Contemporary Clinical Trials 122 (2022) 106962

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

Should studies with no events in both arms be excluded in evidence synthesis?

Chang Xu [a], Luis Furuya-Kanamori [b], Nazmul Islam [c], Suhail A. Doi [d,*]

[a] Ministry of Education Key Laboratory for Population Health Across-life Cycle & School of Public Health, Anhui Medical University, Anhui, China
[b] UQ Centre for Clinical Research, Faculty of Medicine, The University of Queensland, Herston, Australia
[c] Department of Public Health, College of Health Sciences, QU Health, Qatar University, Doha, Qatar
[d] Department of Population Medicine, College of Medicine, QU Health, Qatar University, Doha, Qatar

ARTICLE INFO

Keywords:
Meta-analysis
Double-zero studies
Continuity correction
IVhet

ABSTRACT

Objectives: In safety assessment, studies with no events are a frequent occurrence when conducting meta-analyses. The current approach in meta-analysis is to exclude double-zero studies from the synthesis. In this study, we compared the performance of excluding and including double-zero studies.

Methods: A simulation with 5000 iterations was conducted based on the real-world dataset from Cochrane reviews. The true distribution of the rare events rather than normal distribution for the effects were used in the data generating mechanism to simulate aggregate meta-analysis data. We used Doi's inverse variance heterogeneity (IVhet) model for the meta-analyses with continuity correction (of 0.5) to include double-zero studies and used the odds ratio effect size. The performance of including versus excluding double-zero studies were then compared.

Results: Generally, there was much larger mean squared error when double zero studies were excluded than when double-zero studies were included. The coverage when studies were excluded rapidly deteriorates as heterogeneity increased, while remained at or above the nominal level when double-zero studies were included. When there were very few double-zero studies, the performances was almost the same when including or excluding these studies. Subgroup analysis showed that, even for meta-analyses with unbalanced sample size across the two arms, including double-zero studies improved performance compared to when they were excluded.

Conclusions: Including double-zero studies in meta-analysis improved performance substantively when compared to excluding them, especially when the proportion of double-zero studies was large. Continuity correction with use of the IVhet model is therefore a good solution to deal with double-zero studies and should be considered in future meta-analyses.

1. Introduction

Studies with no events are a frequent occurrence when conducting meta-analyses of binary outcomes. Typically, a study with no event in one of the arms is referred to as a single-zero study, while with no events in both arms is referred as double-zero study. Suppose we consider a two-arm trial with interventions A and B and say the the outcome of interest is fracture; if investigators did not observe any fracture in arm A, but two fractures in arm B, the study is then considered a single-zero study; if no fracture events occurred in both arms A and B, it is then a double-zero study. As estimated by a previous survey, about 34% of the Cochrane Reviews on binary outcomes contain studies with no events; this means, on average, for every three meta-analyses of binary outcomes there would be one containing studies with no events in one or both arms [1].

Methods based on a two-stage framework are the most commonly utilized methods in dealing with studies with no events. These include the continuity correction, the Peto's odds ratio (OR), and the Mantel-Haenszel (MH) method [2,3]. The problem with the Peto's and MH method is that when using relative measures of effect (e.g. the OR), the point and variance estimators cannot be defined for double-zero studies [4,5]. Therefore, in most of the existing meta-analytical software (e.g. RevMan) and packages (e.g. *metan* for Stata) studies with no events in both arms are by default excluded from the meta-analysis, even though a

* Corresponding author at: Department of Population Medicine, College of Medicine, Qatar University, Al Jamiaa Street, P. O. Box 2713, Doha, Qatar.
E-mail address: sardoi@gmx.net (S.A. Doi).

https://doi.org/10.1016/j.cct.2022.106962
Received 20 March 2022; Received in revised form 5 October 2022; Accepted 7 October 2022
Available online 13 October 2022
1551-7144/

continuity correction can be applied [6,7]. The main reason for exclusion of these studies from meta-analysis is that these studies are considered to contain no information in terms of the conditional likelihood theory [8,9], but this ignores potential evidence that suggests that excluding studies with no events in both arms could lead to some statistical problems and also selection bias (e.g., publication bias) [7,10]. For example, in trials with a randomization ratio of 1:1, when there are no events in both arms, these studies provide evidence that there is no difference in the risk of the outcome (e.g. an adverse event) between interventions [5]. Thus, ignoring the evidence from such studies may adversely impact evidence informed decision-making.

In a previous study, Xu et al. used a one-stage method to demonstrate that double-zero studies may not necessarily be non-informative and excluding such studies may alter conclusions [10]. However, two-stage methods are more commonly used in meta-analyses [7], and the continuity correction for studies with no events is one of the most common methods used to estimate the OR and its variance [11]. This could be even more attractive as a solution since the continuity correction may be the second order approximation of the true OR [12], and the OR has been strongly recommended as the relative measure of choice for meta-analysis [13,14]. The only other study that investigated this was the study by Cheng et al [15]. who used a continuity correction with the MH and Peto's OR methods for studies with no events, and concluded that it is recommended to include studies with no events when there is no treatment effect (OR = 1), but exclude studies with no events when there is a treatment effect (OR≠1). However, their simulation was not convincing because it assumed a normal distribution for the expected study effects, which for meta-analyses of rare events, is not a reasonable assumption, especially when zero-events are involved and is unlikely to be valid in practice [16,17].

Therefore, we repeated the simulation study using a different data-generation mechanism that makes no such assumptions about the distribution of the expected study effects (instead, we used their true distribution). Since the distributional assumption is made to try to mitigate overdispersion when studies are heterogeneous, we used a meta-analytic model to synthesize data that avoided the need for such an assumption when heterogeneity was present. We now present the comparison of performance between excluding and including studies with no events, the latter after applying a continuity correction and compare the performance of the meta-analytic estimator.

2. The continuity correction

The continuity correction is one of the most widely used methods to deal with studies with no events in published meta-analyses [7,18,19]. Suppose we denote a, b, c, d in a 2 by 2 table, where $a + b = n_1$, $c + d = n_0$, the OR is then estimated as ad/bc, with the variance of *LnOR* as $1/a + 1/b + 1/c + 1/d$. A correction element, Δ, is added to each of the cells, so that we then have an approximate estimate of the *LnOR* as follows [12]:

$$LnOR_{\Delta} = Ln\left[\left(\frac{a+\Delta}{b+\Delta}\right)\Big/\left(\frac{c+\Delta}{d+\Delta}\right)\right] = Ln\left(\frac{a+\Delta}{b+\Delta}\right) - Ln\left(\frac{c+\Delta}{d+\Delta}\right)$$

Haldane suggested that use of 0.5 as the correction (Δ) applied to each cell can eliminate the first order bias term and thus when Δ takes the value of 0.5, it could be the second order approximation of *LnOR* [18]. Under the continuity correction, the approximate estimates are:

$$Ln(OR_{cc}) = Ln\frac{(a+0.5)(d+0.5)}{(b+0.5)(c+0.5)}$$

$$Var(LnOR_{cc}) = 1/(a+0.5) + 1/(b+0.5) + 1/(c+0.5) + 1/(d+0.5)$$

This method works when $a = 0$ or $c = 0$ (i.e., single-zero-events), and even when $a = c = 0$ (i.e., double-zero-events). The Peto's OR, does not require a continuity correction for single-zero studies because of the nature of the Peto's estimator of effect and variance [2]. These procedure work well when single-zero studies are involved, while for double-zero studies, in principle, it is not applicable as the OR and variance cannot be defined with the MH or Peto methods [20,21]. Thus, in most of meta-analytical software and packages, double-zero studies are automatically dropped from the synthesis. We therefore do not consider the Peto's and MH methods in this paper.

3. Simulation

3.1. Data generating mechanism

In Cheng's study [15], for the data generation they assumed a normal distribution of the true effect; in brief, they first set the event risk (p_0) in the control arm, and assumed the treatment effect $\theta \sim (\mu, \tau^2)$, where both μ (the expected effect) and τ^2 (the between-study variance) were fixed and the variance (within-study error) of μ was determined by the sample size they assumed for each study. Based on p_0 and the sample size, they obtained the events in the control arm (r_0); and based on θ, they obtained the event risk in the treatment arm (p_1) and further obtained the events in the treatment arm (r_1).

As we stated before, the normal distribution assumption of varying but similar expected effects is an assumption made to mitigate overdispersion in the variance estimation, but which is not appropriate [16]. Therefore, in the current study we used a different simulation strategy where we first computed a fixed value for p_0 and the OR and then p_1 was derived from them. For the "true" study event risk of the control arm p_0, we set three scenarios (0.01, 0.05, and 0.1), and the study event risk of treatment arm p_1 was calculated from the fixed simulation true effect (log OR) and p_0.

The "study" sample size was generated from a uniform distribution based on the real-world data from the Cochrane Database of Systematic Reviews. The 25th and the 75th percentile of the real-world study arm sizes were used and the sample size of the treatment and control arms were obtained from a uniform distribution as follows $n_1 \sim uniform(15, 58)$; $n_0 \sim uniform(15, 50)$ [10,20]. Next we computed the cell counts for a 2 × 2 table for the study based on the generated p_1, p_0, n_1, and n_0. Next, we introduce systematic and random error simultaneously by subjecting the rescaled cell counts to a beta distribution as follows: $(p_{1_e} \sim beta(a/f, b/f))$ & $(p_{0_e} \sim beta(c/f, d/f))$ to generate new values for study event risks (p_{1_e}, p_{0_e}) and where f is a scaling factor that adds systematic error when used to generate event risks via the beta distribution. The new event risks are then used to generate the final cell counts for the study from the sample size already defined. The scaling factor used above was derived from a hypothetical study quality count (Qi) (see details in the supplementary material) [22].

Using these parameters, we simulated the 2 by 2 table data for each study included in a meta-analysis. The following simulation settings were applied:

i. number of studies included in the meta-analysis, k, set at 5 or 10;
ii. event rate for control arm, p_0, set at 0.01 or 0.05 or 0.1;
iii. and two simulations were run with either true OR set at 1 or 2;

Therefore, there were 2*3*2 = 12 scenarios and for each scenario, we had 5000 iterations (simulated 5000 meta-analyses) in each of 10 runs, the latter with different amount of heterogeneity as determined by the scaling factor f. Each run thus included different amounts of between-study variance. We conducted the simulation study and compared the performance of either: 1) excluding studies with no events; 2) including studies with no events after applying a continuity correction.

In addition, we calculated the sample size ratio of each study and defined a study with the ratio ≥ 2 or ≤ 0.5 as an "unbalanced" study [19,23]. Previous methodological studies defined a meta-analysis of unbalanced studies when all the included studies were unbalanced in the simulation setting [15,19,24]. However, this is unrealistic since

balanced and unbalanced studies co-exist in a meta-analysis [25]. Therefore, in this study, we defined a meta-analysis of unbalanced studies when ≥50% of the studies were unbalanced. Based on this definition, we further divided the meta-analyses into balanced or unbalanced, and compared the performance of excluding vs including double-zero studies in these two subgroups. We did not consider Peto's OR and MH OR since these two methods are not feasible for studies with double-zero studies. In Cheng's study, the DerSimonian and Lairds random-effect (RE) model [26] was used to pool data within meta-analysis, but as stated above this assumes that expected study effects are normally distributed, and given that Cheng et al.'s simulation also generated data under this assumption this creates a self-fulfilling prophecy where data are generated the way they will be analyzed. To avoid the latter as well as to avoid the pitfalls of such an assumption (see above) resulting in poor performance of random effects models under heterogeneity [16,27,28], we only consider the use of Doi's inverse variance heterogeneity (IVhet) model to synthesize the intervention effects across studies [29]. Let's assume w_i indicate the inverse variance weight for the ith study, θ_i the treatment effect of the ith study, ${s_i}^2$ the within-study variance of the ith study, and τ^2 the method of moments based between-study variance, then the IVhet model effect estimate and variance are computed as:

$$\widehat{\theta}_{\text{IVhet}} = \sum_{i=1}^{k} w_i\theta_i, \text{where } w_i = \frac{1}{{s_i}^2} \Big/ \sum_{i=1}^{k} \frac{1}{{s_i}^2}$$

$$\text{Var}(\widehat{\theta}_{\text{IVhet}}) = \sum_{i=1}^{k} w_i^2 {s_i}^2 \psi_i = \sum_{i=1}^{k} w_i^2 \left({s_i}^2 + \tau^2\right)$$

Here ψ_i is the study-specific overdispersion correction which can be calculated as $\frac{{s_i}^2+\tau^2}{{s_i}^2}$.

3.2. Statistical performance

We consider four parameters to measure the statistical performance of the meta-analytic methods, they were i) $bias^2 = (\widehat{\theta} - \theta)^2$, ii) mean squared error (MSE=$bias2 + Var_{\widehat{\theta}}$), iii) width of 95% confidence interval ($ub - lb$), and iv) coverage proportion [30]. The method with the smallest MSE and closest to nominal coverage was considered to have the best performance. The exact computations used have been reported elsewhere [16].

Here we did not consider bias alone as an important parameter because the primary purpose of a meta-analysis is to trade off increase in bias against reduction in variance. This is also the reason why statisticians do not use the arithmetic mean (i.e., unweighted or natural mean) in meta-analysis, although it is an unbiased estimator [31]. The data generation and analyses were conducted in Stata software (Stata 14/SE, StataCorp, USA) with code in the supplementary material.

4. Results

4.1. Excluding versus including studies with no events

Figs. 1 and 2 present the performance comparison between excluding and including double-zero studies without (OR = 1, Fig. 1) and with (OR = 2, Fig. 2) a treatment effect, with true risk in the control group = 0.01, and the number of studies = 10. In both scenarios, the results were similar. There was smaller bias when studies with no events were excluded; however, it presented much larger MSE than when studies with no events were included. Despite having a wider confidence

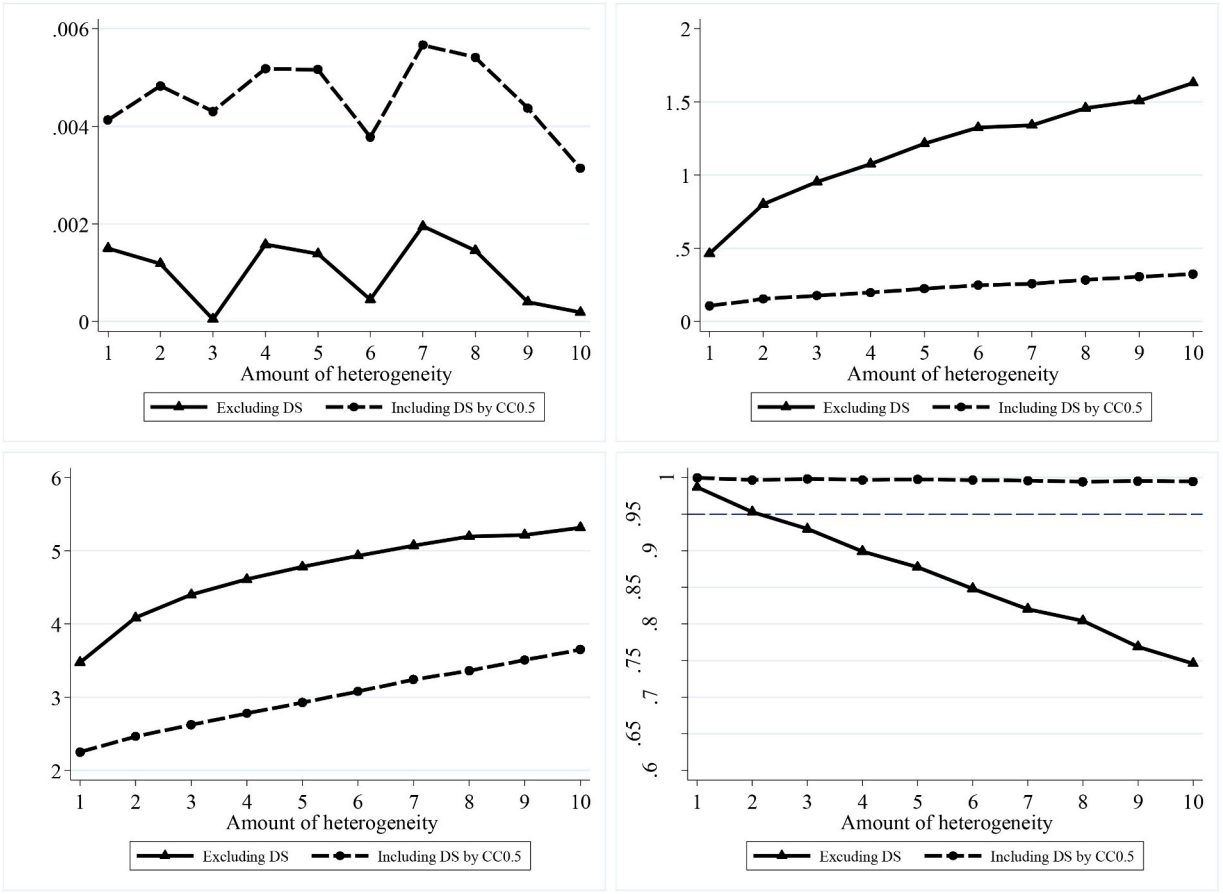

Fig. 1. Performance comparison between excluding (solid lines) and including (dashed lines) studies with no events when there is no treatment effect (OR = 1).

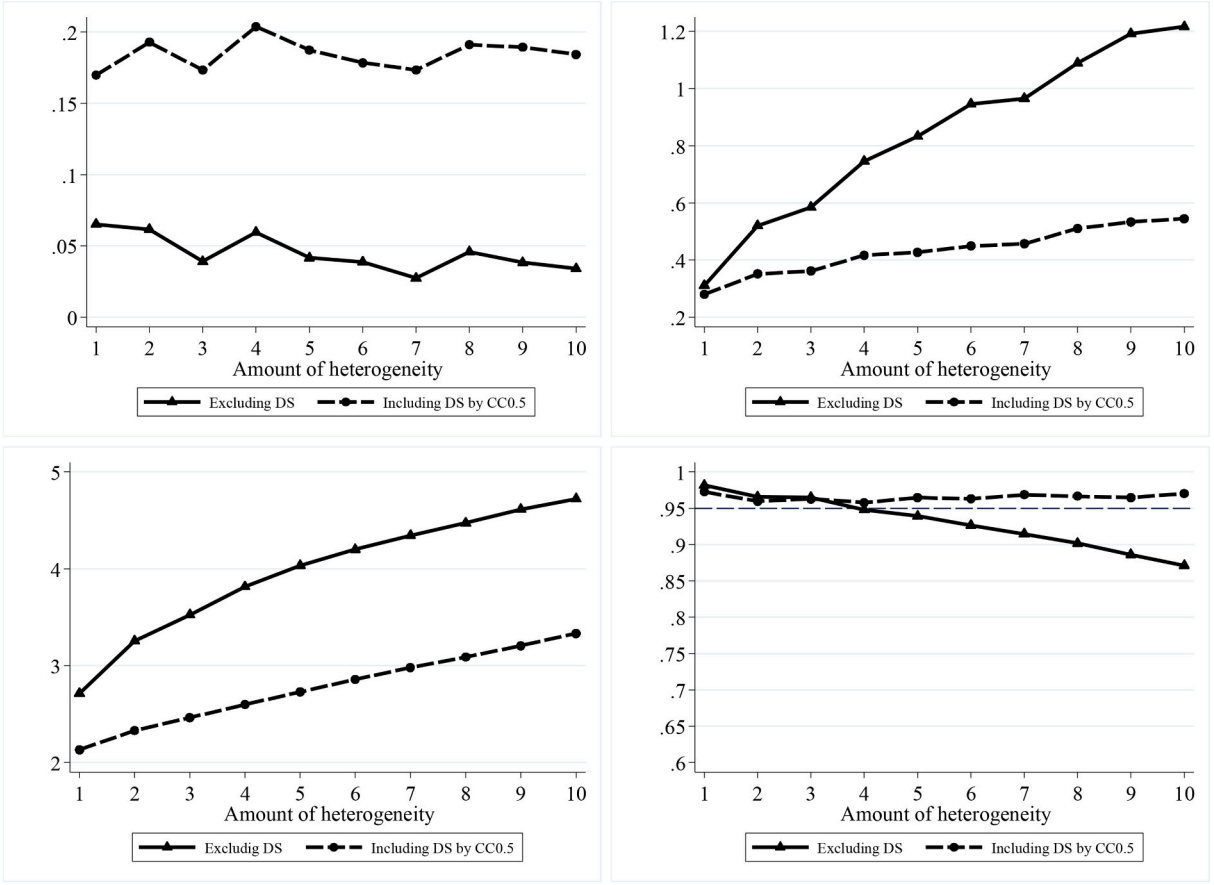

Fig. 2. Performance comparison between excluding (solid lines) and including (dashed lines) studies with no events when there is treatment effect (OR = 2).

interval, the coverage when studies were excluded rapidly deteriorate as heterogeneity increased, while the coverage remained above the nominal level (i.e., 95%) when studies with no events were included. Overall, under these scenarios, including studies with no events in both arms performed substantively better than excluding studies with no events.

4.2. Varying control event rates

Figures S1 (OR = 1) and S2 (OR = 2) present the performance when the control event rate was set at 0.05, while Figs. S3 (OR = 1) and S4 (OR = 2) when the control event rate was set at 0.1. When OR = 1 (S1 and S3), there was almost unbiased point estimation (bias <0.002) for both including and excluding studies with no events. Including studies with no events had a smaller MSE, while the coverage was at the nominal level across the different levels of heterogeneity for including and excluding studies with no events (due to the wider confidence interval for excluding studies). When OR = 2, the bias of including studies with no events increased (bias ranged from 0.01 to 0.03), the coverage remained at the nominal level for including and excluding studies with no events, and again, including studies showed smaller MSE.

To explore the reasons of the discrepancy between the coverage when excluding studies with no events with control event rate of 0.01 (poor coverage) versus 0.05 and 0.1 (adequate coverage), we summarized the median I^2 for the 10 runs (Table 1) and the proportion of studies with no events for the different control group event rates (Table 2). We observed that excluding studies with no events in both arms (compared to including studies) presented larger between-study variance in all the scenarios, and this became more prominent when the control event rate was 0.01 (Table 1). In addition, when the control

Table 1
The mean I^2 from runs 1 to 10.

Run number (increasing amount of heterogeneity)	Mean I^2 (%) (OR = 1)		Mean I^2 (%) (OR = 2)		Mean I^2 (%) (OR = 1)		Mean I^2 (%) (OR = 2)		Mean I^2 (%) (OR = 1)		Mean I^2 (%) (OR = 2)	
	Event risk = 0.01				Event risk = 0.05				Event risk = 0.1			
	Excluding	Including	Excluding	Including	Excluding	Including	Excluding	Including	Excluding	Including	Excluding	Including
1	3.2	0	1.4	0	1.5	1.2	2.2	2.2	4.4	4.4	5.6	5.6
2	6.8	0.1	5.5	0.3	11.7	8.3	14.5	12.8	22.2	21.1	26.7	26.2
3	8.2	0.25	6.6	0.6	17.5	11.6	20.7	17.5	31.2	28.7	36.3	35.1
4	8.2	0.36	8.3	0.8	19.9	13.1	24.7	20.5	36.5	32.7	41.3	39.4
5	8.3	0.43	8.7	1	21.7	13.9	25.9	21.3	39.1	35	45.6	43.4
6	9.1	0.6	8.9	1.2	22.9	14.6	26.7	21.8	41.2	36.4	47.4	44.8
7	8.9	0.83	9.3	1.3	25.2	15.7	28.8	23.5	43.7	38.9	49.6	46.9
8	8.5	0.9	9.5	1.6	26.7	16.6	30.8	24.9	44.1	38.9	50.6	47.8
9	8.8	1	9.6	1.7	27.9	17	31.3	24.9	45	39.6	51.5	48.6
10	7.9	1	9.2	1.9	29.2	18.4	31.9	25.4	46.5	40.8	51.3	48.4

Table 2
The proportion of double-zero studies in simulated meta-analyses.

Proportion of double-zero studies	OR = 1	OR = 2
Baseline risk = 0.01	66.7% (IQR: 55.6% to 80%)	50% (IQR: 40% to 66.7%)
Baseline risk = 0.05	20% (IQR: 10% to 28.6%)	10% (IQR: 0% to 16.7%)
Baseline risk = 0.10	0% (IQR: 0% to 11.1%)	0% (IQR: 0% to 10.0%)

event rate was 0.01, the median proportion of double-zero studies was 67% while it decreased to 20% and 0% when the control event rate increased to 0.05 and 0.1, respectively (Table 2). These observations explained why excluding double-zero studies had poor coverage and large MSE in the simulation when control event rate was the smallest - excluding studies led to substantially larger heterogeneity and as the proportion of double-zero studies in a meta-analysis increased, the increase in heterogeneity became more prominent. This also explained why when the control event rate was 0.05 and 0.1 (small proportion of double-zero studies), the performance of excluding or including double-zero studies was similar.

For the scenarios when number of studies was set as 5, we observed similar results as with 10 studies per meta-analysis (Figs. S5 and S6).

4.3. Varying sample size ratio (balanced and unbalanced studies)

When the baseline risk was 0.05 and 0.1 there were few double-zero studies in most of the meta-analyses (Table 2), thus we focused on the scenarios where the baseline risk was 0.01. Figs. S7 (OR = 1) and S8 (OR = 2) present the performance of excluding versus including studies with no events when studies were balanced in the meta-analysis; while Figs. S9 (OR = 1) and S10 (OR = 2) when studies were unbalanced. Our results suggest that for both, balanced and unbalanced meta-analyses, including studies with no events in both arms outperformed excluding studies.

4.4. Adding 0.5 to all studies

In response to a reviewer's suggestion, we added a simulation that compared an alternative method, namely, adding 0.5 to cells in all studies. This is because adding 0.5 has been proposed for reducing bias in the empirical logit method regardless of the occurrence of zero events [32–34]. Figs. S11 (OR = 1) and S12 (OR = 2) present the results. Our results suggest that adding the 0.5 continuity correction to all studies performed the best amongst all the three methods, which further reinforced our findings that including double-zero studies outperforms excluding them.

5. Example from the literature

We used data from a Cochrane review investigating the effect of antibiotic use for preventing complications in children with measles [35]. Two outcomes, otitis media and mortality, were selected for the review. For otitis media, five studies were included of which one was a double-zero study and one was a single-zero study. For mortality, seven studies were included of which five were double-zero studies and one was a single-zero study. Both could be classified as "MA-MZ" (meta-analysis with a mixture of zero-events study types) according to our framework [1]. We used the continuity correction under Doi's QE [36] and Doi's IVhet [29] and the RE_{DL} [26] models to deal with zero-events for both single-zero and double-zero studies. Figs. 3 and 4 present the results of including and excluding double-zero studies respectively. For otitis media, when including (OR_{IVhet} = 0.43; 95%CI: 0.20, 0.93) and excluding (OR_{IVhet} = 0.42; 95%CI: 0.19, 0.91) the double-zero studies, the pooled ORs were very similar. With the QE model, the results were also similar between including and excluding studies, but with much less certainty than with the IVhet model.

For the mortality outcome, the pooled OR when including (OR_{IVhet} = 1.45; 95%CI: 0.40, 5.20) and excluding the 5 double-zero studies (OR_{IVhet} = 3.12; 95%CI: 0.48, 20.13) significantly differed indicating the problem with excluding double-zero studies, and it is likely that bias was introduced by exclusion of studies. With the QE model, including or excluding the 5 double-zero studies, resulted in a pooled OR of 1.02 (95%CI: 0.14, 7.57) and 3.12 (95%CI: 0.48, 20.13) and again it is

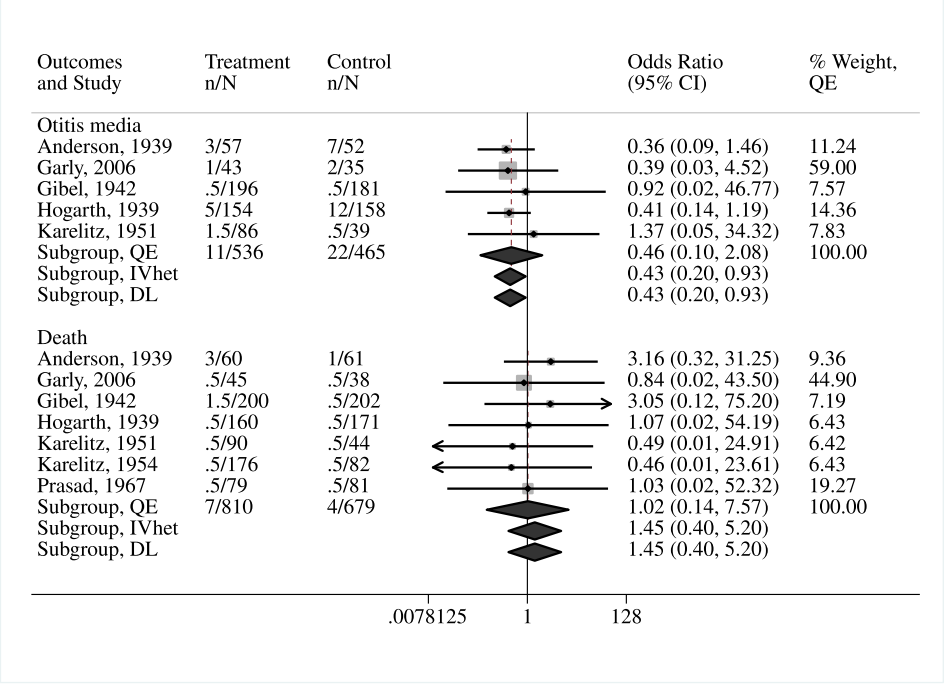

Fig. 3. Forest plot of a real-world example, including double-zero studies.

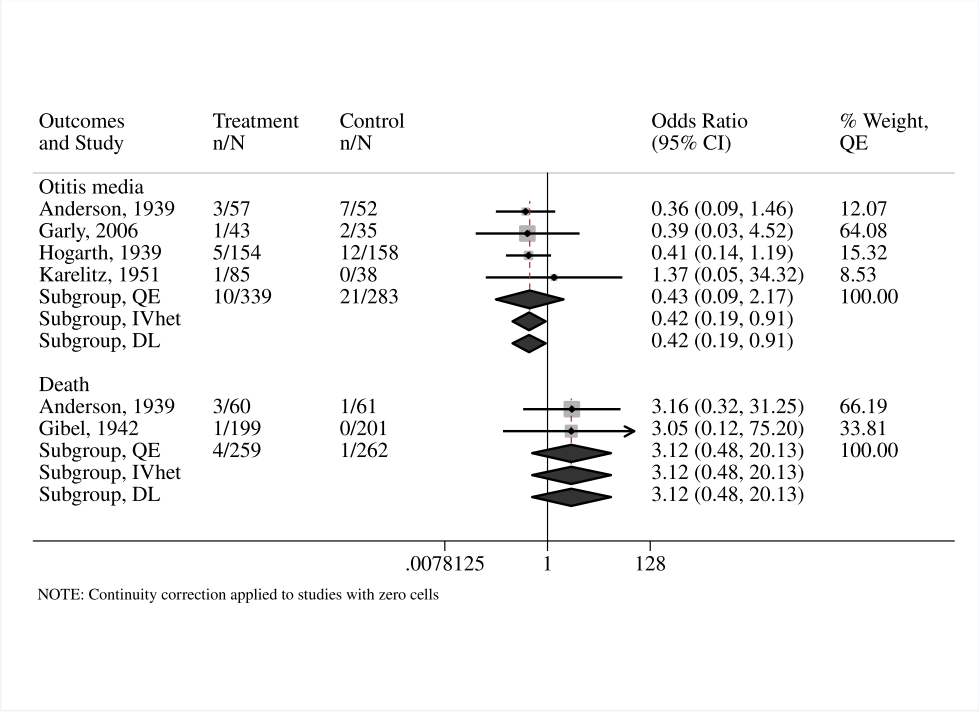

Fig. 4. Forest plot of a real-world example, excluding double-zero studies.

evident that bias adjustment fails when double zero studies are excluded. Excluding double-zero studies from the meta-analysis may have substantial impact on the effect estimates, and the impact is closely associated with the proportions of double-zero studies within a meta-analysis. Of note the IVhet and RE models concur in this example because $I^2 = 0\%$ in all meta-analyses and the limitations we outlined for the RE model kick in as heterogeneity increases.

6. Discussion

In the current study, we replicated the simulation study by Cheng et al. [15] but with a more suitable data generating mechanism to compare the statistical properties of including studies with no events versus excluding them. Our results suggest that including studies with no events in both arms with a continuity correction performed better than excluding them, and this is consistent with Cheng's results. While we further demonstrated that even when there is a treatment effect, including zero-event studies has better performance than excluding them, which differs from Cheng's conclusion. We also found the advantage of including double-zero studies, especially when the proportion of double-zero studies in a meta-analysis is large, as excluding them would increase the between-study variance and further lead to inflated MSE and serious drop in coverage.

Based on our findings, excluding studies with no events in both arms in meta-analysis may not be the most appropriate choice. These findings are consistent with our previous study that applied one-stage methods to compare the statistical performance of including versus excluding such studies [10]. One concern about studies with no events is that the observed events may largely mislead when the risk in both arms is extremely small. For example, suppose there is a 1:1 design trial with a sample size of 100 and true event risk in both arms are 0.005 and 0.001, then the true OR ~ 5. The observed events for both arms are likely to be 0, which makes the estimated OR $= 1$, which then brings random error into the meta-analysis. However, we argue that each study faces the problem of random error, while the method of meta-analysis incorporates variability due to such error into the results. It could be expected that when random and systematic error are well-addressed by appropriate methods, the results of a meta-analysis are still more reliable, regardless of the amount of error brought by a single study.

The use of a continuity correction to deal with zero-event studies is debated for evidence synthesis. This opinion, to some extent, is correct as it would lead to large amount of bias for unbalanced studies [19,37]; While we argue that this opinion is based on the extreme assumption that all included studies are unbalanced, while in the real-world, such a situation is extremely rare as meta-analyses generally contain both balanced and unbalanced studies [25]. Based on a more reasonable definition of "unbalanced" meta-analysis, our simulation results suggest that continuity correction under the IVhet model has good performance, especially when the events are extremely rare.

Another concern with the continuity correction is that although adding 0.5 corrects the first-order bias as compared with not using this correction, it may raise a second-order bias such that the correction may not be sufficient [33]. Statisticians have suggested alternative corrections such as add -0.5 for weighted empirical logit or add 0.25 to minimize the bias of the slope [36,37]. However, none of these corrections suit well for all situations, as Gart et al. [33] pointed out "*It is not possible to recommend a universal correction, …, sometimes 1/2 is best, other times 1/4, 0, -1/2 is best…*". Further review authors may consider sensitivity analysis with divergent correction elements.

For the problem of dealing with double-zero studies in meta-analysis, the current recommendations from the Cochrane Collaboration is to exclude them from the meta-analysis [21], and most meta-analytical software and packages exclude these studies by default. However, the results of the current study suggest that studies with zero-events are not necessarily non-informative and simply excluding them may produce biased pooled estimates. Our findings concur with the results and recommendations from related studies [5,38–47].

7. Conclusion

Based on current evidence, including double-zero studies with a continuity correction performed substantively better than excluding them, especially when the proportion of double-zero studies was large. Therefore, we recommend that double- zero-events studies are included

in future meta-analysis by applying a continuity correction of 0.5 (instead of the current practice of excluding them from the synthesis) and perhaps extending the continuity correction to all studies when there are double-zero studies in the mix.

Ethics statement

We declare that no ethics involved in this study.

Availability of data and simulation code

See supplementary material.

Funding

This work was made possible by the National Natural Science Foundation of China (72204003) and program grant #NPRP-BSRA01-0406-210030 from the Qatar National Research Fund (a member of Qatar Foundation). Luis Furuya-Kanamori was supported by an Australian National Health and Medical Research Council Fellowship (APP1158469). The findings herein reflect the work, and are solely the responsibility of the authors. The funding bodies had no role in any process of the study (i.e., study design, analysis, interpretation of data, writing of the report, and decision to submit the article for publication).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Simulation data was used for the research described in the article and can be generated using the code provided.

Acknowledgment

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cct.2022.106962.

References

[1] C. Xu, L. Furuya-Kanamori, L. Zorzela, L. Lin, S. Vohra, A proposed framework to guide evidence synthesis practice for meta-analysis with zero-events studies, J. Clin. Epidemiol. 135 (2021) 70–78.
[2] S. Yusuf, R. Peto, J. Lewis, R. Collins, P. Sleight, Beta blockade during and after myocardial infarction: an overview of the randomised trials, Prog. Cardiovasc. Dis. 27 (5) (1985) 335–371.
[3] N. Mantel, W. Haenszel, Statistical aspects of the analysis of data from retrospective studies of disease, J. Natl. Cancer Inst. 22 (4) (1959) 719–748.
[4] G. Rücker, G. Schwarzer, J. Carpenter, et al., Why add anything to nothing? The arcsine difference as a measure of treatment effect in meta-analysis with zero cells, Stat. Med. 28 (5) (2009) 721–738.
[5] O. Kuss, Statistical methods for meta-analyses including information from studies without any events-add nothing to nothing and succeed nevertheless, Stat. Med. 34 (7) (2015) 1097–1116.
[6] Higgins JPT, J. Thomas, J. Chandler, et al. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (Updated February 2021), Cochrane, 2021. Available from, www.training.cochrane.org/handbook.
[7] C. Xu, X.Q. Zhou, L. Zorzela, et al., Utilization of the evidence from studies with no events in meta-analyses of adverse events: an empirical investigation, BMC Med. 19 (1) (2021) 141, https://doi.org/10.1186/s12916-021-02008-2.
[8] D. Böhning, P. Sangnawakij, The identity of two meta-analytic likelihoods and the ignorability of double-zero studies, Biostatistics (2020) kxaa004.
[9] M.G. Xie, J. Kolassa, D.G. Liu, et al., Does an observed zero-total-event study contain information for inference of odds ratio in meta-analysis? Stat. Interface 11 (2018) 327–337.
[10] C. Xu, L. Li, L. Lin, et al., Exclusion of studies with no events in both arms in meta-analysis impacted the conclusions, J. Clin. Epidemiol. 123 (2020) 91–99.
[11] R.L. Plackett, The continuity correction in 2 x 2 tables, Biometrika 51 (3–4) (1964) 327–337.
[12] D.K. Bhaumik, A. Amatya, S.L. Normand, et al., Meta-analysis of rare binary adverse event data, J. Am. Stat. Assoc. 107 (498) (2012) 555–567.
[13] S.A. Doi, L. Furuya-Kanamori, C. Xu, L. Lin, T. Chivese, L. Thalib, Questionable utility of the relative risk in clinical research: a call for change to practice, J. Clin. Epidemiol. 142 (2022) 271–279, https://doi.org/10.1016/j.jclinepi.2020.08.019.
[14] S.A. Doi, L. Furuya-Kanamori, C. Xu, et al., The odds ratio is "portable" but not the relative risk: time to do away with the log link in binomial regression, J. Clin. Epidemiol. 142 (2022) 288–293, https://doi.org/10.1016/j.jclinepi.2021.08.003.
[15] J. Cheng, P. Eleanor, J.K. Marshall, et al., Impact of including or excluding both-armed zero-event studies on using standard meta-analysis methods for rare event outcome: a simulation study, BMJ Open 6 (8) (2016), e010983.
[16] S.A.R. Doi, L. Furuya-Kanamori, Selecting the best meta-analytic estimator for evidence-based practice: a simulation study, Int. J. Evid. Based Healthc. 18 (1) (2020) 86–94.
[17] S.A.R. Doi, L. Furuya-Kanamori, L. Thalib, J.J. Barendregt, Meta-analysis in evidence-based healthcare: a paradigm shift away from random effects is overdue, Int. J. Evid. Based Healthc. 15 (4) (2017) 152–160.
[18] J.B. Haldane, The estimation and significance of the logarithm of a ratio of frequencies, Ann. Hum. Genet. 20 (4) (1956) 309–311.
[19] M.J. Sweeting, A.J. Sutton, P.C. Lambert, What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data, Stat. Med. 23 (9) (2004) 1351–1375.
[20] K. Ju, L. Lin, H. Chu, L.L. Cheng, C. Xu, Laplace approximation, penalized quasi-likelihood, and adaptive gauss-Hermite quadrature for generalized linear mixed models: towards meta-analysis of binary outcome with sparse data, BMC Med. Res. Methodol. 20 (1) (2020) 152.
[21] J.J. Deeks, Higgins JPT, D.G. Altman, Chapter 10: analysing data and undertaking meta-analyses, in: Higgins JPT, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, V.A. Welch (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (Updated February 2021), Cochrane, 2021. Available from, www.training.cochrane.org/handbook.
[22] L. Furuya-Kanamori, P. Kostoulas, S.A.R. Doi, A new method for synthesizing test accuracy data outperformed the bivariate method, J. Clin. Epidemiol. 132 (2021) 51–58.
[23] M.J. Bradburn, J.J. Deeks, J.A. Berlin, Localio A. Russell, Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events, Stat. Med. 26 (1) (2007) 53–77.
[24] K. Pateras, S. Nikolakopoulos, K. Roes, Data-generating models of dichotomous outcomes: heterogeneity in simulation studies for a random-effects meta-analysis, Stat. Med. 37 (7) (2018) 1115–1124.
[25] J.C. Dumville, S. Hahn, J.N. Miles, D.J. Torgerson, The use of unequal randomisation ratios in clinical trials: a review, Contemp. Clin. Trials. 27 (1) (2006) 1–12.
[26] R. DerSimonian, N. Laird, Meta-analysis in clinical trials, Control. Clin. Trials 7 (3) (1986) 177–188.
[27] K. Johnson, A. Hayen, M.N.D. Lassere, K. Mengersen, Bias, coverage, and asymptotic behaviour of random effects meta-analysis: a clinically driven simulation study, JBI Evid. Implement. 18 (4) (2020) 355–367.
[28] M. Henmi, J.B. Copas, Confidence intervals for random effects meta-analysis and robustness to publication bias, Stat. Med. 29 (29) (2010) 2969–2983.
[29] S.A. Doi, J.J. Barendregt, S. Khan, L. Thalib, G.M. Williams, Advances in the meta-analysis of heterogeneous clinical trials I: the inverse variance heterogeneity model, Contemp. Clin. Trials 45 (Pt A) (2015) 130–138.
[30] T.P. Morris, I.R. White, M.J. Crowther, Using simulation studies to evaluate statistical methods, Stat. Med. 38 (11) (2019) 2074–2102.
[31] J.J. Shuster, Empirical vs natural weighting in random effects meta-analysis, Stat. Med. 29 (12) (2010) 1259–1265.
[32] S.K. Kabra, R. Lodha, D.J. Hilton, Antibiotics for preventing complications in children with measles, Cochrane Database Syst. Rev. (3) (2008) CD001477, https://doi.org/10.1002/14651858.CD001477.pub3. Update in: Cochrane Database Syst Rev. 2013;8:CD001477.
[33] J.J. Gart, H.M. Pettigrew, D.G. Thomas, The effect of bias, variance estimation, skewness and kurtosis of the empirical logit on weighted least squares analyses, Biometrika. 72 (1) (1985) 179–190.
[34] S. Donnelly, J. Verkuilen, Empirical logit analysis is not logistic regression, J. Mem. Lang. 94 (2017) 28–42.
[35] L.J. Davis, Modification of the empirical logit to reduce bias in simple linear logistic regression, Biometrika. 72 (1) (1985) 199–202.
[36] S.E. Hitchcock, A note on the estimation of the parameters of the logistic function, using the minimum logit X^2 method, Biometrika. 49 (1) (1962) 250–252.
[37] D.R. Cox, E.J. Snell, Analysis of binary data (2nd Edition), 1970. Methuen, London.
[38] S.A. Doi, J.J. Barendregt, S. Khan, L. Thalib, G.M. Williams, Advances in the meta-analysis of heterogeneous clinical trials II: the quality effects model, Contemp. Clin. Trials 45 (Pt A) (2015) 123–129.
[39] O. Efthimiou, Practical guide to the meta-analysis of rare events, Evid. Based Ment. Health 21 (2) (2018) 72–76.
[40] D. Böhning, K. Mylona, A. Kimber, Meta-analysis of clinical trials with rare events, Biom. J. 57 (4) (2015) 633–648.
[41] J.O. Friedrich, N.K. Adhikari, J. Beyene, Inclusion of zero total event trials in meta-analyses maintains analytic consistency and incorporates all available data, BMC Med. Res. Methodol. 7 (2007) 5.

[42] C. Xu, L. Furuya-Kanamori, L. Lin, Synthesis of evidence from zero-events studies: a comparison of one-stage framework methods [published online ahead of print, 2021 Aug 13], Res. Synth. Methods 13 (2) (2022) 176–189.
[43] T. Cai, L. Parast, L. Ryan, Meta-analysis for rare events, Stat. Med. 29 (20) (2010) 2078–2089.
[44] M. Xiao, L. Lin, J.S. Hodges, et al., Double-zero-event studies matter: a re-evaluation of physical distancing, face masks, and eye protection for preventing person-to-person transmission of COVID-19 and its policy impact, J. Clin. Epidemiol. S0895-4356 (21) (2021), 00032-9.
[45] G. Yang, D. Liu, J. Wang, M.G. Xie, Meta-analysis framework for exact inferences with application to the analysis of rare events, Biometrics. 72 (4) (2016) 1378–1386.
[46] D. Liu, R.Y. Liu, M. Xie, Exact meta-analysis approach for discrete data and its application to 2 × 2 tables with rare events, J. Am. Stat. Assoc. 109 (508) (2014) 1450–1465.
[47] S.C. Morton, M.H. Murad, E. O'Connor, et al., Quantitative Synthesis—An Update. Methods Guide for Comparative Effectiveness Reviews (Prepared by the Scientific Resource Center under Contract No. 290–2012-0004-C). AHRQ Publication No. 18-EHC007-EF, Agency for Healthcare Research and Quality, Rockville, MD, February 2018.