zeta_i)p_{i2}$, where $\exp(\theta + \zeta_i)p_{i2} \geq 1$ will be discarded. Finally, we generate X_{i1} and X_{i2} from the binomial distributions with parameters (n_{i1}, p_{i1}) and (n_{i2}, p_{i2}) , respectively. Note that the data will be re-generated if the number of events or non-events in one group are both zero. Finally, with $N = 5000$ repetitions for each setting, we compute the mean squared errors (MSEs) between the estimated RR and the true RR to evaluate the accuracy of the methods.

From the top two panels of Figure 3, it is evident that the three continuity correction methods perform much better than the GLMM in nearly all settings when k is small. Moreover, the hybrid estimator is consistently better than the Haldane and TACC estimators. The middle two panels show that, when k is moderate, the four methods are comparable in most settings. Finally, the bottom two panels indicate that the GLMM performs the best when k is large. To conclude, the accuracy of the different methods depends on the number of studies. In particular, for meta-analysis with few studies, we recommend the random-effects model with the hybrid estimator to handle the zero-event studies; and for meta-analysis with large studies, we recommend the GLMM to handle the random-effects meta-analysis.

[width=15.5cm]Figure3.eps

Figure 3 Comparison of the four methods with $k = 3, 6$ or 12 , $\tau^2 = 0.5$ or 1 . “1” represents the results of the random-effects model with the Haldane estimator, “2” represents the results of the random-effects model with the TACC estimator, “3” represents the results of the random-effects model with the hybrid estimator, and “4” represents the results of the GLMM.

Application to COVID-19 data

As mentioned earlier, Chu et al. [8] conducted a systematic review that revealed the connections of physical distancing, face masks, and eye protection with the transmission of SARS, MERS, and COVID-19. It is noteworthy that their analytical results have attracted more and more attention. As an evidence, their paper has received a total of 1236 citations in Google Scholar as of 16 March 2021. In this section, we propose to reanalyze COVID-19 data and compare the performance of

the different methods with or without the double-zero-event studies, including the Haldane estimator, the TACC estimator, the hybrid estimator, and the GLMMs.

Note that the treatment group represents a further physical distance and the control group represents a shorter physical distance. As shown in the top panel of Figure 4, [8] applied the random-effects model with the Haldane estimator and removed the double-zero-event studies from their meta-analysis. The overall effect size of 0.15 with the 95% CI being $[0.03, 0.73]$ indicates that the infection risk will be significantly reduced with a further physical distance. The middle panel of Figure 4 reports that the random-effects model with the TACC estimator yields the overall effect size of 0.12 with the 95% CI being $[0.03, 0.50]$. Moreover, the bottom panel of Figure 4 shows that the random-effects model with the hybrid estimator yields the overall effect size of 0.13 with the 95% CI being $[0.03, 0.72]$. Note also that the study-specific CIs here are always narrower than the CIs in the top panel, which coincides with the simulation results that the expected lengths of the CI associated with the hybrid estimator are shorter than the Haldane estimator. In addition, the GLMM in (16) does not provide the estimates of the study-specific effect sizes, so the results are listed as follows. By the bootstrap resampling with 1000 replicates, the GLMM with the logit link yields the overall effect size of 0.20 with the 95% bootstrap CI being $[0.05, 0.55]$. Also, the GLMM with the probit link yields the overall effect size of 0.18 with the 95% CI being $[0.04, 0.55]$.

[width=16cm]Figure4.pdf

Figure 4 Meta-analyses of COVID-19 data without the double-zero-event studies by applying the Haldane estimator (top), the TACC estimator (middle), and the hybrid estimator (bottom).

To reanalyze COVID-19 data, we now include the double-zero-event studies. The top panel of Figure 5 shows that the random-effects model with the Haldane estimator yields the overall effect size of 0.22 with 95% CI being $[0.06, 0.82]$. The middle panel of Figure 5 presents that the random-effects model with the TACC estimator provides the overall effect size of 0.18 with the 95% CI being $[0.06, 0.57]$. While for the hybrid estimator, it is shown by the bottom panel that the overall effect size is 0.21 with 95% CI being $[0.05, 0.81]$. At last, the GLMM with the logit link provides the overall effect size of 0.29 with the 95% CI being $[0.10, 0.64]$, and the GLMM with the probit link provides the overall effect size of 0.28 with the 95% CI being $[0.10, 0.56]$.

[width=16cm]Figure5.pdf

Figure 5 Meta-analyses of COVID-19 data with the double-zero-event studies by applying the Haldane estimator (top), the TACC estimator (middle), and the hybrid estimator (bottom).

Discussion

To handle the zero-event studies in meta-analysis of binary data, researchers often apply the random-effects model with the continuity correction, or instead, the GLMMs. From the simulation results, we note that the performance of the different

methods depends on the number of studies. When the number of studies is large, the GLMM is preferable to the random-effects model in terms of the MSEs. While for meta-analysis with few studies, we note that the random-effects model with the continuity correction is able to perform better than the GLMM. In addition to the methods we have compared, it is noteworthy that there are also other methods in the literature that can handle the zero-event studies in meta-analysis including, for example, the beta-binomial model [32, 33, 34]. Furthermore, it has been shown in the literature that the fixed-effects model can serve as another convincing method for meta-analysis with few studies [10, 35, 36, 37]. As an interesting future work, one may consider to apply the fixed-effects model to cope with meta-analysis of binary data with few studies.

For the double-zero-event studies in meta-analysis, we have shown by reanalyzing COVID-19 data that they do impact on the estimate of the mean effect size, and so they may not be uninformative. As noted by Friedrich et al. [38], including the double-zero-event studies moves the mean effect size estimate toward the direction of the null hypothesis. If one arbitrarily excluded the informative double-zero-event studies, there would be a risk of overstating the treatment effect such that the conclusion would be less reliable. As recommended by the literature [7, 11] and the references therein, we suggest including the double-zero-event studies in meta-analysis.

Conclusion

In this paper, we revisited the existing methods that are widely used to handle the zero-event problem in meta-analysis of binary data, in particular with the relative risk (RR) as the effect size which is also known as the risk ratio. For the methods with the continuity correction, we reviewed four existing estimators of RR and also introduced a new hybrid estimator with their applications to the random-effects model. Apart from those, the GLMM was also included which is another state-of-the-art method without the continuity correction. By a comparative study and also a real data analysis on COVID-19 data, we recommend the random-effects model with the hybrid estimator to handle the zero-event studies when there are only few studies in the meta-analysis, or instead use the GLMM when the number of studies is large. This paper also provides a useful addition to Chu et al [8], and meanwhile, it also calls for further observational studies in this field.

Appendix

Appendix 1

Note that X_1 is a random variable from the binomial distribution with parameters n_1 and p_1 , and so we have

$$\begin{aligned}
 \text{Bias} \left(\frac{X_1 + c_1}{n_1 + 2c_1} \right) &= E \left(\frac{X_1 + c_1}{n_1 + 2c_1} \right) - p_1 \\
 &= \frac{n_1 p_1 + c_1}{n_1 + 2c_1} - \frac{n_1 p_1 + 2c_1 p_1}{n_1 + 2c_1}, \\
 &= \frac{c_1 - 2c_1 p_1}{n_1 + 2c_1}; \\
 \text{Bias} \left(\frac{X_1 + c_1}{n_1 + c_1} \right) &= E \left(\frac{X_1 + c_1}{n_1 + c_1} \right) - p_1 \\
 &= \frac{n_1 p_1 + c_1}{n_1 + c_1} - \frac{n_1 p_1 + c_1 p_1}{n_1 + c_1}, \\
 &= \frac{c_1 - c_1 p_1}{n_1 + c_1}.
 \end{aligned}$$

Then we propose to solve

$$\left(\frac{c_1 - 2c_1 p_1}{n_1 + 2c_1} \right)^2 < \left(\frac{c_1 - c_1 p_1}{n_1 + c_1} \right)^2.$$

For $p_1 \leq 0.5$, it is evident that this inequality always holds. For $p_1 > 0.5$, equivalently we have

$$\begin{aligned}
 \frac{2c_1 p_1 - c_1}{n_1 + 2c_1} &< \frac{c_1 - c_1 p_1}{n_1 + c_1} \\
 2c_1 p_1 n_1 + 2c_1^2 p_1 - c_1 n_1 - c_1^2 &< c_1 n_1 - c_1 n_1 p_1 + 2c_1^2 - 2c_1^2 p_1 \\
 3c_1 p_1 n_1 + 4c_1^2 p_1 &< 2c_1 n_1 + 3c_1^2 \\
 p_1 &< \frac{2n_1 + 3c_1}{3n_1 + 4c_1}.
 \end{aligned}$$

Since $(2n_1 + 3c_1)/(3n_1 + 4c_1) > 0.5$, the squared bias of $(X_1 + c_1)/(n_1 + 2c_1)$ is lower than that of $(X_1 + c_1)/(n_1 + c_1)$ over a half of settings. We also note that the variance of $(X_1 + c_1)/(n_1 + 2c_1)$ is always smaller than the variance $(X_1 + c_1)/(n_1 + c_1)$. Hence, MSE of $(X_1 + c_1)/(n_1 + 2c_1)$ is smaller than the MSE of $(X_1 + c_1)/(n_1 + c_1)$ in most settings.

Appendix 2: Comparison of the p_1 estimates

In this simulation study, we propose to explore the effect of c_1 on the family of estimators $\tilde{p}_1(c_1) = (X_1 + c_1)/(n_1 + 2c_1)$ in terms of coverage probability and expected length. To generate the simulation data, we let p_1 range from 0.01 to 0.99 and $c_1 = 0.1, 0.3, 0.5, 0.7$ and 0.9 . We also consider $n_1 = 10$ and 50 as the different numbers of samples. With $N = 100,000$ repetitions for each setting, we generate random numbers from the binomial distribution with parameters p_1 and n_1 to yield the estimates of p_1 and the CIs. By the delta method, the variance of $\ln(\tilde{p}_1(c_1))$ can be approximated as $\text{var}[\ln(\tilde{p}_1(c_1))] \approx 1/(X_1 + c_1) - 1/(n_1 + 2c_1)$. Hence, the 95%

confidence interval of p_1 is given by

$$\exp \left\{ \ln \left(\frac{X_1 + c_1}{n_1 + 2c_1} \right) \pm 1.96 \sqrt{\frac{1}{X_1 + c_1} - \frac{1}{n_1 + 2c_1}} \right\}. \quad (18)$$

Then we compute the frequencies of the true RR falling in the CIs as the coverage probability estimates.

From Figure 6, it is evident that the CIs with $c_1 = 0.1$ and 0.3 yield low coverage probabilities when p_1 is close to 1. On the other side, the CI with $c_1 = 0.9$ always has a deep downward spike when p_1 is close to 0. In addition, the CIs with $c_1 = 0.5$ and 0.7 have comparable coverage probabilities, and so we choose $c_1 = 0.5$ as it is more representative.

[width=15.5cm]Figure6.eps

Figure 6 Comparison of the CIs of p_1 with $c_1 = 0.1, 0.3, 0.5, 0.7$ or 0.9 , and $n_1 = 10$ and 50 . The dotted lines represent the simulation results for $c_1 = 0.1$, the short dashed lines represent the simulation results for $c_1 = 0.3$, the solid lines represent the simulation results for $c_1 = 0.5$, the dot-dashed lines represent the simulation results for $c_1 = 0.7$, and the long dashed lines represent the simulation results for $c_1 = 0.9$.

Appendix 3: Simulation study for the estimators within family (13)

By the results from Appendix 2, we let $c_1 = 0.5$ and consider $c_2 = 0.5$ or 1 for estimator (6) and (13). To generate the simulation data, we let $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $p_1 = p_2 \times \text{RR}$ with RR ranging from 0.2 to $\min\{5, 1/p_2\}$. We also consider the numbers of samples as $n_1 = n_2 = 10$ and 160 . With $N = 100,000$ repetitions for each setting, we generate random numbers from the binomial distributions with parameters (p_1, n_1) and (p_2, n_2) . We then compute the frequencies of the true RR falling in the CIs as the coverage probability estimates.

From the top two panels of Figures 7 or 8, it is evident that the CIs associated with $\widetilde{\text{RR}}(0.5, 0.5)$ and $\widehat{\text{RR}}(0.5, 0.5)$ are comparable, and they perform better than the other CIs. On the other side, the bottom two panels of Figures 7 or 8 with $p_2 = 0.85$ and 0.95 show that the CI associated with $\widehat{\text{RR}}(0.5, 0.5)$ still provides the best performance among all the candidates in most settings. However, the CI associated with $\widetilde{\text{RR}}(0.5, 0.5)$ fails to provide the nominal coverage probability in most settings, in particular when p_1 is not large.

[width=15.5cm]Figure7.eps

Figure 7 Comparison of the four CIs of RR with $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $n_1 = n_2 = 10$. The dot-dashed lines represent the simulation results of the CI associated with $\widetilde{\text{RR}}(0.5, 0.5)$, the dashed lines represent the simulation results of the CI associated with $\widetilde{\text{RR}}(0.5, 1)$, the solid lines represent the simulation results of the CI associated with $\widehat{\text{RR}}(0.5, 0.5)$, and the dotted lines represent the simulation results of the CI associated with $\widehat{\text{RR}}(0.5, 1)$.

Appendix 4: Simulation study for unbalanced n_1 and n_2

In this simulation study, we compare the performance of the four existing intervals and the hybrid interval for unbalanced n_1 and n_2 . We let $n_1 = 20$ or 40 and

[width=15.5cm]Figure8.eps

Figure 8 Comparison of the four CIs of RR with $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $n_1 = n_2 = 50$. The dot-dashed lines represent the simulation results of the CI associated with $\widehat{RR}(0.5, 0.5)$, the dashed lines represent the simulation results of the CI associated with $\widehat{RR}(0.5, 1)$, the solid lines represent the simulation results of the CI associated with $\widehat{RR}(0.5, 0.5)$, and the dotted lines represent the simulation results of the CI associated with $\widehat{RR}(0.5, 1)$.

$n_2 = 0.5n_1$ or $2n_1$. The settings of the other parameters are the same as Section 2.4. The simulation results show that the coverage probabilities of the Haldane interval and the TACC interval are almost equal in most settings, but the TACC interval yields larger expected lengths when p_2 is small. The Carter interval does not provide a satisfactory performance due to its poor coverage probabilities in the left or right tails. From the bottom four panels of Figures 10 and 12, the coverage probabilities of the Pettigrew interval will be lower than 95%. Otherwise, the performance of the Pettigrew interval is comparable to the hybrid interval. In most settings, the hybrid interval alleviates the conservative performance of the Haldane interval, and meanwhile, it avoids the problem of low coverage probabilities that the Carter interval and the Pettigrew interval suffer. Hence, the hybrid interval still performs the best when n_1 and n_2 are unbalanced.

[width=15.5cm]Figure9.eps

Figure 9 Comparison of the five CIs of RR with $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $n_1 = 20, n_2 = 10$. The dot-dashed lines represent the simulation results of the Haldane interval, the long dashed lines represent the simulation results of the TACC interval, the short dashed lines represent the simulation results of the Carter interval, the dotted lines represent the simulation results of the Pettigrew interval, and the solid lines represent the simulation results of the hybrid interval.

[width=15.5cm]Figure10.eps

Figure 10 Comparison of the five CIs of RR with $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $n_1 = 20, n_2 = 40$. The dot-dashed lines represent the simulation results of the Haldane interval, the long dashed lines represent the simulation results of the TACC interval, the short dashed lines represent the simulation results of the Carter interval, the dotted lines represent the simulation results of the Pettigrew interval, and the solid lines represent the simulation results of the hybrid interval.

[width=15.5cm]Figure11.eps

Figure 11 Comparison of the five CIs of RR with $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $n_1 = 40, n_2 = 20$. The dot-dashed lines represent the simulation results of the Haldane interval, the long dashed lines represent the simulation results of the TACC interval, the short dashed lines represent the simulation results of the Carter interval, the dotted lines represent the simulation results of the Pettigrew interval, and the solid lines represent the simulation results of the hybrid interval.

Appendix 5: Comparison of the random-effects models with the existing correction methods

Appendix 6: R code for analyzing COVID-19 data

Without the double-zero-event studies

[width=15.5cm]Figure12.eps

Figure 12 Comparison of the five CIs of RR with $p_2 = 0.05, 0.15, 0.85$ or 0.95 , and $n_1 = 40$, $n_2 = 80$. The dot-dashed lines represent the simulation results of the Haldane interval, the long dashed lines represent the simulation results of the TACC interval, the short dashed lines represent the simulation results of the Carter interval, the dotted lines represent the simulation results of the Pettigrew interval, and the solid lines represent the simulation results of the hybrid interval.

[width=15.5cm]Figure13.eps

Figure 13 Comparison of the four methods with $k = 3, 6$ or 12 , $\tau^2 = 0.5$ or 1 . "1" represents the results of the random-effects model with the Haldane estimator, "2" represents the results of the random-effects model with the TACC estimator, "3" represents the results of the random-effects model with the Carter estimator, and "4" represents the results of the random-effects model with the Pettigrew estimator.

```
# Random-effects model with the Haldane estimator
library(meta)
Covid1 = read.table(header = TRUE, as.is = TRUE, text = "
      author      event.e  n.e  event.c  n.c
      Bai         0        76   12        42
      Burke       0        13    2         2
      Liu         0        17    2         3
      Cheng       5        47    7        36
      Heinzerling 0         4     3        33
    ")
meta1 = metabin(event.e, n.e, event.c, n.c, data = Covid1, sm = "RR",
  incr = 0.5, studlab = author, RR.Cochrane = 1, comb.fixed = F)
forest(meta1)

# Random-effects model with the TACC estimator
Covid2 = read.table(header=TRUE, as.is=TRUE, text="
      author      event.e  n.e  event.c  n.c
      Bai         0        76   12        42
      Burke       0        13    2         2
      Liu         0        17    2         3
      Cheng       5        47    7        36
      Heinzerling 0         4     3        33
    ")
meta2 = metabin(event.e, n.e, event.c, n.c, data=Covid2, sm = "RR",
  incr = "TACC", studlab = author,
  RR.Cochrane = 1, comb.fixed = F)
forest(meta2)

# Random-effects model with the hybrid estimator
Covid3 = read.table(header = TRUE, as.is = TRUE, text = "
      author      event.e  n.e  event.c  n.c
```

```

      Bai          0.5      77   12.5    42.5
      Burke        0.5      14    2.5     2.5
      Liu          0.5      18    2.5     3.5
      Cheng        5.5      48    7.5    36.5
      Heinzerling  0.5       5    3.5    33.5
    ")
meta3 = metabin(event.e, n.e, event.c, n.c, data = Covid3, sm = "RR",
               studlab = author, incr = 0, comb.fixed = F)
forest(meta3)

# The GLMM
library(altmeta)
sid = c(1,1,2,2,3,3,4,4,5,5)      # study id
tid = c(0,1,0,1,0,1,0,1,0,1)     # 0/1: control/experiment group
e = c(12,0,2,0,2,0,7,5,3,0)      # the number of events
n = c(42,76,2,13,3,17,36,47,33,4) # the number of samples
data_glmm = data.frame(sid, tid, e, n)
glmm_logit = meta.biv(sid, tid, e, n, data_glmm, link = "logit",
                     alpha = 0.05, b.iter = 1000)
glmm_probit = meta.biv(sid, tid, e, n, data_glmm, link = "probit",
                     alpha = 0.05, b.iter = 1000)

# With the double-zero-event studies

# Random-effects model with the Haldane estimator
library(meta)
Covid1 = read.table(header = TRUE, as.is = TRUE, text = "
      author      event.e  n.e  event.c  n.c
      Bai         0        76   12        42
      Burke       0        13    2         2
      Liu         0        17    2         3
      Cheng       5        47    7        36
      Heinzerling 0         4     3        33
      Burke       0        50    0        76
      Burke       0        41    0        37
    ")
meta1 = metabin(event.e, n.e, event.c, n.c, data = Covid1,
               sm="RR",incr = 0.5, allstudies = T, studlab = author,
               RR.Cochrane = 1, comb.fixed = F)
forest(meta1)

# Random-effects model with the hybrid estimator
Covid2 = read.table(header=TRUE, as.is=TRUE, text="
      author      event.e  n.e  event.c  n.c
      Bai         0        76   12        42
      Burke       0        13    2         2

```



```

      Liu          0          17   2          3
      Cheng        5          47   7         36
      Heinzerling  0          4    3         33
      Burke        0          50   0         76
      Burke        0          41   0         37
    ")

meta2 = metabin(event.e, n.e, event.c, n.c, data = Covid2,
               sm="RR",incr = "TACC", allstudies = T, studlab = author,
               RR.Cochrane = 1, comb.fixed = F)
forest(meta2)

# Random-effects model with the hybrid estimator
Covid3 = read.table(header = TRUE, as.is = TRUE, text = "
      author      event.e  n.e  event.c  n.c
      Bai         0.5      77   12.5     42.5
      Burke       0.5      14    2.5      2.5
      Liu         0.5      18    2.5      3.5
      Cheng       5.5      48    7.5     36.5
      Heinzerling 0.5       5    3.5     33.5
      Burke       0.5      51    0.5     76.5
      Burke       0.5      42    0.5     37.5
    ")

meta3 = metabin(event.e ,n.e, event.c, n.c, data = Covid3,
               sm = "RR", studlab = author, incr = 0, comb.fixed = F)
forest(meta3)

# The GLMM
# study id
sid = c(1,1,2,2,3,3,4,4,5,5,6,6,7,7)
# 0/1: control/experiment group
tid = c(0,1,0,1,0,1,0,1,0,1,0,1,0,1)
# the number of events
e = c(12,0,2,0,2,0,7,5,3,0,0,0,0,0)
# the number of samples
n = c(42,76,2,13,3,17,36,47,33,4,76,50,37,41)
data_glmm = data.frame(sid, tid, e, n)
glmm_logit = meta.biv(sid, tid, e, n, data_glmm, link = "logit",
                     alpha = 0.05, b.iter = 1000)
glmm_probit = meta.biv(sid, tid, e, n, data_glmm, link = "probit",
                      alpha = 0.05, b.iter = 1000)

```

Acknowledgements

Text for this section...

Funding

Text for this section...

Abbreviations

OR: odds ratio, RR: relative risk, RD: risk difference, SARS: severe acute respiratory syndrome, MERS: Middle East respiratory syndrome, COVID-19: coronavirus disease 2019, CI: confidence interval, MSE: mean squared error, GLMMs: generalized linear mixed models.

Availability of data and materials

Not applicable

Ethics approval and consent to participate

Not applicable

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable

Authors' contributions

TT, JW, EL, and XZ reviewed the literature and designed the methods. TT, JW, and JS conducted the simulation studies. TT, JW, KY, and ZH conducted the experiments and analyzed the read data. All authors contributed to the manuscript preparation.

Authors' information

Text for this section. . .

Author details

¹Department of Mathematics, Hong Kong Baptist University, Hong Kong, China. ²Shenzhen Research Institute of Big Data, Shenzhen, China. ³College of Mathematics and Statistics Shenzhen University, Shenzhen, China.

⁴Center for Evidence-Based and Translational Medicine Zhongnan Hospital of Wuhan University, Wuhan, China.

References

- Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to Meta-Analysis. Hoboken: John Wiley and Son; 2011.
- Ma LL, Wang YY, Yang ZH, Huang D, Weng H, Zeng XT. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? *Military Medical Research*. 2020;7:1–11.
- Davey J, Turner RM, Clarke MJ, Higgins JP. Characteristics of meta-analyses and their component studies in the Cochrane Database of Systematic Reviews: a cross-sectional, descriptive analysis. *BMC Medical Research Methodology*. 2011;11:1–11.
- Jackson D, Law M, Stijnen T, Viechtbauer W, White IR. A comparison of seven random-effects models for meta-analyses that estimate the summary odds ratio. *Statistics in Medicine*. 2018;37:1059–1085.
- Ren Y, Lin L, Lian Q, Zou H, Chu H. Real-world performance of meta-analysis methods for double-zero-event studies with dichotomous outcomes using the Cochrane database of systematic reviews. *Journal of General Internal Medicine*. 2019;34:960–968.
- Vandermeer B, Bialy L, Hooton N, Hartling L, Klassen TP, Johnston BC, et al. Meta-analyses of safety data: a comparison of exact versus asymptotic methods. *Statistical Methods in Medical Research*. 2009;18:421–432.
- Kuss O. Statistical methods for meta-analyses including information from studies without any events—add nothing to nothing and succeed nevertheless. *Statistics in Medicine*. 2015;34:1097–1116.
- Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet*. 2020;395:1973–1987.
- Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*. 2010;1:97–111.
- Lin E, Tong T, Chen Y, Wang Y. Fixed-effects model: the most convincing model for meta-analysis with few studies. Preprint at <https://arxiv.org/abs/2002.04211>; 2020.
- Xu C, Li L, Lin L, Chu H, Thabane L, Zou K, et al. Exclusion of studies with no events in both arms in meta-analysis impacted the conclusions. *Journal of Clinical Epidemiology*. 2020;123:91–99.
- Agresti A. *Categorical Data Analysis*, 2nd Edition. Hoboken: John Wiley and Son; 2003.
- Haldane JBS. The estimation and significance of the logarithm of a ratio of frequencies. *Annals of Human Genetics*. 1956;20:309–311.
- Schwarzer G. meta: An R package for meta-analysis. *R News*. 2007;7:40–45.
- Weber F, Knapp G, Ickstadt K, Kundt G, Glass d. Zero-cell corrections in random-effects meta-analyses. *Research Synthesis Methods*. 2020;11:913–919.
- Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in Medicine*. 2004;23:1351–1375.
- Carter RE, Lin Y, Lipsitz SR, Newcombe RG, Hermayer KL. Relative risk estimated from the ratio of two median unbiased estimates. *Journal of the Royal Statistical Society: Series C*. 2010;59:657–671.
- Chu H, Nie L, Chen Y, Huang Y, Sun W. Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: methods for the absolute risk difference and relative risk. *Statistical Methods in Medical Research*. 2012;21:621–633.
- Chen Y, Hong C, Ning Y, Su X. Meta-analysis of studies with bivariate binary outcomes: a marginal beta-binomial model approach. *Statistics in Medicine*. 2016;35:21–40.
- DerSimonian R, Laird NM. Meta-analysis in clinical trials. *Controlled Clinical Trials*. 1986;7:177–188.

21. Laird NM, Mosteller F. Some statistical methods for combining experimental results. *International Journal of Technology Assessment in Health Care*. 1990;6:5–30.
22. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evidence-based Mental Health*. 2019;22:153–160.
23. Pettigrew HM, Gart JJ, Thomas DG. The bias and higher cumulants of the logarithm of a binomial variate. *Biometrika*. 1986;73:425–435.
24. Hartung J, Knapp G. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Statistics in Medicine*. 2001;20:3875–3889.
25. Fattorini L. Applying the Horvitz-Thompson criterion in complex designs: a computer-intensive perspective for estimating inclusion probabilities. *Biometrika*. 2006;93:269–278.
26. Seber GA. *Statistical Models for Proportions and Probabilities*. Heidelberg: Springer; 2013.
27. Bakbergenuly I, Kulinskaya E. Meta-analysis of binary outcomes via generalized linear mixed models: a simulation study. *BMC Medical Research Methodology*. 2018;18:1–18.
28. McCullagh P. Sampling bias and logistic models. *Journal of the Royal Statistical Society: Series B*. 2008;70:643–677.
29. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*. 1988;44:1049–1060.
30. Lin L, Chu H. Meta-analysis of proportions using generalized linear mixed models. *Epidemiology*. 2020;31:713–717.
31. Lin L, Chu H. *altmeta: Alternative Meta-Analysis Methods*. <https://CRAN.R-project.org/package=altmeta>; 2020.
32. Sarmanov OVE. Generalized normal correlation and two-dimensional Fréchet classes. *Soviet Mathematics Doklady*. 1966;7:596–599.
33. Chen Y, Luo S, Chu H, Su X, Nie L. An empirical bayes method for multivariate meta-analysis with an application in clinical trials. *Communications in Statistics—Theory and Methods*. 2014;43:3536–3551.
34. Luo S, Chen Y, Su X, Chu H. *mmeta: an R package for multivariate meta-analysis*. *Journal of Statistical Software*. 2014;56:1–26.
35. Bender R, Friede T, Koch A, Kuss O, Schlattmann P, Schwarzer G, et al. Methods for evidence synthesis in the case of very few studies. *Research Synthesis Methods*. 2018;9:382–392.
36. Rice K, Higgins JP, Lumley T. A re-evaluation of fixed effect(s) meta-analysis. *Journal of the Royal Statistical Society: Series A*. 2018;181:205–227.
37. Yang K, Kwan HY, Yu Z, Tong T. Model selection between the fixed-effects model and the random-effects model in meta-analysis. *Statistics and Its Interface*. 2020;13:501–510.
38. Friedrich JO, Adhikari NK, Beyene J. Inclusion of zero total event trials in meta-analyses maintains analytic consistency and incorporates all available data. *BMC Medical Research Methodology*. 2007;7:1–6.

Figures

Figure 1

please see the manuscript file for the full caption

Figure 2

please see the manuscript file for the full caption

Figure 3

please see the manuscript file for the full caption

Figure 4

Meta-analyses of COVID-19 data without the double-zero-event studies by applying the Haldane estimator (top), the TACC estimator (middle), and the hybrid estimator (bottom).

Figure 5

Meta-analyses of COVID-19 data with the double-zero-event studies by applying the Haldane estimator (top), the TACC estimator (middle), and the hybrid estimator (bottom).

Figure 6

please see the manuscript file for the full caption

Figure 7

please see the manuscript file for the full caption

Figure 8

please see the manuscript file for the full caption

Figure 9

please see the manuscript file for the full caption

Figure 10

please see the manuscript file for the full caption

Figure 11

please see the manuscript file for the full caption

Figure 12

please see the manuscript file for the full caption

Figure 13

please see the manuscript file for the full caption