Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
22(2) 133–158
© The Author(s) 2011
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0962280211432219
smm.sagepub.com

SAGE

A practical introduction to multivariate meta-analysis

Dimitris Mavridis[1,2] and Georgia Salanti[1]

Abstract

Multivariate meta-analysis is becoming increasingly popular and official routines or self-programmed functions have been included in many statistical software. In this article, we review the statistical methods and the related software for multivariate meta-analysis. Emphasis is placed on Bayesian methods using Markov chain Monte Carlo, and codes in WinBUGS are provided. The various model-fitting options are illustrated in two examples and specific guidance is provided on how to run a multivariate meta-analysis using various software packages.

Keywords

Bayesian methods, correlated outcomes, random effects, software, structural equation models

1 Introduction

Scientific experiments are typically designed to evaluate the impact of an intervention or risk factor on various outcomes and their published reports usually contain a non-random selection of these outcomes. A typical example is in hypertension trials where both systolic and diastolic blood pressures are measured. In the area of educational research, scholastic aptitude test scores may observe assessments on more than one subject, i.e. mathematics and English. Performance of a diagnostic test is often measured by paired indices such as sensitivity and specificity. When the results of many similar studies are considered, investigators typically perform separate meta-analyses to obtain summary estimates of the treatment effect on each outcome separately.

However, outcomes measured on the same population are correlated and by meta-analysing each outcome separately, any possible correlation structure is ignored. The impact of ignoring within-study correlation has been investigated via simulation studies in two recent papers.[1,2] Riley[2] showed that separate meta-analyses of correlated outcomes can lead to overestimated variance of the summary effect size and biased estimates. In addition, independent testing of treatment effects on multiple outcomes, which are indicators of the general effectiveness of an intervention, increases the chances of finding spuriously significant treatment effects, and adjustment for multiple comparisons might be needed.[3]

[1]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece
[2]Department of Primary Education, University of Ioannina, Ioannina, Greece

Corresponding author:
Georgia Salanti, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina 45110, Greece.
Email: gsalanti@cc.uoi.gr

Multivariate meta-analysis provides an elegant solution to the multiplicity problem by summarizing simultaneously all outcomes of interest instead of conducting many separate univariate meta-analyses. The data needed for conducting a multivariate meta-analysis are the study-specific effect sizes and their variance–covariance matrix. Multivariate meta-analysis, although introduced over 20 years ago, is not a popular approach. There are various reasons why researchers still prefer to conduct univariate analyses including tradition, lack of available software, lack of understanding of the correlation structure of the study effects and underestimation of the impact of ignoring correlation. Various statistical routines and packages to conduct multivariate meta-analysis have been recently developed and they give a great boost to the approach by making it more accessible to healthcare researchers.[4–7]

The aim of this article is to help popularizing multivariate meta-analysis by providing an up-to-date review of the methods for fitting multivariate meta-analysis models while giving specific technical guidance about the relevant software and routines. A recent paper addressed multivariate meta-analysis and synopsized the literature on methodology, providing also illustrative examples from areas of applications that show both its advantages and limitations.[8] We review some of the work suggested in this paper, placing also emphasis on Bayesian methods and introducing the reader and the medical scientist to the methodology of structural equation modelling (SEM) which is very popular in areas of social sciences such as psychology and education. In Section 3, we shortly discuss a simple framework for combining the outcomes into a single summary effect. In Section 4, we describe the general bivariate random-effects meta-analysis model that can be easily extended to the multivariate case (Section 4.1) and methods for fitting this model (Section 4.2). In Section 5, we describe SEM in connection with multivariate meta-analysis. The software available for conducting multivariate meta-analysis is presented in Section 7 and worked examples in Section 8.

2 Notation

Let us consider n studies and a total of p outcomes. There are no limitations to the type of the data considered (e.g. continuous, binary, count or time to event data) as long as the same effect measure (e.g. a standardized mean difference, the logarithm of odds ratio, rate ratio or hazard ratio) is computed for all outcomes. Therefore, each study i yields a subset of the p outcomes denoted as $\boldsymbol{y_i} = (y_{i1}, y_{i2}, \ldots, y_{ij}, \ldots, y_{ip})'$ where some y_{ij} are missing and $j = 1, \ldots, p$. With $\boldsymbol{y}$, we denote the $n \times p$ matrix of all study outcomes. Each y_{ij} is estimated with variance σ_{ij}^2 (typically assumed to be equal to the sample variance), and the correlation between pairs of outcomes (jj') within each study is $\text{corr}(y_{ij}, y_{ij'}) = \rho_{ijj'}$. In the systematic review that precedes the meta-analysis, we extract from each study i, the observation vector $\boldsymbol{y_i}'$ and the variance–covariance matrix $\boldsymbol{S}_i$ with diagonal elements σ_{ij}^2 and off-diagonal elements the covariances $\rho_{ijj'}\sigma_{ij}\sigma_{ij'}$. The matrix S is the block-diagonal $np \times np$ matrix for the variances and covariances for all outcomes and studies. It should be pointed out that correlations between outcomes are rarely available from published trials. Possible solutions to overcome the non-availability of within-study correlations are discussed in Section 4.3 and can also be found elsewhere.[2,8] Some studies may not report all p outcomes; however, estimation often requires the same dimensions in matrices across studies. In such cases, it is often assumed that the missing outcomes enter the data with zero values, very large variances and zero correlations to reflect our ignorance regarding their true values and to ensure that the imputed values have negligible contribution to the meta-analysis result.

In a univariate random-effects meta-analysis, the overall variation is partitioned to within-study and between-study variations. Between-study variance, also called heterogeneity, allows the model to take into account the variation in the underlying treatment effects across studies. In a multivariate random-effects meta-analysis not only the overall variation but also the overall correlation is partitioned to within-study and between-study correlations. The latter indicates how the multiple outcomes are related across studies. This will be further discussed in the random-effect multivariate meta-analysis model in Section 4.1.

3 Combining multiple dependent outcomes from each study

The first approach employed to address the problem of multiple outcomes was to combine these outcomes into a single summary effect size for each study and then synthesize these composite effect sizes across studies. Initially, some measure of the location such as the arithmetic or the geometric mean was used. The simplicity of such measures is their main advantage but correlation between the dependent effect sizes is not taken into account.

Two procedures to estimate a weighted average while taking into account the correlation structure have been proposed.[9,10] Rosenthal and Rubin[10] suggested a method to combine multiple continuous outcomes into a single measure. Assuming that we have large studies with few missing outcomes per study and fixed correlations between pairs of the effect sizes (i.e. $\rho_{ijj'} = \rho$), the composite measure (CM) for each study is:

$$\mathrm{CM}_i^{\mathrm{RR}} = \frac{\sum_{j=1}^{p} \lambda_j y_{1j}}{\sqrt{\rho\left(\sum_{j=1}^{p} \lambda_j\right)^2 + (1-\rho)\left(\sum_{j=1}^{p} \lambda_j^2\right)}} \tag{1}$$

where λ_j are predetermined weights for each outcome. In the absence of any prior information or belief, equal weights are assigned to all dependent outcomes.

An alternative estimation of CM when the outcomes are measured as standardized mean differences has been suggested by Hedges and Olkin.[9]

$$\mathrm{CM}_i^{\mathrm{HO}} = \frac{\mathbf{1}'\mathbf{S}_i^{-1}}{\mathbf{1}'\mathbf{S}_i^{-1}\mathbf{1}}\, \boldsymbol{y_i} \tag{2}$$

where $\mathbf{1}$ is a $p \times 1$ unit vector. It has been argued that $\mathrm{CM}^{\mathrm{RR}}$ (equation (1)) and $\mathrm{CM}^{\mathrm{HO}}$(equation (2)) have different purposes and estimate different quantities; so, the choice of the statistic should be guided by the aims of the synthesis.[11]

Combining multiple outcomes into a single measure is not a very popular approach as it entails significant loss of information and it yields a summary estimate which is not directly interpretable and useful for clinicians. The information loss is more severe when results differ systematically across meaningfully different outcomes.[12] The various outcomes often measure different aspects of the treatment's effectiveness and combining these outcomes may result in an incomprehensible index with a low reliability.

4 Models for multiple-outcomes meta-analysis

4.1 Multivariate fixed- and random-effects meta-analysis

For simplicity, the model for bivariate meta-analysis is presented first. Suppose that $j = 1, 2$ so that each study i provides two estimated treatment effects, i.e. $\boldsymbol{y}_i = (y_{i1}, y_{i2})'$. In a fixed-effects setting, the study-specific estimates follow a bivariate normal distribution:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 & \rho_i\sigma_{i1}\sigma_{i2} \\ \rho_i\sigma_{i1}\sigma_{i2} & \sigma_{i2}^2 \end{pmatrix}\right) \tag{3}$$

where $\rho_i = \rho_{i12}$ denotes the within-study correlation for study i between outcomes 1 and 2, $\boldsymbol{\mu} = (\mu_1, \mu_2)'$ the vector of means for each outcome $j = 1, 2$ and matrix $\mathbf{S}_i = \begin{pmatrix} \sigma_{i1}^2 & \rho_i\sigma_{i1}\sigma_{i2} \\ \rho_i\sigma_{i1}\sigma_{i2} & \sigma_{i2}^2 \end{pmatrix}$ the within-study covariance matrix. In a two-dimensional setting, it is assumed that not only the overall variation but also the overall correlation is divided into within-study correlation (ρ_i) and between-study correlation (ρ_τ). More specifically, it is assumed that:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} \Bigg| \begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 & \rho_i\sigma_{i1}\sigma_{i2} \\ \rho_i\sigma_{i1}\sigma_{i2} & \sigma_{i2}^2 \end{pmatrix}\right) \tag{4}$$

with $\boldsymbol{\theta} = (\theta_{i1}, \theta_{i2})'$ being the underlying study-specific effects for each outcome which are also normally distributed

$$\begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \tau_1^2 & \rho_\tau\tau_1\tau_2 \\ \rho_\tau\tau_1\tau_2 & \tau_2^2 \end{pmatrix}\right) \tag{5}$$

where τ_j is the between-study variation (heterogeneity) for effect size j. When outcomes are being jointly meta-analysed, along with matrix $\mathbf{S}_i$, we also have a between-study covariance matrix, $\boldsymbol{\Delta} = \begin{pmatrix} \tau_1^2 & \rho_\tau\tau_1\tau_2 \\ \rho_\tau\tau_1\tau_2 & \tau_2^2 \end{pmatrix}$. Marginally, by combining equations (4) and (5) we have:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 + \tau_1^2 & \rho_i\sigma_{i1}\sigma_{i2} + \rho_\tau\tau_1\tau_2 \\ \rho_i\sigma_{i1}\sigma_{i2} + \rho_\tau\tau_1\tau_2 & \sigma_{i2}^2 + \tau_2^2 \end{pmatrix}\right) \tag{6}$$

The models are easily extended for the case of p outcomes, involving $\binom{p}{2}$ correlation parameters $\boldsymbol{\rho}_\tau = (\rho_{\tau(1,2)}, \rho_{\tau(1,3)}, \ldots, \rho_{\tau(p,p-1)})'$ and equation (6) becomes:

$$\begin{pmatrix} y_{i1} \\ \vdots \\ y_{ip} \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 + \tau_1^2 & \cdots & \rho_i\sigma_{i1}\sigma_{ip} + \rho_{\tau(1,p)}\tau_1\tau_p \\ \vdots & \ddots & \vdots \\ \rho_i\sigma_{i1}\sigma_{ip} + \rho_{\tau(1,p)}\tau_1\tau_p & \cdots & \sigma_{ip}^2 + \tau_p^2 \end{pmatrix}\right)$$

An equivalent presentation of the random-effects model using matrix notation is:

$$\boldsymbol{y}_i = \boldsymbol{\mu} + \boldsymbol{\delta}_i + \boldsymbol{e}_i$$

where $\boldsymbol{\delta}_i$ is a vector of random effects associated with study i, $\boldsymbol{\delta}_i \sim \text{MVN}(0, \boldsymbol{\Delta})$ and $\boldsymbol{e}_i$ a vector of random sampling errors associated with study i which is independent of $\boldsymbol{\delta}_i$ and normally distributed,

$\boldsymbol{e}_i \sim \text{MVN}(0, \mathbf{S}_i)$. The matrix $\boldsymbol{\Delta}$ is the between-study variance–covariance matrix involving τ_j^2 and $\rho_{\tau(jj')}$ unknown parameters:

$$\boldsymbol{\Delta} = \begin{pmatrix} \tau_1^2 & \cdots & \rho_{\tau(1,p)}\tau_1\tau_p \\ \vdots & \ddots & \vdots \\ \rho_{\tau(1,p)}\tau_1\tau_p & \cdots & \tau_p^2 \end{pmatrix} \tag{7}$$

It follows that the variance–covariance matrix ofy_i is $\boldsymbol{\Delta} + \mathbf{S}_i$. We may also include l covariates in the model so that $\boldsymbol{\mu} = \mathbf{X}_i\boldsymbol{\beta}$

$$\boldsymbol{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\delta}_i + \boldsymbol{e}_i \tag{8}$$

where $\mathbf{X}_i$ is the $p \times (l+1)$ matrix with the observed covariate values (including a column of ones allowing for intercepts for the different effect sizes) for each study and $\boldsymbol{\beta}$ the vector of l coefficients and the constant term. The general random-effects multiple outcomes meta-regression model (equation (8)) is also known as a mixed effects model.

4.2 Fitting multivariate meta-analysis models

Several methods exist for estimating the parameters of the model. The parameters of interest are $\boldsymbol{\mu}$, the vector of the summary effects for the p outcomes, the uncertainty in the estimation of $\boldsymbol{\mu}$ expressed by a $p \times p$ variance–covariance matrix $\mathbf{C}$ and the matrix $\boldsymbol{\Delta}$ involving the heterogeneities τ_j^2 and between-studies correlations $\boldsymbol{\rho}_\tau$ (equation (**7**)). It should be noted that the various methods to fit the models within the frequentist framework differ primarily in the estimation of $\boldsymbol{\Delta}$ and hence they impact on the summary estimate and its uncertainty only in the random-effects case. In the following paragraphs, we focus on the random-effects model as the fixed effects estimates can be obtained as a special (and computationally simpler) case.

4.2.1 Likelihood methods

As the studies are assumed independent experiments, likelihood methods can be employed to estimate model parameters.[13] The likelihood is given by:

$$L \approx -\frac{1}{2}\sum_{i=1}^{n}\log|\boldsymbol{\Delta} + \mathbf{S}_i| - \frac{1}{2}\sum_{i=1}^{n}\boldsymbol{e}_i'(\boldsymbol{\Delta} + \mathbf{S}_i)^{-1}\boldsymbol{e}_i$$

The likelihood of the model can be maximized numerically using the expectation–maximization (EM), Newton–Raphson or Fisher scoring algorithms subject to the constraint that $\boldsymbol{\Delta}$ is positive semi-definite. The main difficulty of likelihood-based methods is that they become computationally intensive and time demanding as the dimension of the data (number of studies and outcomes) increases.

Maximum likelihood (ML) estimates: Assuming that all studies report the same outcomes and there are no missing values, the summary estimates obtained by maximizing the likelihood are:

$$\hat{\boldsymbol{\mu}} = \left(\sum_{i=1}^{n}(\hat{\boldsymbol{\Delta}} + \mathbf{S}_i)^{-1}\sum_{i=1}^{n}(\hat{\boldsymbol{\Delta}} + \mathbf{S}_i)\right)^{-1}\boldsymbol{y} \tag{9}$$

The estimates are approximately normally distributed with variance–covariance matrix:

$$\hat{\mathbf{C}} = \left(\sum_{i=1}^{n} (\hat{\boldsymbol{\Delta}} + \mathbf{S}_i)^{-1} \right)^{-1} \tag{10}$$

The parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Delta}$ are estimated iteratively in the EM algorithm with equation (9) being one of the two steps.[14] An approximate $(1-\alpha)\%$ confidence interval can be obtained for μ_j as $\hat{\boldsymbol{\mu}}_j \pm Z_{a/2}\sqrt{\hat{\mathbf{C}}_{jj}}$ where $\hat{\mathbf{C}}_{jj}$ is the j-diagonal element of the $\hat{\mathbf{C}}$ matrix. The use of quantiles from the t-distribution in the calculations has been also suggested.[15] It is very common that not all trials would report on all outcomes. In such a situation, equation (9) cannot be computed as the dimension of the matrices $\boldsymbol{\Delta}$ and $\mathbf{S}_i$ will not be the same across studies. To overcome this problem, we may impute the missing entries in the covariance matrices by allocating very large within-study variances to the missing outcomes and zero within-study correlations, ensuring that missing outcomes are replaced with estimates with negligible weight and information.[8] The main difficulty lies in the estimation of the between-study covariance matrix $\boldsymbol{\Delta}$ which is also used in the estimation of summary effect sizes in equation (9).

Restricted ML (REML): REML is very popular in literature for mixed models because it produces unbiased estimates of variance and covariance parameters. It has also been applied to multivariate meta-analysis problems to estimate $\boldsymbol{\Delta}$,[16] maximizing an expression of the likelihood given in Jennrich and Schluchter.[17] The restricted likelihood that is maximized numerically is:

$$\mathrm{RL} \approx L - \frac{1}{2}\left| \sum_{i=1}^{n} (\hat{\boldsymbol{\Delta}} + \mathbf{S}_i)^{-1} \right|$$

Positive definiteness in $\boldsymbol{\Delta}$ is ensured by maximizing the likelihood functions in terms of its Cholesky decomposition. Equations (9) and (10) are then employed to estimate $\boldsymbol{\mu}$ and $\mathbf{C}$ using the estimated $\hat{\boldsymbol{\Delta}}$.

4.2.2 Method of moments

The estimates for $\boldsymbol{\mu}$ and $\mathbf{C}$ are as in equations (9) and (10) but $\boldsymbol{\Delta}$ is estimated employing a multivariate extension of the Q statistic and the method of moments (MM).[18] In the univariate case for a fixed outcome j, it is $Q_j = \sum_{i=1}^{n} w_{ij}(y_{ij} - \bar{y}_j)^2$ where $\bar{y}_j$ is the weighted mean of y_{ij} across studies and the weights are $w_{ij} = 1/\sigma_{ij}^2$. Then, Q_j is equated to its expectation to yield an estimate of τ_j and μ_j is estimated as a weighted average with weights $w_{ij}^* = 1/(\sigma_{ij}^2 + \tau_j^2)$.

For p outcomes, the Q-statistic becomes:

$$\boldsymbol{Q} = \begin{bmatrix} Q_{11} & \cdots & Q_{1p} \\ \vdots & Q_{jj'} & \vdots \\ Q_{p1} & \cdots & Q_{pp} \end{bmatrix} \tag{11}$$

with elements $Q_{jj} = \sum_{i \in N_{jj'}} (y_{ij} - \bar{y}_j)^2/\sigma_{ij}^2$ and $Q_{jj'} = \sum_{i \in N_{jj'}} (y_{ij} - \bar{y}_{jj'})(y_{ij'} - \bar{y}_{j'j})/\sigma_{ij}\sigma_{ij'}$ where N_{jj} denotes the set of studies where only outcome j is reported and $N_{jj'}$ is the set of studies where both outcomes j and j' are reported. The quantity $\bar{y}_j$ is the weighted average of outcome j over the studies that report only the outcome j (with weight w_{ij}) and the quantity $\bar{y}_{jj'}$ the weighted average of outcome j over the studies that report the outcomes j and j' (with weight $\sqrt{w_{ij}w_{ij'}}$). Missing values can be handled by replacing them with arbitrary values with large within-study variances.

The estimation of $\boldsymbol{\Delta}$ is carried out by equating $\boldsymbol{Q}$ (equation (11)) with its expected value $E(\boldsymbol{Q}) = \boldsymbol{E}$. It was shown that the diagonal elements of $\boldsymbol{E}_{jj}$ are functions of σ_{ij} and τ_j and the off-diagonal elements $E_{jj'}$ are functions of $\sigma_{ij}, \sigma_{ij'}, \tau_j, \tau_{j'}, \rho_\tau$. Therefore, by equating $\boldsymbol{Q}_{jj} = \boldsymbol{E}_{jj}$, all heterogeneity parameters τ_j are estimated and by equating an off-diagonal element $\boldsymbol{Q}_{jj'} = \boldsymbol{E}_{jj'}$, the correlation ρ_τ can be estimated. Hence, all elements of $\boldsymbol{\Delta}$ are identifiable. However, the estimation as described above does not necessarily yield estimates of $\boldsymbol{\Delta}$ that are positive semi-definite and a truncated symmetric version has been suggested as $\hat{\boldsymbol{\Delta}} = \sum_{j=1}^{p} \max(0, \lambda_j)\, \varepsilon_j \varepsilon_j'$ where λ_j is the j eigenvalue and ε_i the corresponding normalized eigenvector. For more details, refer to Jackson et al.[18]

The main advantage of this method is that no numerical maximization or iteration is needed and the method is not time consuming. This could be a major advantage when the number of effect sizes per study is large. The method has been found to perform relatively well in comparison with likelihood methods both with simulated and real data sets.[18] Results are expected to be similar to those obtained by likelihood methods when there are only few outcomes and moderate to large heterogeneity. However, it should be noted that likelihood methods are often preferred to MM as the former have higher probability of being close to the quantities to be estimated. MM estimates can be used as starting values for iterations in likelihood methods.

4.2.3 Generalized least squares

We assume that each outcome can be modelled by a regression line. The regression models are not independent and correlations need to be taken into account. The $n \times p$ matrix of effect sizes $\boldsymbol{y}$ is reshaped into a $np \times 1$ vector where the first p elements refer to the first study, the following p to the next, etc. When there are no covariates in the model, matrix X is:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X_1} \\ \vdots \\ \mathbf{X_n} \end{bmatrix}$$

with $\boldsymbol{X_i}$ being an identity $p \times p$ matrix. The generalized least squares (GLS) criterion to be minimized is $(\boldsymbol{y} - \mathbf{X}\boldsymbol{\mu})'\mathbf{S}^{-1}(\boldsymbol{y} - \mathbf{X}\boldsymbol{\mu})$. The GLS fixed-effects estimate of the summary effect is:

$$\hat{\boldsymbol{\mu}} = (\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{S}^{-1}\boldsymbol{y} \tag{12}$$

with variance–covariance matrix $\hat{\mathbf{C}} = (\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}$. The fixed-effects estimate using GLS is the same as the ML estimate (equations (9) and (10) with $\hat{\boldsymbol{\Delta}}$ being replaced by a zero $p \times p$ matrix). The random-effects estimator computes $\hat{\boldsymbol{\mu}}$ (equation (12)) iteratively and $\hat{\boldsymbol{\Delta}} = 1/(n-2)\boldsymbol{e}'\boldsymbol{e} - \frac{1}{n}\sum_{i=1}^{n} \mathbf{S}_i$ where $\boldsymbol{e} = \boldsymbol{y} - \hat{\boldsymbol{\mu}}$, until convergence is attained.[14]

4.2.4 Bayesian approaches for multivariate meta-analysis using Markov chain Monte Carlo

The advent of high-speed computers popularized Markov chain Monte Carlo (MCMC) methods for estimating model parameters in all fields of statistics. A further boost in MCMC methods was given in mid-1990s with the development of WinBUGS.[19] Bayesian methods have been used to perform multivariate meta-analyses, but they are not as popular as in the univariate case, mainly because of the difficulties encountered on placing priors for the between-studies covariance matrix that ensure its positive semi-definiteness.[20]

In a Bayesian framework, all parameters are treated as random variables and by placing suitable prior distributions, they can be estimated using MCMC. This means that in contrast to the methods

presented so far, Bayesian fitting accounts for the fact that τ_j and ρ_τ are estimated rather than assumed fixed. For each parameter of interest, its posterior distribution is being estimated using MCMC. WinBUGS[19] provides us with posterior distributions for all parameters which can be interpreted directly as the distributions of the quantities of interest. Summaries of the distributions such as posterior means, posterior medians, variances and $(1-\alpha)\%$ credible intervals are often derived. A credible interval can be interpreted as the interval within which the parameter lies with probability $(1-\alpha)\%$. Exploring the shape of the posterior distribution is often useful, particularly for between-study variances; in cases of asymmetry, the posterior median may be used instead of the posterior mean to convey the central location of the parameter.

A main advantage of the fitting multivariate meta-analysis in a Bayesian framework is that external evidence or information from historical data can be easily incorporated in the model via informative priors. With a large number of studies, the choice of the prior distribution impacts less on the results since data dominate the analysis. However, with a small number of studies, estimates of the heterogeneities and between-study correlation can be sensitive to prior selection. In a simulation study that compared 13 different prior distributions for the heterogeneity parameter, it was shown that results may vary substantially when the number of studies is small[21] and we expect that this applies to multivariate meta-analysis as well. A solution to overcome this problem is to conduct a sensitivity analysis to the choice of prior distribution.[21,22] It should be noted that poor estimation of the between-studies variance is not encountered in MCMC context only; with few studies, the ML based methods can give biased results as well as the large-sample optimality properties do not apply.

It has been previously suggested that an inverse-gamma prior should be avoided when heterogeneities are close to 0 because of computational problems[22] and one needs to specify parameter values for which the non-informativeness is difficult to defend. A natural choice for the prior distribution for the $\mathbf{\Delta}$ matrix is the Wishart distribution which is a multivariate generalization of the gamma distribution. In analogy to the gamma prior, the Wishart prior has computational problems with variance and covariance values close to zero and does not allow differential prior knowledge for these parameters across outcomes.

An extra challenge for prior specification in the multivariate setting is that the priors for the between-study variance–covariance matrix $\mathbf{\Delta}$ should ensure non-negative definiteness. The spherical parameterization based on Cholesky decomposition has been suggested to generate a positive-definite matrix for the correlation matrix $\mathbf{R}$, which in turn can be used to obtain the covariance matrix $\mathbf{\Delta} = \mathbf{V}^{1/2}\mathbf{R}\mathbf{V}^{1/2}$ where $\mathbf{V}^{1/2}$ is a diagonal matrix with elements τ_j.[23] The correlation matrix is decomposed into the product of a lower triangular matrix and its conjugate transpose $\mathbf{R} = \mathbf{L}\mathbf{L}'$ where for the case of two outcomes $\mathbf{L} = \left(\begin{smallmatrix} 1 & \cos\varphi \\ 0 & \sin\varphi \end{smallmatrix}\right)$ and hence $\mathbf{R} = \left(\begin{smallmatrix} 1 & \cos\varphi \\ \sin\varphi & 1 \end{smallmatrix}\right)$. By generating the angle φ from a uniform distribution on $[0, \pi]$, it is ensured that the correlation $(\cos\varphi)$ lies in $[-1, 1]$. Variances are easily sampled from a variety of candidate distributions including the uniform, half-normal and half-t distributions. The method is easily applied to higher dimensions.[23]

4.3 Handling unknown within-study covariance matrices

A key feature in meta-analysis with multiple outcomes is the possible dependence between them and the incorporation of this dependence into a statistical model. Therefore, sample estimates of the within-study correlations $\rho_{ijj'}$ are of crucial importance. With individual patient data, variances and covariances are estimated directly but in meta-analyses of aggregated data, it often occurs to have missing $\rho_{ijj'}$.

Simulation studies restricted to cases where the between-studies variance σ_{ij} is much smaller than the heterogeneity τ_j showed that the estimation of the summary treatment effect $\boldsymbol{\mu}$ is not affected by inaccurate approximations of the within-study correlation $\rho_{ijj'}$.[1] Further simulations on a wider spectrum of parameter values established that when τ_j is of comparable magnitude to σ_{ij} or when there are large differences between $\mathbf{S}_i$'s, ignoring or inaccurately estimating within-study correlations increases the mean-square error and standard error of the summary estimate.[2] An alternative model (AM) for fitting bivariate meta-analysis that does not require the within-study correlations has been suggested.[24] The general bivariate random-effects meta-analysis model, as presented in Section 4.1, partitions the observed variation of the effect sizes to within-study and between-study variations using a fully hierarchical structure. The AM partitions the overall variation but not the overall correlation. Instead, a single parameter ρ_ψ is proposed to model the overall correlation:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{i1}^2 + \psi_1^2 & \rho_\psi\sqrt{(\sigma_{i1}^2 + \psi_1^2)(\sigma_{i2}^2 + \psi_2^2)} \\ \rho_\psi\sqrt{(\sigma_{i1}^2 + \psi_1^2)(\sigma_{i2}^2 + \psi_2^2)} & \sigma_{i2}^2 + \psi_2^2 \end{pmatrix}\right)$$

The additional variation beyond sampling error is denoted by $\psi_j^2, j = 1, 2$. However, ψ_j^2 is not directly equivalent to τ_j^2 in the fully hierarchical model in equation (6).

The main advantage of the approach is that estimates of the within-study correlation are not needed to estimate the model parameters $\rho_\psi, \psi_1, \psi_2, \mu_1$ and μ_2. When the within-study variation σ_{ij}^2 is relatively large or the number of studies is small, both the fully hierarchical and the AM do not converge or estimate the between-study correlation as 1 or -1. When $|\rho| < 0.95$, the AM produced estimates with little bias.[24]

Another approach to the problem of unknown within-study correlation is to consider external information, e.g. from a group of relevant studies included in the review that do report $\rho_{ijj'}$. It is also customary to undertake a sensitivity analysis over the entire correlation range. An advantage with models fitted within Bayesian framework is that unknown quantities can be given priors rather than a series of specific values. When correlation is unknown, it can be given suitable priors; for example, $\rho_{ijj'} \sim \text{beta}(5, 1)$ if it is believed that outcomes are likely to be highly correlated within studies. Several informative and non-informative distributions (e.g. uniform in $[-1, 1]$) shall be used and the change in the estimates can be monitored before drawing conclusions.[20]

5 Multivariate meta-analysis formulated as an SEM

Although SEM and meta-analysis are seemingly unrelated techniques that use a different terminology and approach, the estimation of parameters from different perspectives are closely related. Recent efforts to unify the two methodologies showed that meta-analysis is a special case of SEM.[25]

The key feature of SEM is that it places emphasis on modelling the covariances or the correlation vector under a hypothesized model. If the data $\boldsymbol{y}$ yield an observed correlation vector $\boldsymbol{r}$ and under an assumed model (e.g. the outcome is a function of some parameter $\boldsymbol{\mu}$) we get a correlation vector $\boldsymbol{\rho}(\boldsymbol{\mu})$, then the SEM model focuses on minimizing the function:

$$(\boldsymbol{r} - \boldsymbol{\rho}(\boldsymbol{\mu}))' \mathbf{V}_r (\boldsymbol{r} - \boldsymbol{\rho}(\boldsymbol{\mu})) \tag{13}$$

Assuming that the model is a linear regression model for observations that are independent and identically distributed, minimization of expression (13) yields the least-squares estimate. Several

other commonly used statistical models can be seen as special cases of SEM, such as factor analysis and multilevel models. Meta-analysis can be viewed as a weighted regression with no covariates; so, implementation in SEM context comes naturally and there have been several attempts in this direction.[26–28] Equation (13) can be extended for 'multiple' groups n[29]

$$\sum_{i=1}^{n} (\boldsymbol{r} - \boldsymbol{\rho}(\boldsymbol{\mu})_i)' \mathbf{V}_{r_i} (\boldsymbol{r} - \boldsymbol{\rho}(\boldsymbol{\mu})_i) \tag{14}$$

The first step of a SEM-based meta-analysis is to transform the trial observations so that they have unit variances since it is a fundamental assumption in SEM that all observations are distributed with the same variance. This is achieved by multiplying the effect sizes with $\sqrt{\mathbf{W}_i} = \text{diag}\left[\frac{1}{s_{i1}}, \frac{1}{s_{i2}}, \ldots, \frac{1}{s_{ip}}\right]$. The fixed-effects meta-analysis model becomes:

$$\sqrt{\mathbf{W}}\boldsymbol{y} = \sqrt{\mathbf{W}}\boldsymbol{\mu} + \sqrt{\mathbf{W}}\boldsymbol{e} \Leftrightarrow \boldsymbol{y}^* = \boldsymbol{\mu}^* + \boldsymbol{e}^* \text{ so that } \boldsymbol{e}^* \sim N_n(0, \mathbf{I}_n)$$

Then, the effect sizes $\boldsymbol{y}^*$ and their transformed within-study variance–covariance matrix are used instead of $\boldsymbol{r}_i$ and $\mathbf{V}_{r_i}$, respectively in equation (14). Note that the discrepancy function can be minimized also using ML or REML to deliver equivalent fixed-effects estimates with GLS. Random-effects estimates can be obtained using any of the methods presented so far. The main advantage of expressing multivariate meta-analysis as SEM is that all SEM-related methodological developments can be used, such as the ability to handle missing covariates or translate the model into path diagrams and use SEM software.[29]

6 Applications of multivariate meta-analysis in diagnostic accuracy tests and networks of interventions

Synthesis of diagnostic test accuracy studies is a popular application area of multivariate meta-analysis.[30] Diagnostic test accuracy studies typically report the number of true positives (diseased people correctly diagnosed), false positives (non-diseased people incorrectly diagnosed as diseased), true negatives and false negatives. Sensitivity (the conditional probability of testing positive in diseased subjects) and specificity (the conditional probability of testing negative in non-diseased subjects) can be seen as the outcomes of interest in each study and they can either be extracted from the individual studies or modelled as parameters in binomial distributions for the numbers of true positives and negatives. Then, a bivariate random-effects meta-analysis model for the logit transformation of sensitivity and specificity is employed (hierarchical logistic regression).[31] As sensitivity and specificity are estimated from different samples in each study (diseased and non-diseased patients), they can be assumed to be independent so that the within-study correlations ρ_i are set to zero. However, there may be a non-zero between-studies correlation ρ_τ which should be accounted for.

Another special application of multivariate meta-analysis is when synthesizing studies with multiple arms in a network meta-analysis (or multiple-treatments meta-analysis). Assume the simplest case where we have only one outcome but three treatments A, B and C. A three-arm trial yields estimates y_{ij} where now $j = AB, AC, BC$ denotes the three possible comparisons. However, because it holds that:

$$\boldsymbol{y}_{iAB} = \boldsymbol{y}_{iAC} - \boldsymbol{y}_{iBC} \tag{15}$$

we only need to include two out of the three estimates in the model. The two estimates to be modelled, say $\boldsymbol{y}_{iAC}, \boldsymbol{y}_{iBC}$ are correlated as they include the same reference arm C and so are their underlying random effects $\theta_{iAC}, \theta_{iBC}$. A basic assumption of the network meta-analysis (also called multiple-treatments meta-analysis or mixed-treatment comparison) is that equation (15) holds for the first and second moments. This assumption has important implications for the estimation of matrix $\boldsymbol{\Delta}$ as the between-studies covariance turns out to be a function of the between-studies variances. In simple cases where it is assumed that $\tau_j = \tau$, it turns out that $\rho_\tau = 0.5$[32] but in the general case, the form of $\boldsymbol{\Delta}$ depends on the total number of treatments involved; see recent work by Lu and Ades[23] for the case when τ_j varies across comparisons. Consequently, whereas in principal multivariate meta-regression can be used to fit a network meta-analysis, one has to build into the model a specific structure for the $\boldsymbol{\Delta}$ matrix where ρ_τ is a function of the comparison-specific heterogeneities τ_j.[33]

7 Software for multivariate meta-analysis

Multivariate meta-analysis can be conducted using various software including STATA,[34] SAS,[35] WinBUGS,[19] R[36] and SEM such as LISREL[37] and Mplus.[38]. In this section, we focus on STATA, R and WinBUGS which are very popular with medical researchers conducting meta-analyses. We also show how SEM-based meta-analysis can be conducted using R.

7.1 STATA

The `mvmeta` command in STATA performs fixed- and random-effects multivariate meta-regression analysis. For the latter, there is a variety of fitting methods available, such as ML, REML and MM.[2,6]. Recently, an updated version of the `mvmeta` command was published and made available online (by typing in STATA `net from http://www.mrc-bsu.cam.ac.uk/IW_Stata/meta`).[7] The updated command allows for meta-regression and for network meta-analysis.

The syntax of the command is

```
mvmmeta y S, options
```

where `y` is the matrix $\boldsymbol{y}$ represented in the data by p columns each one named using the same first letter and some index (e.g. for three outcomes the data has columns named y1, y2, y3). The argument `S` is the variance covariance matrix $\mathbf{S}$ which is represented in the data by a set of $p + \binom{p}{2}$ columns that are named using the first starting letter and appropriate combination of the indexes as used in $\boldsymbol{y}$; for example, the columns S11, S22, S33, S23, S12 and S13 contain the variances and the co-variances of the three outcomes. The `options` can used to specify the method to estimate the $\boldsymbol{\Delta}$ matrix or to define the fixed-effects model as the synthesis method.

As within-study covariances are usually unknown, a common approach is to impute an assumed correlation value and then perform sensitivity analysis over the entire correlation range; this can be done using the option `wscorr(expression)`. There is also the option to fit the AM suggested by Riley et al.[24] and presented in Section 4.3 by typing `wscorr(riley)`. It should be noted that `mvmeta` does not require all studies to report all outcomes.

Finally, the command `metandi` performs bivariate meta-analysis of diagnostic test accuracy studies using hierarchical logistic regression.[39] The function takes as arguments the number of true positives (`tp`), the number of false positives (`fp`), the number of true negatives (`tn`) and the number of false negatives (`fn`).

7.2 WinBUGS

Although not very popular among non-statisticians, WinBUGS has the advantage of great flexibility as MCMC easily allows the inclusion of studies with missing outcomes and extensions to multivariate meta-regression follow naturally. We provide in the Appendix and also in www.mtm.uoi.gr WinBUGS codes for fitting random-effects bivariate meta-analysis models as well as the AM that does not require within-study correlation estimates. To model the variance–covariance matrix $\mathbf{\Delta}$, we used the Cholesky decomposition as described in section 4.2.4 and we carried out a sensitivity analysis to the prior distribution of the dispersion parameter τ. The default prior put on the precision was $\frac{1}{\tau^2} \sim U[0, 100]$ and other choices include the inverse gamma $\frac{1}{\tau^2} \sim IG(:, :)$, half-normal prior on heterogeneity standard deviation $\tau \sim \text{HN}(0, 1)$ and uniform prior distribution on $\log(\tau) \sim U(-10, 10)$.[40] Convergence should be monitored via several diagnostics.[41] In the applications presented, we run 100 000 cycles after a burn-in period of 10 000, and two chains according to different initial values were monitored to ensure convergence. To minimize autocorrelation, summary statistics were obtained after thinning intervals of 10 draws producing a final MCMC sample of size 10 000 cycles.

7.3 R

R software can be used to fit SEM models. The R package `metaSem` can be downloaded from the CRAN project (http://cran.r-project.org/) and can be used to conduct univariate and multivariate meta-analyses using SEM via the `OpenMx` package[4] The `meta` command for univariate or multivariate SEM meta-analyses uses ML methods and can include covariates. The syntax of the command is

```
meta(y,v,x)
```

where `y` is a $n \times p$ matrix with the effect sizes, `v` the $n \times \left(p + \binom{p}{2}\right)$ matrix with the variances and covariances and `x` the (optional) matrix of the covariates. The output of `meta` consists of the summary estimates for the outcomes (e.g μ_1, μ_2), the corresponding standard errors and confidence intervals, z-statistics and p-values. A major drawback of the command is that it does not allow for missing outcomes in studies. However, these could be imputed by hand.

Recently, a new R package called `mvmeta` has been released.[5] The program allows the user to perform fixed- and random-effects multivariate meta-analysis and meta-regression. Just as in STATA, the main command is called `mvmeta` and it has a similar syntax with the `meta` command in the `metaSEM` package. Estimation methods for random-effects models are based on REML using quasi-Newton methods. Fixed-effect estimates are computed with GLS given the within-study covariance matrices $\mathbf{S}_i$.

The syntax of the command is

```
meta(y~1,v, method)
```

where y is a $n \times p$ matrix with the effect sizes and the formula `y~1` determines that a multivariate meta-analysis is conducted with no covariates and `v` the $n \times \left(p + \binom{p}{2}\right)$ matrix with the variances and covariances. Alternatively, specifying `y~x` computes multivariate meta-regression with `x` being the covariates matrix. The option `method` defines the parameter estimation technique `('reml','ml','fixed')`.

We could not identify any R routine or package for fitting the random-effects multivariate meta-analysis model using iterative GLS as described in section 4.2.3. For this purpose, we developed an R routine which is given in the Appendix and can be also found in http://www.mtm.uoi.gr/.

The author of the `metafor` package in R is currently working in extending the package to the multivariate case, which however was not published at the time of writing this article.[42]

7.4 Other software: LISREL, SAS and MATLAB

The SAS PROC MIXED was one the first routines that popularized multivariate meta-analyses over a decade ago.[16] More recently, the SAS command METADAS was made available to fit bivariate meta-analysis models for diagnostic test accuracy studies.[43]

LISREL can be used to fit multivariate meta-analysis using SEM; the code is provided in a recent paper by Cheung.[29] However, LISREL is not likely to become very popular with medical scientists as it is primarily syntax based.

Note that all methods described here can be programed in any statistical software. To assist MATLAB[44] users, we provide in the Appendix and in http://www.mtm.uoi.gr/ a code for fitting the bivariate random-effects meta-analysis model using iterative GLS.

8 Examples

8.1 Surgical versus non-surgical treatment of periodontal disease

The first data set consists of five studies assessing the relative effectiveness of surgical interventions versus non-surgical procedures for treating periodontal disease with respect to two outcomes: improvement in probing depth (outcome 1) and improvement in attachment level (outcome 2). Both outcomes were measured as mean differences (Table A.1). The data was first presented in a paper introducing the multivariate meta-analysis model fitted with GLS[14] and has been also used in further methodological publications[2,4,5] and in illustrative examples showing potential of relevant software.[4,5]

The variables were inserted in a STATA database; the mean differences as variables y1 (y_1) and y2 (y_2) and the variances and covariances as S11 (S_1^2), S22 (S_2^2) and S12 ($\rho S_1 S_2$), respectively. Random-effects (using ML, RELM and MM) and fixed-effects multivariate meta-analysis can be obtained as

```
mvmeta y S,ml
mvmeta y S,reml
mvmeta y S,mm
mvmeta y S,fixed
```

Using the same variable names, a multivariate meta-analysis in R can be carried out as SEM

```
library(metaSEM)
summary(meta(y=cbind(y1,y2), v=cbind(S11, S12, S22)))
```

The `mvmeta` command in R can be also used

```
library(mvmeta)
summary(mvmeta(y=cbind(y1,y2)~1, v=cbind(S11, S12, S22),method='reml'))
```

Figure 1 and Table 1 present the results using the methods discussed so far. Univariate random-effects meta-analysis was performed using the `metareg` command in STATA and heterogeneity estimated with ML. There are large differences between fixed- and random-effect estimates and this is attributed to the small number of studies and the large heterogeneity observed in the second outcome. The differences between univariate and multivariate results of the overall summary estimates are primarily associated with increased precision for all methods, particularly for the second outcome. This comes as a result of estimating with smaller heterogeneity values in the multivariate approach (Table 1).

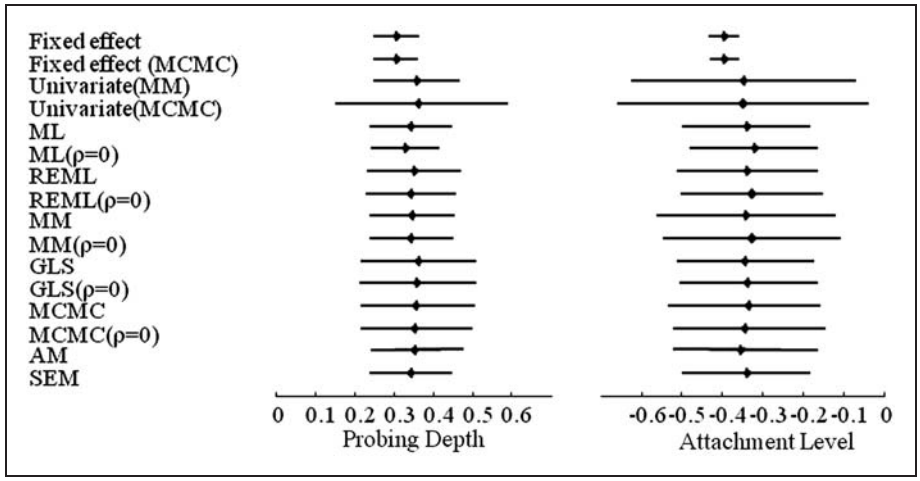

Figure 1. Bivariate meta-analysis summary estimates of the mean difference for probing depth and attachment level between surgical and non-surgical treatment of periodontal disease. Notes: ML, maximum likelihood; REML, restricted maximum likelihood; MM, method of moments, GLS, generalized least squares; MCMC, Markov chain Monte Carlo, AM, alternative model that does not require data on within-study correlations; and SEM, structural equation modelling. *The fixed-effects estimates using REML, GLS, MM and SEM are identical to the ML estimation.

Table 1. Bivariate meta-analysis estimates of heterogeneity and covariance for surgical versus non-surgical treatment of periodontal disease

Estimation method	τ_1^2	τ_2^2	$\rho_\tau \tau_1 \tau_2$
Univariate (ML)	0.010	0.057	
Univariate (MCMC)	0.023 (0.010,0.165)	0.039 (0.012,0.294)	
ML	0.007	0.026	0.009
ML ($S_{12} = 0$)	0.005	0.027	0.012
REML	0.011	0.033	0.012
REML ($S_{12} = 0$)	0.010	0.033	0.014
MM	0.010	0.057	0.018
MM ($S_{12} = 0$)	0.010	0.057	0.020
GLS	0.023	0.032	0.011
GLS ($S_{12} = 0$)	0.023	0.032	0.013
MCMC	0.015 (0.010,0.065)	0.028 (0.011,0.123)	0.001 (−0.016,0.039)
MCMC ($S_{12} = 0$)	0.015 (0.010,0.058)	0.028 (0.011,0.121)	0.011 (−0.013,0.043)
SEM	0.007	0.026	0.009

Notes: For MCMC methods medians and 95% credible intervals are displayed. Abbreviations are as in Figure 1.

Bayesian and frequentist estimation of the heterogeneity parameters differ remarkably when each outcome is analysed separately. As highlighted in Section 4.2, the use of priors for heterogeneity is problematic when only a few studies are included and different priors can lead to different results and this applies to multivariate meta-analysis as much as the univariate case. Sensitivity to priors can be important when true heterogeneity is close to zero as MCMC tends to give heterogeneity

estimates which are biased upwards.[21,22] The results of the sensitivity analysis for the multivariate model fitted with MCMC are presented in Table A.2. The summary point estimates of the mean difference are not affected by the choice of the prior distributions, but their precision decreases with higher values of estimated heterogeneity. The gamma prior on precision seems to overestimate the heterogeneity. This is in line with the results from simulation studies. Although the inverse gamma is a conjugate distribution, it may yield misleading results when true heterogeneity is close to zero.

8.2 Comparing chemotherapy regimens for colorectal cancer

The second example is taken from a multiple-treatments meta-analysis that combined information from direct and indirect evidence about the relative effectiveness of chemotherapy regimens in colorectal cancer.[45] The primary endpoint in the analysis was overall survival (outcome 1) and the secondary outcome was disease progression (outcome 2). Both outcomes were measured using the logarithm of the hazard ratio. Out of the 37 studies included in the review, we selected nine studies that compared the same treatments; the combination of fluorouracil and leucovorin versus fluorouracil and leucovorin plus irinotecan. The data are shown in Table A.3.

The within-study correlations are unknown and we explored different scenarios. We assumed the same correlation across studies which takes values $\rho_i = \rho = (0, 0.4, 0.8, 0.99)$. For each ρ value, we synthesized the studies using ML. For comparison, the univariate meta-analysis was carried out using `metareg` with the `ml` option. Subsequently, we fit the MCMC model treating ρ as an unknown parameter and assuming three prior distributions that are indicative of different prior beliefs regarding the magnitude of the correlation. We assumed a uniform prior $\rho \sim U(0, 1)$ (equal probability to all correlation values), $\rho \sim \text{beta}(5, 1)$ (higher probability to larger positive correlation values) and $\rho \sim \text{beta}(1, 5)$ (higher probability to smaller positive correlation values). The results are shown in Figure 2 and Table 2. The precision in the estimates obtained from multivariate meta-analysis is smaller compared to the univariate meta-analysis in both frequentist and Bayesian setting. Although in principal, standard errors obtained from multivariate meta-analysis are smaller compared to those from separate univariate models,[2,46] the opposite can be observed when the within-study and between-study variances are all equal to zero as it the case in this example.[47]

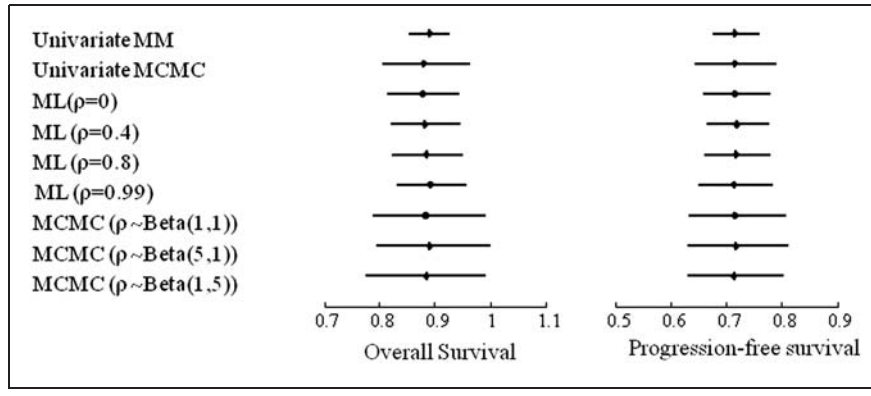

Figure 2. Bivariate meta-analysis summary estimates of the hazard ratio for overall and progression-free survival. Notes: Abbreviations are as in Figure 1.

Table 2. Bivariate random-effects meta-analysis estimates of heterogeneity and covariance for overall and progression-free survival

Estimation method	τ_1^2	τ_2^2	$\rho_\tau \tau_1 \tau_2$
Univariate (ML)	0.000	0.000	
Univariate (MCMC)	0.006 (0,0.066)	0.013 (0,0.119)	
ML ($\rho = 0$)	0.001	0.003	0.002
ML ($\rho = 0.4$)	0.000	0.000	0.000
ML ($\rho = 0.8$)	0.000	0.002	0.001
ML ($\rho = 0.99$)	0.000	0.008	0.001
MCMC $\rho \sim$ beta (1,1)	0.013 (0.010,0.032)	0.015 (0.010,0.047)	0.012 (−0.011,0.028)
MCMC $\rho \sim$ beta (5,1)	0.013 (0.010,0.030)	0.016 (0.010,0.050)	0.013 (−0.007,0.029)
MCMC $\rho \sim$ beta (1,5)	0.013 (0.010,0.037)	0.015 (0.010,0.048)	0.012 (−0.014,0.028)

Notes: See 'Notes' of Table 1.

Table 3. Comparative features of software packages to fit multivariate meta-analysis

	Package		
	`Mvmeta` (STATA)	`mvmeta` (R)	`metaSEM` (R)
MM estimator	✓		
ML estimator	✓		
REML estimator	✓	✓	✓
Iterative GLS estimator			
AM	✓		
Meta-regression	✓	✓	✓
Can deal with missing outcomes	✓	✓	
Open source		✓	✓

The estimated heterogeneity in the frequentist approaches is close to zero and MCMC gives considerably larger estimates, resulting in wide credible intervals for the summary estimates. The ML, MM and REML estimates of the between-study correlation ρ_τ was one for all within-study correlations and under this scenario, it is known that the estimates are biased.[8] Note that in the Bayesian multivariate approach, the results account for the uncertainty in the unknown within-study correlation which leads to larger credible intervals compared to the frequentist estimation.

9 Conclusions

The advantages and limitations of multivariate meta-analysis have been discussed in various research articles and demonstrated theoretically and empirically.[1,2,8,16] The main limitation of the approach is its computational complexity which, combined with the fact that medical scientists trust and understand univariate meta-analysis better, have made multivariate meta-analysis unpopular.

Whereas it will take its time toll before the medical community acknowledge the importance of taking correlation between multiple outcomes into account,[8] the rapid development of related software is anticipated to help towards this direction. In this review, we fit the models using a variety of software packages and routines and provide worked examples and self-programmed

routines to assist popularizing multivariate meta-analysis. The three major packages identified, two of which are freely available in the open-source R software, were easy to use. We summarize their features in Table 3. A major limitation is that within-study correlations are usually unknown and in this case, the AM[24] is an option to consider. This model can be fitted using the `mvmeta` command in STATA.[7] The problem of unknown within-study correlations can also be tackled by conducting a sensitivity analysis; such an option is incorporated in STATA's `mvmeta` command and is easily accounted for in a Bayesian context either by assuming fixed values for the correlations or prior distributions reflecting different prior beliefs. When the meta-analysis includes studies that do not report all outcomes, then `mvmeta` (either in STATA or R) can be used or the missing values can be inserted (zeros for effect sizes and within-study correlations and large values, i.e. 1000 for within-study variances). We could not identify any package that includes the iterative GLS algorithm to estimate the random-effects variance covariance matrix; so, we provided self-programmed routines in R and MATLAB.

WinBUGS may prove to be a powerful tool for researchers as they will be able to write problem-specific codes for complex models. However, with few studies, the choice of prior may affect the results considerably, especially when the heterogeneity is close to zero. The use of priors that give support to unrealistically large values of heterogeneity should be avoided and it is suggested to investigate vague prior distributions within a realistic range of values for the data set under consideration in a sensitivity analysis.[47]

A major difficulty encountered in practice is that typically only a few studies are available. This is a common problem in univariate meta-analyses often associated with poor estimation of the heterogeneity; in a multivariate setting with p outcomes, along with the p treatment effects and p heterogeneities, additionally $\binom{p}{2}$ correlation parameters need to be estimated and a substantial amount of data is required to do so efficiently. Particularly problematic are situations where the between-study correlation is estimated on the boundaries 1 or -1 leading to an overestimation or underestimation of the variances of the effect sizes. Jackson et al.[8] nicely illustrated this problem in an example with six studies and two outcomes (estimating thus five parameters) in which REML estimates of ρ_τ where found to be equal to 1 whereas MM estimates were estimated at -1.

In a multivariate meta-analysis setting, we borrow strength from the reported outcomes to estimate the ones not reported. This is an important advantage because trial investigators typically compute many outcomes but decide to publish only those indicating a significant treatment.[48] As systematic reviewers and meta-analysts increasingly recognize the selective reporting bias as a major threat to the validity of the meta-analysis result,[49] the use of multivariate meta-analysis could provide a viable solution.[50] An extensive simulation study to compare the various fitting methods and the impact of priors under different scenarios will further give insight to the limitations and advantages of the method, particularly for situation when the between-study correlation is close to one and/or the heterogeneities are zero.

Acknowledgement

Both authors received research funding from the European Research Council (IMMA 260559).

References

1. Ishak KJ, Platt RW, Joseph L and Hanley JA. Impact of approximating or ignoring within-study covariances in multivariate meta-analyses. *Stat Med* 2008; **27**: 670–686.
2. Riley RD. Multivariate meta-analysis: the effect of ignoring within-study correlation. *J R Stat Soc, Ser. A* 2009; **172**(4): 789–811.

3. Bender R, Bunce C, Clarke M, Gates S, Lange S, Pace NL, et al. Attention should be given to multiplicity issues in systematic reviews. *J Clin Epidemiol* 2008; **61**: 857–865.
4. Cheung, MWL. 'metaSEM: Meta-analysis: A Structural Equation Modeling Approach', R package version 0.5-3, http://courses.nus.edu.sg/course/psycwlm/Internet/metaSEM/ (2011) (accessed 3 January 2012)
5. Gasparrini A. 'Package 'mvmeta'. Multivariate Meta-analysis and Meta-regression', http://cran.r-project.org/web/packages/mvmeta/mvmeta.pdf (2011) (accessed 3 January 2012).
6. White I. Multivariate random-effects meta-analysis. *STATA J* 2009; **9**(1): 40–56.
7. White IR. Multivariate random-effects meta-regression: updates to mvmeta. *STATA J* 2011; **11**(2): 255–270.
8. Jackson D, Riley R and White IR. Multivariate meta-analysis: potential and promise. *Stat Med* 2011; **30**(20): 2481–2498.
9. Hedges LV and Olkin I. *Statistical methods for meta-analysis*. Orlando, FL: Academic Press, 1985.
10. Rosenthal R and Rubin DB. Meta-analytic procedures for combining studies with multiple effect sizes. *Psychol Bull* 1986; **99**: 400–406.
11. Marin-Martinez F and Sanchez-Meca J. Averaging dependent effect sizes in meta-analysis: a cautionary note about procedures. *Span J Psychol* 1999; **2**: 32–38.
12. Hartung J, Knapp G and Sinha B. *Statistical meta-analysis with applications*. New York: Wiley, 2008.
13. Houwelingen HC, Zwinderman K and Stijnen T. A bivariate approach to meta-analysis. *Stat Med* 1993; **12**: 2272–2284.
14. Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F and Colditz GA. Meta-analysis of multiple outcomes by regression with random effects. *Stat Med* 1998; **17**: 2537–2550.
15. Riley RD, Abrams KR, Sutton AJ, Lambert PC and Thompson JR. Bivariate random effects meta-analysis and the estimation of between-study correlation. *BMC Med Res Methodol* 2007; **7**(3): 1–15.
16. Houwelingen HC, Arends LR and Stijnen T. Advanced methods in meta analysis: multivariate approach and meta-regression. *Stat Med* 2002; **21**: 589–624.
17. Jennrich RI and Schluchter MD. Unbalanced repeated-measures models with structured covariance matrices. *Biometrics* 1986; **42**: 805–820.
18. Jackson D, White IR and Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Stat Med* 2010; **29**(12): 1282–1297.
19. Lunn DJ, Thomas A, Best N and Spiegelhalter D. WinBugs – a Bayesian modeling framework: concepts, structure and extensibility. *Stat Comput* 2011; **10**: 325–337.
20. Nam IS, Mengersen K and Garthwaite P. Multivariate meta-analysis. *Stat Med* 2003; **22**: 2309–2333.
21. Lambert PC, Sutton AJ, Burton PR, Abrams KR and Jones DR. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Stat Med* 2005; **24**: 2401–2428.
22. Gelman A. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal* 2006; **1**(3): 515–533.
23. Lu G and Ades AE. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics* 2009; **10**(4): 792–805.
24. Riley RD, Thompson JR and Abrams KR. An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics* 2008; **9**: 172–186.
25. Cheung MWL. A model for integrating fixed-, random and mixed-effects meta-analyses into structural equation modeling. *Psychol Methods* 2008; **13**: 182–202.
26. Becker BJ. Using results from replicated studies to estimate linear models. *J Educ Stat* 1992; **17**: 341–362.
27. Cheung MWL and Chan W. Meta-analytic structural equation modeling: a two stage approach. *Psychol Methods* 2005; **10**: 40–64.
28. Viswesvaran C and Ones DS. Theory testing: combining psychometric meta-analysis and structural equation modeling. *Pers Psychol* 1995; **48**: 865–885.
29. Cheung MWL. Fixed-effects meta-analyses as multiple-group structural equation models. *Struct Equ Model* 2011; **17**: 481–509.
30. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM and Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; **58**(10): 982–990.
31. Harbord RM, Deeks JJ, Egger M, Whiting P and Sterne JAC. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007; **8**: 239–251.
32. Higgins JPT and Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med* 1996; **15**(24): 2733–2749.
33. Salanti G, Higgins JP, Ades AP and Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; **17**(3): 279–301.
34. StataCorp. Statistical Software: Release 10.0. College Station: Stata Corporation, 2007.
35. SAS Institute. SAS/STAT User's guide', version 8. Cary, NC: SAS Publishing, 2000.
36. R. The R foundation for statistical computing, R version 2.13.0, 2011.
37. Joreskog KG, Sorbom D, Du Toit S and Du Toit M. *{LISREL 8}: new statistical features*. Chicago, IL: Scientific Software International, 2001.
38. Muthen BO and Muthen L. Mplus: The comprehensive modeling program to applied Researchers, 11965 Venice Boulevard, Suite 407, Los Angeles, CA 90066, 2000.
39. Harbord RM and Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *STATA j* 2011; **9**(2): 211–229.
40. Lambert PC, Sutton AJ, Burton PR, Abrams KR and Jones KR. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Stat Med* 2011; **24**: 2401–2428.
41. Cowles MK and Carlin BP. Markov chain Monte Carlo convergence diagnostics: a comparative review. *J Am Stat Assoc* 1996; **91**(434): 883–904.
42. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Software* 2010; **36**(3): 1–48.
43. Takwoingi Y. 'METADAS: A SAS Macro for Meta-analysis of Diagnostic Accuracy Studies', http://srdta.cochrane.org/software-development (2008) (accessed 3 January 2012).
44. MATLAB. The language of technical computing, version 7.0.1.24704 (R14), 2004.
45. Golfinopoulos V, Salanti G, Pavlidis N and Ioannidis JPA. Survival and disease-progression benefits with treatment regimens for advanced colorectal cancer: a meta-analysis. *Lancet Oncol* 2007; **8**: 898–911.
46. Riley RD, Abrams KR, Lambert PC, Sutton AJ and Thompson JR. An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes. *Stat Med* 2007; **26**: 78–97.
47. Smith TC, Spiegelhalter DJ and Thomas A. Bayesian approaches to random effects meta-analysis: a comparative study. *Stat Med* 1995; **14**: 2685–2699.
48. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd C, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *Br Med J* 2010; **15**: 340–365.

49. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coelllo P, et al. GRADE guidelines: 4. Rating the quality of evidence–study limitations (risk of bias). *J Clin Epidemiol* 2011; **64**(4): 407–415.
50. Kirkham JJ, Riley R and Williamson PR. Is multivariate meta-analysis a solution for reducing the impact of outcome reporting bias in systematic reviews?, Proceedings of the 18th Cochrane Colloquium, 18–22 October, Keystone, CO, 2010.

Appendix

Table A.1. Results from five published trials comparing surgical and non-surgical treatments for medium-severity periodontal disease, 1 year after treatment

	Probing depth		Attachment level		
Study	y_1	S_1^2	y_2	S_2^2	ρ
1	0.47	0.0075	−0.32	0.0077	0.39
2	0.20	0.0057	−0.60	0.0008	0.42
3	0.40	0.0021	−0.12	0.0014	0.41
4	0.26	0.0029	−0.31	0.0015	0.43
5	0.56	0.0148	−0.39	0.0304	0.34

Table A.2. Parameter values and standard errors using MCMC with different prior distributions for the heterogeneities

Prior distribution	Probing depth (μ_1)	Attachment level (μ_2)	τ_1^2	τ_2^2	$\rho_\tau \tau_1 \tau_2$
$\tau \sim HN(0, 1)$	0.358 (0.112)	−0.349 (0.163)	0.058 (0.123)	0.128 (0.201)	0.024 (0.079)
$\tau \sim HN(0, 100)$	0.350 (0.055)	−0.346 (0.082)	0.010 (0.010)	0.028 (0.015)	0.006 (0.008)
$\tau \sim U(0, 1)$	0.355 (0.114)	−0.350 (0.164)	0.059 (0.107)	0.130 (0.156)	0.029 (0.085)
$\tau \sim U(0, 100)$	0.360 (0.145)	−0.342 (0.207)	0.096 (0.401)	0.196 (0.620)	0.051 (0.232)
$\tau^2 \sim HN(0, 1)$	0.365 (0.183)	−0.347 (0.230)	0.160 (0.245)	0.266 (0.308)	0.072 (0.206)
$\tau^2 \sim HN(0, 100)$	0.361 (0.099)	−0.342 (0.124)	0.043 (0.038)	0.072 (0.044)	0.020 (0.033)
$\tau^2 \sim U(0, 1)$	0.361 (0.179)	−0.347 (0.223)	0.151 (0.187)	0.242 (0.229)	0.067 (0.178)
$\tau^2 \sim U(0, 100)$	0.369 (0.608)	−0.331 (0.748)	1.831 (7.244)	2.730 (8.826)	0.875 (5.089)
$1/\tau^2 \sim HN(0, 1)$	0.364 (0.422)	−0.336 (0.489)	0.891 (0.607)	1.157 (1.088)	0.552 (0.843)
$1/\tau^2 \sim HN(0, 100)$	0.372 (1.391)	−0.356 (1.468)	9.327 (20.54)	10.72 (8.130)	4.48 (8.936)
$1/\tau^2 \sim U(0, 1)$	0.367 (0.571)	−0.326 (0.610)	1.609 (1.109)	1.863 (1.595)	1.173 (1.280)
$1/\tau^2 \sim U(0, 100)$	0.357 (0.075)	−0.342 (0.095)	0.021 (0.016)	0.036 (0.033)	0.007 (0.012)
$\log(\tau^2) \sim U(-4, 4)$	0.365 (0.119)	−0.341 (0.146)	0.063 (0.176)	0.104 (0.238)	0.022 (0.112)
$\log(\tau^2) \sim U(-10, 10)$	0.350 (0.072)	−0.349 (0.123)	0.019 (0.082)	0.069 (0.135)	0.010 (0.041)
$\log(\tau) \sim U(-10, 10)$	0.350 (0.055)	−0.354 (0.120)	0.011 (0.045)	0.065 (0.100)	0.006 (0.020)
$\log(\tau) \sim U(-4, 4)$	0.352 (0.076)	−0.348 (0.124)	0.022 (0.052)	0.068 (0.113)	0.011 (0.033)

Table A.3. Log-hazards ratio for overall and progression-free survival

Studies	y_1	S_1^2	y_2	S_2^2
1	0.0068	0.0284	−0.2114	0.0234
2	0.069	0.0714	−0.1823	0.0551
3	−0.2103	0.01	−0.5276	0.0173
4	−0.0488	0.04	−0.1128	0.0376
5	−0.1278	0.0138	−0.4308	0.0104
6	−0.1398	0.026	−0.5108	0.0577
7	−0.0834	0.0064	−0.2614	0.0052
8	−0.2485	0.0121	−0.4463	0.0125
9	−0.1222	0.0041	−0.2995	0.0037

Codes for fitting multivariate meta-analysis in WinBUGS, R and MATLAB

WinBUGS

Data. # n is the number of studies and in the codes
p is the number of outcomes
y is a $k \times p$ matrix with effect sizes
s1 is the vector of variances for effect size y_1
s2 is the vector of variances effect size y_2
s12 is the covariance of y_1 and y_2

Periodontal disease data. list(n=5, p=2,y=structure(.Data=c(0.4700, -0.3200, 0.2000, -0.6000, 0.4000, -0.1200, 0.2600, -0.3100, 0.5600, -0.3900),.Dim=c(5,2)), s2=c(0.0077, 0.0008, 0.0014, 0.0015, 0.0304),s1=c(0.0075, 0.0057, 0.0021, 0.0029, 0.0148),s12=c(0.003, 0.0009, 0.0007, 0.0009, 0.0072))

Colorectal cancer data. list(n=9, p=2, y=structure(.Data=c(0.0068, -0.2114, 0.069, -0.1823,-0.2103, -0.5276,-0.0488, -0.1128,-0.1278, -0.4308,-0.1398,-0.5108,-0.0834, -0.2614,-0.2485, -0.4463,-0.1222, -0.2995),.Dim=c(9,2)), s2=c(0.0234, 0.0551,0.0173,0.0376, 0.0104, 0.0577, 0.0052, 0.0125, 0.0037), s1=c(0.0284,0.0714,0.0100, 0.0400,0.0138, 0.0260, 0.0064, 0.0121, 0.0041))

WinBUGS code 1: random-effects bivariate meta-analysis with known within-study correlations (equation (6))

```
model{
  for (i in 1:n){
  w1[i]<-1/s1[i]
  rho[i]<-s12[i]/sqrt(s1[i]*s2[i])
  y[i,1]~dnorm(m[i,1],w1[i])
  y[i,2]~dnorm(my[i],w2[i])
  my[i]<-m[i,2]+rho[i]*sqrt(s2[i])/sqrt(s1[i])*(y[i,1]-m[i,1]) # E(y_i2|y_i1)
  vary[i]<-s2[i]*(1-pow(rho[i],2)) # V(y_i2|y_i1)
  w2[i]<-1/vary[i]
  }
  for (i in 1:n){
  m[i,1]~dnorm(mean[1],prec[1])
  m[i,2]~dnorm(m.re[i],prec.vre)
```

```
  m.re[i]<-mean[2]+rho.tau*sqrt(tau.sq[2])/sqrt(tau.sq[1])*(m[i,1]-mean[1])
  # E(θ_i2|θ_i1)
  }
  v.re<-tau.sq[2]*(1-pow(rho.tau,2)) # V(θ_i2|θ_i1)
  prec.vre<-1/v.re
  pi<-3.14
  a~dunif(0, pi) #spherical parameterization
  rho.tau<-cos(a)
  for (j in 1:p) {
  tau.sq[j]<-1/prec[j]
  }
  for (j in 1:p) {
  mean[j]~dnorm(0,0.001)
  prec[j]~dunif(0,100)} #1/τ_j^2 ~ U(0,100).
  cov.tau<-sqrt(tau.sq[1])*sqrt(tau.sq[2])*rho.tau
  }
```

WinBUGS code 2: random-effects bivariate meta-analysis with unknown within-study correlations

```
model{
  for (i in 1:n){
  rho[i]~dbeta(1,1# a uniform prior on [0,1]
  #rho[i]~dbeta(1,5) # a Beta(1,5) prior
  #rho[i]~dbeta(5,1) # a Beta(5,1) prior
  }
  for (i in 1:n){
  w1[i]<-1/s1[i]
  y[i,1]~dnorm(m[i,1],w1[i])
  y[i,2]~dnorm(my[i],w2[i])
  my[i]<-m[i,2]+rho[i]*sqrt(s2[i])/sqrt(s1[i])*(y[i,1]-m[i,1])
  vary[i]<-s2[i]*(1-pow(rho[i],2))
  w2[i]<-1/vary[i]
  }
  for (i in 1:n){
  m[i,1]~dnorm(mean[1],prec[1])
  m[i,2]~dnorm(m.re[i],prec.vre)
  m.re[i]<-mean[2]+rho.tau*sqrt(tau.sq[2])/sqrt(tau.sq[1])*(m[i,1]-mean[1])
  }
  v.re<-tau.sq[2]*(1-pow(rho.tau,2))
  prec.vre<-1/v.re
  pi<-3.14
  a~dunif(0, pi)
  rho.tau<-cos(a)
  for (j in 1:p) {
  mean[j]~dnorm(0,0.001)
  }
  for (j in 1:p) {
  tau.sq[j]<-1/prec[j]
```

```
  prec[j]~dunif(0,100)
  }
  cov.tau<-sqrt(tau.sq[1])*sqrt(tau.sq[2])*rho.tau}
```

WinBUGS code 3: AM that does not require within-study correlations

```
model{
  for (i in 1:n){
  w1[i]<-1/(s1[i]+psi[1])
  y[i,1]~dnorm(mean[1],w1[i])
  y[i,2]~dnorm(my[i],w2[i])
  my[i]<-mean[2]+rho*sqrt((psi[1]+s1[i])*(psi[2]+s2[i]))*w1[i]*(y[i,1]-mean[1])
  vary[i]<-psi[2]+s2[i]-
rho*sqrt((psi[1]+s1[i])*(psi[2]+s2[i]))*w1[i]*sqrt((psi[1]+s1[i])*(psi[2]+s2[i]))
  w2[i]<-1/vary[i]
  }
  rho~dunif(-1,1)
  for (j in 1:p) {
  mean[j]~dnorm(0,0.001)
  psi[j]~dunif(0,100)
  }}
```

WinBUGS code for assigning different prior distributions on heterogeneity

```
  # inverse gamma prior on precision
  for (j in 1:p) {
  prec[j]~dgamma(0.1,0.1)
  tau.sq[j]<-1/prec[j]
  tau[j]<-sqrt(tau.sq[j])
  }
  # Uniform [−10,10] on heterogeneity
  for (j in 1:p) {
  lts[j]~dunif(-10,10)
  log(tau.sq[j])<-lts[j]
  tau[j]<-sqrt(tau.sq[j])
  prec[j]<-1/tau.sq[j]
  }
  # Half-Normal [0,100] on heterogeneity
  for (j in 1:p) {
  tau.sq[j]~dnorm(0,1)I(0,)
  tau[j]<-sqrt(tau.sq[j])
  prec[j]<-1/tau.sq[j]
  }
```

R code

GLS estimation for a fixed-effects model

```
glsfe=function(y1,y2,s1,s2,s12){
  missing1=is.na(y1) # we check if there are any missing values in y1
```

```
  y1[missing1]=0# replace the missing outcome with mean zero with
  s12[missing1]=0 # a very large variance and correlation equal to zero
  s1[missing1]=1000
  missing2=is.na(y2)
  y2[missing2]=0 # the same is repeated for the second outcome
  s12[missing2]=0
  s1[missing1]=1000
  yy=cbind(y1,y2)
  size=dim(yy)
  n=size[1]
  p=size[2]
  y=NULL
  for (i in 1:n) {for (j in 1:p) {
  y[p*(i-1)+j]=yy[i,j]}}
  tt=matrix(0,n,p^2)
  V=rep(0,n*p^2)
  dim(V)=c(p,p,n)
  for (i in 1:n)
  {tt[i,1:p^2]=c(s1[i], s12[i],s12[i], s2[i])
  V[,i]=matrix(tt[i,],2,2)}
  C=matrix(rep(0,n*p),n*p,n*p)
  for (i in 1:size[1]) {
  C[((i-1)*2+1):((i-1)*2+p),((i-1)*2+1):((i-1)*2+p)]=V[,i];
  }
  X=matrix(rep(c(1,0,0,1),5),n*p,p,byrow=T)
  beta=solve(t(X)%*%solve(C)%*%X)%*%t(X)%*%solve(C)%*%y
  varbeta=solve(t(X)%*%solve(C)%*%X)
  result <- list(beta=beta,varbeta=varbeta)
  result
  }
```

GLS estimation for a random-effects model

```
glsre=function(y1,y2,s1,s2,s12,tol){
  missing1=is.na(y1) # we check if there are any missing values in y1
  y1[missing1]=0 # replace the missing outcome with a zero value with
  s12[missing1]=0 # a very large variance and correlation equal to zero
  s1[missing1]=1000
  missing2=is.na(y2)
  y2[missing2]=0 # the same is done for the second outcome
  s12[missing2]=0
  s1[missing1]=1000
  yy=cbind(y1,y2)
  size=dim(yy)
  n=size[1]
  p=size[2]
  y=NULLv
  for (i in 1:n) {
```

```
  for (j in 1:p) {y[p*(i-1)+j]=yy[i,j]}}
  tt=matrix(0,n,p^2)
  V=rep(0,n*p^2)
  dim(V)=c(p,p,n)
  for (i in 1:n)
  {tt[i,1:p^2]=c(s1[i], s12[i],s12[i], s2[i])
  V[,i]=matrix(tt[i,],2,2)}
  C=matrix(rep(0,n*p),n*p,n*p)
  for (i in 1:n) {
  C[((i-1)*2+1):((i-1)*2+p),((i-1)*2+1):((i-1)*2+p)]=V[,i];
  }
  X=matrix(rep(c(1,0,0,1),5),n*p,p,byrow=T)
  beta=matrix(nrow=2,ncol=10)
  beta[,1]=c(0,0)
  VN=rep(0,n*p^2)
  dim(VN)=c(p,p,n)
  sumV=matrix(rep(0,4),2,2)
  Delta=matrix(rep(0,4),2,2)
  for (i in 1:n) {sumV=sumV+V[,i]}
  for (j in 1:100) {
  et=y-X%*%beta[,j]
  e=matrix(et,n,p,byrow=T)
  Delta=1/(n-2)*t(e)%*%e-1/n*sumV
  for (i in 1:n)
  {VN[,i]=V[,i]+Delta;}
  for (i in 1:n) {
  C[((i-1)*2+1):((i-1)*2+p),((i-1)*2+1):((i-1)*2+p)]=VN[,i];}
  beta[,j+1]=solve(t(X)%*%solve(C)%*%X)%*%t(X)%*%solve(C)%*%y
  if (max(abs(beta[,j+1]-beta[,j]))<tol) {break}
  }
  beta=beta[,j+1]
  varbeta=solve(t(X)%*%solve(C)%*%X)
  result=list(beta=beta,varbeta=varbeta,Delta=Delta)
  result
  }
```

MATLAB

GLS estimation

```
function[beta,varbeta,Delta]=glsmma(y1,y2,s1,s2,s12,tol)
  % This function computes the Generalized Least Squares estimates for
  % bivariate meta-analysis for the xed-e␣ect model and the random % e␣ects model
  %Input variables should be row vectors, the two e␣ect sizes y1 and %y2, the
corresponding variances s1 and s2
  % their covariance s12 and a tolerance level for the Newton-Raphson % algorithm to stop
when a random-e␣ects model is assumed
  % Example (Berkey data)
```

```
% s2=[0.0077 0.0008 0.0014 0.0015 0.0304];
% s1=[0.0075 0.0057 0.0021 0.0029 0.0148];
% s12=[0.003 0.0009 0.0007 0.0009 0.0072];
% y1=[0.47 0.2 0.4 0.26 0.56];
% y2=[-0.32 −0.6 −0.12 −0.31 −0.39];
% For a xed-e␣ect model [beta,varbeta]=glsmma(y1,y2,s1,s2,s12,tol)
% and for a random-e␣ects model %[beta,varbeta,Delta]=glsmma(y1,y2,s1,s2,s12,tol)
if sum(isnan(y1))~=0
y1(nd(isnan(y1)))=0;
s1(nd(isnan(s1)))=1000;
s12(nd(isnan(s12)))=0;
end
if sum(isnan(y2))~=0
y2(nd(isnan(y2)))=0;
s2(nd(isnan(s2)))=1000;
s12(nd(isnan(s12)))=0;
end
y=[y1' y2'];
[n,p]=size(y);
y=reshape(y',n*p,1);
for i=1:n
V(:,:,i)=[s1(i) s12(:,i);s12(:,i) s2(i)];
end
C=zeros(n*p,n*p);
for i=1:n
C((i-1)*2+1:(i-1)*2+p,(i-1)*2+1:(i-1)*2+p)=V(:,:,i);
end
X=[repmat(eye(p),n,1)];
if nargout==2
beta=pinv(X'*pinv(C)*X)*X'*pinv(C)*y;
varbeta=pinv(X'*pinv(C)*X);
else
C_se=diag(diag(C));
beta=pinv(X'*pinv(C_se)*X)*X'*pinv(C_se)*y;
et=y-X*beta;
e=reshape(et,n,p);
sumV=sum(V,3);
for j=1:100
et=y-X*beta(:,j);
e=reshape(et,p,n)';
Delta(:,:,j)=1/(n-2)*e'*e-1/n*sumV;
for i=1:n
VN(:,:,i)=V(:,:,i)+Delta(:,:,j);
end
for i=1:n
C((i-1)*2+1:(i-1)*2+p,(i-1)*2+1:(i-1)*2+p)=VN(:,:,i);
end
```

```
beta(:,j+1)=pinv(X'*pinv(C)*X)*X'*pinv(C)*y;
if min(abs(beta(:,j+1)-beta(:,j)))<=tol
break
end
end
beta=beta(:,j+1);
Delta=Delta(:,:,j);
varbeta=pinv(X'*pinv(C)*X);
end
end
```

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.