Efficient two-step multivariate random effects meta-analysis of individual participant data for longitudinal clinical trials using mixed effects models

Hisashi Noma 1*, Kazushi Maruo 2, Masahiko Gosho 2, Stephen Z. Levine 3, Yair Goldberg 4, Stefan Leucht 5 and Toshi A. Furukawa 6

Abstract

Background: Mixed effects models have been widely applied in clinical trials that involve longitudinal repeated measurements, which possibly contain missing outcome data. In meta-analysis of individual participant data (IPD) based on these longitudinal studies, joint synthesis of the regression coefficient parameters can improve efficiency, especially for explorations of effect modifiers that are useful to predict the response or lack of response to particular treatments.

Methods: In this article, we provide a valid and efficient two-step method for IPD meta-analyses using the mixed effects models that adequately addresses the between-studies heterogeneity using random effects models. The two-step method overcomes the practical difficulties of computations and modellings of the heterogeneity in the one-step method, and enables valid inference without loss of efficiency. We also show the two-step method can effectively circumvent the modellings of the between-studies heterogeneity of the variance-covariance parameters and provide valid and efficient estimators for the regression coefficient parameters, which are the primary objects of interests in the longitudinal studies. In addition, these methods can be easily implemented using standard statistical packages, and enable synthesis of IPD from different sources (e.g., from different platforms of clinical trial data sharing systems).

Results: To assess the proposed method, we conducted simulation studies and also applied the method to an IPD meta-analysis of clinical trials for new generation antidepressants. Through the numerical studies, the validity and efficiency of the proposed method were demonstrated.

Conclusions: The two-step approach is an effective method for IPD meta-analyses of longitudinal clinical trials using mixed effects models. It can also effectively circumvent the modellings of the between-studies heterogeneity of the variance-covariance parameters, and enable efficient inferences for the regression coefficient parameters.

Keywords: Individual participant data meta-analysis, Mixed effects models, Multivariate random effects model, Longitudinal studies, Heterogeneity
Background
In clinical trials for drug developments, the outcome variables often involve longitudinal repeated measurements. In these trials, the primary statistical analyses usually compare the responses from experimental treatments with a control treatment at the end of a follow-up period. However, during the follow-up period, dropouts or missing data often arise, and may seriously influence the validity and precision of the statistical inference. Conventionally, simple imputation methods such as LOCF (last observation carried forward) had been adopted to handle these problems, but these ad-hoc approaches have been known to provide biased results and possibly produce erroneous conclusions [1, 2]. Recently, regulatory guidelines concerning missing data problems in clinical trials have been issued [1, 3, 4], and these invalid practices are rapidly being abandoned by more sophisticated methods. The mixed effects model [5, 6], especially known as the MMRM (mixed effects model for repeated measures) [7], is one of these methods and has been widely adopted for primary analyses of recent clinical trials in drug developments, which allows for valid statistical inference based on the direct likelihood approach under the missing at random (MAR) assumptions [5, 6].

The mixed effects model is increasingly used in the context of evidence synthesis as well. Recent advances of clinical trials data sharing systems (e.g., ClinicalStudyDataRequest.com, the YODA project) have enabled evidence synthesis of large individual participant level datasets and explorations of effect modifiers that are useful to predict the response or lack of response to particular treatments, because the powers of treatment-by-covariate interaction tests increasingly gain using rich statistical information from multiple clinical trials [8, 9]. However, there are additional complexities in the individual participant data (IPD) meta-analysis to synthesize longitudinal trial datasets based on the mixed effects models, due to the longitudinal nature of the datasets. The most important issue is about addressing correlations among different time points in the repeated measurements and the between-studies heterogeneity. In the IPD meta-analyses of ordinary univariate regression models such as analysis of covariance (ANCOVA) models, the synthesizing analyses can be simply implemented by multivariate-meta-analysis [10, 11], but for the mixed effects models, the between-studies heterogeneity should be adequately considered not only for the fixed effects coefficients, but also for the variance-covariance parameters, simultaneously. In addition, to use the statistical information among multiple trials maximally, we should consider the correlations among the estimators of model parameters to gain efficiency. In the IPD meta-analyses using mixed effects models, Ishak et al. [12] considered to aggregate only summarized data of each time point, but their methods do not use the correlation information. Also, Jones et al. [13] proposed to implement IPD meta-analyses of longitudinal data using mixed effects models, but their methods only assumed fixed effects models for the regression coefficient parameters, and the heterogeneity among the studies was not considered. In existing IPD meta-analyses, these problems have also not been addressed and fixed effects models that assume no heterogeneity have been usually used.

In this article, we provide a valid and efficient method for IPD meta-analyses using the mixed effects models that adequately address the between-studies heterogeneity using random effects models. We especially show the ordinary one-step methods have practical difficulties on computations and modellings of the heterogeneity, and we provide a two-step approach as an alternative effective procedure to overcome these problems. We will show that the two-step method can effectively circumvent the modellings of the between-studies heterogeneity of the variance-covariance parameters and provides valid and efficient estimators for the regression coefficient parameters, which are the primary objects of interests. In addition, these methods can be easily implemented using standard statistical software, and have advantages to synthesize IPD from different sources of clinical trial datasets (e.g., from different platforms of clinical trial data sharing systems). We also demonstrate the effectiveness of these methods via simulation studies and a real data application to an IPD meta-analysis of clinical trials for new generation antidepressants [14].

Methods
Mixed effects models for longitudinal data
We consider a meta-analysis of \( N \) longitudinal clinical trials with \( n_i \) individual participants within the \( i \)th trial (\( i = 1, \ldots, N \)). Let \( Y_{ij} \) be continuous repeated measurement outcome of \( j \)th time point of participant \( j \) in trial \( i \), (\( j = 1, \ldots, n_i ; t = 1, \ldots, T \)), possibly to be missing at some points. Note that the time points planned to be measured the outcome variables are uncommon across the \( N \) trials, we can formally regard these unmeasured time point data as ‘missing’ without loss of generality.

At first, we discuss linear mixed effect models [5, 6] for the analysis of individual trials separately,

\[
Y_{ij} = X_{ij} \beta_i + Z_{ij} b_j + \epsilon_{ij}
\]  

where \( Y_{ij} = (Y_{ij1}, \ldots, Y_{ijT})^T \) is a \( T \)-vector of repeated measurements for participant \( j \) in trial \( i \), \( X_{ij} = (x_{ij1}, \ldots, x_{ijT})^T \) is a \( T \times p \) matrix of fixed effect explanatory covariates, \( \beta_i \) is a \( p \)-vector of fixed regression coefficients, \( Z_{ij} = (z_{ij1}, \ldots, z_{ijq})^T \) is a \( T \times q \) matrix of random-effect covariates of participant \( j \) in trial \( i \) (usually a subset of \( X_{ij} \)), and \( b_j \) is a \( q \)-vector of random effects, assumed to follow a multivariate normal distribution \( b_j \sim \text{MVN}(0, D_j) \), where \( D_j \) is
the $q \times q$ covariance matrix of the random effects distribution. $\epsilon_{ij}$ is a $T$-vector of random error, assumed to follow a multivariate normal distribution $\epsilon_{ij}\sim MVN(0, \Sigma)$, where $\Sigma$ is the $T \times T$ covariance matrix of the random error distribution. The random effects $b_{ij}$ and the error $\epsilon_{ij}$ are assumed to be independent, and thus the outcome variables among different participants are also independent. $Y_{ij}$ marginally follows $MVN(X_{ij}\beta_i,V_{ij})$, where $V_{ij} = Z_{ij}DZ_{ij}' + \Sigma_i$. The model parameters can be estimated by maximum likelihood (ML) or restricted maximum likelihood (REML) methods [6]; we suppose to use the REML estimator here owing to its theoretical property [6, 15]. If missing outcome data are available, these likelihood-based inference methods provide valid statistical inference results under MAR assumptions [5, 6], and thus the linear mixed models have been widely adopted in longitudinal data analyses. In addition, recently, owing to the developments of regulatory guidelines concerning missing data treatments in clinical trials [3, 4], the mixed effects models have been increasingly popular. Especially, an alternative parameterized model named as MMRM [7] has been adopted for primary analyses of many clinical trials, which directly models the marginal distribution of $Y_{ij}$'s and parameterize the covariance matrix $V_{ij}$ not specifying $D$ and $\Sigma$. The primary objects of interests are usually the fixed effects parameters, and we have rather little interests in the variance-covariance parameters in these clinical trials [16, 17]. Thus, through the re-parameterization to the marginal covariance matrix $V_{ij}$, possible model misspecifications concerning the random effects structures might be effectively circumvented in the statistical inference. The procedures of statistical analyses and their theoretical properties are identical between the two parameterized models, so we conduct the following discussions using the former model (1) in this article.

**IPD meta-analyses based on the mixed effects models**

Based on the mixed effects model (1), IPD meta-analysis can be implemented in a straightforward manner. Jones et al. [13] proposed a fixed effects model with an assumption the regression coefficients $\beta_i$'s are common across the $N$ trials using both of one-step and two-step approaches. However, the between-studies heterogeneity is an essential issue to be addressed in modern evidence synthesis researches [9], so the random effects models are preferred in general [9, 18]. While, there are several crucial problems to be addressed in applying them to the mixed effects models for IPD meta-analysis:

1. When we consider the heterogeneity across studies, we should adequately address the heterogeneities for both of the fixed effects and variance-covariance parameters in the synthesizing analyses. A straightforward approach would be to model these parameters as random effects, the joint random effects model for $b_{ij}$ and the components of $V_{ij}$. This strategy would also require complicated computations that cannot be implemented in standard statistical packages, and we should completely specify the statistical model without any misspecifications in order to assure the consistency of parameter estimation for fixed effects. Besides, we have little interests in the heterogeneities of variance-covariance parameters, i.e., the components of $V_{ij}$ in the longitudinal clinical trials [16, 17].

2. In most of meta-analyses of longitudinal trials, the time-points to be planned to collect outcome measurements are not common across studies. In the mixed effects model analyses, the time variables are often treated as factor variables (especially for the MMRM analyses [19]), and then all of the outcome variables at several time points are possibly missing (unmeasured) in some studies. In these cases, when we conduct a one-step random meta-analysis using standard statistical packages (e.g., PROC MIXED in SAS, lme4 [20] for R), the statistical model can be partially not identifiable, i.e., we cannot obtain the model estimates.

Given these problems, there are practical difficulties in implementing ordinary one-step IPD meta-analyses. Note that for the first problem, Jones et al. [13] proposed another approach to modelling the variance-covariance parameters as different parameters across studies in their fixed effects model framework, but these approaches are not implementable using standard packages. In addition, this problem has not been usually considered in many existing IPD meta-analyses, possibly due to its practical difficulties, and common variance-covariance structures have usually been assumed. Besides, the second issue can be resolved by making an original program that maximises the (restricted) likelihood function that assumes different regression functions across the studies (and possibly different random effects structures), but it requires usually large efforts in making the complicated original programs in case-by-case practices.

**Two-step marginal multivariate random effects analyses**

To circumvent the complicated modelling problems and computational difficulties, we provide an alternative two-step approach for IPD meta-analyses based on the mixed effects models. Our proposition is to divide the model parameters into two components, fixed effects parameters and variance-covariance parameters, and to conduct synthesis analyses after obtaining study-specific aggregated statistics separately. Because we have little interests in the random effects in these clinical trials and the mixed effects models are typically designed to assess fixed effects [16, 17], here we specifically focus on the synthesis analyses for the fixed effects parameters. As noted later, the same method can be adapted to the synthesis analyses.
of the variance-covariance parameters. The computations are easily implementable, but we can show the resultant inference is valid and efficient, and that its theoretical properties do not depend on any heterogeneous structures of within-studies variance-covariance parameters.

We describe the procedure as follows.

**Step-1.** Analyse the IPD from $N$ longitudinal trials separately, using the mixed effects models (1), and obtain the REML estimates of the model parameters $(\hat{\beta}_i, \hat{D}_i, \hat{\Sigma}_i)$, $i = 1, 2, \ldots, N$. In addition, obtain the variance-covariance matrix estimates of $\hat{\beta}_i, \hat{S}_i(\hat{\beta}_i)$ for $i = 1, 2, \ldots, N$.

**Step-2.** Conduct synthesizing analyses for the regression coefficient parameters $(\hat{\beta}_i, \hat{S}_i(\hat{\beta}_i))$ that are the primary objects of interests in these meta-analyses, dropping statistical information about the within-studies covariance parameters $(\hat{D}_i, \hat{\Sigma}_i)$.

$$\hat{\beta}_i \sim \text{MVN}\left(\hat{\beta}_i, \hat{S}_i(\hat{\beta}_i)\right)$$

$$\beta_i \sim \text{MVN}(\mu, \Omega)$$

Inference concerning the model parameters $(\mu, \Omega)$ can be conducted using the log likelihood or the restricted log likelihood functions [21, 22],

$$\ell(\mu, \Omega) = \frac{1}{2} \sum_{i=1}^{N} \left( \log |\Omega + \hat{S}_i(\hat{\beta}_i)| + (\beta_i - \mu)^T W_i (\beta_i - \mu) + p\log 2\pi \right),$$

$$\ell_{RL}(\mu, \Omega) = \ell(\mu, \Omega) - \frac{1}{2} \sum_{i=1}^{N} W_i + \frac{1}{2} p \log 2\pi$$

where $W_i = (\Omega + \hat{S}_i(\hat{\beta}_i))^{-1}$.

The computation of Step-1 is straightforwardly conducted using standard packages for linear mixed models, e.g., PROC MIXED in SAS, lme4 [20] for R. Also, Step-2 can be also simply implemented using existing packages for multivariate meta-analyses, e.g., mvmeta for Stata [23, 24] and R [25].

The resultant estimators of $(\mu, \Omega)$ have the following properties.

**Remark A**

There are no efficiency losses regardless of dropping the within-studies variance-covariance information of $(\hat{D}_i, \hat{\Sigma}_i)$ in the multivariate meta-analysis models (2) for inferences of $(\mu, \Omega)$, i.e., the ML and REML estimators of $(\mu, \Omega)$ on the multivariate meta-analysis models (2) are the most efficient estimators, and their efficiency bound is equivalent to that of the one-step estimators for the completely specified correct model.

This property is assured because the ML and REML estimators of $\hat{\beta}_i$ and $(\hat{D}_i, \hat{\Sigma}_i)$ are orthogonal [6] in the sense of Cox and Reid [26], i.e., the non-diagonal submatrices of the information matrix of mixed model (1) are zero-matrices. Therefore, the statistical information of $(\hat{D}_i, \hat{\Sigma}_i)$ does not contribute to synthesizing $\hat{\beta}_i$’s when we conduct one-step IPD meta-analyses under the completely specified correct model. Thus, fully efficient estimator of $(\mu, \Omega)$ can be obtained via the two-step method above, even if the statistical information of $(\hat{D}_i, \hat{\Sigma}_i)$ are dropped. This property is confirmed in the simulation studies in Section 3.

**Remark B**

When the heterogeneity for the variance-covariance parameters of $(\hat{D}_i, \hat{\Sigma}_i)$ is modelled by a random effects model $G(D, \Sigma)$, we can obtain the consistent and fully efficient estimator of $(\mu, \Omega)$ using the two-step method under any distribution form of $G(D, \Sigma)$.

Note this property is important that it assures the validity and optimality of the two-step method without modelling the random effects model $G(D, \Sigma)$ explicitly, even if there exists substantial heterogeneity for the variance-covariance parameters. This property is assured because the ML and REML estimators of $\hat{\beta}_i$ and $(\hat{D}_i, \hat{\Sigma}_i)$ are orthogonal as the same arguments with Remark A. Besides, if we adopt the one-step approach, we should completely specify the joint random effects distributions of $\hat{\beta}_i$ and $(\hat{D}_i, \hat{\Sigma}_i)$ correctly to assure the consistency. These parametric assumptions are also not substantial objects of interests in the evaluations of treatment effects. Through the two-step approach, we can effectively circumvent the unnecessary additional complicated assumptions as well as the computational efforts in practices.

In addition, in IPD meta-analyses of longitudinal clinical trials, the time-points of outcome measurements are possibly uncommon across studies. As mentioned regarding the second issue in the previous section, when the time variables are often treated as factor variables (e.g., for the typical MMRM analyses [19]), the one-step approach has practical and computational difficulties.

**Remark C**

For synthesis of longitudinal clinical trials in which time-points of outcome measurements are possibly uncommon, we can apply the two-step approach formally via dropping the corresponding time-points variables in Step-1. Then, in Step-2, the corresponding components of $\hat{\beta}_i$ should be treated as missing outcomes variables.

The missing problems are common not only in two-step IPD meta-analysis, but in various multivariate
meta-analysis, and have been extensively discussed in previous papers [21, 22]. Accordingly, in many software packages for multivariate meta-analysis [23–25], the standard modules can treat these partially missing outcomes adequately. Usually, these packages conduct available data analyses based on direct likelihood methods [23, 24].

In addition, in some cases, we can obtain the IPD for multiple clinical trials from different platforms of clinical trial databases. Recently, the clinical trial data sharing has been increasingly advanced [27, 28], and researchers can access the IPD of clinical trials for scientific research purposes, e.g., via ClinicalStudyDataRequest.com. However, in many platforms, the users cannot download the datasets and can only analyse on the remote accesses on the platform, in which case the IPD can only be analysed separately. In such cases, the one-step IPD meta-analyses cannot be conducted, but the two-step approaches can still be applied as valid and efficient procedure. In analysing longitudinal trial datasets, the two-step marginal synthesis method would be an effective strategy for evaluating treatment effects and interactions of the candidates of effect modifiers.

Results
Simulation studies
We conducted simulation studies to evaluate the validity and efficiency of the proposed method under practical IPD meta-analysis scenarios. In this context, validity corresponds to the accuracy of point estimation, standard error (SE) evaluation, and the coverage rate of the confidence interval (CI). In addition, the efficiency corresponds to the precision of estimation and the power of statistical test.

Simulation 1
First, we conducted simulations to evaluate the efficiency of two-step marginal multivariate meta-analysis approach. In particular, we would evaluate the equivalence of the two-step marginal method and the fully efficient one-step IPD meta-analysis. Simulation data were generated as three time-points longitudinal data, following the mixed effects model involving two-way interactions,

\[
Y_{ijt} = b_{ij} + \beta_{1t} + \beta_{2t}treat_{ij} + \beta_{3t}time_{2ij} + \beta_{4t}time_{3ij} + \beta_{5t}treat_{ij} \times time_{2ij} + \beta_{6t}time_{3ij} \times z_{ij} \times \epsilon_{ijt}
\]

where \(b_{ij}\) is a random intercept for the \(i\)th participant of the \(j\)th trial following \(N(0,4.5^2)\) and \(\epsilon_{ijt}\) is the random error for the \(t\)th visit of \(j\)th participant of \(i\)th trial following \(N(0,3.2^2)\), \(treat_{ij}\) is a binary variable indicating treatment group. \(time_{2ij}\) and \(time_{3ij}\) are dummy variables indicating time-points of 2nd and 3rd visits, respectively (reference is the 1st visit). The parameter settings were mimicked to the Furukawa et al. [14]'s example in Section 4. \(z_{ij}\) is a covariate variable that is a candidate to be an effect modifier, which is a measurement that is associated with response or lack of response to a particular therapy. The two-way interaction terms represent the potential usefulness of \(z_{ij}\) as an effect modifier. Following the mixed effects model (3), we generated \(N\) longitudinal clinical trial datasets, for \(N = 5, 10, 15, \text{and} 20\). In addition, the number of participants of each trial was generated from a discrete uniform distribution on \([50, 500]\).

In these first simulation studies, the primary purpose is to assess the validity and efficiency of the two-step approach provided in Section 2. Also, we would evaluate the equivalence of the two-step marginal method and the fully efficient one-step IPD meta-analysis. At first, to circumvent the discussions to be complicated, we assumed there was no heterogeneity for \(\beta_i\) and \((D_i, \Sigma_i)\) among the \(N\) trials, which is a special case of the fixed effects models considered by Jones et al. [13]. In other words, all \(\beta_i\) \((i = 1, \ldots, N)\) are equal to the grand mean \(\mu\).

The parameter settings of \(\mu\) were also mimicked to the Furukawa et al. [14], \(\mu_1 = -0.04, \mu_2 = -0.14, \mu_3 = 1.56, \mu_4 = 0.83, \mu_5 = -0.06, \mu_6 = 0.14, \mu_7 = 0.15, \mu_8 = 0.05, \mu_{10} = -0.06\). Under these settings, we conducted 10,000 simulations, and we analyzed the simulated datasets using the one-step method by the correctly specified mixed effects model and the two-step multivariate meta-analysis approach in Section 2.

The results are presented in Table 1. We presented empirical means and SE of the regression coefficient estimates, means of the SE estimates (SE), empirical coverage rates of the 95% CIs, and empirical powers for the tests of individual regression coefficients. The contexts of the results are too rich, so we present the ones for selected coefficients, \(\mu_2, \mu_4, \mu_5, \mu_7, \mu_9, \mu_{10}\). The results for the remaining parameters were similar. These results consistently showed that the fully efficient one-step approach was equivalent to the marginal two-step approach under all of the scenarios. Both the mean and SE of 10,000 estimates were mostly the same, and the SE estimate could accurately estimate the actual SE of the estimates. In addition, coverage rates of 95% CIs were accurate and the power of the two approaches were equivalent. As a whole, the two-step approach was shown to provide accurate results without loss of efficiency, regardless of dropping the variance-covariance information. The efficiency was also retained for the interaction tests.

Simulation 2
Second, we conducted other simulations to evaluate the validity of the two-step random effects multivariate
meta-analysis approaches under substantial heterogeneity of \( \beta \) as well as \((D_i, \Sigma_i)\), among \( N \) longitudinal trials. We considered the mixed effects models (3) for individual trials, and additionally considered heterogeneity of \( \beta_i \) and \((D_i, \Sigma_i)\). We considered a compound symmetry structure for \( \Omega \), which assumes a common variance among all of the components of \( \beta \), \( \beta_{i2} \sim N(\mu_2, \tau^2), k = 1, \ldots, 10 \). We assumed no correlations among \( \beta_{i2} \), and also assumed heterogeneity among the random effects distributions of \( b_{ij} \sim N(0, \zeta^2) \), where \( \zeta \) is generated from a continuous uniform distribution on \([2.5, 9.0]\). We considered two settings: (a) outcome

### Table 1: Results for the simulation 1: fully efficient one-step approach and the two-step marginal multivariate meta-analysis approach were compared

| \( \mu_2 = 0.14 \) | N = 5 \( \mu_2 = 0.14 \) | N = 10 \( \mu_2 = 0.14 \) | N = 15 \( \mu_2 = 0.14 \) | N = 20 \( \mu_2 = 0.14 \) |
|---|---|---|---|---|
| One-step | Two-step | One-step | Two-step | One-step | Two-step | One-step | Two-step |
| Mean | 0.135 | 0.135 | 0.146 | 0.146 | 0.135 | 0.136 | 0.130 | 0.130 |
| SE | 0.866 | 0.867 | 0.612 | 0.613 | 0.495 | 0.496 | 0.434 | 0.435 |
| \( SE \) | 0.867 | 0.866 | 0.608 | 0.606 | 0.495 | 0.494 | 0.428 | 0.427 |
| Coverage Rate | 0.949 | 0.950 | 0.948 | 0.947 | 0.949 | 0.947 | 0.950 | 0.948 |
| Power | 0.052 | 0.053 | 0.056 | 0.058 | 0.058 | 0.059 | 0.071 | 0.070 |
| \( \mu_4 = 0.83 \) | Mean | 0.827 | 0.826 | 0.836 | 0.836 | 0.828 | 0.829 | 0.825 | 0.824 |
| SE | 0.415 | 0.416 | 0.287 | 0.287 | 0.236 | 0.236 | 0.206 | 0.205 |
| \( SE \) | 0.410 | 0.410 | 0.287 | 0.287 | 0.234 | 0.234 | 0.202 | 0.202 |
| Coverage Rate | 0.950 | 0.950 | 0.951 | 0.951 | 0.948 | 0.947 | 0.939 | 0.941 |
| Power | 0.533 | 0.533 | 0.825 | 0.825 | 0.936 | 0.937 | 0.978 | 0.978 |
| \( \mu_5 = -0.46 \) | Mean | -0.461 | -0.461 | -0.460 | -0.460 | -0.460 | -0.460 | -0.461 | -0.461 |
| SE | 0.032 | 0.032 | 0.023 | 0.023 | 0.018 | 0.018 | 0.016 | 0.016 |
| \( SE \) | 0.032 | 0.032 | 0.022 | 0.022 | 0.018 | 0.018 | 0.016 | 0.016 |
| Coverage Rate | 0.949 | 0.948 | 0.949 | 0.947 | 0.953 | 0.955 | 0.951 | 0.949 |
| Power | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| \( \mu_7 = 0.14 \) | Mean | 0.139 | 0.139 | 0.141 | 0.141 | 0.143 | 0.143 | 0.144 | 0.144 |
| SE | 0.178 | 0.178 | 0.122 | 0.122 | 0.100 | 0.100 | 0.088 | 0.088 |
| \( SE \) | 0.176 | 0.176 | 0.123 | 0.123 | 0.100 | 0.100 | 0.087 | 0.087 |
| Coverage Rate | 0.948 | 0.947 | 0.949 | 0.949 | 0.949 | 0.949 | 0.940 | 0.942 |
| Power | 0.127 | 0.128 | 0.211 | 0.215 | 0.295 | 0.297 | 0.382 | 0.381 |
| \( \mu_9 = 0.05 \) | Mean | 0.050 | 0.050 | 0.050 | 0.050 | 0.050 | 0.050 | 0.050 | 0.050 |
| SE | 0.020 | 0.020 | 0.014 | 0.014 | 0.011 | 0.011 | 0.010 | 0.010 |
| \( SE \) | 0.020 | 0.020 | 0.014 | 0.014 | 0.011 | 0.011 | 0.010 | 0.010 |
| Coverage Rate | 0.947 | 0.947 | 0.951 | 0.951 | 0.949 | 0.952 | 0.941 | 0.940 |
| Power | 0.734 | 0.735 | 0.947 | 0.947 | 0.995 | 0.995 | 0.999 | 0.999 |
| \( \mu_{10} = -0.06 \) | Mean | -0.060 | -0.060 | -0.060 | -0.060 | -0.060 | -0.060 | -0.059 | -0.059 |
| SE | 0.042 | 0.042 | 0.030 | 0.030 | 0.024 | 0.024 | 0.021 | 0.021 |
| \( SE \) | 0.042 | 0.042 | 0.029 | 0.029 | 0.024 | 0.024 | 0.021 | 0.021 |
| Coverage Rate | 0.949 | 0.949 | 0.948 | 0.947 | 0.949 | 0.950 | 0.946 | 0.945 |
| Power | 0.301 | 0.302 | 0.540 | 0.543 | 0.700 | 0.701 | 0.811 | 0.811 |
variables are measured on all three time-points, and (b) outcome variables are measured on two time-points for 40\% of trials and on three time-points for the remaining 60\% trials.

Under these conditions, the two-step approach can be easily implemented using standard statistical packages. Again, note that the two-step method does not require modelling of the heterogeneity structures of (D\_i, \Sigma\_i), and the validity and efficiency of the inference are assured under any heterogeneous structures of (D\_i, \Sigma\_i). Besides, to implement valid inference using one-step approach, we should adopt the correct statistical model involving the random effects distributions for both of \( \beta_i \) and (D\_i, \Sigma\_i). One ad-hoc approach uses a Bayesian hierarchical modelling involving random effects models, but it requires complicated computations using special software (e.g., OpenBUGS [29]).

The simulation results are presented in Tables 2 and 3, which correspond to the settings (a) and (b) respectively. We conducted 10,000 simulations and provided the same outputs as Table 1. We considered 6 scenarios, setting \( N = 5, 10, \) or 15 and \( r^2 = 0.10 \) or 0.20. The results also showed unbiasedness of the two-step estimates under all of the scenarios. In addition, the validities of SE estimation and CIs (inversely, the sizes of the corresponding statistical tests) could be confirmed. Again, note that we only took parametric assumptions for the random effects distribution of \( \beta_i \) and no restrictive assumptions concerning the heterogeneity of (D\_i, \Sigma\_i). The validity of the two-step method was demonstrated under the flexible and weak assumptions. The efficiency is not compared with those of the one-step methods here, but the equivalence is theoretically assured (Remark B).

Applications to IPD meta-analysis for new generation antidepressants

To illustrate our method, we analyzed the dataset from an IPD meta-analysis for new generation antidepressants by Furukawa et al. [14] The authors conducted IPD meta-analysis from four placebo-controlled, double-blinded randomized clinical trials (1482 participants) for patients undergoing acute phase treatment for major depression. The outcome variable was the change score on the 17-item Montgomery-Asberg Depression Rating Scale (MADRS). The outcome variable was the change score on the 17-item Montgomery-Asberg Depression Rating Scale (MADRS). The outcome variable was the change score on the 17-item Montgomery-Asberg Depression Rating Scale (MADRS). The outcome variable was the change score on the 17-item Montgomery-Asberg Depression Rating Scale (MADRS).

We considered the following mixed effects model on 1st, 2nd, 4th and 6th weeks involving two-way interaction terms based on the model (1),

\[
Y_{ijt} = b_{ij} + \beta_{1i1} + \beta_{1i2}treat_1 + \beta_{1i3}week_{1,ij} + \beta_{1i4}week_{2,ij} + \beta_{2i1}week_{3,ij} + \beta_{2i2}week_{4,ij} + \beta_{3i1}treat_2 \times week_{1,ij} + \beta_{3i2}treat_2 \times week_{4,ij} + \beta_{4i1}week_{1,ij} \times z_{ij} + \beta_{4i1}week_{2,ij} \times z_{ij} + \beta_{4i2}week_{4,ij} \times z_{ij} + \beta_{4i3}treat_2 \times z_{ij} + \epsilon_{ijt}
\]

where \( week_{1, ij}, week_{2, ij} \) and \( week_{4, ij} \) are dummy variables for the time points (reference: 6th week). We first attempted a one-step random effects meta-analysis that modelled all of the regression coefficients as random effects among the 4 trials, but the REML estimate was not computable using standard statistical packages. There might also be between-studies heterogeneity for the variance-covariance parameters, which was not modelled in this analysis. Thus, we provided the results of a one-step fixed effects meta-analysis only consider correlations of repeated measured outcomes among different time points within individual participants. The REML estimates, 95\% CI, and \( P \)-values are presented in Table 4.

Next, we conducted IPD meta-analyses using the two-step approach based on the REML estimates of the fixed effects coefficients of the 4 trials. In the multivariate random effects meta-analysis models, we considered several assumptions for between-studies variance-covariance structures \( \Omega \), and the fitting of the candidate models was evaluated using the Bayesian information criterion (BIC) [31]. The best fitting model was the compound symmetry model (BIC = 125.73), whereas for the fixed effects model, BIC = 316.89.

In Table 3, we also provided the results of two-step multivariate meta-analysis of the fixed effects model and the random effects model using the compound symmetry model. In the random effects model, the between-studies SD estimate was 0.047 and the between-studies correlation coefficient was 0.064. The results for the fixed effects model analyses using one-step and two-step approaches were consistent, which was expected as described in Remark A. However, note that even if there was between-studies heterogeneity for the variance-covariance parameters, the two-step approach provides valid inference results. Besides, comparing the fixed and random effects two-step analyses, the coefficient estimates were somewhat different. In addition, the SE estimates of the fixed effects coefficients were larger for the random effects model, indicating that uncertainty might be larger when using this model. Considering the simulation results, even when there is substantial heterogeneity in \( \beta_s \), the two-step random effects approach provides valid inference results. The interaction tests of \( \beta_{1i3} \) were not significant by significance level of 5\% for any of the three models,
consistent with the original analyses of Furukawa et al. [14], but these tests should be the most powerful under the assumption that the statistical models were correct.

**Discussion**

IPD meta-analysis is becoming increasingly popular in recent medical studies, especially as an effective approach for exploring effect modifiers for precision medicine [8]. Given recent trends in clinical trial data sharing [27, 28], the practical values of these methodologies are bound to increase in future studies. The efficiencies of the one-step and two-step approaches have also been discussed for IPD meta-analyses based on the conventional ANCOVA, logistic and Cox regression models [32]. Besides, as noted above, IPD

| Table 2 | Results for the simulation 2(a): outcome variables were measured on all three time-points |
|---------|---------------------------------------------------------------|
|         | N = 5              | N = 10             | N = 15             |
|         | $r^2 = 0.10$       | $r^2 = 0.20$      | $r^2 = 0.10$       | $r^2 = 0.20$      | $r^2 = 0.10$       | $r^2 = 0.20$      |
| $\mu_2 = 0.14$ |                 |                    |                    |                    |                    |                    |
| Mean    | 0.141             | 0.144              | 0.141              | 0.144              | 0.145              | 0.143              |
| SE      | 0.694             | 0.709              | 0.694              | 0.709              | 0.564              | 0.565              |
| $\hat{SE}$ | 0.680            | 0.692              | 0.680              | 0.692              | 0.553              | 0.562              |
| Coverage Rate | 0.946         | 0.947              | 0.946              | 0.947              | 0.947              | 0.950              |
| Power   | 0.061             | 0.057              | 0.061              | 0.057              | 0.059              | 0.057              |
| $\mu_4 = 0.83$ |                 |                    |                    |                    |                    |                    |
| Mean    | 0.828             | 0.831              | 0.828              | 0.831              | 0.832              | 0.833              |
| SE      | 0.306             | 0.323              | 0.306              | 0.323              | 0.252              | 0.268              |
| $\hat{SE}$ | 0.307            | 0.325              | 0.307              | 0.325              | 0.251              | 0.265              |
| Coverage Rate | 0.951         | 0.950              | 0.951              | 0.950              | 0.948              | 0.947              |
| Power   | 0.773             | 0.725              | 0.773              | 0.725              | 0.912              | 0.877              |
| $\mu_6 = -0.46$ |                 |                    |                    |                    |                    |                    |
| Mean    | -0.459            | -0.458             | -0.459             | -0.458             | -0.461             | -0.460             |
| SE      | 0.104             | 0.143              | 0.104              | 0.143              | 0.084              | 0.117              |
| $\hat{SE}$ | 0.103            | 0.143              | 0.103              | 0.143              | 0.084              | 0.117              |
| Coverage Rate | 0.944         | 0.947              | 0.944              | 0.947              | 0.949              | 0.949              |
| Power   | 0.992             | 0.890              | 0.992              | 0.890              | 1.000              | 0.976              |
| $\mu_7 = 0.14$ |                 |                    |                    |                    |                    |                    |
| Mean    | 0.142             | 0.141              | 0.142              | 0.141              | 0.138              | 0.140              |
| SE      | 0.163             | 0.193              | 0.163              | 0.193              | 0.134              | 0.157              |
| $\hat{SE}$ | 0.163            | 0.192              | 0.163              | 0.192              | 0.133              | 0.157              |
| Coverage Rate | 0.947         | 0.948              | 0.947              | 0.948              | 0.947              | 0.951              |
| Power   | 0.140             | 0.115              | 0.140              | 0.115              | 0.181              | 0.146              |
| $\mu_9 = 0.05$ |                 |                    |                    |                    |                    |                    |
| Mean    | 0.051             | 0.050              | 0.051              | 0.050              | 0.048              | 0.050              |
| SE      | 0.101             | 0.140              | 0.101              | 0.140              | 0.082              | 0.116              |
| $\hat{SE}$ | 0.100            | 0.141              | 0.100              | 0.141              | 0.082              | 0.116              |
| Coverage Rate | 0.944         | 0.945              | 0.944              | 0.945              | 0.947              | 0.947              |
| Power   | 0.091             | 0.071              | 0.091              | 0.071              | 0.091              | 0.076              |
| $\mu_{10} = -0.06$ |                |                    |                    |                    |                    |                    |
| Mean    | -0.061            | -0.058             | -0.061             | -0.058             | -0.060             | -0.061             |
| SE      | 0.106             | 0.145              | 0.106              | 0.145              | 0.087              | 0.119              |
| $\hat{SE}$ | 0.105            | 0.145              | 0.105              | 0.145              | 0.086              | 0.119              |
| Coverage Rate | 0.945         | 0.944              | 0.945              | 0.944              | 0.945              | 0.947              |
| Power   | 0.096             | 0.074              | 0.096              | 0.074              | 0.110              | 0.084              |
meta-analyses using the mixed effects models involve more complicated problems, which should address the heterogeneity of variance-covariance structures of random effects distributions. In this paper, we presented a flexible two-step method that can effectively circumvent these difficulties, and demonstrated its validity and efficiency.

In the simulation studies, the coverage rates of 95% CIs for the random effects model retained nearly at the nominal level, but previous studies showed that SE are possibly underestimated under small \(N\) settings in random effects meta-analyses, and the standard CIs can have serious undercoverage properties [33–36].

| \(\mu_2 = 0.14\) | \(N = 5\) | \(N = 10\) | \(N = 15\) |
|-----------------|----------|----------|----------|
| \(\tau^2 = 0.10\) | Mean 0.140, SE 1.026, \(\hat{SE} 0.989\), Coverage Rate 0.947, Power 0.057 | Mean 0.144, SE 0.693, \(\hat{SE} 0.681\), Coverage Rate 0.948, Power 0.058 | Mean 0.136, SE 0.702, \(\hat{SE} 0.690\), Coverage Rate 0.946, Power 0.058 |
| \(\tau^2 = 0.20\) | Mean 0.142, SE 1.004, \(\hat{SE} 1.004\), Coverage Rate 0.950, Power 0.052 | Mean 0.136, SE 0.690, \(\hat{SE} 0.690\), Coverage Rate 0.946, Power 0.058 | Mean 0.139, SE 0.559, \(\hat{SE} 0.561\), Coverage Rate 0.953, Power 0.053 |

| \(\mu_4 = 0.83\) | \(N = 5\) | \(N = 10\) | \(N = 15\) |
|-----------------|----------|----------|----------|
| \(\tau^2 = 0.10\) | Mean 0.825, SE 0.546, \(\hat{SE} 0.537\), Coverage Rate 0.948, Power 0.347 | Mean 0.823, SE 0.377, \(\hat{SE} 0.376\), Coverage Rate 0.950, Power 0.316 | Mean 0.873, SE 0.283, \(\hat{SE} 0.257\), Coverage Rate 0.907, Power 0.822 |
| \(\tau^2 = 0.20\) | Mean 0.822, SE 0.584, \(\hat{SE} 0.574\), Coverage Rate 0.950, Power 0.316 | Mean 0.828, SE 0.404, \(\hat{SE} 0.402\), Coverage Rate 0.947, Power 0.328 | Mean 0.836, SE 0.329, \(\hat{SE} 0.317\), Coverage Rate 0.947, Power 0.717 |

| \(\mu_5 = -0.46\) | \(N = 5\) | \(N = 10\) | \(N = 15\) |
|-----------------|----------|----------|----------|
| \(\tau^2 = 0.10\) | Mean -0.459, SE 0.147, \(\hat{SE} 0.143\), Coverage Rate 0.931, Power 0.874 | Mean -0.460, SE 0.103, \(\hat{SE} 0.102\), Coverage Rate 0.944, Power 0.947 | Mean -0.461, SE 0.073, \(\hat{SE} 0.073\), Coverage Rate 0.947, Power 0.945 |
| \(\tau^2 = 0.20\) | Mean -0.459, SE 0.204, \(\hat{SE} 0.200\), Coverage Rate 0.932, Power 0.625 | Mean -0.459, SE 0.143, \(\hat{SE} 0.143\), Coverage Rate 0.944, Power 0.890 | Mean -0.461, SE 0.118, \(\hat{SE} 0.117\), Coverage Rate 1.000, Power 0.974 |

| \(\mu_7 = 0.14\) | \(N = 5\) | \(N = 10\) | \(N = 15\) |
|-----------------|----------|----------|----------|
| \(\tau^2 = 0.10\) | Mean 0.145, SE 0.295, \(\hat{SE} 0.291\), Coverage Rate 0.943, Power 0.088 | Mean 0.136, SE 0.207, \(\hat{SE} 0.206\), Coverage Rate 0.947, Power 0.105 | Mean 0.138, SE 0.245, \(\hat{SE} 0.245\), Coverage Rate 0.950, Power 0.082 |
| \(\tau^2 = 0.20\) | Mean 0.145, SE 0.359, \(\hat{SE} 0.346\), Coverage Rate 0.936, Power 0.083 | Mean 0.135, SE 0.245, \(\hat{SE} 0.245\), Coverage Rate 0.950, Power 0.082 | Mean 0.141, SE 0.202, \(\hat{SE} 0.201\), Coverage Rate 0.949, Power 0.110 |

| \(\mu_9 = 0.05\) | \(N = 5\) | \(N = 10\) | \(N = 15\) |
|-----------------|----------|----------|----------|
| \(\tau^2 = 0.10\) | Mean 0.051, SE 0.185, \(\hat{SE} 0.178\), Coverage Rate 0.925, Power 0.086 | Mean 0.051, SE 0.133, \(\hat{SE} 0.128\), Coverage Rate 0.935, Power 0.084 | Mean 0.049, SE 0.093, \(\hat{SE} 0.091\), Coverage Rate 0.940, Power 0.089 |
| \(\tau^2 = 0.20\) | Mean 0.052, SE 0.263, \(\hat{SE} 0.252\), Coverage Rate 0.927, Power 0.080 | Mean 0.055, SE 0.184, \(\hat{SE} 0.181\), Coverage Rate 0.940, Power 0.084 | Mean 0.050, SE 0.151, \(\hat{SE} 0.149\), Coverage Rate 0.943, Power 0.070 |

| \(\mu_{10} = -0.06\) | \(N = 5\) | \(N = 10\) | \(N = 15\) |
|-----------------|----------|----------|----------|
| \(\tau^2 = 0.10\) | Mean -0.062, SE 0.150, \(\hat{SE} 0.147\), Coverage Rate 0.934, Power 0.084 | Mean -0.060, SE 0.106, \(\hat{SE} 0.105\), Coverage Rate 0.942, Power 0.074 | Mean -0.060, SE 0.075, \(\hat{SE} 0.074\), Coverage Rate 0.943, Power 0.118 |
| \(\tau^2 = 0.20\) | Mean -0.063, SE 0.205, \(\hat{SE} 0.203\), Coverage Rate 0.937, Power 0.079 | Mean -0.060, SE 0.146, \(\hat{SE} 0.145\), Coverage Rate 0.943, Power 0.074 | Mean -0.059, SE 0.118, \(\hat{SE} 0.118\), Coverage Rate 0.948, Power 0.083 |

Table 3 Results for the simulation 2(b): outcome variables were measured on two time-points for 40% trials and on three time-points for 60% trials.
Conclusions

The two-step approach is an effective method for synthesizing the regression coefficient parameters efficiently for IPD meta-analyses of longitudinal clinical trials. Using this method, modelling of heterogeneities of the within-studies variance-covariance parameters can be effectively circumvented without additional parametric assumptions, which are not of primary interests in the analyses of these clinical trials [16, 17, 19]. If we adopt the one-step approach to modelling the heterogeneity, complicated modelling is required and the validity is possibly violated if the adopted model is misspecified. Based on the relaxed assumptions of the two-step approach, we can implement efficient inferences for the treatment effects and interactions. In addition, the two-step approach has advantages for implementations because it can be easily performed using standard statistical packages [23–25]. It can also be effectively applied to IPD meta-analyses using clinical trials data sharing systems in which individual trials can only be analyzed separately.

Abbreviations

ANOVA: analysis of variance; BIC: Bayesian information criterion; CI: confidence interval; HRSD: Hamilton Rating Scale for Depression; IPD: individual participant data; MADRS: Montgomery-Asberg Depression Rating Scale; MAR: missing at random; ML: maximum likelihood; MMRM: mixed effects model for repeated measures; REML: restricted maximum likelihood; SE: standard error

Acknowledgements

Not applicable.

Funding

This study was supported by Grant-in-Aid from the Japan Agency for Medical Research and Development (Grant number: JP17dk0307072) and Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (Grant numbers: JP15K15954, JP17K19808) in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.
Availability of data and materials
The original datasets were available upon agreements between this research group and the sponsor companies, so are not publicly available.

Authors’ contributions
HN, KM, MG contributed to the developments of methods. HN, KM, TAF contributed to the acquisition of data. HN, KM, MG, SZL, YG, SL, TAF contributed to the analysis and interpretation of data. HN drafted the manuscript. All authors reviewed, read and approved the final version to be submitted.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
HN has received lecture fees from Boehringer Ingelheim and Kyowa Hakko Kirin, and research support from Kyowa Hakko Kirin and GSK. He is an associate editor of this journal, MG received lecture and/or consultant fees from Daichii Sanyo, Ferring Pharmaceuticals, Novartis, and Tiho, received travel fees from Takeda, and received manuscript fees from Kowa; SZL received honoraria or research support or consultancy fees or travel support from Shire Pharmaceuticals, F. Hoffmann-La Roche and Eli Lilly, SL has received honoraria or support from LB Pharma, Lundbeck, Otsuka, TEVA, Geodon Richter, Recordati, LTS Lohmann, Boehringer Ingelheim, and for lectures from Janssen, Lilly, Lundbeck, Otsuka, SanofiAvventis and Servier; TAF has received lecture fees from Janssen, Price MJ, Jackson D, Wardle M, Guayffier F, Wang J, Staessen JA, White IR. Multivariate meta-analysis using individual participant data. Res Synth Methods. 2015;6(2):157–74.

Jones AP, Riley RD, Williamson PR, Whitehead A. Meta-analysis of individual patient data versus aggregate data from longitudinal clinical trials. Clin Trials. 2009;6(1):16–27.

Furukawa TA, Maruo K, Noma H, Tanaka S, Imai H, Shinohara K, Ikeda K, Yamawaki S, Levine SZ, Goldberg Y, et al. Initial severity of major depression and efficacy of antidepressants: individual-participant data meta-analysis. Acta Psychiatr Scand. 2018;137(6):450–8.

McKilloch CE, Searle SR, Generalized NJM. Linear, and mixed models. 2nd ed. New York: Wiley; 2008.

Mallinckrodt CH, Kaiser CJ, Watkin JG, Moltenberghs G, Carroll RJ. The effect of correlation structure on treatment contrasts estimated from incomplete clinical trial data with likelihood-based repeated measures compared with last observation carried forward ANOVA. Clin Trials. 2004;1(6):477–89.

Mallinckrodt CH, Lane PW, Schnell D, Peng Y, Mancuso JP. Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. Drug Inf J. 2008;42(4):303–19.

Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc A. 2009;172(1):137–59.

Gosho M, Hirakawa A, Noma H, Maruo K, Sato Y. Comparison of biased-corrected covariance estimators for MMRM analysis in longitudinal data with dropouts. Stat Methods Med Res. 2017;26(5):2389–406.

Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft. 2015;67(1).

Jackson D, Riley R, White IR. Multivariate meta-analysis: potential and promise. Stat Med. 2011;30(20):2481–98.

Mavridis D, Salanti G. A practical introduction to multivariate meta-analysis. Stat Methods Med Res. 2013;22(2):133–58.

White IR. Multivariate Random-effects meta-analysis. J Stat Soc B. 2009:40–56.

White IR. Multivariate random-effects meta-regression: updates to mvmeta. Stat Med. 2012;31(29):3821–39.

Cox DR, Reid N. Parameter orthogonality and approximate conditional inference. J R Stat Soc B. 1986;49(1):1–39.

Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, Hong ST, Hallemaak L, Gollogly L, Godlee F, et al. Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal. Ann Intern Med. 2017;167(1):63–5.

Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, Hong ST, Hallemaak L, Gollogly L, Godlee F, et al. Data sharing statements for clinical trials. BMJ. 2017;357:j2372.

Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. Stat Med. 2009;28(25):3049–67.

Carmödy TJ, Rush AJ, Bernstein I, Warden D, Brannan S, Burnham D, Woo A, Trivedi MH, The Montgomery Asberg and the Hamilton ratings of depression: a comparison of measures. Eur Neuropsychopharmacol. 2006;16(8):601–11.

Schwarz G. Estimating the dimension of a model. Ann Stat. 1978;6(2):461–4.

Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. Stat Med. 2017;36(5):855–75.

Brockwell SE, Gordon R. A comparison of statistical methods for meta-analysis. Stat Med. 2001;20(6):825–40.

Jackson D, Riley RD. A refined method for multivariate meta-analysis and meta-regression. Stat Med. 2014;33(4):541–54.

Noma H. Confidence intervals for a random-effects meta-analysis based on Bartlett-type corrections. Stat Med. 2011;30(8):3304–12.

Bender R, Friede T, Koch A, Kuss O, Schlattmann P, Schwarzer G, Skipka G. Methods for evidence synthesis in the case of very few studies. Res Synth Methods. 2018. https://doi.org/10.1002/jrsm.1297.

Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. Stat Methods Med Res. 2008;17(3):279–301.

White IR. Network meta-analysis. Sta J. 2015;15:951–85.
39. Morris TP, Fisher DJ, Kenward MG, Carpenter JR. Meta-analysis of Gaussian individual patient data: two-stage or not two-stage? Stat Med. 2018;37(9):1419–38.
40. Noma H, Nagashima K, Maruo K, Gosho M, Furukawa TA. Bartlett-type corrections and bootstrap adjustments of likelihood-based inference methods for network meta-analysis. Stat Med. 2018;37(7):1178–90.
41. Fitzmaurice GM, Laird NM, Ware J. Applied longitudinal analysis. Hoboken, NJ: John Wiley & Sons; 2011.
42. Duchateau L, Janssen P. The frailty model. New York: Springer; 2008.