Robust inference for the unification of confidence intervals in meta-analysis

Wei Liang, Haicheng Huang, *Xiamen University, China,*

Hongsheng Dai*, *Newcastle University, Newcastle upon Tyne, U.K.*

hongsheng.dai@newcastle.ac.uk,

and Yinghui Wei, *University of Plymouth, U.K.*

April 23, 2024

Abstract

Traditional meta-analysis assumes that the effect sizes estimated in individual studies follow a Gaussian distribution. However, this distributional assumption is not always satisfied in practice, leading to potentially biased results. In the situation when the number of studies, denoted as K, is large, the cumulative Gaussian approximation errors from each study could make the final estimation unreliable. In the situation when K is small, it is not realistic to assume the random-effect follows Gaussian distribution. In this paper, we present a novel empirical likelihood method for combining confidence intervals under the meta-analysis framework. This method is free of the Gaussian assumption in effect size estimates from individual studies and from the random-effects. We establish the large-sample properties of the non-parametric estimator, and introduce a criterion governing the relationship between the number of studies, K, and the sample size of each study, n_i. Our methodology supersedes conventional meta-analysis techniques in both theoretical robustness and computational efficiency. We assess the performance of our proposed methods using simulation studies, and apply our proposed methods to two examples. confidence Interval; Empirical Likelihood; Meta-analysis; Random-effect Model.

1 Introduction

Meta-analysis is a statistical method to combine results from multiple studies, facilitating the inference-making process concerning specific parameters of interests, such as odds ratio.

---

[0]To whom correspondence should be addressed.

Parameters estimated from individual studies can be combined using either a fixed-effect or a random-effects meta-analysis model (Higgins and Thompson, 2009; Deeks *and others*, 2022). A fixed-effect model assumes that a single parameter value is common to all studies, and a random-effects model assumes that parameters of the underlying studies follow some distributions.

Let the random variable Y_i $(i = 1, 2, \cdots, K)$ denote an effect size estimate from study i. Here Y_i could be the mean difference between treatment groups, log-odds ratios, log-hazard ratios, or other effect size estimates from study i. The conventional random-effects meta-analysis model is defined as

$$Y_i = \theta_i + \epsilon_i, \quad \theta_i = \theta + \xi_i \tag{1}$$

where θ_i is the true effect size of study i, θ is the common overall effect size, ϵ_i is the within-study random error and ξ_i is a random variable reflecting study-specific deviation from the common overall effect. In the literature, it is often assumed that ϵ_i and ξ_i are independent Gaussian variables, with $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ and $\xi_i \sim \mathcal{N}(0, \tau^2)$, where σ_i^2 is the within-study error variance and depends on the sample size n_i of study i. For instance, σ_i^2 is usually in the order of n_i^{-1}. The between-study variance τ^2 reflects the heterogeneity in the effect sizes across studies. Under the conventional random-effects meta-analysis model, the marginal distribution of Y_i is a Gaussian distribution with mean θ and variance $\sigma_i^2 + \tau^2$. The overall effect θ can be estimated by the maximum likelihood method (Brockwell and Gordon, 2001; Normand, 1999); and the between-study variance τ^2 can also be estimated by the (profile) maximum likelihood method (Brockwell and Gordon, 2001; Hardy and Thompson, 1996; Partlett and Riley, 2017; Langan *and others*, 2019), DerSimonian and Laird's method of moments (DerSimonian and Laird, 1986) or quantile approximation (Brockwell and Gordon, 2007). Equation (1) becomes a fixed-effect meta-analysis model when $\xi_i = 0$ ($\mathbb{E}\,\xi_i = 0$ and $\tau^2 = 0$). Within the existing literature, both models have leaned upon a Gaussian assumption. Nonetheless, practical reality often deviates from this parametric assumption, as demonstrated by (Lee and Thompson, 2008; Brockwell and Gordon, 2001; Sutton and Higgins, 2008).

One important objective in meta-analysis is to estimate a confidence interval for the overall effect size. The confidence interval can be constructed using the asymptotic distribution for the estimator $\hat{\theta}$, derived from a maximum likelihood approach (Brockwell and Gordon, 2001) or a profile likelihood approach (Brockwell and Gordon, 2001; Hardy and Thompson, 1996). Both approaches relied on the assumption that the estimator $\hat{\theta}$ approximately follows a Gaussian distribution, which however may not be true if n_i and/or K is small. Constructing a confidence interval for θ based on alternative parametric distributions was proposed, including t-distribution, other parametric distributions which allow for poten-

tially skewed and heavy tails in random-effects distributions (Sidik and Jonkman, 2002; Lee and Thompson, 2008; Baker and Jason, 2008; Ozturk and Balakrishnan, 2020), and quantile approximation (Brockwell and Gordon, 2007; Jackson and Bowden, 2009). Confidence distributions (CDs), which can be viewed as "distribution estimators" (Schweder and Hjort, 2002), is a technique which can be used to construct point estimators, confidence intervals and p-values. The fusion of CD functions from individual studies have been explored for constructing a confidence interval for the overall effect size (Singh $and\ others$, 2005; Xie $and\ others$, 2011). Nevertheless, the CD approach still relies on Gaussian assumptions for the integration of CD functions from individual studies into a unified CD function. For example, Singh $and\ others$ (2005) and Xie $and\ others$ (2011) used Gaussian distribution to construct CD functions and Nagashima $and\ others$ (2019) proposed a parametric bootstrap algorithm based on Gaussian and t-distributions. We refer the reader to Veroniki $and\ others$ (2018) for a review of methods (all depend on parametric assumptions) to calculate a confidence interval for the estimated overall effect size from a random-effects meta-analysis.

Different from the existing meta-analysis methods, we consider a more challenging situation where only the $1-\alpha$ confidence interval for the effect size of study i, $[L_i, U_i]$, is observed, and individual study effect size estimates Y_i and within-study error variance estimate $\hat{\sigma}_i^2$ are not observed. Further, we seek to relax the Gaussian assumption. The aim of this paper is to propose a novel non-parametric meta-analysis method based on empirical likelihood (EL) to make robust inferences on the confidence interval of the overall effect size θ. We relax the Gaussian distributional assumption for Y_i and θ_i. In other words, we consider

$$Y_i = \theta_i + \epsilon_i, \ \ \epsilon_i \sim F_i(\cdot); \qquad \theta_i = \theta + \xi_i, \ \ \xi_i \sim G(\cdot), \tag{2}$$

where the distributions F_i, G are not subject to any parametric assumption. The proposed non-parametric method allows $K \to \infty$, which is more reasonable than many existing literature where K was treated as a constant. This innovation empowers the non-parametric meta-analysis to effectively manage a substantial number of studies.

The rest of the paper is organised as follows. The proposed non-parametric meta-analysis method is presented in Section 2 and 3. We evaluate the performance of the proposed method using simulation studies in Section 4. Two real example applications are presented in Section 5. The paper concludes with a discussion and theoretical proofs for the large sample theory are provided in Web Appendix A of the Supplementary Material.

2   Model-based meta-analysis approach

For the parametric Gaussian models presented in Section 1, Singh $and\ others$ (2005) and Xie $and\ others$ (2011) developed a general framework for combining results from indi-

vidual studies in a meta-analysis. They used a Confidence Distribution (CD) function $H(\boldsymbol{X}, \theta)$ of the data $\boldsymbol{X}$ and parameter θ, from which inference on confidence intervals, point estimates or p-values can be derived. Their method amalgamates the individual study's CD functions $H_i(\boldsymbol{X}_i, \theta)$ into a unified CD function $H_c(\theta)$, through a monotonic function $g_c(H_1(\boldsymbol{X}_1, \theta), \cdots, H_K(\boldsymbol{X}_K, \theta))$. Under the Gaussian assumption in the random-effects meta-analysis model, and assuming the estimated effect size Y_i and the within-study error variance σ_i^2 for studiy i are available, a typical CD function for θ is

$$H_i(\theta) = \Phi\left(\frac{\theta - Y_i}{s_i}\right), \tag{3}$$

where $\Phi(\cdot)$ is the cumulative Gaussian distribution function, $s_i^2 = \sigma_i^2 + \hat{\tau}^2$ in the random-effects model ($s_i^2 = \sigma_i^2$ in the fixed-effect model), and $\hat{\tau}^2$ is the estimated between-study variance.

Singh *and others* (2005) and Xie *and others* (2011) summarized methodologies used in the literature in group decision analysis, where linear combination of individual study result is commonly used as the final outcome. They introduced the following linear combination function

$$g_c(u_1, \cdots, u_K) = \sum_{k=1}^{K} w_k \, \Phi^{-1}(u_k), \tag{4}$$

where the weight $w_k = 1/s_k$. Then the combination of CDs for θ is defined as

$$H_c(\theta) = \Phi\left(\frac{g_c(H_1(\theta), \cdots, H_K(\theta))}{\sqrt{\sum_{i=1}^{K} w_i^2}}\right) = \Phi\left(\frac{\sum_{k=1}^{K} w_k \frac{\theta - Y_k}{s_k}}{\sqrt{\sum_{i=1}^{K} w_i^2}}\right).$$

Inference on θ can then be obtained using the combined CD function $H_c(\theta)$. For instance, the $(1 - \beta)$ confidence interval for θ can be constructed as

$$[L_C, U_C] = [H_c^{-1}(\beta/2), \, H_c^{-1}(1 - \beta/2)]. \tag{5}$$

The above approach involves two assumptions: 1. The model error terms in model (1) in Section 1 follow a Gaussian distribution; 2. The linear combination in (4) follows a zero-mean Gaussian distribution. If each of the sample size n_i is large enough, the model error term in model (1) can be assumed to be Gaussian, i.e. the above assumption 1 is reasonable, but it fails if n_i is small. However, this is a strong assumption. If the first assumption is violated, the second assumption will also be false. If the Gaussian assumption is not held true for some individual studies, in the presence of a large number of studies, i.e. large K, the accumulated errors by combining all different studies might not be bounded. For instance, under model (2) without the Gaussian assumptions, if we consider $\theta_i \sim N(\theta = 0, 1)$ but $y_{ij}|\theta_i \sim \chi^2(4) - 4 + \theta_i$, $Y_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij}$, then we will not be able to use the combined

statistic $Z = K^{-\frac{1}{2}} \sum_{i=1}^{K} (\theta - Y_i)/s_i$ to construct the meta-analysis confidence interval, since Z will not converge to 0 when $K \to \infty$. This can be seen from Figure 1. Furthermore, the relationship between the number of studies K and sample size n_i in meta-analysis has not been extensively discussed. We will propose a non-parametric method and relax this Gaussian assumption in the following section and discuss the relationship between K and n_i.
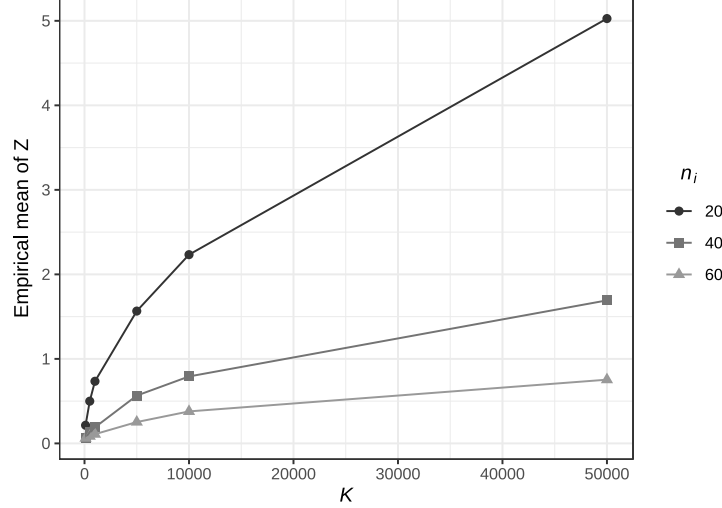

Figure 1: The empirical mean of Z with $\theta_i \sim N(0, 1)$, $y_{ij}|\theta_i \sim \chi^2(4) - 4 + \theta_i$ and $Y_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij}$. We consider the setting with large $K = 100, 500, 1000, 5000, 10000, 50000$ and small $n_i = 20, 40, 60$ for all i.

3 Non-parametric meta-analysis based on empirical likelihood

In this section, we propose a robust inference procedure for conducting meta-analysis that does not make distributional assumptions on effect size estimates for individual studies. Hence, this new approach is more reliable even if some studies have a small sample size n_i. We first present a generalized theorem and necessary assumptions.

Let $\boldsymbol{W}_i(\theta)$ be an r-dimensional function of θ for study i. If $\boldsymbol{W}_i(\theta)$ is asymptotically unbiased to 0 at the true parameter value θ_0, we can construct the estimating equation $\frac{1}{K} \sum_{i=1}^{K} \boldsymbol{W}_i(\theta) = 0$ for the unknown parameter θ. Following standard conventions in the literature, individual studies are independent of each other. Hence, we assume that the following assumptions hold.

ASSUMPTION 1. **Asymptotic Unbiaseness**. *For any i, at the true parameter θ_0,*

$$\mathbb{E}\, \boldsymbol{W}_i(\theta_0) = O(n_i^{-1/2}),$$

which means that each component of $\mathbb{E}\, \boldsymbol{W}_i$ is in the order of $n_i^{-1/2}$.

ASSUMPTION 2. **Independence**. *$\boldsymbol{W}_i(\theta), i = 1, \cdots, K$ are independent.*

We then consider the Empirical Likelihood (EL) approach and define the EL function

$$\mathcal{L}(\theta) = \sup \left\{ \prod_{i=1}^{K} p_i : \sum_{i=1}^{K} p_i\, \boldsymbol{W}_i(\theta) = 0, \sum_{i=1}^{K} p_i = 1 \right\}.$$

At the true parameter θ_0, the corresponding EL ratio statistic is

$$\mathcal{R}(\theta_0) = \frac{\mathcal{L}(\theta_0)}{\sup_\theta \mathcal{L}(\theta)}. \tag{6}$$

We further need the following assumptions to discuss the large-sample properties of the EL ratio statistic.

ASSUMPTION 3. **Finite Variance**. *For any i, at the true parameter θ_0, $\boldsymbol{W}_i(\theta_0)$ has a full rank finite covariance matrix $\boldsymbol{M}_i = var\,(\boldsymbol{W}_i(\theta_0))$, and $\boldsymbol{M}_i, i = 1, \cdots, K$ are uniformly bounded by $\boldsymbol{M}$, which means for any vector $\boldsymbol{e} \in \mathbb{R}^r$, $\sup_i \boldsymbol{e}^\mathrm{T}\, \boldsymbol{M}_i\, \boldsymbol{e} \le \boldsymbol{e}^\mathrm{T}\, \boldsymbol{M}\, \boldsymbol{e}$.*

Note that Assumption 3 always holds in practice, since typical classical statistical estimation can always be formalised as a sequence of asymptotically unbiased estimating equations, such as the likelihood inference and generalised estimating equation approaches. The finite variance of $\mathbf{M}_i$ is guaranteed when we use standardised estimating equations.

The last assumption is about the relationship between the number of studies K and the individual study sample size n_i, which is necessary when we discuss the large sample theories of meta-analysis under a non-parametric approach. To illustrate this assumption, we introduce the following notation and assumption. Let $n^{(K)}$ be the minimum number of individual sample sizes, that is $n^{(K)} = \min\{n_i, i = 1, \cdots, K\}$.

ASSUMPTION 4. **Large Sample Order Condition**. *$K/n^{(K)} \to 0$, as $K \to \infty$ or $n^{(K)} \to \infty$.*

Assumption 4 includes the scenario in the literature, where K is treated as a fixed number and sample sizes n_i may goes to infinity. First, the above assumption allows K to be a fixed number and in this case, as all n_i goes to infinity, the above assumption is always true. In addition, the above assumption is more general since it allows allows K, the number of studies, to go to infinity. Assumption 4 implies that for large values of K, we only need K to be in a lower order than all n_i, the sample sizes of individual studies. It gives a sufficient

condition required for that meta-analyses can be conducted with a substantial multitude of studies. Recently, Cai *and others* (2023) discussed the case of large K but finite n_i.

Under these assumptions, we have the main result as presented in the following theorem, with the proofs given in Web Appendix A of the Supplementary Material.

Theorem 5. *Assume that $\mathcal{R}(\theta)$ is an EL ratio statistic defined in (6). Under Assumptions 1-4, as $K \to \infty$, we have $-2 \log \mathcal{R}(\theta_0) \to \chi^2(r)$ in distribution.*

For a fixed confidence level β, we can construct the $(1 - \beta)$-level confidence interval for θ as

$$\{\theta : -2 \log \mathcal{R}(\theta) \leq \chi^2_{1-\beta}(r)\}. \tag{7}$$

Next, we will apply this theorem to two different meta-analysis models. The key point in applying the EL inference is to construct an estimating equation to define the EL ratio statistic.

3.1 Random-effects meta-analysis

Consider the random-effects meta-analysis model (2). Assume that we have collected $1 - \alpha$ confidence intervals $[L_i, U_i]$ for the parameter θ_i from K studies. The sample size for each study is denoted as n_i, $i = 1, 2, \cdots, K$. Based on the confidence intervals $[L_i, U_i]$, we define

$$W_i(\theta) = \frac{U_i + L_i - 2\theta}{U_i - L_i}, \tag{8}$$

which can be written as $W_i(\theta) = \frac{\overline{G}_i(\theta) - \underline{G}_i(\theta)}{\overline{G}_i(\theta) + \underline{G}_i(\theta)}$, with $\overline{G}_i(\theta) = U_i - \theta$ and $\underline{G}_i(\theta) = \theta - L_i$. Then

$$
\begin{aligned}
\mathbb{E}(\underline{G}_i(\theta_0)) &= \mathbb{E}(\theta_0 - Y_i + Y_i - L_i) = \mathbb{E}(\mathbb{E}(\theta_0 - Y_i)|\theta_i) + \mathbb{E}(Y_i - L_i) \\
&= \mathbb{E}(\theta_0 - \theta_i) + \mathbb{E}(Y_i - L_i) = \mathbb{E}(Y_i - L_i),
\end{aligned}
$$

and $\mathbb{E}(\overline{G}_i(\theta_0)) = \mathbb{E}(U_i - Y_i)$. Because of the large sample property of Y_i, it is generally true that

$$\mathbb{E}(Y_i - L_i) = O(n_i^{-1/2}), \qquad \mathbb{E}(U_i - Y_i) = O(n_i^{-1/2}).$$

On the other hand, when using empirical likelihood method to construct confidence intervals, we have $\mathbb{E}(Y_i - L_i) - \mathbb{E}(U_i - Y_i) \leq c \cdot n_i^{-1}$, for some constant c. This holds as well when confidence intervals are constructed via symmetric distributions, such as t-distribution or Gaussian distribution, since $\mathbb{E}(Y_i - L_i) - \mathbb{E}(U_i - Y_i) = 0$. Therefore, we here focus on the confidence intervals such that,

$$\mathbb{E}(Y_i - L_i) - \mathbb{E}(U_i - Y_i) \leq c \cdot n_i^{-1}.$$

Now we have

$$\mathbb{E}(\overline{G}_i(\theta_0)) - \mathbb{E}(\underline{G}_i(\theta_0)) = O(n_i^{-1}), \qquad \mathbb{E}(\overline{G}_i(\theta_0)) + \mathbb{E}(\underline{G}_i(\theta_0)) = O(n_i^{-1/2}), \qquad (9)$$

and thus Assumption 1 is satisfied, i.e.,

$$\mathbb{E}\, W_i(\theta_0) = O(n_i^{-1/2}). \qquad (10)$$

Now we define

$$\mathcal{L}(\theta) = \sup\left\{ \prod_{i=1}^{K} p_i : \sum_{i=1}^{K} p_i \left(\frac{L_i + U_i - 2\theta}{U_i - L_i} \right) = 0, \sum_{i=1}^{K} p_i = 1 \right\},$$

then the corresponding EL ratio statistic is $\mathcal{R}(\theta_0) = \mathcal{L}(\theta_0)/\sup_\theta \mathcal{L}(\theta)$. According to Theorem 5, we have $-2\log\mathcal{R}(\theta_0) \to \chi^2(1)$ in distribution.

3.2 Fixed-effect meta-analysis

We now consider a fixed-effect model, i.e. for all $i = 1, \ldots, K$, $\theta_i = \theta$ in equation (2). Since this is a special case of the random-effects model, we can still use the method presented in the previous subsection to make inferences. In the fixed-effect model, we can also use the confidence level to construct the estimating equation. First, we define

$$V_i(\theta) = I\{L_i \le \theta \le U_i\} - (1 - \alpha). \qquad (11)$$

Most of the time, the confidence interval $[L_i, U_i]$ is constructed based on an asymptotic distribution rather than an accurate distribution, hence, the expectation of (11) is approximately equal to 0. For example, when the confidence interval $[L_i, U_i]$ was constructed according to the central limit theorem, we have

$$\mathbb{E}\, V_i(\theta_0) = \mathbb{E}\left(I\{L_i \le \theta_0 \le U_i\} - (1 - \alpha) \right) = O(n_i^{-1/2}).$$

However, in the random-effects meta-analysis model the above $V_i(\theta_0)$ is not asymptotically unbiased to 0. This is demonstrated by the following result under model (1) with the Gaussian assumption for ϵ_i and ξ_i,

$$\mathbb{E}\, V_i(\theta_0) = \mathbb{E}\left(I\left\{ Y_i - z_{1-\frac{\alpha}{2}} \cdot \hat{\sigma}_i \le \theta_0 \le Y_i + z_{1-\frac{\alpha}{2}} \cdot \hat{\sigma}_i \right\} \right) - (1 - \alpha)$$
$$= \mathbb{E}\left(\Phi\left(z_{1-\frac{\alpha}{2}} - \frac{\theta_i - \theta_0}{\hat{\sigma}_i} \right) - \Phi\left(-z_{1-\frac{\alpha}{2}} - \frac{\theta_i - \theta_0}{\hat{\sigma}_i} \right) \right) - (1 - \alpha).$$

This expectation does not converge to 0, and it cannot be used to construct the asymptotically unbiased estimating equation in the random-effects model. Therefore, we consider the following three ways to define EL ratio statistic in the fixed-effect meta-analysis model. Denote $\boldsymbol{V}_i(\theta)$ as the estimating equation for study i, and r as the dimension of $\boldsymbol{V}_i(\theta)$.

- Case I: the estimating equation based on the confidence level, i.e. the following function $V_i(\theta) = I\{L_i \le \theta \le U_i\} - (1 - \alpha)$.

- Case II: the estimating equation based on symmetry (this is the same as the random-effects model), i.e. using the following function

$$V_i(\theta) = W_i(\theta) = z_{1-\frac{\alpha}{2}} \frac{L_i + U_i - 2\theta}{U_i - L_i}.$$

- Case III: the estimating equation based on confidence level and symmetry. In this case, $r = 2$, and $V_i(\theta) = (V_{i1}(\theta), \ V_{i2}(\theta))^{\mathsf{T}}$, where

$$V_{i1}(\theta) = I\{L_i \le \theta \le U_i\} - (1 - \alpha), \quad V_{i2}(\theta) = z_{1-\frac{\alpha}{2}} \frac{L_i + U_i - 2\theta}{U_i - L_i}.$$

By choosing an appropriate $V_i(\theta)$, we can define three different empirical likelihood statistics, which allow us to make inferences on the confidence interval in meta-analysis,

$$\mathcal{L}_{f1}(\theta) = \sup \left\{ \prod_{i=1}^{K} p_i : \sum_{i=1}^{K} p_i I\{L_i \le \theta \le U_i\} = 1 - \alpha, \ \sum_{i=1}^{K} p_i = 1 \right\},$$

$$\mathcal{L}_{f2}(\theta) = \sup \left\{ \prod_{i=1}^{K} p_i : \sum_{i=1}^{K} p_i \left(\frac{L_i + U_i - 2\theta}{U_i - L_i} \right) = 0, \ \sum_{i=1}^{K} p_i = 1 \right\}, \tag{12}$$

$$\mathcal{L}_{f3}(\theta) = \sup \left\{ \prod_{i=1}^{K} p_i : \sum_{i=1}^{K} p_i I\{L_i \le \theta \le U_i\} = 1 - \alpha, \ \sum_{i=1}^{K} p_i \left(\frac{L_i + U_i - 2\theta}{U_i - L_i} \right) = 0, \ \sum_{i=1}^{K} p_i = 1 \right\}.$$

Then the corresponding empirical likelihood ratio statistics are

$$\mathcal{R}_{fk}(\theta_0) = \frac{\mathcal{L}_{fk}(\theta_0)}{\sup_\theta \mathcal{L}_{fk}(\theta)}, \qquad k = 1, 2, 3$$

respectively. Following Theorem 5, we have $-2 \log \mathcal{R}_{f1}(\theta_0) \to \chi^2(1)$, $-2 \log \mathcal{R}_{f2}(\theta_0) \to \chi^2(1)$ and $-2 \log \mathcal{R}_{f3}(\theta_0) \to \chi^2(2)$, in distribution. Then for a fixed confidence level β, we can construct the $(1 - \beta)$-level confidence interval for θ based on (7).

4   Simulation Studies

To demonstrate the performance of the proposed non-parametric method, we conduct a comprehensive comparison against two established methods: the existing CD method and the conventional meta-analysis method. To be specific, DerSimonian & Laird method is used to estimate τ^2 in CD method, and 'rma' in the 'metafor' R package where we use REML to estimate τ^2.

We consider four different scenarios in this section, which corresponds to model (2) but have different distributions for the zero-mean error terms,

$$y_{ij} \sim F(\cdot \mid \theta_i, \sigma^2), \ \ Y_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij},$$

$$\theta_i = \theta + \xi_i, \ \ \xi_i \sim G(\cdot \mid \tau^2), \tag{13}$$

where $\mathbb{E}(y_{ij}) = \theta_i$, i.e. θ_i is the expected effect size of study i, Y_i is the estimated effect size of study i, and θ is the overall effect size. The parameter σ^2 is the within-study variance. This model setting is the same as (2), where $\text{var}(Y_i)$ is in the order of $\sigma_i^2 = \sigma^2/n_i$. For all scenarios, we set $\theta = 0$, and $\sigma^2 = 1$, and we consider different values of K, $20, 60$ and 100.

We set the four simulation scenarios as follows, with different distributions F and G.

- Scenario 1 is the Gaussian distribution case, where $F = \mathcal{N}(\theta_i, \sigma^2)$ and $G = \mathcal{N}(0, \tau^2)$. This is equivalent to model (1).

- Scenario 2 sets $F = \mathcal{N}(\theta_i, \sigma^2)$ and G to log-normal distribution with mean 0 and variance τ^2. The individual study effect θ_i is not Gaussian. This is common in practice. For example, a study's effect size may have a positive and heavy tail. We choose $\xi_i = c_2 \cdot \text{log-normal}(1, 1) + d_2$, where c_2 and d_2 are constants such that the distribution of ξ_i has mean 0 and variance as a pre-specified value τ^2.

- Scenario 3 sets F to be non-Gaussian, but G to be Gaussian $\mathcal{N}(0, \tau^2)$. When an individual study has a small sample size, the individual studies' effect size estimate Y_i does not follow a Gaussian distribution. In our simulation, we first generate $\theta_i \sim \mathcal{N}(0, \tau^2)$, $i = 1, 2, \cdots, K$. Given the sample size n_i and θ_i, draw $y_{ij} \mid \theta_i \sim a_{3,i} \cdot \chi^2(4) + b_{3,i}$ such that $\mathbb{E}(y_{ij} \mid \theta_i) = \theta_i, \text{var}(y_{ij} \mid \theta_i) = \sigma^2$.

- Scenario 4 is the most general case where we set both F and G to be non-Gaussian. We choose G to be the distribution of $c_4 \cdot \chi^2(4) + d_4$ to generate θ_i, and we obtain random sample $Y_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij}$ with $y_{ij} \mid \theta_i \sim a_{4,i} \cdot \text{log-normal}(1, 1) + b_{4,i}$, where $a_{4,i}$ and $b_{4,i}$ are such that $\mathbb{E}(y_{ij} \mid \theta_i) = \theta_i$, and $\text{var}(y_{ij} \mid \theta_i) = \sigma^2$.

For each scenario, we obtain the estimated effect size and the estimated within-study variance by $Y_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij}$ and $\hat{\sigma}_i^2 = \sum_{j=1}^{n_i} (y_{ij} - Y_i)^2/(n_i(n_i - 1))$, respectively. Then the $1 - \alpha = 95\%$ two-side confidence interval $[L_i, U_i]$ of θ is generated by either

$$L_i = Y_i - z_{1-\alpha/2} \cdot \hat{\sigma}_i, \quad U_i = Y_i + z_{1-\alpha/2} \cdot \hat{\sigma}_i, \qquad n_i \geq 30, \qquad (14)$$

for large n_i, or

$$L_i = Y_i - t_{1-\alpha/2}(n_i - 1) \cdot \hat{\sigma}_i, \quad U_i = Y_i + t_{1-\alpha/2}(n_i - 1) \cdot \hat{\sigma}_i, \qquad n_i < 30, \qquad (15)$$

for small n_i, where $t_\beta(n)$ is the β-quantile for the t-distribution with degree of freedom n.

We consider the settings of different values of τ^2, ranging from 0 to 100. When $\tau^2 = 0$, it becomes a fixed-effect meta-analysis model. Based on the observations $[L_i, U_i]$, we compare $1 - \beta = 95\%$ confidence intervals which are constructed via the conventional meta-analysis method, the CD method and the non-parametric EL method.

4.1 Scenarios 1, 2 and 3

We consider the conventional settings of random-effects meta-analysis model ($\tau^2 \neq 0$) and fixed-effect meta-analysis model ($\tau^2 = 0$). Figure 2 shows the results coverage probability against the ratio of between-study variance and within-study variance for Scenarios 1 to 3, based on different values of K, 20, 60, and 100, $\sigma^2 = 1$, and different values of τ^2, $0, 0.001, 0.01, 0.1, 1, 10$ and 100. For each study, the number of observations n_i is drawn randomly from the uniform distribution $0.2 \cdot \mathcal{U}[100, 500]$.
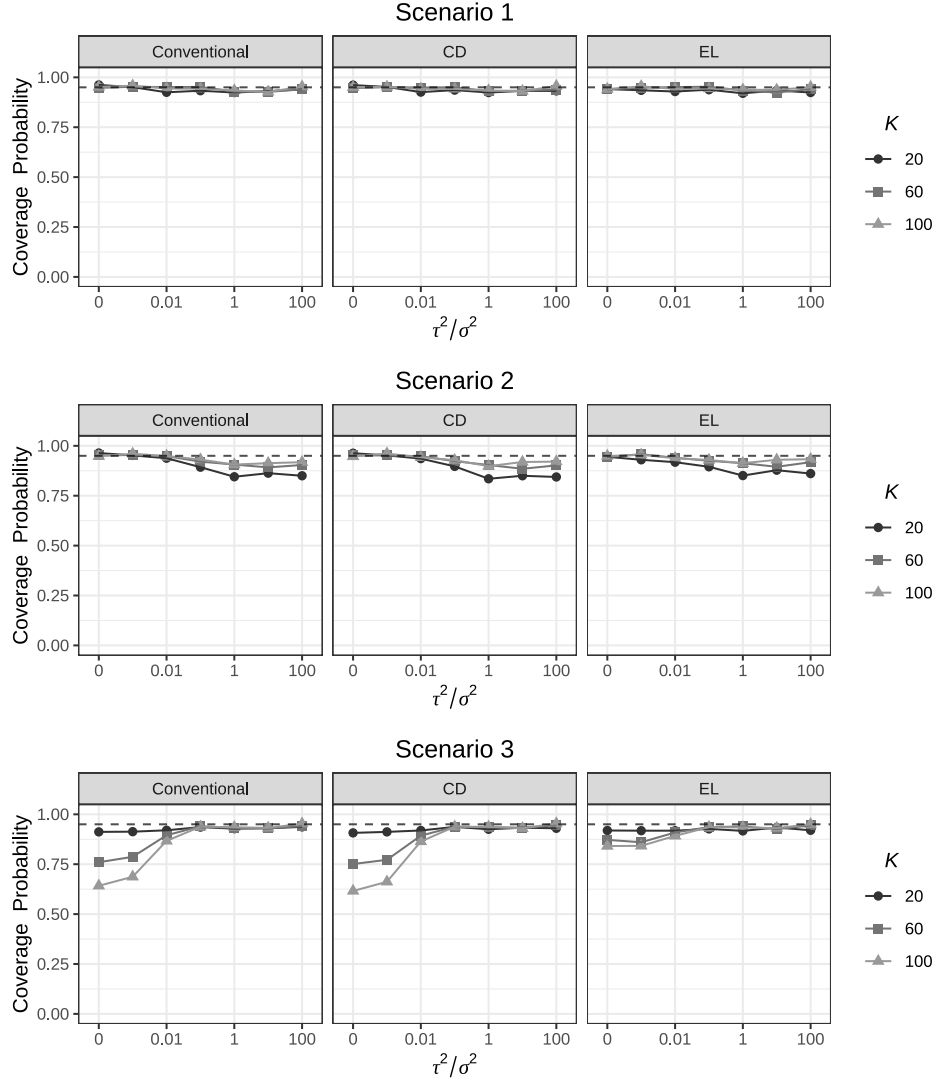

Figure 2: The coverage probabilities of $1 - \beta = 95\%$ confidence intervals which are constructed via the conventional meta-analysis method, CD method and EL method respectively under the setting of different K and τ^2 in Scenario 1. In each setting the n_i is sampled from $0.2 \cdot \mathcal{U}[100, 500]$.

The simulation results for Scenario 1 are shown in the first row of plots in Figure 2. For all three methods the coverage probabilities of the confidence interval for the overall effect are close to the nominal level 0.95, under all different values of τ^2. Under the Gaussian assumptions, the non-parametric EL method is as good as the conventional meta-analysis and the CD method.

In Scenario 2, the individual study's effect size estimate is Gaussian but the random-effect is a log-normal distribution. From the second row of plots in Figure 2, we can see that all three methods, the conventional meta-analysis, the CD approach and the non-parametric EL method give underestimated coverage probabilities, for large values of τ^2 (with values $1, 10, 100$) and small values of K. When $K = 20$, the larger τ^2 is, the worse performance for the conventional meta-analysis and the CD method. This is because the Gaussian assumption for the individual study's effect size is invalid and the error is mainly from the non-Gaussian random-effects. When τ^2 becomes larger, the error of using a Gaussian distribution to approximate a log-normal distribution becomes larger, especially when K is small. The performance of all three methods improves as K increases. For example. when $K = 100$, all three methods provide reasonably well coverage probability for overall effect estimate.

The proposed non-parametric EL method exhibits a notable and significant advantage over the other methods in Scenario 3, where the study effect size is a χ^2 distribution but the random-effect is a Gaussian distribution. The results are presented in the third row of plots in Figure 2. The coverage probabilities of confidence intervals derived from both the conventional and CD methods exhibit a rapid decline as the number of studies (K) increases, especially when the true τ^2 value is small. This decline implies a notable departure from the anticipated large sample properties for these methods.

In cases where τ^2 is small, it is understood that the primary source of error originates from the within-study approximation. Given that the true distribution for each study in Scenario 3 follows a χ^2 distribution, the within-study error, under the Gaussian assumption, becomes considerably inflated. Consequently, the conventional and CD methods manifest ineffectiveness. However, as the τ^2 value increases, errors in the overall effect estimation stem mainly from the random effects. This scenario proves beneficial for the conventional and CD methods, as the Gaussian approximation aligns more closely with the true distribution of θ_i, resulting in smaller between-study approximation errors. In contrast, the non-parametric EL method is more robust and perform better for both small and large values of τ^2. Just as with the conventional and CD methods, there is a potential for convergence challenges within the non-parametric EL method. Notably, as K increases, the convergence probability fails to align with the true nominal value of 0.95. This deviation can be attributed to

the insufficiency of the sample size (n_i) within each study, as governed by Assumption 4. Further exploration of the effect of augmenting the sample size (n_i) will be undertaken in the subsequent subsection, focusing on Scenario 4.

4.2 Scenario 4 and results based on different n_i

We now study the performance of different methods under Scenario 4, that both the individual study effect size and random-effects are non-Gaussian. The top-row plots within Figure 3 present results for relatively small n_i values, drawn from $0.2 \cdot \mathcal{U}[100, 500]$. It's evident that neither the conventional meta-analysis method nor the CD method performs well under small τ^2 values. Although the non-parametric EL method exhibits an improved coverage probability, it still falls considerably short of the target nominal value of 0.95. This performance gap can be attributed to the relatively modest n_i values presented in this scenario. Significantly larger values of n_i are needed to guarantee the convergence of the log-likelihood ratio statistic. The bottom row of plots in Figure 3 shows the performance with large values of n_i, which are random values from $\mathcal{U}[4K, 5K]$. We can see that the coverage probabilities based on the conventional meta-analysis method and the CD method converge very slowly to the true nominal value 0.95, even if we have a large number of n_i and a large number of K. For instance, when $\tau^2 = 0, K = 100$ and n_i is drawn from $\mathcal{U}[400, 500]$, both the conventional method and the CD method give coverage probabilities about 0.4. Remarkably, the non-parametric EL method exhibits significantly improved coverage probabilities, outperforming the conventional and CD methods. Under the same conditions, the non-parametric EL method approximates a coverage probability of about 0.8, converging much faster to the true nominal value.

4.3 The relationship between K and n_i in the Scenario 4

Further, we study the relationship between K and n_i in different methods under Scenario 4 with fixed $\tau^2 = 0.01$. The simulation results are show in Figure 4. When K is fixed, the coverage probability of each method will converge to the true nominal value as n_i increases, but our EL method performs better and robustly especially in the small n_i. When n_i is fixed, the coverage probability of each method will not converge as K increases since Assumption 4 is not satisfied.

4.4 Convergence of the non-parametric EL estimator

Assumption 4 was not well addressed in the meta-analysis literature, but it was the key condition to guarantee the consistency of the meta-analysis estimate. Here we further use
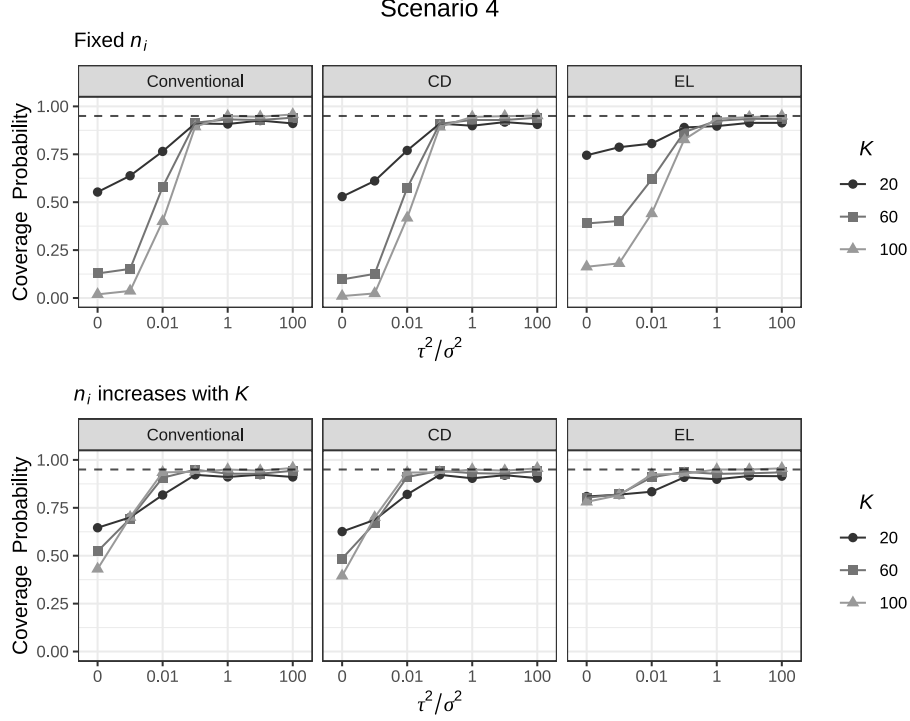
Figure 3: The coverage probabilities of $1 - \beta = 95\%$ confidence intervals which are constructed via conventional meta-analysis method, CD method and EL method respectively under the setting of different K and τ^2 in Scenario 4. In each setting the n_i is sampled from $0.2 \cdot \mathcal{U}[100, 500]$ (top panel) or sampled from $\mathcal{U}[4K, 5K]$ (bottom panel).

a numerical study to demonstrate the convergence of the distribution of the log-likelihood function at the true parameter, under different settings of K and n_i, i.e. numerical justification of Theorem 5. For brevity, here we only consider the convergence of different methods mentioned in the fixed-effect model. We use QQ plots to empirically examine the accuracy of the $\chi^2(r)$ approximation to the distribution of the log-likelihood function at the true parameter, $-2 \log \mathcal{R}(\theta_0)$.

Taking $F(\cdot \mid \theta_i) = \chi^2(4)$, $\theta_i = \theta_0 = 4$ for the fixed-effect meta-analysis model, and $n_1 = n_2 = \cdots = n_K = n$, we compare the sample quantiles obtained from 1000 replicated values of $-2 \log \mathcal{R}(\theta_0)$ under different methods. The theoretical asymptotic distributions of $-2 \log \mathcal{R}(\theta_0)$ under EL1, EL2 and EL3 are $\chi^2(r)$ distributions, with $r = 1$ for EL1 and EL2, and $r = 2$ for EL3. The sample quantiles are plotted in Figure 5. The proximity of sample quantiles to the 45-degree straight line in QQ plots indicates the degree of resemblance between the distribution of these samples and the $\chi^2(r)$ distribution. From these QQ plots we can conclude that, when $K/n = o(1)$, except for extreme upper quantiles, the

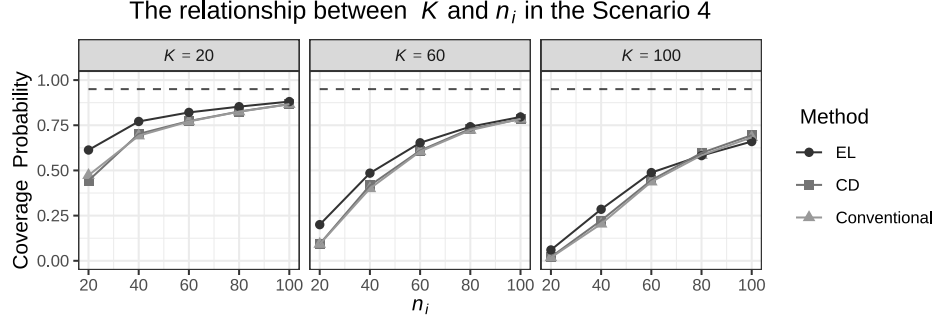The relationship between K and n_i in the Scenario 4

Figure 4: The coverage probabilities of $1 - \beta = 95\%$ confidence intervals which are constructed via conventional meta-analysis method, CD method and EL method respectively under the setting of different K and n_i in Scenario 4. In each setting the n_i is the same for each i and $\tau^2 = 0.01$.

distributions of $-2 \log \mathcal{R}(\theta_0)$ under different methods are reasonably well approximated by a $\chi^2(r)$ distribution. A specific instance can be observed in the top-right panel of Figure 5, where $n = 800$ and $K = 200$. Here, the alignment between sample quantiles and theoretical quantiles is strikingly close. When K/n increases, especially when $K = 800$ and $n = 200$, the convergence of $-2 \log \mathcal{R}(\theta_0)$ to $\chi^2(r)$ diminishes. It is shown by the left-bottom plot of Figure 5, that $-2 \log \mathcal{R}(\theta_0)$ does not converge to $\chi^2(r)$, since Assumption 4 may be violated. When $n = K$ (the top-left plot and the bottom-right plot), $-2 \log \mathcal{R}(\theta_0)$ seems still to converge to $\chi^2(r)$, but its performance is not as good as the result with $n = 800$ and $K = 200$.

5   Application

We apply our proposed methods to two real examples. Example 1 has a large number of studies ($K = 74$), whilst Example 2 has a small number of studies ($K = 6$). The assumption for K to be bounded in the existing methodologies may not be valid for a large value K such as in Example 1, since the accumulated errors from each Gaussian approximation within each study may be too large. For a small number of K such as in Example 2, the Gaussian assumption for θ_i in the conventional meta-analysis methods may not be valid.

5.1   Example 1: Mortality in patients with non-COVID illnesses during and before pandemic

During the COVID-19 pandemic, healthcare providers struggled to manage and deliver quality health care services. Quantifying the effect of COVID-19 pandemic on patients
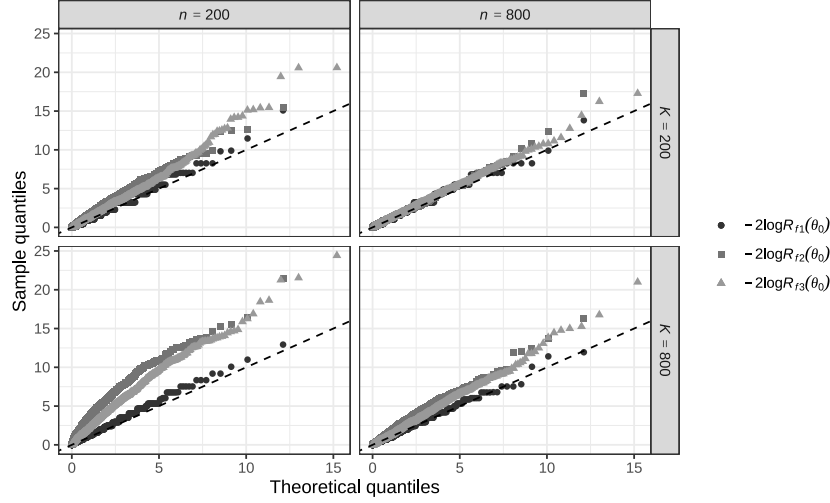
Figure 5: A QQ plot with sample quantiles obtained from 1000 replicated values of $-2 \log \mathcal{R}(\theta_0)$ under different methods (EL1, EL2 and EL3) and the corresponding the quantiles of the $\chi^2(r)$ distribution ($r = 1$ for EL1 and EL2, $r = 2$ for EL3) under different settings: $(n, K) = (200, 200), (200, 800), (800, 200), (800, 800)$. The 45-degree line is included in each plot.

with non-COVID illnesses and on healthcare systems compared to pre-pandemic times can help to understand the impact and disruption of pandemic on health care. A systematic review and meta-analysis of mortality in patients with non-COVID illness during and before the COVID-19 pandemic was conducted (Lau *and others*, 2022), which presented a meta-analysis (74 observational studies with 491,862 patients), and showed an increased mortality rate during COVID-19 pandemic compared with a pre-pandenmic period for patients with non-COVID illnesses. They studied the log-risk ratio (RR), where

$$\text{risk ratio} = \frac{\text{risk during the pandemic}}{\text{risk before the pandemic}},$$

where risk is defined as the ratio of the number of death and the number of patients; and the log-odds ratio (OR), where

$$\text{odds ratio} = \frac{\text{odds during the pandemic}}{\text{odds before the pandemic}},$$

where the odds is the ratio of the number of death and the number of alive.

Table 1 shows the confidence intervals for the estimated overall effect. We also used the the CD function method and our proposed non-parametric empirical likelihood method. The pandemic is associated with an increased risk of mortality, with consistent results across

the conventional meta-analysis, CD and EL methods as well as fixed-effect and random-effects meta-analysis. Our EL approach resulted in larger CIs (RR: 1.206 $(1.104, 1.469)$, OR: 1.537, $(1.293, 2.070)$), which is expected as our method is non-parametric.

Due to the heterogeneity of different studies, a random-effects model is more realistic. Table 1 shows that, when the CD method is used, the risk ratio is 1.384 (95%CI: 1.278 to 1.500) from the random-effects model, and is attenuated to 1.075 (95%CI: 1.068 to 1.082) when we used a fixed-effect model. However, when using our EL method, the risk ratios are closer to each other between the random-effects and fixed-effect meta-analyses. In particular, the non-parametric approach with empirical likelihood $\mathcal{L}_{f2}$ in (12) can be used in both random-effects model and fixed-effect model. The results are similar despite under different model assumptions. We observed similar patterns for OR in this example.

Table 1: Example 1. The combined confidence interval under each method based on the confidence interval of log-risk ratio or log-odds ratio in each study reported in Lau *and others* (2022). Combined confidence intervals were estimated by the conventional meta-analysis, CD method, EL methods in random-effects model and fixed-effect model. The conventional meta-analysis was fitted by using 'rma' in the 'metafor' **R**, and for the conventional random-effect meta-analysis we used REML method to estimate the between-study variance τ^2.

	Fixed-effect meta-analysis			Random-effects meta-analysis		
	Method	Est.	95% CI	Method	Est.	95% CI
	Conventional	1.075	$(1.068, 1.082)$	Conventional	1.455	$(1.270, 1.668)$
	CD	1.075	$(1.068, 1.082)$	CD	1.384	$(1.278, 1.500)$
Risk Ratio	EL1	1.166	$(1.086, 1.443)$	EL	1.206	$(1.104, 1.469)$
	EL2	1.206	$(1.104, 1.469)$			
	EL3	1.175	$(1.117, 1.447)$			
	Conventional	1.318	$(1.284, 1.353)$	Conventional	1.647	$(1.407, 1.929)$
	CD	1.318	$(1.284, 1.353)$	CD	1.650	$(1.401, 1.944)$
Odds Ratio	EL1	1.397	$(1.385, 1.547)$	EL	1.537	$(1.293, 2.070)$
	EL2	1.537	$(1.293, 2.070)$			
	EL3	1.423	$(1.385, 1.577)$			

5.2 Example 2: Stillbirths in pregnancy people with and without COVID-19 vaccination

During the pandemic, COVID-19 vaccination was developed and rolled out at an unprecedented scale and speed addressing the global public health emergency. However, COVID-19 vaccination hesitancy has been identified as an issue. Apprehension surrounding pregnancy was identified as one contributing source of vaccination hesitancy (Bhattacharya *and others*, 2022; Golder *and others*, 2023; Razai *and others*, 2021). Evidence of the effectiveness of COVID-19 vaccination in pregnant women can help to improve the uptake of the vaccination in this vulnerable group. Stillbirth is an important perinatal outcome, which is the death or loss of a baby before or during delivery. The impact of COVID-19 mRNA vaccination on stillbirth was evaluated in a systematic review (Prasad *and others*, 2022) and a meta-analysis of six studies (66,067 vaccinated v.s. 424,624 unvaccinated pregnant people).

Table 2 showed that results are inconsistent across methods. From the results obtained from the CD method, COVID-19 vaccination is associated with lower risk of stillbirths. However, in other methods (conventional random-effects meta-analysis, EL random-effects meta-analysis and EL fixed-effect meta-analysis), we found no evidence for the association of COVID-19 vaccination with stillbirths. In summary, the CD method implies that the COVID-19 vaccination is associated with a reduced risk of still-birth. However, the results from meta-analysis using the non-parametric EL method and the conventional meta-analysis method cannot confirm this.

6   Discussion

This paper introduces a novel non-parametric method for conducting meta-analysis of confidence intervals. The approach is rooted in empirical likelihood, applicable to both fixed-effect and random-effects meta-analysis models. Notably, the non-parametric EL method departs from the conventional Gaussian assumption often relied upon, while still yielding large sample theoretical outcomes. Our contribution provides the solutions in addressing the requirements regarding the number of studies (K) and the sample size (n_i) within each study, to ensure the consistency and asymptotic normality of the overall effect estimate. This theoretical result fills a crucial gap that has not been addressed in the current landscape of meta-analysis literature.

The distinguish feature of the non-parametric EL approach is in its reliance on an asymptotically unbiased estimating equation for the unknown parameter, derived from the confidence intervals obtained from individual studies. As long as the function $\boldsymbol{W}_i(\theta)$ within the estimating equation $\boldsymbol{W}_i(\theta) = \boldsymbol{0}$ remains asymptotically unbiased at a rate

Table 2: Example 2. Stillbirths in pregnancy people with and without COVID-19 vaccination. confidence intervals were estimated by using the conventional random-effects and fixed-effect meta-analysis, the CD method and different non-parametric empirical likelihood methods (EL), based on the confidence interval of log-risk ratio or log-odds ratio in each study reported in Prasad *and others* (2022).

	Fixed-effect meta-analysis			Random-effects meta-analysis		
	Method	Est.	95% CI	Method	Est.	95% CI
	Conventional	0.859	$(0.739, 0.997)$	Conventional	0.859	$(0.737, 1.001)$
	CD	0.859	$(0.739, 0.997)$	CD	0.859	$(0.739, 0.997)$
Risk Ratio	EL1	0.644	$(0.644, 1.163)$	EL	0.867	$(0.762, 1.030)$
	EL2	0.867	$(0.762, 1.030)$			
	EL3	0.854	$(0.736, 1.080)$			
	Conventional	0.858	$(0.738, 0.997)$	Conventional	0.859	$(0.736, 1.001)$
	CD	0.858	$(0.738, 0.997)$	CD	0.858	$(0.738, 0.997)$
Odds Ratio	EL1	0.643	$(0.643, 1.163)$	EL	0.867	$(0.761, 1.030)$
	EL2	0.867	$(0.761, 1.030)$			
	EL3	0.856	$(0.735, 1.080)$			

$\mathbb{E} \boldsymbol{W}_i(\theta_0) = O(n_i^{-1/2})$ (as per Assumption 1), the non-parametric EL approach guarantees a consistent estimate. This assumption, being quite mild, holds true for a wide array of existing statistical estimators, such as the maximum likelihood estimator derived from score functions, the estimating equations in Cox partial likelihood for survival analysis, and the generalized estimating equation approach.

In our illustrative examples, our non-parametric EL method produced wider 95%CI for the overall effect. In example 2, the confidence intervals estimated from the EL method are consistent across methods, in contrast, confidence intervals estimated by CD methods or the conventional meta-analysis are not consistent between the fixed-effect random-effect models. In particular, it reveals distinctive insights that diverge from existing meta-analysis techniques. In Example 2, the non-parametric EL method leads to different conclusion - there is insufficient evidence that COVID-19 vaccination is associated with the risk of stillbirths. This exemplifies the effectiveness of the proposed approach in addressing the underlying challenge in meta-analysis, particularly when individual study estimates violate the Gaussian assumptions.

References

BAKER, ROSE AND JASON, DAN. (2008). A new approach to outliers in meta-analysis. *Health Care Management Science* **11**, 121–131.

BHATTACHARYA, OASHE, SIDDIQUEA1, BODRUN NAHER, SHETTY, AISHWARYA, AFROZ, AFSANA AND BILLAH, BAKI. (2022). Covid-19 vaccine hesitancy among pregnant women: a systematic review and meta-analysis. *BMJ Open*.

BROCKWELL, S.E. AND GORDON, I.R. (2001). A comparison of statistical methods for meta-analysis. *Statistics in Medicine* **20**(16), 825–840.

BROCKWELL, SARAH E. AND GORDON, IAN R. (2007). A simple method for inference on an overall effect in meta-analysis. *Statistics in Medicine* **26**(25), 4531–4543.

CAI, CHENCHENG, CHEN, RONG AND XIE, MIN-GE. (2023). Individualized group learning. *Journal of the American Statistical Association* **118**(541), 622–638.

DEEKS, JONATHAN J., RILEY, RICHARD D. AND HIGGINS, JULIAN P.T. (2022). *Combining Results Using Meta-Analysis*, Chapter 9. John Wiley & Sons, Ltd, pp. 159–184.

DERSIMONIAN, R. AND LAIRD, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials* **7**(3), 177–188.

GOLDER, S., MCROBBIE-JOHNSON, A. C. E., KLEIN, A., POLITE, F. G. AND G., GONZALEZ HERNANDEZ. (2023). Social media and covid-19 vaccination hesitancy during pregnancy: a mixed methods analysis. *Scandinavian Journal of Statistics* **29**, 309–332.

HARDY, R.J. AND THOMPSON, S.R. (1996). A likehood approach to meta-analysis with random effects. *Statistics in Medicine* **15**(6), 619–629.

HIGGINS, J.P.T. AND THOMPSON, S.G. (2009). A re-evaluation of random-effects meta-analysis. *Journal of Royal Statistical Society Series A* **172**(1), 137–159.

JACKSON, DAN AND BOWDEN, JACK. (2009). A re-evaluation of the 'quantile approximation method' for random effects meta-analysis. *Statistics in Medicine* **28**(2), 338–348.

LANGAN, DEAN, HIGGINS, JULIAN P.T., JACKSON, DAN, BOWDEN, JACK, VERONIKI, ARETI ANGELIKI, KONTOPANTELIS, EVANGELOS, VIECHTBAUER, WOLFGANG AND SIMMONDS, MARK. (2019). A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Research Synthesis Methods* **10**(1), 83–98.

LAU, VINCENT ISSAC, DHANOA, SUMEET, CHEEMA, HARLEEN, LEWIS, KIMBERLEY, GEERAERT, PATRICK, LU, DAVID, MERRICK, BENJAMIN, VANDER LEEK, AARON, SEBASTIANSKI, MEGHAN, KULA, BRITTANY, CHAUDHURI, DIPAYAN, AGARWAL, ARNAV, NIVEN, DANIEL J., FIEST, KIRSTEN M., STELFOX, HENRY T., ZUEGE, DANNY J., REWA, OLEKSA G. *and others*. (2022, June). Non-COVID outcomes associated with the coronavirus disease-2019 (COVID-19) pandemic effects study (COPES): A systematic review and meta-analysis. *PLOS ONE* **17**(6), e0269871.

LEE, KATHERINE J. AND THOMPSON, SIMON G. (2008). Flexible parametric models for random-effects distributions. *Statistics in Medicine* **27**(3), 418–434.

NAGASHIMA, KENGO, NOMA, HISASHI AND FURUKAWA, TOSHI A. (2019). Prediction intervals for random-effects meta-analysis: A confidence distribution approach. *Statistical Methods in Medical Research* **28**(6), 1689–1702.

NORMAND, S.L. (1999). Meta-analysis: formulating, evaluating, combining, and reporting. *Statistics in Medicine* **18**(3), 321–359.

OZTURK, OMER AND BALAKRISHNAN, NARAYANASWAMY. (2020). Meta-analysis of quantile intervals from different studies with an application to a pulmonary tuberculosis data. *Statistics in Medicine* **39**(29), 4519–4537.

PARTLETT, CHRISTOPHER AND RILEY, RICHARD D. (2017). Random effects meta-analysis: Coverage performance of 95% confidence and prediction intervals following reml estimation. *Statistics in Medicine* **36**(2), 301–317.

PRASAD, SMRITI, KALAFAT, ERKAN, BLAKEWAY, HELENA, TOWNSEND, ROSEMARY, O'BRIEN, PAT, MORRIS, EDWARD, DRAYCOTT, TIM, THANGARATINAM, SHAKILA, LE DOARE, KIRSTY, LADHANI, SHAMEZ, VON DADELSZEN, PETER, MAGEE, LAURA A., HEATH, PAUL *and others*. (2022, May). Systematic review and meta-analysis of the effectiveness and perinatal outcomes of COVID-19 vaccination in pregnancy. *Nature Communications* **13**(1), 2414.

RAZAI, MOHAMMAD S, CHAUDHRY, UMAR A R, DOERHOLT, KATJA, BAULD, LINDA AND MAJEED, AZEEM. (2021). Covid-19 vaccination hesitancy. *BMJ* **373**.

SCHWEDER, T. AND HJORT, N. L. (2002). Confidence and likelihood. *Scandinavian Journal of Statistics* **29**, 309–332.

SIDIK, K. AND JONKMAN, J. N. (2002). A simple confidence interval for meta-analysis. *Statistics in Medicine* **21**, 3153–3159.

Singh, K., Xie, M. and Strawderman, W.E. (2005). Combining information from independent sources through confidence distributions. *The Annals of Statistics* **33**(1).

Sutton, Alexander J. and Higgins, Julian P. T. (2008). Recent developments in meta-analysis. *Statistics in Medicine* **27**(5), 625–650.

Veroniki, A. A., Jackson, D., Bender, R., Kuss, O., Langan, D., Higgins, J.P.T., Knapp, G. and Salanti, G. (2018). Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. *Research Synthesis Methods* **10**(1), 23–43.

Xie, M., Singh, K. and Strawderman, W.E. (2011). Confidence distributions and a unifying framework for meta-analysis. *Journal of the American Statistical Association* **106**(493), 320–333.