

An alternative pseudolikelihood method for multivariate random-effects meta-analysis

Yong Chen,^{a,*†} Chuan Hong^a and Richard D. Riley^b

Recently, multivariate random-effects meta-analysis models have received a great deal of attention, despite its greater complexity compared to univariate meta-analyses. One of its advantages is its ability to account for the within-study and between-study correlations. However, the standard inference procedures, such as the maximum likelihood or maximum restricted likelihood inference, require the within-study correlations, which are usually unavailable. In addition, the standard inference procedures suffer from the problem of singular estimated covariance matrix. In this paper, we propose a pseudolikelihood method to overcome the aforementioned problems. The pseudolikelihood method does not require within-study correlations and is not prone to singular covariance matrix problem. In addition, it can properly estimate the covariance between pooled estimates for different outcomes, which enables valid inference on functions of pooled estimates, and can be applied to meta-analysis where some studies have outcomes missing completely at random. Simulation studies show that the pseudolikelihood method provides unbiased estimates for functions of pooled estimates, well-estimated standard errors, and confidence intervals with good coverage probability. Furthermore, the pseudolikelihood method is found to maintain high relative efficiency compared to that of the standard inferences with known within-study correlations. We illustrate the proposed method through three meta-analyses for comparison of prostate cancer treatment, for the association between paraoxonase 1 activities and coronary heart disease, and for the association between homocysteine level and coronary heart disease. © 2014 The Authors. *Statistics in Medicine* Published by John Wiley & Sons Ltd.

Keywords: composite likelihood; correlation; multivariate meta-analysis; singular estimated covariance matrix problem; pseudolikelihood

1. Introduction

The rapid growth of evidence-based medicine has led to dramatically increasing attention to meta-analysis, which combines statistical evidence from multiple studies. In many randomized clinical trials or observational studies, multiple and possibly correlated outcomes of interest need to be meta-analyzed. Conventionally, univariate methods are used to investigate one outcome at a time, such as pooling the summary measures from all studies through fixed or random effects model [1]. These methods gain their popularity because of their usefulness and simplicity. However, if the objective of inference is to compare the overall effects between types of outcomes, such as comparing the risks of a certain disease between two groups, calculating the weighted sums of the estimated sensitivity and specificity in diagnostic tests, or comparing the effect sizes with respect to two endpoints, univariate methods are not sufficient because they cannot account for the possible correlations between multiple outcomes [2–4].

Recently, multivariate meta-analysis (MMA) has received a great deal of attention [5]. MMA jointly analyzes multiple and possibly correlated outcomes, which can account for two types of correlations:

^aDivision of Biostatistics, University of Texas School of Public Health, 1200 Pressler St, Houston, Texas 77030, U.S.A.

^bSchool of Health and Population Science, and School of Mathematics, University of Birmingham, Birmingham, B15 2TT, U.K.

*Correspondence to: Yong Chen, Division of Biostatistics, University of Texas School of Public Health, 1200 Pressler St, Houston, Texas 77030, U.S.A.

†E-mail: yong.chen@uth.tmc.edu

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

the within-study correlation and the between-study correlation. The within-study correlation ρ_{W_i} exists because the effects are sometimes measured using the same set of subjects, such as the overall and disease-free survival of cancer patients. The between-study correlation ρ_B allows the underlying effects to be correlated. In MMA, a ‘two-stage’ approach for inference is often adopted. At the first stage, analyses of each study are performed, and estimates of effects, referred to as summary measures, are obtained. These estimates are then combined at the second stage by a random effects model. The overall effect sizes or their comparative measures can be inferred using maximum likelihood (ML) or restricted ML (REML) inference based on the marginal distribution of summary measures. For an excellent review of MMA, see Van Houwelingen *et al.* (2002) [2] and Jackson *et al.* (2011) [5].

Although conceptually straightforward, conducting MMA faces at least two practical challenges, even when only two outcomes are considered. The first challenge is the lack of knowledge on the within-study correlations ρ_{W_i} , which are generally required in standard inference procedures for MMA. However, in practice, the within-study correlations are often not reported and are difficult to obtain even on request [6, 7]. In addition, calculation of the within-study correlation may not be easy and sometimes requires more computationally intensive methods [8]. Such challenge is acknowledged by the excellent review paper of Jackson *et al.* [5], ‘perhaps the greatest difficulty applying the MMA model in practice is that the within-study correlations are required by the model and are typically unknown.’ In such situations, sensitivity analysis with imputed within-study correlations and Bayesian methods have been proposed [9, 10]. Wei and Higgins [11] recently proposed a practical method for MMA when the within-study correlations are missing. Specifically, they used the information on possible correlations between the underlying outcomes to impute any missing within-study covariances, and then conducted the inference by REML estimation. The second challenge is the (restricted) ML estimate of the between-study correlation ρ_B is often at or very close to the boundary of its parameter space, that is, ± 1 . In that case, the estimated covariance matrix is singular. The singular estimated covariance matrix can lead to biased estimates of standard errors and is most severe when the number of studies is small or the within-study variation is relatively large [5, 12]. As a consequence, the confidence intervals may be too wide or too narrow.

To overcome the aforementioned challenges, Riley *et al.* [13] proposed a method for bivariate meta-analysis (BMA). Specifically, Riley *et al.* [13] postulated a synthesis correlation parameter to describe the overall marginal correlation between outcomes. This method is attractive because it does not require within-study correlations and can alleviate the singular estimated covariance matrix problem. However, the prevalence of singular estimated covariance matrix problem can still be substantial in some situations, and the extension from BMA to MMA is not straightforward in order to deal with more than two outcomes. In this paper, we propose an alternative method along the line of the work by Riley *et al.* [13]. The idea is to construct a pseudolikelihood function for overall effect sizes using a *working* independence assumption. An immediate advantage of the proposed method is the within-study correlations are not required in the construction of pseudolikelihood. The singular estimated covariance matrix problem is resolved because there is no correlation parameter involved in the pseudolikelihood. In fact, the proposed pseudolikelihood belongs to the family of composite likelihoods [14–16]. Hence, the pseudolikelihood enjoys the established properties of composite likelihood [14, 16–18]. In particular, when a working independence assumption is adopted, the pseudolikelihood is called independence likelihood [19], where the covariance between estimates of overall effect sizes can be consistently estimated by the Huber–White standard error estimates, also known as ‘sandwich’ variance estimator [20, 21]. Another advantage of the pseudolikelihood method is the simplicity of the extension to MMA where more than two outcomes are analyzed, and to missing data situations where some of the multiple outcomes are missing completely at random (MCAR). In this paper, we present the proposed method in the meta-analysis of bivariate outcomes and describe the extension to three or more outcomes that is provided in Section 2.4.

This paper is organized as follows. In Section 2, we describe the standard likelihood inferences for MMA, the method proposed by Riley *et al.* [13] and the proposed pseudolikelihood method. In Section 3, we extend the proposed method to the missing data situation where only a proportion of studies have all outcomes reported, and so remaining studies have at least one outcome missing. In Section 4, we conduct simulation studies to compare the pseudolikelihood method with the current methods and investigate the prevalence of singular covariance matrix problems, bias, coverage probability (CP), and relative efficiency (RE). We apply the proposed method to three meta-analyses in Section 5. Finally, we provide a brief discussion in Section 6.

2. Statistical methodology

2.1. Bivariate random-effects meta-analysis model

We consider a meta-analysis with m studies where two outcomes in each study are of interest. For the i th study, denote Y_{ij} and s_{ij} the summary measure for the j th outcome of interest and associated standard error, respectively, both assumed known, $i = 1, \dots, m$, and $j = 1, 2$. Each summary measure Y_{ij} is an estimate of the true effect size θ_{ij} . To account for heterogeneity in effect size across studies, we assume θ_{ij} to be independently drawn from a common distribution with overall effect size β_j and between study variance τ_j^2 , $j = 1, 2$. Under normal distribution assumption for Y_{ij} and θ_{ij} , the general bivariate random-effects meta-analysis model can be written as [2]

$$\begin{aligned} \begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} &\sim N \left(\begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix}, \Delta_i \right), & \Delta_i &= \begin{pmatrix} s_{i1}^2 & s_{i1}s_{i2}\rho_{wi} \\ s_{i1}s_{i2}\rho_{wi} & s_{i2}^2 \end{pmatrix}, \\ \begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix} &\sim N \left(\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \Omega \right), & \Omega &= \begin{pmatrix} \tau_1^2 & \tau_1\tau_2\rho_B \\ \tau_1\tau_2\rho_B & \tau_2^2 \end{pmatrix}, \end{aligned} \quad (1)$$

where Δ_i and Ω are the respective within-study and between-study covariance matrices, and ρ_{wi} and ρ_B are the respective within-study and between-study correlations. When the within-study correlations ρ_{wi} are known, inference on the overall effect sizes β_1 and β_2 , or their comparative measures (e.g. $\beta_1 - \beta_2$), can be based on the marginal distribution of (Y_{i1}, Y_{i2})

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \sim N \left(\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \mathbf{V}_i \right), \mathbf{V}_i = \Delta_i + \Omega = \begin{pmatrix} s_{i1}^2 + \tau_1^2 & s_{i1}s_{i2}\rho_{wi} + \tau_1\tau_2\rho_B \\ s_{i1}s_{i2}\rho_{wi} + \tau_1\tau_2\rho_B & s_{i2}^2 + \tau_2^2 \end{pmatrix}.$$

We note that the variance of Y_{ij} is partitioned into two parts s_{ij}^2 and τ_j^2 as in analysis of variance for univariate random effects model, and the covariance, $\text{cov}(Y_{i1}, Y_{i2}) = s_{i1}s_{i2}\rho_{wi} + \tau_1\tau_2\rho_B$, is also partitioned into two parts as the sum of within and between study covariances.

2.2. Restricted maximum likelihood method

For simplicity of notation, denote $\mathbf{Y}_i = (Y_{i1}, Y_{i2})^T$, $\boldsymbol{\beta} = (\beta_1, \beta_2)^T$, $\boldsymbol{\eta}_1 = (\beta_1, \tau_1^2)^T$, and $\boldsymbol{\eta}_2 = (\beta_2, \tau_2^2)^T$. The restricted likelihood of $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \rho_B)$ can be written as

$$\begin{aligned} &\log L(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \rho_B) \\ &= -\frac{1}{2} \left[\log \left(\left| \sum_{i=1}^m \mathbf{V}_i^{-1} \right| \right) + \sum_{i=1}^m \left\{ \log |\mathbf{V}_i| + (\mathbf{Y}_i - \boldsymbol{\beta})^T \mathbf{V}_i^{-1} (\mathbf{Y}_i - \boldsymbol{\beta}) \right\} \right]. \end{aligned}$$

The parameters $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \rho_B)$ can be estimated by the REML approach as described in Van Houwelingen *et al.* [2]. The between-study variances (τ_1^2, τ_2^2) are usually modeled in their log-scale, so that they are forced to be non-negative. REML can be implemented using Newton–Raphson or quasi-Newton methods, as in ‘mvmeta’ package in STATA [22] or R [23]. Such methods of implementing REML rarely suffer convergence issues. However, two practical challenges in the standard likelihood inference have been reported [5, 13]. The first is the lack of knowledge on ρ_{wi} . The standard likelihood inference based on the bivariate random-effects meta-analysis requires within-study correlation estimates ρ_{wi} , but these are rarely available. The second is the singular estimated covariance matrix problem [13, 24]. Specifically, a singular covariance matrix is sometimes obtained (i.e., the estimate of the between-study correlation ρ_B is close to ± 1). Consequently, estimates of standard errors are often biased and can lead to confidence intervals that are too conservative or too liberal. As will be illustrated in Section 4, the singular estimated covariance matrix problem can be substantial (e.g., greater than 25% of 1000 samples for meta-analysis with 10 studies) when REML approach is applied. Such a problem is more severe when the ML approach is taken.

2.3. Riley method

To account for the aforementioned practical issues of REML method, Riley *et al.* [13] proposed a new method, which we refer as Riley method hereafter. Specifically, instead of partition, the overall covariance

into within and between study covariances, Riley *et al.* (2008) [13] proposed a synthesis correlation parameter ρ_S to account for the *marginal* correlation between Y_{i1} and Y_{i2} . Following the notations in Riley *et al.* [13], the new marginal model for (Y_{i1}, Y_{i2}) is

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \sim N \left(\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \Phi_i \right),$$

$$\Phi_i = \begin{pmatrix} \psi_1^2 + s_{i1}^2 & \rho_S \sqrt{(\psi_1^2 + s_{i1}^2)(\psi_2^2 + s_{i2}^2)} \\ \rho_S \sqrt{(\psi_1^2 + s_{i1}^2)(\psi_2^2 + s_{i2}^2)} & \psi_2^2 + s_{i2}^2 \end{pmatrix}, \quad (2)$$

where ψ_j^2 's account for the additional variation beyond the within-study variances s_{ij}^2 , $j = 1, 2$, and ρ_S accounts for the *marginal* correlation between Y_{i1} and Y_{i2} . Note that ψ_j^2 is not equivalent to the between-study variation τ_j^2 because the model (2) does not have a fully hierarchical structure. Riley *et al.* [13] proposed to base the inference on the restricted log-likelihood defined as

$$\begin{aligned} & \log L_{\text{Riley}}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \rho_S) \\ &= -\frac{1}{2} \left[\log \left(\left| \sum_{i=1}^m \Phi_i^{-1} \right| \right) + \sum_{i=1}^m \left\{ \log |\Phi_i| + (\mathbf{Y}_i - \boldsymbol{\beta})^T \Phi_i^{-1} (\mathbf{Y}_i - \boldsymbol{\beta}) \right\} \right]. \end{aligned}$$

The method based on the model (2) has the attractive features of not requiring knowledge on within-study correlations ρ_{Wi} and being able to account for the correlation between Y_{i1} and Y_{i2} . In addition, Riley method was found to be less prone to the estimation problem compared to standard likelihood inferences based on REML [13]. However, as acknowledged by Riley *et al.* [13], the problem of unstable estimates when estimate of ρ_S is close to 1 or -1 is still an issue especially when the between-study heterogeneity is small relative to the within-study variation. In addition, the extension from BMA to MMA with more than two outcomes is not straightforward because multiple synthesis correlation parameters may be required to describe the pairwise marginal correlations between pairs of outcomes. In fact, the 'mvmeta' package in STATA by Ian White has recently extended Riley method to multivariate data with more than two outcomes [22]. However, it is expected that increasing the number of outcomes is likely to increase the chance of not well-defined correlation estimates because of the need of estimating increased number of correlation parameters.

2.4. Pseudo-restricted maximum likelihood method

Now we propose an alternative method to overcome the aforementioned challenges, which we refer as Pseudo-REML method hereafter. Our strategy is to base the inference of $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$ on a pseudolikelihood function constructed as follows. Under the *working* independence assumption (i.e., setting $\rho_{Wi} = \rho_B = 0$ in the joint distribution of Y_{i1} and Y_{i2}), we obtain the following pseudolikelihood

$$\log L_p(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = \log L_1(\boldsymbol{\eta}_1) + \log L_2(\boldsymbol{\eta}_2), \quad (3)$$

where

$$\log L_j(\boldsymbol{\eta}_j) = -\frac{1}{2} \sum_{i=1}^m \left\{ \log \left(s_{ij}^2 + \tau_j^2 \right) + \frac{(Y_{ij} - \beta_j)^2}{s_{ij}^2 + \tau_j^2} \right\} \quad \text{for } j = 1, 2. \quad (4)$$

We note that $\log L_j(\boldsymbol{\eta}_j)$ is simply the log likelihood for $\boldsymbol{\eta}_j$ when a univariate random effects model for meta-analysis is assumed. Therefore, the Pseudo-REML method indeed provides the same point and variance estimates as that from the univariate method as far as *only* β_1 (or β_2) is concerned. The main difference between the Pseudo-REML method and the univariate method is that the model-based standard error of the difference between β_1 and β_2 (or any functions of β_1 and β_2) can be correctly estimated by the Pseudo-REML method, but not by two separate univariate meta-analyses due to the ignored covariance between the estimated β_1 and β_2 . In other words, the Pseudo-REML method can be considered as an inferential strategy for REML or Riley method using working independence assumption.

Note that the score equation for η_j , $\partial \log L_p(\eta_1, \eta_2)/\partial \eta_j = 0$ can be calculated as $\partial \log L_j(\eta_j)/\partial \eta_j = 0$, hence is unbiased. The maximum pseudolikelihood estimator $(\tilde{\eta}_1, \tilde{\eta}_2)$, defined as a solution of the score equation, can be shown to be consistent and asymptotically normal with covariance matrix

$$\Sigma = \begin{pmatrix} \mathbf{I}_{11}^{-1} & \mathbf{I}_{11}^{-1} \mathbf{I}_{12} \mathbf{I}_{22}^{-1} \\ (\mathbf{I}_{11}^{-1} \mathbf{I}_{12} \mathbf{I}_{22}^{-1})^T & \mathbf{I}_{22}^{-1} \end{pmatrix},$$

where

$$\mathbf{I}_{jj} = E \left\{ -\frac{\partial^2 \log L_j(\eta_j)}{\partial \eta_j^2} \right\} \quad \text{and} \quad \mathbf{I}_{12} = E \left[\left\{ \frac{\partial \log L_1(\eta_1)}{\partial \eta_1} \right\} \left\{ \frac{\partial \log L_2(\eta_2)}{\partial \eta_2} \right\}^T \right] \quad \text{for } j = 1, 2.$$

The general asymptotic results of composite likelihood have been provided by Kent [17], Lindsay [14], and Molenberghs and Verbeke [18]. For the interest of readers, an outline of derivation for the covariance is provided in the Appendix. We note that the pseudolikelihood $\log L_p(\eta_1, \eta_2)$ is generally not a true likelihood function unless all correlations are truly zero (i.e., setting $\rho_{Wi} = \rho_B = 0$ in the joint distribution of Y_{i1} and Y_{i2}). Consequently, the conventional covariance matrix estimator $\mathcal{I}_p(\tilde{\eta}_1, \tilde{\eta}_2)$, where $\mathcal{I}_p(\eta_1, \eta_2) = -\partial^2 \log L_p(\eta_1, \eta_2)/\partial(\eta_1, \eta_2)^2$, is no longer valid because $E\{\mathcal{I}_p(\eta_1, \eta_2)\}$ is not the covariance of $\partial L_p(\eta_1, \eta_2)/\partial(\eta_1, \eta_2)$ in the presence of correlations.

The information matrices \mathbf{I}_{11} , \mathbf{I}_{22} , and \mathbf{I}_{12} can be empirically estimated as

$$\hat{\mathbf{I}}_{11} = \frac{1}{m} \sum_{i=1}^m \mathbf{U}_{1i}(\tilde{\eta}_1) \mathbf{U}_{1i}(\tilde{\eta}_1)^T, \quad \hat{\mathbf{I}}_{22} = \frac{1}{m} \sum_{i=1}^m \mathbf{U}_{2i}(\tilde{\eta}_2) \mathbf{U}_{2i}(\tilde{\eta}_2)^T$$

$$\hat{\mathbf{I}}_{12} = \frac{1}{m} \sum_{i=1}^m \mathbf{U}_{1i}(\tilde{\eta}_1) \mathbf{U}_{2i}(\tilde{\eta}_2)^T,$$

and

$$\mathbf{U}_{ji}(\tilde{\eta}_j) = \begin{pmatrix} \frac{Y_{ij} - \tilde{\beta}_j}{s_{ij}^2 + \tilde{\tau}_j^2}, -\frac{1}{2(s_{ij}^2 + \tilde{\tau}_j^2)} + \frac{(Y_{ij} - \tilde{\beta}_j)^2}{2(s_{ij}^2 + \tilde{\tau}_j^2)^2} \end{pmatrix}^T, \quad \text{for } j = 1, 2 \quad \text{and } i = 1, \dots, m.$$

The Pseudo-REML method is simple to implement in practice. Computationally, the point estimate of overall effect size $\tilde{\eta}_j$ is the same as that from univariate meta-analysis of (Y_{ij}, s_{ij}^2) using a random effects model based on REML [1]. The covariance matrix of $(\tilde{\eta}_1, \tilde{\eta}_2)$ can be easily calculated using the aforementioned closed-form formulas. For example, if one was interested in the function $\beta_1 - \beta_2$, then one would take the difference in the univariate pooled estimates for outcomes 1 and 2, and the variance of $\beta_1 - \beta_2$ can then be calculated by $a^T \hat{\Sigma} a$ where $a = (1, 0, -1, 0)^T$. $\hat{\Sigma}$ is a 4×4 matrix here, generally with dimensions $2K \times 2K$ and K as the number of outcomes. We implement the Pseudo-REML method in R. We use the ‘mvmeta’ function in the R (R Development Core Team, Version 2.14.1) package ‘mvmeta’ to obtain the REML estimates [23]. R codes are attached in the Supplemental Materials. We note that the matrix $\hat{\mathbf{I}}_{jj}^{-1}/m, j = 1, 2$ is the same as the covariance estimated from the univariate meta-analysis of $(Y_{ij}, s_{ij}^2), j = 1, 2$, whereas the matrix $\hat{\mathbf{I}}_{11}^{-1} \hat{\mathbf{I}}_{12} \hat{\mathbf{I}}_{22}^{-1}/m$ accounts for the covariance between the estimated overall effect sizes $\tilde{\eta}_1$ and $\tilde{\eta}_2$. Properly accounting such covariance is not available if investigators conduct separate univariate meta-analyses on (Y_{ij}, s_{ij}^2) .

It is easy to see that the Pseudo-REML method resolves the two practical issues in standard inference of REML method. Specifically, the within-study correlations ρ_{Wi} are not required in the construction of pseudolikelihood. And the singular estimated covariance matrix problem is resolved because there is no correlation parameter (i.e., ρ_B or ρ_S) involved in the pseudolikelihood. Furthermore, the Pseudo-REML method can be easily extended to MMA where more than two outcomes are analyzed. Specifically, for MMA with K outcomes, the corresponding pseudolikelihood becomes $\log L_p(\eta_1, \dots, \eta_K) = \sum_{k=1}^K \log L_k(\eta_k)$ where $\log L_k(\eta_k)$ is defined in equation (4). The corresponding covariance matrix can be derived similarly. However, one potential drawback of the Pseudo-REML method when assuming the data are complete or MCAR is that because the point estimates of overall effect sizes are the same as

that from separate univariate meta-analyses, there is potential efficiency loss compared to the methods accounting for the correlations between dependent outcomes. We investigate this potential drawback in Section 4.

3. Extension to missing data where only a subset of outcomes is reported

In Section 2, we assumed that all outcomes of each study are observed. In practice, however, it is common that only a proportion of studies have all outcomes reported, and the remaining studies have some of outcomes missing. In this section, we discuss extension of the Riley method and Pseudo-REML method for missing data situations. In this paper, we only consider the situation that the outcomes are MCAR, which has been considered in [13]. Although such assumption often does not hold in practice, it is instructive to consider MCAR setting as a step toward missing at random (MAR) and missing not at random (MNAR) situation. Extension of the Pseudo-REML method to MAR and MNAR is beyond the scope of this paper and is discussed in Section 6.

For simplicity of notations, we assume that two endpoints are of interest. Consider a meta-analysis of m studies where the first m_1 studies reported both endpoints, the next m_2 studies reported the first endpoint only, and the remaining m_3 studies reported the second endpoint only. The restricted log-likelihood for the model proposed by Riley *et al.* [13] in the missing data situation can be written as

$$\begin{aligned} & \log L_{\text{Riley}}^{\text{mis}}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \rho_S) \\ &= -\frac{1}{2} \left[\log \left(\left| \sum_{i=1}^{m_1} \boldsymbol{\Phi}_i^{-1} \right| \right) + \sum_{i=1}^{m_1} \left\{ \log |\boldsymbol{\Phi}_i| + (\mathbf{Y}_i - \boldsymbol{\beta})^T \boldsymbol{\Phi}_i^{-1} (\mathbf{Y}_i - \boldsymbol{\beta}) \right\} \right] \\ & \quad - \frac{1}{2} \left[\log \left(\sum_{i=m_1+1}^{m_1+m_2} \frac{1}{\psi_1^2 + s_{i1}^2} \right) + \sum_{i=m_1+1}^{m_1+m_2} \left\{ \log (\psi_1^2 + s_{i1}^2) + \frac{(Y_{i1} - \beta_1)^2}{\psi_1^2 + s_{i1}^2} \right\} \right] \\ & \quad - \frac{1}{2} \log \left[\left(\sum_{i=m_1+m_2+1}^m \frac{1}{\psi_2^2 + s_{i2}^2} \right) + \sum_{i=m_1+m_2+1}^m \left\{ \log (\psi_2^2 + s_{i2}^2) + \frac{(Y_{i2} - \beta_2)^2}{\psi_2^2 + s_{i2}^2} \right\} \right]. \end{aligned}$$

The pseudolikelihood in the missing data situation can be defined as

$$\log L_p^{\text{mis}}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = \log L_1^{\text{mis}}(\boldsymbol{\eta}_1) + \log L_2^{\text{mis}}(\boldsymbol{\eta}_2), \tag{5}$$

where

$$\begin{aligned} \log L_1^{\text{mis}}(\boldsymbol{\eta}_1) &= -\frac{1}{2} \sum_{i=1}^{m_1+m_2} \left\{ \log (s_{i1}^2 + \tau_1^2) + \frac{(Y_{i1} - \beta_1)^2}{s_{i1}^2 + \tau_1^2} \right\} \\ \log L_2^{\text{mis}}(\boldsymbol{\eta}_2) &= -\frac{1}{2} \left[\sum_{i=1}^{m_1} \left\{ \log (s_{i2}^2 + \tau_2^2) + \frac{(Y_{i2} - \beta_2)^2}{s_{i2}^2 + \tau_2^2} \right\} + \sum_{i=m_1+m_2+1}^{m_1+m_2+m_3} \left\{ \log (s_{i2}^2 + \tau_2^2) + \frac{(Y_{i2} - \beta_2)^2}{s_{i2}^2 + \tau_2^2} \right\} \right]. \end{aligned}$$

We note that when the full data are available, the parameter estimators reduce to those in Section 2.4. The score equation for $\log L_{\text{Riley}}^{\text{mis}}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \rho_S)$ and $\log L_p^{\text{mis}}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$ is unbiased if the data are MCAR [25]. Assume the cumulative proportions $m_1/m \rightarrow r_1 > 0, (m_1 + m_2)/m \rightarrow r_2 > 0, (m_1 + m_3)/m \rightarrow r_3 > 0$ when $m \rightarrow \infty$, the maximum pseudolikelihood estimator $(\tilde{\boldsymbol{\eta}}_1, \tilde{\boldsymbol{\eta}}_2)$, can be shown to be consistent and asymptotically normal with covariance matrix

$$\boldsymbol{\Sigma}^{\text{mis}} = \begin{pmatrix} r_2^{-1} \mathbf{I}_{11}^{\text{mis}^{-1}} & r_1^{-1} \mathbf{I}_{11}^{\text{mis}^{-1}} \mathbf{I}_{12}^{\text{mis}^{-1}} \mathbf{I}_{22}^{\text{mis}^{-1}} \\ r_3^{-1} \mathbf{I}_{22}^{\text{mis}^{-1}} \end{pmatrix},$$

where

$$\mathbf{I}_{jj}^{\text{mis}} = E \left\{ -\frac{\partial^2 \log L_j^{\text{mis}}(\boldsymbol{\eta}_j)}{\partial \boldsymbol{\eta}_j^2} \right\} \quad \text{and} \quad \mathbf{I}_{12}^{\text{mis}} = E \left[\left\{ \frac{\partial \log L_1^{\text{mis}}(\boldsymbol{\eta}_1)}{\partial \boldsymbol{\eta}_1} \right\} \left\{ \frac{\partial \log L_2^{\text{mis}}(\boldsymbol{\eta}_2)}{\partial \boldsymbol{\eta}_2} \right\}^T \right] \quad \text{for } j = 1, 2.$$

An outline of derivation for the asymptotic covariance is provided in the Appendix. The covariance matrix Σ^{mis} can be empirically estimated as detailed in Appendix.

4. Simulation study

4.1. Methods under comparison

In this section, we evaluate the finite sample performance of the Pseudo-REML method and compare it to that of REML and Riley methods. The data are generated from a two-stage procedure as specified by equation (1). To cover a wide spectrum of scenarios, we vary the values for four factors that are considered important in practice. Specifically, the number of studies is set to 10 or 25 to represent meta-analysis of a moderate number or large number of studies, respectively. We consider both the complete data and missing data scenarios. For missing data scenario, we assume 40% of studies have one of two outcome MCAR. To reflect the heterogeneity in standard error of summary measure across studies, we sample s_{ij}^2 from the square of $N(0.25, 0.50)$ distribution, which leads to a median value of 0.25 for s_{ij}^2 . The size of the within-study variation relative to the between-study variation may have a substantial impact on the performance of the methods. To this end, we let the ratio of the within-study variation relative to the between study variation to be relatively small, comparable, and large, corresponding to $\tau_1^2 = \tau_2^2 = 0.5, 0.25, \text{ or } 0.1$, respectively. For within-study correlations, we consider two different settings: one is ρ_{Wi} being constant with value $-0.5, 0, \text{ or } 0.5$, the other is ρ_{Wi} varying across studies and being randomly sampled from Uniform $(-0.8, 0.8)$, Uniform $(-0.8, 0)$ and Uniform $(0, 0.8)$ to reflect arbitrary, negative, and positive within-study correlations, respectively. The between-study correlation ρ_B is set to $-0.5, 0, \text{ or } 0.5$. We set the overall effect sizes to be $\beta_1 = \beta_2 = 0$ and denote the difference in effect size between two outcomes by $\delta = \beta_1 - \beta_2$. For each simulation setting, we generate 5000 samples. The samples are simulated in R using the ‘mvrnorm’ function. Each of the simulated data is analyzed by three methods, namely, REML, Riley method, and the Pseudo-REML method. We note that REML method, although not applicable in practice due to the unknown within-study correlations, can serve as a gold standard for comparison. The results from REML method are obtained from the ‘mvmeta’ function in the R package ‘mvmeta’.

4.2. Simulation results

Table I summarizes empirical bias (Bias), empirical standard error (ESE), model-based standard error estimates, and the CP of the confidence intervals for δ when the within-study/between-study variation ratio is close to 1 (i.e., median of s_{ij}^2 is 0.25 and $\tau_1^2 = \tau_2^2 = 0.25$) for complete and missing data settings. The ESE is calculated as the standard deviation of point estimates, and model-based standard error estimate is calculated as the average of model-based standard errors. As shown in the upper panel of Table I, for the complete data settings, all three methods provide unbiased estimates of δ . When the number of studies is moderate ($m = 10$), Riley method tends to underestimate the standard error, leading to confidence intervals with coverage probabilities that are less than the nominal level (range of CP: 89.1 ~ 92.0%). The possible reason is that Riley method performs well only when estimated $\hat{\rho}_s$ is relatively small. In contrast, the standard error of estimate for δ using the Pseudo-REML method is well estimated, and the coverage probabilities of confidence intervals are close to the nominal level (range of CP: 94.9 ~ 95.8%). To assess the singular covariance matrix problem, we display the percentage of the singular covariance matrix (SP) under different settings when $m = 10$ in Figure S2 of the Supplemental Materials. We find that the REML method suffers greatly from the singular estimated covariance matrix problem (SP > 25% under all settings). Riley method alleviates this problem to some extent but still has a sizable proportion of singular estimated covariance matrix in some settings (range of SP: 0 ~ 20%). In contrast, there is no singular estimated covariance matrix problem for the Pseudo-REML method. We further examine the performance of estimates under three different methods when the estimated covariance matrix using REML method is singular. When REML method has singular estimated covariance matrix problem, the point estimates from three methods are similar, but their standard error estimates are quite different due to the singularity of estimated covariance matrix, in that the standard errors are larger in the REML method. This was also noted by Riley *et al.* (2007) [12]. The details can be found in Figure S3 of the Supplemental Materials. When the number of studies is relatively large ($m = 25$), both Riley method and the Pseudo-REML method perform well in that estimates are unbiased and confidence intervals have coverage probabilities close to the nominal level.

Table I. Estimates of bias, empirical standard error (ESE), model-based standard error (MBSE), coverage probability (CP) of $\delta = \beta_1 - \beta_2$ in 5000 simulations based on data generated from restricted maximum likelihood (REML) model, with the within-study/between-study variation ratio close to 1, for different number of studies m and different between-study correlation ρ_B and within-study correlations ρ_{Wi} . All entries in the table are multiplied by 100.

m	Scenarios	ρ_{Wi}	ρ_B	REML					Riley method					Pseudo-REML				
				Bias	ESE	MBSE	CP	Bias	ESE	MBSE	CP	Bias	ESE	MBSE	CP	Bias	ESE	MBSE
10	Complete	-0.5	-0.5	-0.7	38.0	33.5	87.5	87.5	-0.6	38.8	35.8	89.9	89.9	-0.8	37.8	43.4	95.2	
		0.0	0.0	0.6	35.3	30.6	86.5	86.5	0.5	35.2	33.6	91.0	91.0	0.5	34.5	40.3	95.2	
		0.5	0.5	0.2	30.7	26.6	87.6	87.6	0.3	30.9	30.2	92.0	92.0	0.3	30.2	36.9	95.8	
	Missing	0	-0.5	-0.5	-0.5	35.9	32.4	88.5	88.5	-0.4	36.7	33.7	90.0	90.0	-0.6	35.8	41.1	95.0
			0.0	0.0	0.4	33.1	29.0	87.5	87.5	0.4	32.9	31.0	90.9	90.9	0.5	32.3	38.0	95.0
			0.5	0.5	0.3	28.2	24.6	88.2	88.2	0.4	28.1	27.1	91.6	91.6	0.4	28.0	34.5	95.7
		0.5	-0.5	-0.5	-0.4	33.4	30.5	88.9	88.9	-0.4	34.6	31.4	89.1	89.1	-0.5	34.0	38.7	94.9
			0.0	0.0	0.4	29.8	26.8	88.4	88.4	0.4	30.2	28.1	90.7	90.7	0.5	29.9	35.7	95.3
			0.5	0.5	0.4	25.1	21.9	87.5	87.5	0.5	25.0	23.5	90.7	90.7	0.4	25.6	31.6	95.4
Missing	-0.5	-0.5	-0.5	-0.3	43.1	36.9	86.0	86.0	0.1	45.9	39.3	86.2	86.2	-0.4	43.2	53.9	95.7	
		0.0	0.0	0.4	42.2	35.9	86.5	86.5	0.5	44.8	39.0	87.2	87.2	0.5	41.4	52.2	95.9	
		0.5	0.5	0.1	39.2	33.9	87.9	87.9	-0.1	41.0	37.1	88.9	88.9	0.0	37.9	50.1	96.3	
	0	-0.5	-0.5	-0.5	0.1	42.4	36.0	85.9	85.9	0.0	44.3	38.7	87.6	87.6	-0.1	41.7	52.5	96.0
			0.0	0.0	0.4	41.0	34.4	85.3	85.3	0.6	42.9	37.9	88.0	88.0	0.5	39.9	50.7	95.8
			0.5	0.5	0.3	37.2	31.7	87.2	87.2	0.0	38.8	35.4	89.1	89.1	0.1	36.3	48.6	96.6
		0.5	-0.5	-0.5	-0.1	41.1	34.5	85.2	85.2	0.0	42.6	37.5	88.1	88.1	-0.1	40.3	50.9	96.1
			0.0	0.0	0.7	38.9	31.9	85.1	85.1	0.5	40.5	36.0	88.4	88.4	0.7	38.2	49.2	95.8
			0.5	0.5	0.4	34.8	28.5	85.2	85.2	0.1	36.1	33.1	89.3	89.3	0.0	34.8	46.5	96.6

Table I. Continued.

m	Scenarios	ρ_{wi}	ρ_B	REML					Riley method					Pseudo-REML						
				Bias	ESE	MBSE	CP	CP	Bias	ESE	MBSE	CP	CP	Bias	ESE	MBSE	CP			
25	Complete	-0.5	-0.5	0.0	23.3	22.1	92.1	92.1	92.1	92.1	-0.1	23.3	22.5	92.9	92.9	0.1	23.5	24.8	95.0	
		0.0	0.0	-0.1	21.0	19.9	91.4	91.4	91.4	91.4	-0.1	21.0	21.2	94.0	94.0	-0.1	20.9	22.6	95.2	
		0.5	0.5	0.3	18.8	16.7	89.3	89.3	89.3	89.3	0.1	18.9	18.9	94.1	94.1	0.1	18.9	20.1	95.5	
	Missing	0	-0.5	-0.1	22.2	21.3	92.3	92.3	92.3	92.3	-0.1	22.2	21.2	92.6	92.6	-0.1	22.2	23.5	95.0	
		0.0	0.0	-0.1	19.9	18.9	91.8	91.8	91.8	91.8	0.0	19.7	19.4	93.4	93.4	0.0	19.6	21.2	95.0	
		0.5	0.5	0.2	17.3	15.6	90.1	90.1	90.1	90.1	0.1	17.1	17.0	94.0	94.0	0.1	17.2	18.6	95.8	
	25	Complete	-0.5	-0.5	-0.2	20.5	20.0	92.7	92.7	92.7	92.7	-0.2	20.9	19.8	92.2	92.2	-0.2	20.8	22.2	94.8
			0.0	0.0	0.0	18.2	17.5	92.3	92.3	92.3	92.3	0.0	18.2	17.5	92.6	92.6	0.0	18.3	19.7	94.6
			0.5	0.5	0.2	15.2	14.2	91.4	91.4	91.4	91.4	0.1	15.2	14.7	93.7	93.7	0.0	15.5	17.0	95.7
Missing		-0.5	-0.5	-0.2	26.4	24.3	90.7	90.7	90.7	90.7	-0.1	26.6	25.0	91.7	91.7	-0.1	26.6	29.0	95.1	
		0.0	0.0	-0.3	25.5	23.0	90.2	90.2	90.2	90.2	-0.1	25.5	24.5	92.7	92.7	-0.2	24.8	27.7	95.6	
		0.5	0.5	0.0	24.2	21.1	88.8	88.8	88.8	88.8	-0.2	24.1	23.4	93.4	93.4	-0.1	23.6	26.2	95.9	
Missing		0	-0.5	-0.4	26.0	23.8	90.6	90.6	90.6	90.6	-0.3	26.0	24.6	91.7	91.7	-0.3	25.8	28.1	95.1	
		0.0	0.0	-0.2	24.6	22.2	89.6	89.6	89.6	89.6	-0.1	24.4	23.7	93.0	93.0	-0.1	23.9	26.8	95.8	
		0.5	0.5	-0.1	22.9	20.1	89.1	89.1	89.1	89.1	-0.2	22.5	22.2	93.9	93.9	-0.2	22.4	25.3	96.4	
Missing	0.5	-0.5	-0.5	25.3	23.0	90.3	90.3	90.3	90.3	-0.4	25.3	23.8	91.7	91.7	-0.5	25.0	27.3	94.9		
	0.0	0.0	-0.1	23.3	21.2	89.6	89.6	89.6	89.6	-0.1	23.2	22.5	92.3	92.3	-0.1	23.1	25.9	95.5		
	0.5	0.5	-0.2	21.3	18.5	88.4	88.4	88.4	88.4	-0.2	20.9	20.8	94.4	94.4	-0.3	21.3	24.4	96.6		

Table II. Estimates of bias, empirical standard error (ESE), model-based standard error (MBSE), coverage probability (CP) of $\delta = \beta_1 - \beta_2$ in 5000 simulations based on data generated from restricted maximum likelihood (REML) model, with the within-study/between-study variation ratio close to 2.5, for different number of studies m and different between-study correlation ρ_B and within-study correlations ρ_{Wi} . All entries in the table are multiplied by 100.

m	Scenarios	ρ_{Wi}	ρ_B	REML					Riley method					Pseudo-REML				
				Bias	ESE	MBSE	CP	CP	Bias	ESE	MBSE	CP	CP	Bias	ESE	MBSE	CP	
10	Complete	-0.5	-0.5	-0.5	28.2	24.3	86.4	86.4	-0.2	29.2	26.7	89.0	89.0	-0.6	28.2	34.1	96.1	
		0.0	0.0	0.3	27.1	23.2	86.4	86.4	0.4	27.8	26.2	91.0	91.0	0.3	26.4	32.5	95.9	
		0.5	0.5	0.2	24.7	21.7	88.3	88.3	0.3	25.5	25.0	91.6	91.6	0.1	24.2	30.9	96.6	
	Missing	0.0	-0.5	-0.3	26.5	23.3	87.5	87.5	0.0	27.2	25.0	89.8	89.8	-0.4	26.3	31.8	96.0	
		0.0	0.0	0.1	25.0	21.8	87.4	87.4	0.1	25.4	23.8	90.8	90.8	0.2	24.5	30.2	95.9	
		0.5	0.5	0.2	22.5	19.8	88.7	88.7	0.3	22.8	22.0	91.3	91.3	0.2	22.2	28.6	96.1	
	Missing	0.5	-0.5	-0.5	-0.2	24.0	21.4	87.8	87.8	0.0	25.0	22.4	88.4	88.4	-0.2	24.5	29.4	95.5
			0.0	0.0	0.1	22.0	19.4	87.9	87.9	0.1	22.4	20.7	90.2	90.2	0.2	22.4	27.8	95.7
			0.5	0.5	0.2	19.4	17.0	87.8	87.8	0.3	19.4	18.3	89.8	89.8	0.2	20.0	25.8	96.1
0.0		-0.5	-0.5	-0.2	32.6	28.0	86.6	86.6	0.5	35.6	30.5	86.4	86.4	-0.3	33.0	44.4	96.9	
		0.0	0.0	0.1	32.2	27.9	87.1	87.1	0.2	35.2	30.6	87.3	87.3	0.1	31.8	43.6	97.2	
		0.5	0.5	0.0	30.8	27.4	88.4	88.4	-0.1	33.2	29.9	87.9	87.9	-0.1	30.1	43.1	97.6	
0.5		-0.5	-0.5	0.1	31.8	27.4	86.5	86.5	-0.1	34.3	30.0	87.5	87.5	-0.1	31.5	43.1	97.0	
		0.0	0.0	0.1	31.1	26.8	86.6	86.6	0.2	33.2	29.6	88.0	88.0	0.2	30.5	42.0	97.1	
		0.5	0.5	0.1	29.0	25.7	88.2	88.2	0.1	31.3	28.3	88.9	88.9	0.0	28.5	41.4	97.7	
0.5	-0.5	-0.5	0.2	30.2	25.7	86.1	86.1	0.0	32.1	28.5	87.7	87.7	0.1	29.9	41.2	97.2		
	0.0	0.0	0.2	29.0	24.5	85.7	85.7	0.4	31.0	27.8	87.9	87.9	0.3	29.0	40.8	97.1		
	0.5	0.5	0.2	26.8	22.8	86.5	86.5	0.0	28.6	26.1	88.6	88.6	-0.1	27.1	39.3	97.6		

Table II. Continued.

m	Scenarios	ρ_{wi}	ρ_b	REML				Riley method				Pseudo-REML				
				Bias	ESE	MBSE	CP	Bias	ESE	MBSE	CP	Bias	ESE	MBSE	CP	
25	Complete	-0.5	-0.5	-0.1	17.1	15.6	90.1	0.0	17.3	16.3	91.8	0.1	17.3	18.2	94.6	
		0.0	0.0	-0.1	16.0	14.5	89.9	-0.2	16.0	15.9	93.6	-0.2	15.8	17.2	95.3	
		0.5	0.5	0.2	14.8	13.0	89.2	0.0	15.0	15.1	94.1	0.1	14.9	15.9	95.1	
	Missing	0.0	-0.5	-0.1	16.1	15.0	91.0	-0.1	16.2	15.2	91.6	-0.1	16.1	17.1	94.5	
		0.0	0.0	-0.1	15.0	13.6	89.8	-0.1	14.8	14.4	92.6	-0.1	14.8	15.9	94.9	
		0.5	0.5	0.1	13.6	12.0	89.3	0.0	13.4	13.3	93.8	0.0	13.5	14.6	95.8	
	Missing	0.5	-0.5	-0.5	-0.2	14.5	13.8	91.8	-0.1	14.9	13.8	91.0	-0.2	14.8	15.9	94.7
			0.0	0.0	0.0	13.3	12.4	90.7	0.0	13.4	12.6	91.8	0.0	13.6	14.6	94.6
			0.5	0.5	0.1	11.7	10.6	90.4	0.0	11.6	11.2	93.3	0.0	12.1	13.2	95.4
-0.5		-0.5	-0.1	19.7	17.6	89.8	-0.1	19.9	18.5	90.7	0.0	19.8	21.8	95.1		
		0.0	0.0	-0.2	19.3	17.0	88.8	-0.1	19.4	18.3	91.5	-0.2	18.9	21.1	95.1	
		0.5	0.5	0.0	18.6	16.4	89.6	-0.1	18.9	18.0	92.3	-0.1	18.3	20.4	95.8	
0.0	-0.5	-0.2	19.2	17.2	90.1	-0.2	19.3	18.1	91.5	-0.2	19.1	21.1	95.5			
	0.0	0.0	-0.2	18.5	16.4	89.3	-0.1	18.5	17.6	91.9	-0.1	18.2	20.4	95.4		
	0.5	0.5	-0.1	17.6	15.6	89.6	-0.1	17.6	17.0	93.1	-0.2	17.4	19.7	96.1		
0.5	-0.5	-0.4	18.4	16.4	89.6	-0.3	18.6	17.3	91.3	-0.4	18.4	20.3	95.4			
	0.0	0.0	-0.1	17.3	15.3	88.6	-0.1	17.4	16.6	92.0	-0.1	17.4	19.5	95.4		
	0.5	0.5	-0.2	16.1	14.1	89.1	-0.2	16.1	15.8	93.1	-0.2	16.5	18.8	96.5		

Table II summarizes the similar results as in Table I when the within-study/between-study variation ratio is close to 2.5 (i.e., median of s_{ij}^2 is 0.25 and $\tau_1^2 = \tau_2^2 = 0.1$). For the complete data settings when the number of studies is moderate ($m = 10$), the ranges of CPs are 86.4 ~ 88.7% for REML method, 88.4 ~ 91.6% for Riley method, and 95.5 ~ 96.6% for the Pseudo-REML method, respectively. When the number of studies is relatively large ($m = 25$), both Riley method and the Pseudo-REML method perform well. We find that when the between-study heterogeneity is relatively small (i.e., the within-study/between-study variation ratio is relatively large), there is a substantial increase in the percentage of singular estimated covariance matrix problem in both REML and Riley methods.

For missing data settings, there is more advantage in avoiding the singular estimated covariance matrix problem for the Pseudo-REML method when the number of studies is moderate. For example, as shown in the lower panel of Table I when $m = 10$, the ranges of CPs are 85.1 ~ 87.9% for REML method, 86.2 ~ 89.3% for Riley method, and 95.7 ~ 96.6% for the Pseudo-REML method, respectively.

We also consider the settings when the number of studies is relatively small ($m = 5$) and settings when the within-study/between-study variation ratio is close to 0.5 (i.e., median of s_{ij}^2 is 0.25 and $\tau_1^2 = \tau_2^2 = 0.5$). The findings are similar to the setting with moderate sample size ($m = 10$) and hence are not presented here for limited space (see Table S1 in Supplemental Material). In summary, for meta-analysis with moderate number of studies, REML method suffers greatly from the singular estimated covariance matrix problem, even when within-study correlations are known. Riley method alleviates such problem to some extent but still has a sizable proportion of singular estimated covariance matrix. The Pseudo-REML method does not suffer from the singular estimated covariance matrix problem and can produce confidence intervals with coverage probabilities close to the nominal level under all settings considered.

One interesting phenomenon shown in Tables I and II is that the ESE from REML method is very close to that from the Pseudo-REML method for the complete data settings when number of studies is moderate or large. This suggests that the efficiency gain in the individual pooled estimates through the joint analysis of Y_{i1} and Y_{i2} may not be large for complete data, even when the within-study correlations ρ_{Wi} are available. Such finding is consistent with the literature; for example, see [26], [24], and [27]. To compare the efficiency in the estimation of δ , we consider the RE defined by the square of the ESE of the estimates from REML method, divided by that of Riley method or the Pseudo-REML method. We plot the RE against the between-study correlation ρ_B in Figure 1 (for within-study/between-study variation ratio close to 1). To cover a wide range of between-study correlations and to include the rare situations of extreme between-study correlations, we let ρ_B vary from -0.99 to 0.99 . As shown in Figure 1, in complete data settings when $m = 10$, the relative efficiencies of the both Riley method and Pseudo-REML method compared to REML method are close to 100% except the extreme situation where ρ_B is greater than 0.8.

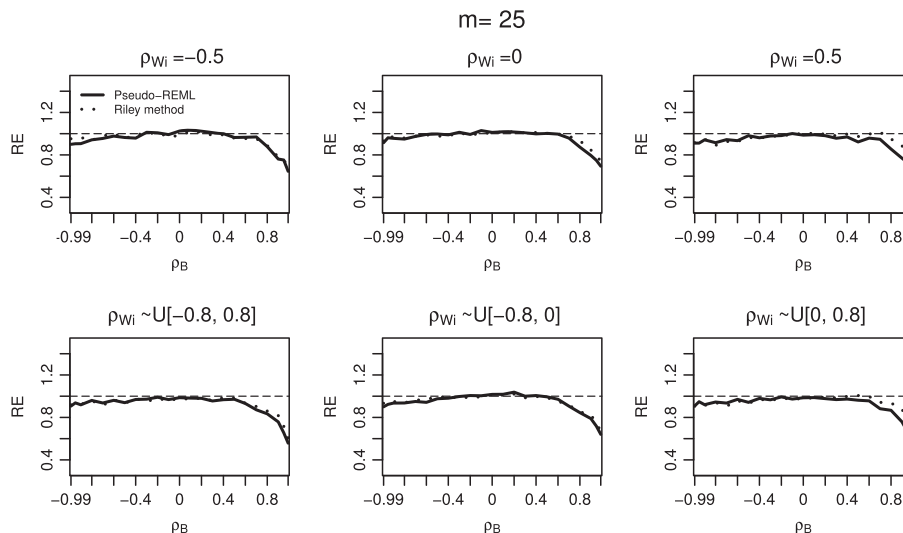


Figure 1. Relative efficiency of estimator of $\delta = \beta_1 - \beta_2$ based on Riley method and Pseudo-restricted maximum likelihood (REML) method comparing to the estimator based on REML model, with the within-study/between-study variation ratio close to 1 for different between-study correlation ρ_B and within-study correlations ρ_{Wi} . The number of studies $m = 10$ and number of simulations is 5000.

Table III. Type I error (in %) for restricted maximum likelihood (REML), Riley method, and Pseudo-REML considered at 5% significance level when the within-study/between-study variation ratio is close to 1.

m	ρ_{Wi}	ρ_B	REML	Riley method	Pseudo-REML
10	-0.5	-0.5	12.5	10.1	4.8
		0.0	13.5	9.0	4.8
		0.5	12.4	8.0	4.2
	0.0	-0.5	11.5	10.0	5.0
		0.0	12.5	9.1	5.0
		0.5	11.8	8.4	4.3
	0.5	-0.5	11.1	10.9	5.1
		0.0	11.6	9.3	4.7
		0.5	12.5	9.3	4.6
25	-0.5	-0.5	7.9	7.1	5.0
		0.0	8.6	6.0	4.8
		0.5	10.7	5.9	4.5
	0.0	-0.5	7.7	7.4	5.0
		0.0	8.2	6.6	5.0
		0.5	9.9	6.0	4.2
	0.5	-0.5	7.3	7.8	5.2
		0.0	7.7	7.4	5.4
		0.5	8.6	6.3	4.3

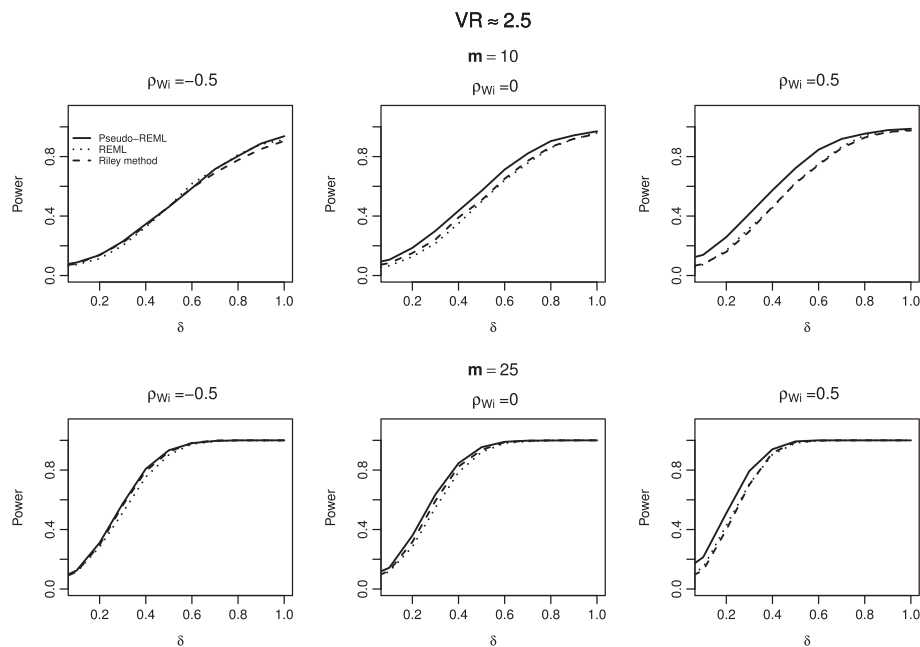


Figure 2. The power curves of restricted maximum likelihood (REML) method, Riley method, and Pseudo-REML method, with the within-study/between-study variation ratio close to 1, for different number of studies m and within-study correlations ρ_{Wi} . The between-study correlation ρ_B is 0.5. The number of simulations is 5000.

We evaluated the Type I error and power of REML, Riley method, and Pseudo-REML method for the complete data settings. Table III summarizes the Type I error at 5% significance level when the within-study/between-study variation ratio is close to 1. We can see that Pseudo-REML method controls the Type I errors well at all settings considered, whereas the REML and Riley methods have inflated type I errors. It is worth mentioning that if only assessing on results with no singularity problem, the results from the REML method improve substantially in that the type I errors are close to the nominal level. Figure 2 displays the corresponding powers of these three methods when between study correlation $\rho_B = 0.5$. The critical region that is used to calculate the power is adjusted so that the corresponding Type I error is 5%. The powers of three methods are similar with Pseudo-REML method being slightly more powerful than the other two.

In summary, our simulation studies suggest that both Riley method and the Pseudo-REML method perform well when the number of studies is large and the Pseudo-REML method does not suffer from the singular estimated covariance matrix problem, and maintains good CP of confidence intervals and high RE when number of studies is relatively small or moderate. Pseudo-REML method has well-controlled Type I error and competitive power compared to Riley method.

5. Applications

We illustrate the Pseudo-REML method by three meta-analyses. For these three meta-analyses, the within-study correlations are unavailable, which is common in practice. As a result, REML method cannot be applied. Alternatively, Riley method and the Pseudo-REML method can be used. In the first example, we consider a meta-analysis of small number of studies where both outcomes are reported, while in the second and third examples, we consider meta-analyses of large number of studies where some studies only have one of two outcomes reported.

5.1. Comparison between overall survival and disease-free survival for prostate cancer

Prostate cancer is a malignant tumor that develops in the prostate and is the sixth leading cause of cancer-related deaths in men [28]. Radiotherapy is considered as the most commonly used treatment for locally advanced prostate cancer because of its high survival rate and low morbidity. Androgen deprivation therapy, also called hormone therapy, is a strategy adjuvant to radiotherapy. Recent randomized trials show that androgen deprivation therapy has inhibitory effect on the growth and proliferation of prostate cancer cell. Goserelin acetate, an injectable gonadotropin releasing hormone superagonist, is often used to suppress the production of the sex hormones in the treatment of prostate cancer. Sasse *et al.* [29] compared the hormone therapy using Goserelin acetate combined with radiotherapy (referred to as ‘combined therapy’ hereafter) versus radiotherapy alone in overall survival and disease-free survival. Five randomized clinical trials published between 1988 and 2011 have reported log-hazard ratio estimates comparing combined therapy using Goserelin acetate with radiotherapy with respect to both overall survival and that with respect to disease-free survival, denoted by Y_{i1} and Y_{i2} , respectively. It is of clinically importance to evaluate the difference between the log-hazard ratio with respect to overall survival and disease-free survival [30]. Denote this difference by δ , we conduct a meta-analysis of the five trials by applying both Riley method and the Pseudo-REML method. As shown in the upper left panel of Table IV, the difference δ is estimated as 0.320 (95% CI: (0.054, 0.585); p -value: 0.018) using the Pseudo-REML method and 0.312 (95% CI: (0.104, 0.519); p -value: 0.003) using Riley method. The estimated overall marginal correlation $\hat{\rho}_s = 0.45$ using Riley method. The results from both methods suggest that the log hazard ratio of overall-survival is significantly higher than that of the disease-free survival. The discrepancies between two methods in significance level and in width of confidence interval are consistent with what we observed in the simulation studies where Riley method tends to underestimate the standard error of the difference in pooled estimates in small sample size settings. For illustration purpose, we also calculate the estimated difference using separate univariate meta-analyses ignoring the correlation between the overall log hazard ratio with respect to overall survival and that with respect to disease-free survival. Compared to the results using Pseudo-REML method, the estimated difference δ using univariate analyses has the same point estimate but wider confidence interval. Specifically, the confidence interval is 95% CI: (0.032, 0.607) and p -value is 0.030. We note that the inference using univariate meta-analyses is potentially misleading because the effect with respect to overall survival and that with respect to disease-free survival are likely to be positively correlated. Ignorance of such correlation can lead to the overestimated standard error and inflated p -value.

Table IV. Results of Riley method and the Pseudo restricted maximum likelihood (Pseudo-REML) method applied to the three meta-analyses.

Study	Method	Complete data scenario			Missing data scenario		
		Pooled estimates	95% CI	<i>p</i> -value	Pooled estimates	95% CI	<i>p</i> -value
Sasse 2011 [†]	Riley method	0.312	(0.104, 0.519)	0.003	—	—	—
	Pseudo-REML	0.320	(0.054, 0.585)	0.018	—	—	—
Zhao 2012 [‡]	Riley method	-0.078	(-0.387, 0.232)	0.623	-0.281	(-0.688, 0.126)	0.176
	Pseudo-REML	-0.075	(-0.377, 0.228)	0.627	-0.300	(-0.928, 0.328)	0.349
Thompson 2005 [*]	Riley method	0.055	(-0.050, 0.160)	0.304	0.072	(0.024, 0.121)	0.004
	Pseudo-REML	0.076	(-0.030, 0.181)	0.159	0.070	(0.030, 0.110)	0.001

[†]The pooled estimate is the difference between the log-hazard ratio with respect to the overall survival and disease-free survival.

[‡]The pooled estimate is the difference between the standard mean difference comparing cases and controls of paroxonase activity and that of arylesterase activity.

^{*}The pooled estimate is $\log OR_{TT\text{ vs }CC} / \delta P$, which is the effect of one unit change in homocysteine level on $\log OR_{TT\text{ vs }CC}$.

5.2. Treatment comparison for paraoxonase 1 activities on reducing coronary heart disease risk

Paraoxonase 1 (PON1) is an enzyme synthesized in the liver. PON1 has the ability to inhibit high-density lipoprotein and is cardioprotective. Studies of the relationship between PON1 activity and coronary heart disease (CHD) risk in humans have yielded inconsistent results [31]. To investigate the PON-CHD relationship, Zhao *et al.* [32] conducted a meta-analysis combining 55 studies, which involve 9481 CHD patients and 11,148 controls. It is important to study the difference in the standard mean difference (SMD) comparing cases and controls of *paraoxonase activity* of PON1, denoted by Y_{i1} , and that comparing cases and controls of *arylesterase activity* of PON1, denoted by Y_{i2} [33]. For illustration purpose, we consider 17 studies where both summary measures Y_{i1} and Y_{i2} are reported. As shown in the middle left panel of Table IV, the overall difference δ is estimated as -0.075 (95% CI: $(-0.377, 0.228)$; p -value: 0.627) using the Pseudo-REML method, and -0.078 (95% CI: $(-0.387, 0.232)$; p -value: 0.623) using Riley method. The estimated overall marginal correlation $\hat{\rho}_s = 0.41$ using Riley method. The results from both methods suggest there is no statistically significant differences in SMD between the *paraoxonase* and *arylesterase activities*. The results from both methods are very similar, which agree with the simulation studies with relatively large number of studies.

We note that many of the studies have only one of two outcomes reported. Excluding these studies in the analysis may lead to a substantial loss of efficiency. Under the MCAR assumption, both Riley method and Pseudo-REML method can be applied to all 55 studies, and the results are summarized in the middle right panel of Table IV. The overall difference δ is estimated as -0.300 (95% CI: $(-0.928, 0.328)$; p -value: 0.349) using the Pseudo-REML method and -0.281 (95% CI: $(-0.688, 0.126)$; p -value: 0.176) using Riley method. Comparing with the estimates based on the 17 studies with both endpoints reported, the estimates based on all studies have larger difference in SMD between the *paraoxonase* and *arylesterase* activities of PON 1. But none of the estimates reaches statistically significant difference. However, as suggested by the simulation studies, the confidence interval produced by Riley method tends to be too narrow, which may be corrected by the Pseudo-REML method.

5.3. Methylene tetrahydrofolate reductase gene, homocysteine, and coronary heart disease

The observed effect of blood homocysteine on CHD suggested by many observational studies is questioned by researchers because of potential confounders (e.g., smoking and blood pressure) and reverse causation (i.e., elevations in blood homocysteine may result from atherosclerosis and CHD) [34]. Having a common polymorphism affecting the level of homocysteine in blood, methylene tetrahydrofolate reductase gene can be employed as an instrumental variable to adjust for the bias due to confounding or reverse causation [35]. Thompson *et al.* [36] conducted a meta-analysis involving 66 genetic studies to estimate the unconfounded association of the homocysteine level and CHD. Comparing with the wild-type CC, the mutant genotype TT of methylene tetrahydrofolate reductase gene is associated with both risk increase of CHD and higher level of homocysteine. Let $\log OR_{TT \text{ vs. } CC}$ be the log odds ratio for the association between genotype (TT vs. CC) and CHD, and let δP be the mean difference in homocysteine level between genotype TT and CC. The objective is to estimate the ratio $\log OR_{TT \text{ vs. } CC} / \delta P$, which is the effect of one unit change in homocysteine level on $\log OR_{TT \text{ vs. } CC}$. Out of the 66 studies, 18 studies reported both $\log OR_{TT \text{ vs. } CC}$ and δP , while the remaining studies only reported one of the two outcomes. The estimates of $\log OR_{TT \text{ vs. } CC} / \delta P$ using the Riley method and Pseudo-REML method based on the 18 studies and the estimates based on the 66 studies are shown in the lower panels in Table IV. The Riley method produces smaller pooled estimates based on 18 studies than that based on 66 studies (0.055 vs. 0.072), while the Pseudo-REML method produces similar pooled estimates (0.076 vs. 0.070) based on 18 or 66 studies. The estimated overall marginal correlation $\hat{\rho}_s = 0.46$ using Riley method. The difference between the pooled estimates from Riley method and those from the proposed Pseudo-REML method is due to the change in individual pooled estimates through borrowing of strength across outcomes in the Riley method, but the Pseudo-REML method does not allow this. The length of confidence interval produced by Riley method is similar to that produced by the Pseudo-REML method. The results based on all 66 studies using both methods suggest that there is a statistically significant effect of homocysteine level on CHD.

6. Discussion

In this paper, we propose a Pseudo-REML method for multivariate random-effects meta-analysis. The idea is to base the inference on a working independence model [19]. Such an idea is motivated by the

construction of the generalized estimating equation by Liang and Zeger [37] for the inference of marginal models in longitudinal data and the construction of the pseudo-partial likelihood function by Lin [38] for the analysis of multivariate survival data. This method is applicable when the within-study correlations are unknown, is not suffering from problems related to singularity of estimated covariance matrix, and can maintain a reasonably good inference performance in all the scenarios tested in the simulation studies. In addition, the Pseudo-REML method can be extended to multivariate meta-regression models where study-level covariates are available. More details of the regression extension can be found in the Supplemental Materials.

In this paper, we compare the performance of the REML method, Riley method, and the proposed Pseudo-REML method, and investigate their singular estimated covariance matrix problem. The REML method is the standard approach when the within-study correlations are known. However, it may suffer from the singular covariance matrix problem and can lead to poorly estimated between study correlations. The Riley method or the proposed Pseudo-REML method can be a good alternative in this situation. When within-study correlations are unknown, both Riley method and the proposed Pseudo-REML method are applicable. Riley method performs well when the estimated synthesis correlation parameter ρ_s is relatively small. When the estimated synthesis correlation parameter ρ_s is relatively large, the proposed Pseudo-REML method will be the best choice.

We note that this paper focuses on the method for meta-analysis based on summary data rather than on methods based on individual patient data (IPD). Naturally, meta-analyses based on summary data suffer from limitations such as potentially inconsistent exclusion criterion for patients across studies, and no quality assessment available based on the summary data alone [39–44]. The IPD meta-analysis for pooling data from different studies should be preferred when the IPD are available; for example, see Piedbois *et al.* [45] and Di Leo *et al.* [46]. In this case, the correlations between multiple outcomes would become available.

Because both the REML method and Riley method are presented in the bivariate setting [2, 13], we also presented the Pseudo-REML method in the bivariate setting in this paper to keep the notation simple and convey the main idea. The Pseudo-REML method can be easily applied to MMA where more than two outcomes are analyzed. In contrast, although it is also conceptually straightforward that REML method and Riley method can be extended to MMA with more than two outcomes, we expect that these two methods will experience more severe singular estimated covariance matrix problem as the number of outcomes increases (due to the increased number of between-study correlations that require estimation). The performance of the Pseudo-REML method with more than two outcomes requires further investigation.

In this paper, we considered Pseudo-REML method when data are MCAR. One limitation of this procedure is that the Pseudo-REML method may yield biased estimates if MCAR assumption is violated, for example, in the settings of MAR or in the presence of outcome reporting bias [47] or publication bias. The current version of our Pseudo-REML method is only suitable for MCAR situations. Extension of our method to MAR and MNAR situations might be possible. For example, if publication bias is present, statistical methods such as ‘Trim and Fill’ method may be used to impute the missing studies and then incorporated into the proposed Pseudo-REML [48]. The development of the Pseudo-REML under MAR and MNAR settings and the finite sample performance of regression extension of our method is currently under investigation and will be reported elsewhere.

Appendix A

Derivation of the asymptotic covariance formula for the maximum pseudolikelihood estimator

Denote $\boldsymbol{\eta} = (\boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T)^T$. By Taylor expansion of $\partial L_p(\boldsymbol{\eta})/\partial(\boldsymbol{\eta})$ around $\boldsymbol{\eta}$, we have

$$0 = \frac{1}{\sqrt{m}} \frac{\partial \log L_p(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} + \frac{1}{m} \frac{\partial^2 \log L_p(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^2} \sqrt{m}(\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}) + o_p(1). \quad (\text{A.1})$$

Therefore, we have

$$\sqrt{m}(\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}) \approx A_m^{-1} \frac{1}{\sqrt{m}} B_m, \quad \text{where} \quad A_m = -\frac{1}{m} \frac{\partial^2 \log L_p(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^2} \quad \text{and} \quad B_m = \frac{\partial \log L_p(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}.$$

It is easy to show that $\frac{1}{\sqrt{m}}\mathbf{B}_m \rightarrow N(0, \boldsymbol{\Sigma}^*)$ where

$$\boldsymbol{\Sigma}^* = \begin{pmatrix} I_{11} & I_{12} \\ I_{12}^T & I_{22} \end{pmatrix} \quad \text{and} \quad A_m \rightarrow \begin{pmatrix} I_{11} & 0 \\ 0 & I_{22} \end{pmatrix},$$

as $m \rightarrow \infty$. Note that the off-diagonal terms in $\boldsymbol{\Sigma}^*$ account for the fact that the multivariate outcomes may be dependent. Finally, the asymptotic distribution is immediately followed by Slutskys theorem and recalling equation (A.2).

Derivation of the asymptotic variance formula for the maximum pseudolikelihood estimator under the missing data scenario

Denote $\boldsymbol{\eta} = (\boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T)^T$. By Taylor expansion of $\partial L_p^{\text{mis}}(\boldsymbol{\eta})/\partial(\boldsymbol{\eta})$ around $\boldsymbol{\eta}$, we have

$$0 = \frac{1}{\sqrt{m}} \frac{\partial \log L_p^{\text{mis}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} + \frac{1}{m} \frac{\partial^2 \log L_p^{\text{mis}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^2} \sqrt{m}(\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}) + o_p(1). \quad (\text{A.2})$$

Therefore, we have

$$\sqrt{m}(\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}) \approx \mathbf{A}_m^{\text{mis}^{-1}} \frac{1}{\sqrt{m}} \mathbf{B}_m^{\text{mis}}, \quad \text{where} \quad \mathbf{A}_m^{\text{mis}} = -\frac{1}{m} \frac{\partial^2 \log L_p^{\text{mis}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^2} \quad \text{and} \quad \mathbf{B}_m^{\text{mis}} = \frac{\partial \log L_p^{\text{mis}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}.$$

It is easy to show that $\frac{1}{\sqrt{m}}\mathbf{B}_m^{\text{mis}} \rightarrow N(0, \boldsymbol{\Sigma}^{\text{mis}*})$ where

$$\boldsymbol{\Sigma}^{\text{mis}*} = \begin{pmatrix} r_2 I_{11}^{\text{mis}} & r_1 I_{12}^{\text{mis}} \\ r_3 I_{22}^{\text{mis}} & \end{pmatrix} \quad \text{and} \quad A_m^{\text{mis}} \rightarrow \begin{pmatrix} r_2 I_{11}^{\text{mis}} & 0 \\ 0 & r_3 I_{22}^{\text{mis}} \end{pmatrix},$$

where $m_1/m \rightarrow r_1$, $(m_1+m_2)/m \rightarrow r_2$, and $(m_1+m_3)/m \rightarrow r_3$. The asymptotic distribution is immediately followed by Slutsky's theorem and recalling equation (A.2).

Estimation of the covariance matrix under missing data scenario $\boldsymbol{\Sigma}^{\text{mis}}$

The information matrices $\mathbf{I}_{11}^{\text{mis}}$, $\mathbf{I}_{22}^{\text{mis}}$, and $\mathbf{I}_{12}^{\text{mis}}$ can be empirically estimated as

$$\begin{aligned} \hat{\mathbf{I}}_{11}^{\text{mis}} &= \frac{1}{m_1 + m_2} \sum_{i=1}^{m_1+m_2} \mathbf{U}_{1i}^{\text{mis}}(\tilde{\boldsymbol{\eta}}_1) \mathbf{U}_{1i}^{\text{mis}}(\tilde{\boldsymbol{\eta}}_1)^T \\ \hat{\mathbf{I}}_{22}^{\text{mis}} &= \frac{1}{m_1 + m_3} \left\{ \sum_{i=1}^{m_1} \mathbf{U}_{2i}^{\text{mis}}(\tilde{\boldsymbol{\eta}}_2) \mathbf{U}_{2i}^{\text{mis}}(\tilde{\boldsymbol{\eta}}_2)^T + \sum_{i=m_1+m_2}^{m_1+m_2+m_3} \mathbf{U}_{2i}^{\text{mis}}(\tilde{\boldsymbol{\eta}}_2) \mathbf{U}_{2i}^{\text{mis}}(\tilde{\boldsymbol{\eta}}_2)^T \right\} \\ \hat{\mathbf{I}}_{12}^{\text{mis}} &= \frac{1}{m_1} \sum_{i=1}^{m_1} \mathbf{U}_{1i}^{\text{mis}}(\tilde{\boldsymbol{\eta}}_1) \mathbf{U}_{2i}^{\text{mis}}(\tilde{\boldsymbol{\eta}}_2)^T \end{aligned}$$

where

$$\mathbf{U}_{ji}^{\text{mis}}(\tilde{\boldsymbol{\eta}}_j) = \left(\frac{Y_{ij} - \tilde{\beta}_j}{s_{ij}^2 + \tilde{\tau}_j^2}, -\frac{1}{2(s_{ij}^2 + \tilde{\tau}_j^2)} + \frac{(Y_{ij} - \tilde{\beta}_j)^2}{2(s_{ij}^2 + \tilde{\tau}_j^2)^2} \right)^T.$$

Acknowledgements

Yong Chen was supported by grant number R03HS022900 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality. Whilst undertaking this work, Richard Riley was supported by funding from an MRC Methodology Research Program Grant (MR/J013595/1). The authors want to thank Professor Geert Molenberghs for helpful comments and pointing out several key references in composite likelihoods. The authors are especially grateful to an anonymous reviewer who identified that our original REML estimation procedure could be improved. Yong Chen is grateful for the general advices from Dr. Shengfeng Cheng.

References

1. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; **7**(3):177–188.
2. Van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* 2002; **21**(4):589–624.
3. Jones A, Riley R, Williamson P, Whitehead A. Meta-analysis of individual patient data versus aggregate data from longitudinal clinical trials. *Clinical Trials* 2009; **6**(1):16–27.
4. Hand D. Evaluating diagnostic tests: the area under the roc curve and the balance of errors. *Statistics in Medicine* 2010; **29**(14):1502–1510.
5. Jackson D, Riley R, White I. Multivariate meta-analysis: potential and promise. *Statistics in Medicine* 2011; **30**(20):2481–2498.
6. Borenstein M, Hedges L, Higgins J, Rothstein H. *Introduction to Meta-Analysis*. Wiley Online Library: The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom, 2009.
7. Hartung J, Knapp G, Sinha BK. *Statistical Meta-Analysis with Applications*, Vol. 738. Wiley-Interscience: Hoboken, New Jersey, USA, 2011.
8. Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. *Statistics in Medicine* 1997; **16**(17):1965–1982.
9. Berkey C, Antczak-Bouckoms A, Hoaglin D, Mosteller E, Pihlstrom B. Multiple-outcomes meta-analysis of treatments for periodontal disease. *Journal of Dental Research* 1995; **74**(4):1030–1039.
10. Nam I, Mengersen K, Garthwaite P. Multivariate meta-analysis. *Statistics in Medicine* 2003; **22**(14):2309–2333.
11. Wei Y, Higgins J. Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. *Statistics in Medicine* 2013; **32**(7):1191–1205.
12. Riley R, Abrams K, Sutton A, Lambert P, Thompson J. Bivariate random-effects meta-analysis and the estimation of between-study correlation. *BMC Medical Research Methodology* 2007; **7**(1):15.
13. Riley R, Thompson J, Abrams K. An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics* 2008; **9**(1):172–186.
14. Lindsay B. Composite likelihood methods. *Contemporary Mathematics* 1988; **80**(1):221–39.
15. Cox D, Reid N. A note on pseudolikelihood constructed from marginal densities. *Biometrika* 2004; **91**(3):729–737.
16. Varin C, Reid N, Firth D. An overview of composite likelihood methods. *Statistica Sinica* 2011; **21**(1):5–42.
17. Kent JT. Robust properties of likelihood ratio tests. *Biometrika* 1982; **69**(1):19–27.
18. Molenberghs G, Verbeke G. *Models for Discrete Longitudinal Data*. Springer: New York, NY 10013, USA, 2005.
19. Chandler R, Bate S. Inference for clustered data using the independence loglikelihood. *Biometrika* 2007; **94**(1):167–183.
20. Huber P. The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, University of California, Berkeley, CA, USA, 1967, 221–33.
21. White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society* 1980; **48**:817–838.
22. White I. Multivariate random-effects meta-regression: updates to mvmeta. *Stata Journal* 2011; **11**(2):255–270.
23. Gasparrini A, Gasparrini MA. Package mvmeta, 2014.
24. Riley R, Abrams K, Lambert P, Sutton A, Thompson J. An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes. *Statistics in Medicine* 2007; **26**(1):78–97.
25. Little RJ, Rubin DB. *Statistical Analysis with Missing Data*, Vol. 4. Wiley: New York, 1987.
26. Sohn S. Multivariate meta analysis with potentially correlated marketing study results. *Naval Research Logistics (NRL)* 2000; **47**(6):500–510.
27. Riley R. Multivariate meta-analysis: the effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2009; **172**(4):789–811.
28. Baade P, Youlden D, Krnjacki L. International epidemiology of prostate cancer: geographical distribution and secular trends. *Molecular Nutrition & Food Research* 2009; **53**(2):171–184.
29. Sasse A, Sasse E, Carvalho A, Macedo L. Androgenic suppression combined with radiotherapy for the treatment of prostate adenocarcinoma: a systematic review. *BMC Cancer* 2012; **12**(1):1–11.
30. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Statistics in Medicine* 1998; **17**(24):2815–2834.
31. Getz G, Reardon C. Paraoxonase, a cardioprotective enzyme: continuing issues. *Current Opinion in Lipidology* 2004; **15**(3):261–267.
32. Zhao Y, Ma Y, Fang Y, Liu L, Wu S, Fu D, Wang X. Association between pon1 activity and coronary heart disease risk: a meta-analysis based on 43 studies. *Molecular Genetics and Metabolism* 2011; **105**:141–148.
33. Epstein FH, Diaz MN, Frei B, Vita JA, Keaney JF, Jr. Antioxidants and atherosclerotic heart disease. *New England Journal of Medicine* 1997; **337**(6):408–416.
34. Sterne JA, Egger M, Smith GD. Investigating and dealing with publication and other biases in meta-analysis. *Bmj* 2001; **323**(7304):101–105.
35. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 1996; **91**(434):444–455.
36. Thompson JR, Minelli C, Abrams KR, Tobin MD, Riley RD. Meta-analysis of genetic studies using mendelian randomization—a multivariate approach. *Statistics in Medicine* 2005; **24**(14):2241–2254.
37. Liang K, Zeger S. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; **73**(1):13–22.
38. Lin DY. Cox regression analysis of multivariate failure time data: the marginal approach. *Statistics in Medicine* 1994; **13**(21):2233–2247.

39. Turner R, Omar R, Yang M, Goldstein H, Thompson S. A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine* 2000; **19**:3417–3432.
40. Higgins J, Whitehead A, Turner R, Omar R, Thompson S. Meta-analysis of continuous outcome data from individual patients. *Statistics in Medicine* 2001; **20**(15):2219–2241.
41. Piedbois P, Buyse M. Meta-analyses based on abstracted data: a step in the right direction, but only a first step. *Journal of Clinical Oncology* 2004; **22**(19):3839–3841.
42. Smith C, Williamson P, Marson A. Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Statistics in Medicine* 2005; **24**(9):1307–1319.
43. Clarke M, Stewart L, Tierney J, Williamson P. Individual patient data meta-analyses compared with meta-analyses based on aggregate data. *The Cochrane Library* 2008.
44. Riley R, Lambert P, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ. British Medical Journal* 2010; **340**(7745):521–525.
45. Piedbois P, Buyse M. Meta-analysis based on individual patient data: example of advanced colorectal cancer. *Recherche en Soins Infirmiers* 2010; **101**:25–28.
46. Di Leo A, Desmedt C, Bartlett JM, Piette F, Ejlertsen B, Pritchard KI, Larsimont D, Poole C, Isola J, Earl H, Mouridsen H, O'Malley FP, Cardoso F, Tanner M, Munro A, Twelves CJ, Sotiriou C, Shepherd L, Cameron D, Piccart MJ, Buyse M, HER2/TOP2A Meta-analysis Study Group. Her2 and top2a as predictive markers for anthracycline-containing chemotherapy regimens as adjuvant treatment of breast cancer: a meta-analysis of individual patient data. *The Lancet Oncology* 2011; **12**(12):1134–1142.
47. Rothstein H, Sutton A, Borenstein M. *Publication Bias in Meta-Analysis*. Wiley Online Library: The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom, 2005.
48. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000; **56**(2):455–463.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.