Applying Bivariate Meta-analyses when Within-study Correlations are Unknown

Thilan, A.W.L.P. [*1], Jayasekara, L.A.L.W. [2]

[*1] Corresponding author, Department of Mathematics, Faculty of Science, University of Ruhuna, Matara, Sri Lanka,

[2] Professor, Department of Mathematics, Faculty of Science, University of Ruhuna, Matara, Sri Lanka,

Abstract

Multivariate meta-analysis can be used to combine several outcome measures instead of only concerning one specific outcome measure. We need to quantify the correlations between outcomes across studies when fitting multivariate models. The model requires an estimate of the correlations between treatment effect for each study. These correlations are known as within study correlations and rarely available in published literature. This limits the multivariate approach in practice. Therefore, we have to either approximate them or ignore correlations between effect estimates within the same studies. In this paper, we discuss some widely applicable ways in which this problem can be resolved. We also examine the use of the Bayesian correlation method as a novel approach for dealing with unknown within-study correlations in bivariate meta-analysis. When there is heterogeneity of effects across studies and high correlation within studies, our approach perform quite well. An application to a meta-analysis of treatments for acute stroke data illustrates the use of the approximated correlation in bivariate meta-analysis with correlated outcomes.

Key Words: multivariate meta-analysis, within study correlations, unknown covariance, Bayesian correlation method, Pearson correlation

1. Introduction

Meta-analysis has become an increasingly important technique and it enables researchers to combine the statistical results of many pieces of research on the same research question. Univariate meta-analysis refers to quantitative methods of synthesizing one statistical result from many studies that examine a certain topic. However, having measures on many variables limit the use of univariate meta-analysis and seek methods that allow combining the results of different studies where each contains a multivariate analysis. There are a variety of multivariate techniques for data synthesis across such studies and one needs to choose a specific technique for a given research problem.

Meta-analysis is typically done on independent studies, and consequently their effect sizes are also independent. However, this is not true for each study. Some studies can use multiple treatments compared to a common control group resulting in correlations between effect sizes. When studies sharing a common control group, the estimate of effect sizes may not be independent and such studies are called as multiple-treatment studies (Gleser & Olkin, 1994). Moreover, studies can have combination of several outcome measures instead of only concerning one specific outcome measure. If such outcome measures come from same individual of interest, they are likely to be correlated; corresponding estimated effect sizes for these measures will be correlated within studies. Studies of this type are called as multiple-endpoints studies (Berkey et al., 1996, Gleser & Olkin, 1994).

Multivariate meta-analysis allows a joint synthesis of the multiple end points or multiple treatment effects and it will produce a pooled result for each end point or each treatment effect simultaneously. This can account for any correlation between end points or treatment effects and the correlation may exist both within studies and between studies. The within-study correlation indicates the association between the summary end points or treatment effect estimates within a study. Same individual in a study contributing towards multiple treatment effects or multiple end points is the reason for that type of correlation. The between-study correlation indicates how the true underlying end point or treatment effect summary values are related across studies and is caused by deference across studies in patient characteristics, such as age, or changes in study characteristics, like the threshold level in diagnostic studies (Riley, 2009).

When the summary data per study are multi-dimensional, practitioners usually choose to perform a separate univariate meta-analysis of each end point or each outcome variable. One reason for this may include the increased complexity of the multivariate approach. The requirement of special statistical software is another problem. Perhaps a lack of understanding as to when multivariate meta-analysis is beneficial or a lack of understanding the consequences of ignoring correlation in meta-analysis (Riley, 2009). Another major barrier is that multivariate meta-analysis models require, from each study, the within-study correlation or individual participant data are available. Unfortunately, within-study correlations or individual participant are rarely reported in primary study publications, and conceivably this must put practitioner off the multivariate meta-analysis approach.

The remainder of this paper is organized as follows. In Section 2 we describe the bivariate fixed effect, bivariate random effect, and bivariate marginal models and then extend the discussion with the detail on how to estimate the pooled estimates, the effect of ignoring within-study correlation, a review of dealing with unavailable within-study correlation needed for the multivariate meta-analysis approach, and a description of a new method called Bayesian correlation approach to approximate this correlation when it cannot be obtained

from alternative sources. In Section 3 we illustrate a simulation study to explore how the statistical properties of the pooled effect estimates are affected by (i) the estimation of unknown with-study correlations using Pearson correlation approach and new Bayesian correlation approach, (ii) strength of the estimated within-study correlations, and (iii) strength of the between study and the within-study variances. The simulation allows us to compare the performance of bivariate fixed and random effect models against separate univariate fixed effect meta-analyses models. In Section 4 we apply the methods to the motivating example. In Section 5, we conclude with a discussion of the benefits of using new Bayesian correlation approach to estimate unknown within-study correlations.

2. Material and Methods

2.1 Bivariate Fixed Effect Meta-Analysis Model (BFMA)

In a bivariate fixed effect model, the two correlated end points or treatment effects of interest are assumed to follow a bivariate normal distribution within studies

$$\mathbf{y}_i|\boldsymbol{\mu}_i \sim N(\boldsymbol{\mu}_i, \mathbf{S}_i), \quad \mathbf{S}_i = \begin{pmatrix} S_{11i}^2 & S_{11i}S_{22i}\rho_{S_i} \\ S_{11i}S_{22i}\rho_{S_i} & S_{22i}^2 \end{pmatrix}, \tag{2.1}$$

where N denotes a bivariate normal distribution, $\boldsymbol{\mu}_i = (\mu_{1i}, \mu_{2i})$ is the true underlying effect for the i^{th} study and $\mathbf{S}_i$ is the covariance matrix of $\mathbf{y}_i$. The matrices $\mathbf{S}_i$ are referred to as the within-study covariance matrices; their entries are estimated in practice using the individual patient data (IPD) for each study separately but regarded fixed and known when pooling the results to make inferences. The within-study standard errors are usually reported in primary publications and they can be used to obtain the within-study variances (the diagonal entries of $\mathbf{S}_i$). The model (2.1) requires $\rho_{\mathbf{S}i}$ (off-diagonal entries of $\mathbf{S}_i$) to be available in those studies providing both end points, which is unfortunately unlikely. We will review available methods for dealing with unknown within-study correlations in Section 2.6 and also propose a new approach to conduct multivariate meta-analysis in the absent of within-study correlations in the Section 2.6.8.

2.2 Bivariate Random Effects Meta-Analysis Model (BRMA)

The bivariate random effects model allows the μi to vary from one study to the next and further assumes that

$$\boldsymbol{\mu}_i \sim N(\boldsymbol{\mu}, \mathbf{D}), \quad \mathbf{D} = \begin{pmatrix} D_{11}^2 & D_{11}D_{22}\rho_D \\ D_{11i}D_{22}\rho_D & D_{22}^2 \end{pmatrix}, \tag{2.2}$$

where $\boldsymbol{\mu}$ is the average effect from a normal distribution of study treatment effects and $\mathbf{D}$ is the between-study covariance matrix. The between-study variances D_{11}^2 and D_{22}^2 account for any heterogeneity in μ_{1i} and μ_{2i} across studies, and ρ_D represents their between-study correlation.

2.3 Bivariate Marginal Model

The conventional bivariate random effects meta-analysis model is marginally given by

$$\boldsymbol{y}_i \sim N(\mu, \mathbf{S}_i + \mathbf{D}), \tag{2.3}$$

where the $\mathbf{y}_i$ are further assumed to be independent because they come from separate studies. The model (2.3) reverts to a bivariate fixed effects meta-analysis when $\mathbf{D} = 0$, and to a separate univariate meta-analysis of each end point when $\rho S_i = \rho_D = 0$, i.e. all correlations are 0. Usually the objective of meta-analysis is to estimate $\boldsymbol{\mu}$

and $\mathbf{D}$. Once $\widehat{\mathbf{D}}$ has been calculated, the estimated between-study correlations can be obtained directly as the appropriate entry of $\widehat{\mathbf{D}}$ divided by the corresponding between-study standard deviations, which are obtained as the square roots of the diagonal entries.

2.4 Estimation

A variety of approaches for fitting the random effects model for meta-analysis have been developed. Assuming all studies provide all effects, the pooled estimates $\widehat{\boldsymbol{\mu}}$ are given in terms of $\widehat{\mathbf{D}}$ by

$$\widehat{\boldsymbol{\mu}} = \left(\sum_{\mathbf{i=1}}^{\mathbf{n}} (\mathbf{S_i} + \widehat{\mathbf{D}})^{-1}\right)^{-1} \left(\sum_{\mathbf{i=1}}^{\mathbf{n}} (\mathbf{S_i} + \widehat{\mathbf{D}})^{-1} \mathbf{y_i}\right), \qquad (2.4)$$

where n is the number of studies.

2.5 The Effect of Ignoring Within-study Correlation

Multivariate meta-analysis models require, from each study, the within-study correlation ρ_{Si} (or equivalently S_{12i}) to be available. But, in practice within-study correlations are rarely reported in primary study publications. In some scenarios the within-study correlations can justifiably be assumed zero or close to zero (Reitsma, et al., 2005, Daniels & Hughes, 1997, Korn, et al., 2005, Thompson & Sharp, 1999), such as in diagnostic studies where sensitivity and specificity estimates are independently derived from separate patients. However, in other settings such as the meta-analysis of longitudinal data (Jones, et al., 2009), or for multiple outcomes such as overall and disease-free survival (Riley, et al., 2007), the true within-study correlations are likely to be non-zero. One study in the literature concludes that if interest lies only in the pooled effects one can fit multivariate model with S_{12i} set to 0 without any significant risk of bias or loss of precision in estimates. However, this recommendation has been questioned by analytically assessing the influence of within-study correlation on the pooled estimates and also by some simulation studies (Riley, 2009).

2.6 Estimating the Within-study Correlation

In BFMA, we assume that (y_{1i}, y_{2i}) follow a bivariate normal distribution as indicated in the model (2.1) and in which ρ_{Si} are the within-study correlations. They indicate the dependency between outcome estimates within a study and which is what we want in implementing a BFMA model. Meta-analysts should not ignore the dependence among study outcomes and should use some procedure to deal with dependence. Some widely applicable ways in which this problem can be resolved are described here and we also propose and evaluate a new method, which we termed as Bayesian correlation approach in the Section 2.6.8.

2.6.1 Individual Patient Data Approach

One proposed method for dealing with unknown within-study correlation is individual patient data approach. It involves the collection of raw data from the studies of interest and it allows one to access complete data records from each of the included studies (Dutton, 2011). In multivariate meta-analysis, availability of IPD allows us to calculate the within-study correlation directly in each study. For example, the within-study correlation between the effects of treatment on systolic and diastolic blood pressure can be calculated with the use of individual patient data, by modelling a bivariate regression model between two outcomes jointly in each study (Riley, et al., 2008). Bootstrapping methods may be required to obtain the within-study correlations

using IPD in more complex modelling situations, for instance two different survival outcomes are of interest (Daniels & Hughes, 1997). One issue with this approach is that the individual patient data may not be available in all studies and therefore within-study correlations are still not fully available. In such a situation, one can compute the within-study correlation from a single study where individual data are available and it can be used as a 'likely' within-study correlation estimate for all other studies with unavailable within-study correlations. An average correlation could be used if the IPD is available for more than one study (Kirkham, et al., 2012).

2.6.2 Biological Reasoning (Expert Opinion)

In the absent of IPD, biological reasoning or expert opinion may be used to approximate the within-study correlation. An expert in the clinical field could be asked to suggest a plausible within-study correlation between the estimators for use in all studies. For example, researchers hypothesize that the relationship between all-cause mortality and treatment failure has a positive correlation in the study where beta-lactam is used for the treatment of cancer patients with neutropenia (Kirkham, et al., 2012). In this case even though the direction of the correlation is distinct, the strength might not be determined. However, in some particular situations, an expert can give a numerical value for the correlation based on his prior experience. The incorporation of expert opinion in estimating within study correlation is an alternative in multivariate meta-analysis as it provides a quick and inexpensive alternative when IPD are not available. However, careless use of expert's opinions can result in inaccurate or bias conclusions. Issues such as over confidence, representativeness, translation and linguistic uncertainty can lead experts to provide false or misleading opinions. Hence, the accuracy of the approximation of within-study correlations is an important determinant of the impact of the expert opinion to the outcome of the meta-analysis.

2.6.3 Narrow the Range of Possible Values

In some situations, it is possible to narrow the range of possible values for the unknown within-study correlations. For example, external information has been used for this purpose in the literature (Raudenbush, et al., 1988). Another study has narrowed the range of possible values for the within-study correlation by calculating lower and upper bounds from the 2×2 tables that were available from each study (Berrington & Cox, 2003). The identification of a range of correlation values has similarly helped to inform meta-analysis in other contexts (Abrams, et al., 2005).

2.6.4 Perform Sensitivity Analyses

In the absent of IPD, a sensitivity analysis could be performed by imputing correlations over the entire range of values (i.e. from - 1 to 1), to assess whether and how conclusions depend on the correlation that is imputed. For example, if the eligibility of some studies in the meta-analysis is dubious because unavailability of within study correlations, sensitivity analysis may involve undertaking the meta-analysis twice: first, including all studies and second, only including those that are definitely known to be eligible. In a Bayesian framework, a *uniform (-1, 1)* prior distribution has been used on the within-study correlation and then assessed whether conclusions are robust to change in the specification of this prior (Riley, 2009). However, the use of sensitivity analysis for the estimation of within-study correlations can become problematic in a situation where the dimension is more than two.

2.6.5 Use an Alternative Model

Unavailability of within-study correlations has motivated researchers to build alternative multivariate random effects models for meta-analysis which do not require the within-study correlations. Riley, et al., (2008) has proposed such an alternative model for bivariate random effects meta-analysis and this model includes only one overall correlation parameter ρ which is a combination of the within-study and between-study correlations. This removes the need to know the within-study correlations, and the data that are required to fit the model are the same as those needed for a separate univariate analysis of each outcome, which makes it widely applicable. The alternative BRMA model can be specified as

$$\binom{y_{1i}}{y_{2i}} \sim N\left\{\binom{\mu_1}{\mu_2}, \emptyset_i\right\} \tag{2.5}$$

$$\emptyset_i = \begin{pmatrix} S_{11i}^2 + \varphi_{11}^2 & \rho\sqrt{(S_{11i} + \varphi_{11}^2)(S_{22i} + \varphi_{22}^2)} \\ \rho\sqrt{(S_{11i} + \varphi_{11}^2)(S_{22i} + \varphi_{22}^2)} & S_{22i}^2 + \varphi_{22}^2 \end{pmatrix}, \tag{2.6}$$

where φ_{jj}^2 indicates the additional variation beyond sampling error and ρ denotes the overall correlation. In here, interest lies only in the pooled estimates, or some function of them, and estimation can however become unstable when the estimated correlation ($\hat{\rho}$) is very close to 1 or -1. This alternative model produces pooled estimates that are very similar to those from fitting the general BRMA model where the within-study correlations are known and have better statistical properties than those from separate URMAs, especially given missing data (Riley, et al., 2008).

2.6.6 Pearson Correlation Method

Another approach known as the Pearson correlation method has been proposed with the use of treatment effect estimates in the literature (Kirkham, et al., 2012). In the present of data on both outcomes, the Pearson correlation coefficient can be calculated between the pairs of treatment effect estimates for a bivariate analysis and it can be assumed as a common within-study correlation ρ in each study. Simply, we have to assume that the within-study correlations are the same in each study and that this correlation will be closely reflected in the observed correlation between paired outcome effects across studies. In conclusion, a precise estimate of the Pearson correlation could be obtained with five data points (Abdel-Megeed, 1984).

2.6.7 Use of Formulas

Most imputation approaches have been based on imputing the correlation between treatment effect estimates and assuming this correlation to be identical for every study. Wei and Higgins (2013) propose an approach to approximate the within-study covariance based on information about likely correlations between underlying outcomes instead of treatment effect estimates. In brief, the treatment effect is a quantity that describes the benefit or harm of the treatment, whereas the outcome is the direct measurement on the participants. Evaluating within-study correlation at the outcome level can bring two main advantages. First, these correlations can easily be obtained from external sources, and if not, then plausible values for them could be provided than between treatment effect estimates. Second, these correlations are more natural descriptors of inherent similarities and allow the correlations between treatment effect estimates to vary according to other

measurable features of the study. This paper considers both continuous and dichotomous outcomes, which are the most common in meta-analysis. For instance, the covariance between two estimates of mean differences is given in closed form as

$$COV(MD_1, MD_2) = \frac{n_{12t}}{n_{1t}n_{2t}}\rho S_{1t}S_{2t} + \frac{n_{12c}}{n_{1c}n_{2c}}\rho S_{1c}S_{2c} \quad (2.7)$$

where, n_{1t} , n_{2t} , and n_{12t} denote the number of participants who report outcome 1, of those who report outcome 2, and of those who report both outcome 1 and outcome 2, respectively, in the treatment group. In a similar way, we define n_{1c}, n_{2c}, and n_{12c} for the control group. The correlation between the two outcomes themselves is denoted by ρ and is substituted by a value taken from an external source of information.

2.6.8 Bayesian Correlation Method

In this paper we propose a novel approach and which we call as Bayesian correlation method to deal with unknown within-study correlations in bivariate meta-analysis. The Pearson's correlation approach is quite simple whenever someone wants to get an approximation for the common within study correlation. In practice, Pearson's correlation coefficient is calculated when both variables being studied are normally distributed. This coefficient is highly affected by the outliers, which may exaggerate or dampen the value of the coefficient (Mukaka, 2012). The smaller the number of available treatment effect estimates, the greater the effect of the outliers on value of correlation coefficient. In meta-analysis, the coverage for both outcomes can be slightly worse for the BFMA approach when the true within-study correlation is largely overestimated by the Pearson's correlation approach; this was particularly evident with fewer studies included (Kirkham, et al., 2012).

In Bayesian theory, the estimation of correlation coefficient reduces to estimating the parameters of a bivariate normal distribution given some data. The bivariate normal distribution is not parameterized using ρ, that is, we cannot estimate ρ directly. The model is defined in terms of μ_x and μ_y the means of the two marginal distributions and a covariance matrix Σ which define σ_x^2 and σ_y^2, the variances of the two marginal distributions, and the covariance, how much the marginal distributions vary together. Given correlation as ρ, the covariance can be written down as $\rho\sigma_x\sigma_y$. The model that we want to estimate is

$$[x_i, y_i] \sim N\big([\mu_x, \mu_y], \Sigma\big), \quad \Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}. \quad (2.8)$$

In a Bayesian framework, we need to choose suitable prior distributions to implement the model (2.8). Even though the covariance matrix plays the main role in the model, yet modeling a covariance matrix is often a difficult task in practice due to its dimensionality and the non-negative definite constraint. When modeling correlations, a prior distribution is often chosen for the covariance matrix. The most commonly used prior for the covariance matrix is Inverse-Wishart distribution. In order to model a covariance matrix directly, recent interest has focused on broken down it into variance components. Priors are then assigned separately forσ_x, σ_y and ρ. The main advantage of this approach is the greater flexibility (Barnard, et, al., 2000). In order to make our Bayesian correlation estimate more robust, we can replace the bivariate normal distribution to a bivariate t-distribution. Then, it is required to specify a prior for the degree of freedom that both allow for completely normally distributed data or normally distributed data with an occasional outliers. In practice, it is recommended to use the t-distribution because it also estimates the heaviness of the tails of the bivariate t-

distribution and if there is sufficient evidence in the data for normality the estimated t-distribution will be very close to a normal distribution.

In the model 2.8, the pair (x_i, y_i) represents the two treatment effect estimates. Normal priors were assigned for the mean parameters, μ_x and μ_y. The standard deviation S_μ of the prior on μ was seted as 1,000 times the standard deviation of the pooled data to keep the prior distribution broad relative to the arbitrary scale of the data. To keep the prior scaled appropriately relative to the arbitrary scale of the data, the mean M_μ of the prior on μ is arbitrarily set to the mean of the pooled data. A uniform distribution has been assigned as the prior on the standard deviation and in which low value L_σ, set to one thousandth of the standard deviation of the pooled data, to a high value H_σ , set to one thousand times the standard deviation of the pooled data. The degrees of freedom was assigned that is exponentially distributed, which spreads prior credibility fairly evenly over nearly normal and heavy tailed data (Kruschke, 2013).

2.7 Software Implementation

The model 2.8 can be implemented with *R* and the *JAGS* sampler interfaced with *R* using the rjags package. *JAGS* stands for Just Another Gibbs Sampler. It is a program for the analysis of Bayesian models using Markov Chain Monte Carlo (MCMC). *JAGS* is designed to work closely with the *R* language. The *rjags* package works directly with *JAGS* from within *R*. Running a model refers to generating samples from the posterior distribution of the model parameters.

As the number of iterations tends to infinity, the output from an MCMC sampler converges to the posterior distribution of the model parameters. Usually, the MCMC output is divided into two parts: an initial ”burn-in” period, which is discarded, and the remainder of the run, in which the output is considered to have converged (sufficiently close) to the target distribution. Samples from the second part are used to create approximate summary statistics for the target distribution. A reasonable point estimate for correlation coefficient can be obtained from this summary statistics.

3. Simulation Study

A simulation study was carried out to demonstrate the estimation properties from BRMA when a common within study correlation is estimated using Pearson correlation and also using new Bayesian approach. All BRMA simulations results were compared with corresponding UFMA and URMA results.

The within study variances were generated from a 0.25x χ_1^2 distribution. Two set of values (length n) were obtained as one for within study variance of outcome X and one for within study variance of outcome Y. Values outside the range [0.009, 0.6] were disregarded and a new value was generated. Simulated within study variances were then sorted so that the first study has the largest pair of values ($s_{X_i}^2$ and $s_{Y_i}^2$) and so on, until the last study had the smallest pair of values. For each meta-analysis in the simulation, new set of within study variances were generated. This simulation procedure has been used previously and has been shown to simulate a realistic mixture of study sample sizes (Jackson, et al., 2010, Jamie, et al., 2012). The model (2.3) was used to simulate pairs of X_i and Y_i by taking both the true overall treatment effects μ_1 and μ_2 be zero for all simulation scenarios.

The following criterion was used in order to choose suitable parameter values to investigate in the simulation study. The proportion of marginal variation in X and Y due to heterogeneity is given by $I_X^2 = D_{11}^2/(0.056 +$

D_{11}^2) and $I_Y^2 = D_{22}^2/(0.056 + D_{22}^2)$. Three values of these I^2 terms were considered: 0 (no marginal between-study heterogeneity), 0.3 (mild heterogeneity) and 0.75 (notable heterogeneity), giving nine pairs of I^2 values. In order to investigate the special case where all outcomes are independent, both the between-study correlation ρ_D and all within-study correlations ρ_{S_i} set to 0 in simulation runs 1-9. When one or both of the I^2 values is 0 the correlation ρ_D is not defined, and $\rho_D = 0$ is then taken to mean that the covariance $\rho_D D_{11} D_{22} = 0$ in such instances. Runs 10-17 considered situations where the between and within-study correlations are similar, where ρ_D and all ρ_{S_i} were set to 0.7 or to 0.95; only the combinations of $I_X^2 = (0.3, 0.75)$ and $I_Y^2 = (0.3, 0.75)$ were considered when using these correlations, as values I^2 of zero do not permit such a correlation. Runs 18-25 repeated runs 10-17 with the ρ_{S_i} all set to zero.

3.2 Models Fitted to Each Generated Dataset

Different models were fitted to each meta-analysis dataset generated in the simulation study. The UFMA and URMA models were fitted and then the BRMA model was fitted, first using the Pearson correlation approach and then using the Bayesian correlation approach. In the Pearson correlation approach, a common within-study correlation for each study was estimated by calculating the Pearson correlation between the pairs of available treatment effect estimates for the two outcomes. In Bayesian correlation approach, a common within-study correlation for each study was estimated using the posterior distribution of ρ in the model 2.8.

3.3 Assessment of Performance

The performance of the estimates μ_1 and μ_2 were assessed in terms of bias, standard error, mean square error (MSE) and coverage. The BRMA estimates were obtained for both Pearson correlation approximation and Bayesian correlation approximation for each of the 1000 simulation scenarios. The corresponding 1000 UFMA and URMA estimates were also obtained for each of the scenarios (see Appendix Table 4 and Table 5). The comparison was performed by calculating: (a) the average parameter estimates across all the simulations (to estimate bias), (b) the average standard error and MSE of μ_1 and μ_2 (to assess precision) and (c) the coverage of the 95% confidence intervals (CIs) for μ_1 and μ_2. The percentage of simulated datasets for which the 95% confidence interval for an outcome's treatment effect estimate contained the true effect estimate was taken as the coverage.

3.4 Results of Simulation Scenarios

In simulation runs 1-9, both the between-study correlation ρ_D and all within-study correlations ρ_{S_i} were set to 0 and different combinations of between study variances were considered. Applying the UFMA model leads to produce 0.0805 and 0.0790 average standard error for the first and second outcomes respectively. The corresponding statistics when fitting a URMA model were 0.1260 and 0.0877 respectively. Those values indicate that there is a slight increase of standard error of the URMA model over UFMA model. For the same simulation scenario using UFMA model, the MSEs were observed as 0.0291 and 0.0098 for the first and second outcomes respectively. For the URMA model corresponding statistics were observed as 0.0354 and 0.0118 respectively. Further, the UFMA approach had poorer coverage of 73% compared with the URMA approach where the coverage was 93%. Even though both the between-study correlation ρ_D and all within-study correlations ρ_{S_i} were zero, fitting a UFMA gives pooled results with a severely affected coverage. This inconsistency in results could be explained by the deference in between study variances.

In runs 10-17, we considered the situation in which the within-study correlations and the between study correlation are similar, where ρ_D and all ρ_{S_i}were set to 0.7 or to 0.95. The two values (0.024and 0.168) were considered as the between study variances. Under this simulation scenario, applying UFMA model leads to produce mean square errors of 0.0486 and 0.0354 for the first and second outcomes respectively. For URMA model, the corresponding statistics were observed as 0.0572 and 0.0382 respectively. The coverage percentages were observed as 65 and 92 for UFMA and URMA model respectively. When these simulation results were compared with previous simulation (1-9 runs) results, it is clearly evidence a reduction in coverage and an increment in MSE for both UFMA and URMA models. This situation could be explained as an effect of fitting a univariate model when high within study correlation exists between treatment effect estimates.

In runs 18-25, we considered zero the within study correlations and the high between study correlation. The two values (0.024 and 0.168) were considered as the between study variances. Here applying UFMA model leads to produce mean square errors of 0.0379 and 0.0234 for the first and second outcomes respectively. For URMA model, the corresponding statistics were observed as 0.0461 and 0.0270 respectively. The coverage percentages were observed as 71 and 93 for UFMA and URMA models respectively. There is a clear reduction in MSE and an improvement in coverage for UFMA model over the previous simulation scenario and this could be explained as an advantage of UFMA when the within study correlation is actually zero. The improvement of the URMA model over the previous simulation could also be justified as a consequence zero within study correlation.

For all data simulation scenarios, the pooled estimates were approximately unbiased for both $BRMA_{(Bayesian)}$ and $BRMA_{(Pearson)}$ (see Appendix Table 4). Applying the BRMA approach using the Bayesian correlation method ($BRMA_{(Bayesian)}$) appeared to perform quite well comparing with other three approaches. In most of the runs, the $BRMA_{(Bayesian)}$ produces the smallest mean square error for both outcomes. In BRMA (Bayesian) model, the standard errors were an improvement over the other three approaches for both outcomes. For the first outcome, $BRMA_{(Bayesian)}$ has the smallest standard error over all 25 runs and for the second outcome, both $BRMA_{(Bayesian)}$ and $BRMA_{(Pearson)}$ have comparable standard errors. In $BRMA_{(Bayesian)}$ model, the mean coverage for the first outcome is 91.752% and for the second outcome is 94.568%. In BRMA (Pearson) model, the mean coverage for the first outcome is 91.904% and for the second outcome is 94.428%. Thus, in $BRMA_{(Bayesian)}$ model, there is a little improvement over the mean coverage for the second outcome but not for the first outcome. Although URMA approach had poorer standard errors and mean square errors, its pooled estimates were approximately unbiased and coverage was comparable with other two BRMA approaches. Fitting UFMA produces poorer statistical results under all simulation scenarios.

4. Application to Acute Stroke Data

We now apply URMA model and BRMA model to data (see Appendix Table 2) from 21 studies that assess the effectiveness of hypertension treatment for lowering blood pressure (Geeganage & Bath, 2010). A random effects model was chosen a priori because heterogeneity was significant and the value of I^2 index which can be interpreted as the percentage of the total variability in a set of effect sizes due to true heterogeneity, that is, to between-studies variability (Huedo-Medina, et al., 2006) was obtained as 54.4%. The treatment effects on the systolic and diastolic blood pressures are denoted by SBP and DBP respectively. Negative estimates indicate that the treatment is beneficial. The within-study variances corresponding to treatment effects SBP

and DBP are denoted by $Var(MD_{SBP})$ and $Var(MD_{DBP})$ respectively (see Appendix Table 3). The within-study correlations are unknown in this example but expected to be highly positively correlated (Gavish, et al., 2008). Thus, we selected a *uniform(0,1)* prior for the common within study correlation ρ and performed the proposed Bayesian approach to get an approximation for the common with study correlation.

We conducted a BRMA using restricted maximum likelihood and assuming the within-study correlations are equal to the Bayesian approximation 0.6639 gives a mean difference in SBP as -2.5540 and mean difference in DBP as -2.4467. Again a BRMA was performed assuming the within-study correlations are equal to the Pearson approximation 0.6945 gives a mean difference in SBP as -2.5728 and mean difference in DBP as -2.4363 (Table 1).

Table 1: UFMA, URMA, $BRMA_{Bayesian}$ and $BRMA_{Pearson}$ results.

Parameter	UFMA	URMA	$BRMA_{Bayesian}$	$BRMA_{Pearson}$
μ_1(SBP)	-2.2497 (0.9323)	-2.6276 (1.4677)	-2.5540 (1.3612)	-2.5728 (1.3854)
μ_2(DBP)	-2.0175 (0.5009)	-2.4668 (1.0416)	-2.4467 (1.0293)	-2.4363 (1.0295)

When the Bayesian approximation was applied, there was a slight increase in the precision of the estimated treatment effects over the Pearson correlation approach. The correlations between treatment effects were estimated as 1 and 0.9779 across studies for the Bayesian and Pearson approaches respectively. As evidence in here, very high estimated correlations are a common finding in applications of multivariate meta-analysis. When the number of studies is small and/or the within-study variation is large relative to the between-study variation, the between-study correlation is often poorly estimated as +1 or 1 (Riley, et al., 2008). The URMA model produces estimates with generally less precision statistical properties than those from any BRMA models. Applying the UFMA approach has let to poorer treatment effect estimates severely for both outcomes.

5. Discussion

In practice within-study correlations are rarely reported in primary study publications. One of the main challenges of the multivariate approach is knowing how to obtain the within-study correlations to measure the dependence between the outcomes for each study when these are unknown (Riley, 2009). In this paper we proposed the Bayesian correlation method to estimate this quantity to be the assumed within-study correlation across all studies and it was compared with the previously suggested the Pearson correlation approach. Both are convenient approach to obtain a value as the common within study correlation when no other source is available. The simulations show that both method $BRMA_{(Bayesian)}$ and $BRMA_{(Pearson)}$ generally performed well over UFMA or URMA. Comparison results also indicates that $BRMA_{(Bayesian)}$ produces better statistical properties than $BRMA_{(Pearson)}$.

In general, multivariate meta-analysis can offer advantages over a univariate approach. However, the use of multivariate meta-analysis becomes problematic when within-study correlations cannot be specified. We have derived an approximation for the common within study correlation for situations in which individual patient data are not available but correlations between treatment effects can be specified. Our simulation studies assess whether imputing correlation between treatment effect can improve estimation compared with

alternative approaches. We can conclude that when there is heterogeneity of effects across studies and high correlation within studies, our approach perform quite well.

References

[1] Abdel-Megeed, S.M., 1984. Accuracy of Correlation Coeffficient with Limited Number of Points. *The Journal of Experimental Education*, 52(4). pp.188-191.

[2] Abrams, K.R., Gillies, C.L., and Lambert, P.C., 2005. Meta-analysis of heterogeneously reported trials assessing change from baseline. *Statistics in Medicine*, 24, pp.3823-3844.

[3] Barnard, J., McCulloch, R., and Meng, X.L, 2000. Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Statistica Sinica*, 10, p.1281-1311.

[4] Berkey, C., Anderson, J., and Hoaglin, D., 1996. Multiple-Outcome Meta-Analyses of Clinical Trials. *Statistics in Medicine*, 15, pp.537-557.

[5] Berrington, A., and Cox, D.R., 2003. Generalized least squares for the synthesis of correlated information. *Biostatistics*, 4, pp.423-431.

[6] Daniels, M.J., and Hughes, M.D., 1997. Meta-analysis for the evaluation of potential surrogate markers. *Statistics in Medicine*, 16, pp.1965-1982.

[7] Dutton, M., 2011. Individual patient-level data meta-analysis: *A comparision of methods for the diverse populations contribution data set.* Ph.D. Baylor University.

[8] Gavish, B., Ben-Dov I.Z., Bursztyn, M., 2008. Linear relationship between systolic and diastolic blood pressure monitored over 24 h: assessment and correlates. *Journal of Hypertension*, 26(2): pp.199-209.

[9] Geeganage, C., Bath, PM., 2010. Vasoactive drugs for acute stroke. *Cochrane Database of Systematic Reviews*, (7):CD0028-39.

[10] Gleser, L.J., and Olkin, I., 1994. Stochastically independent effect sizes. *The Handbook of Research Synthesis. New York: Russell Sage Foundation*, pp.339-355.

[11] Huedo-Medina, T.B., Sanchez-Meca, J., and Marin-Martinez, F., 2006. Assessing Heterogeneity in Meta-Analysis: Q Statistic or I 2 Index?. *Psychological Methods*, Vol. 11, No. 2, pp.193-206.

[12] Jackson, D., White, I.R., Thompson, S.G., 2010. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Statistics in Medicine*, 29, pp.1282-1297.

[13] Jones, A.P., Riley, R.D., Williamson, P.R., Whitehead, A., 2009. Meta-analysis of longitudinal data. *Clincal Trials*, 6, pp.16-27.

[14] Kirkham, J.J, Riley R.D., and Williamson, P.R., 2012. A multivariate meta-analysis approach for reducing the impact of outcome reporting bias in systematic reviews. *Statistics in Medicine*, 31, pp. 2179-2195.

[15] Kruschke, J.K., 2013. Bayesian Estimation Supersedes the t Test. *Journal of Experimental Psychology*, Vol. 142, No. 2, pp.573-603.

[16] Korn, E.L., Albert, P.S., and McShane, L.M., 2005. Assessing surrogates as trial endpoints usingmixed models. *Statistics in Medicine*, 24, pp.163-182.

[17] Mukaka, M.M., 2012. Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal*; 24(3): pp.69-71.

[18] Raudenbush, S.W., Becker, B.J., and Kalaian, H., 1988. Modeling multivariate effect sizes. *Psychol. Bull.*, 103, pp.111-120.

[19] Reitsma, J.B., Glas, A.S., Rutjes, A.W., Scholten, R.J., Bossuyt, P.M. and Zwinderman, A.H. 2005. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology*, 58 (10), 982-990.

[20] Riley, R.D., Abrams, K.R., Lambert, P.C., Sutton, A.J., Thompson, J.R., 2007. An evaluation of bivariate random effects meta-analysis for the joint synthesis of two correlated outcomes. *Statistics in Medicine*, 26, 78-97.

[21] Riley, R.D., Thompson, J.R., and Abrams, K.R., 2008. An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics*, 9, pp.172-186.

[22] Riley, R.D., Lambert, P.C., Staessen, J.A., Wang, J., Gueyffier, F., Thijs, L., Boutitie, F., 2008. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Statistics in Medicine*, 27(11), pp.1870-1893.

[23] Riley, R.D., 2009. Multivariate meta-analysis: the effect of ignoring within-study correlation. *Journal of Royal Statistical Society*, 172(4).

[24] Thompson, S.G., and Sharp, S.J., 1999. Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine*, 18, pp.2693-2708.

[25] Wei, Y., and Higgins, J.P., 2013. Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. *Statistics in Medicine*, 32, pp.1191-1205.

6. Appendix

Table 2: Trials of vasoactive drugs for acute stroke (Geeganage & Bath, 2010)

			SBP						DBP					
Study	N_t	N_c	Treatment			Control			Treatment			Control		
			N	Mean	SD	*N*	Mean	SD	*N*	Mean	SD	*N*	Mean	SD
Barer 1988 atenolol	18	11	18	138.06	24.68	10	140.53	23.27	18	77.61	11.34	10	81.05	14.39
Barer 1988 propanolol	16	11	14	142.14	30.93	9	140.53	23.27	14	80.71	16.51	9	81.05	14.39
Barer 1988/50 mg	102	50	95	137.31	25.14	47	145.47	22.58	95	79.63	13.62	47	86.07	15.23
Barer 1988/80 mg	101	50	97	142.3	23.85	48	145.47	22.58	97	82.66	13.93	48	86.07	15.23
ASCLEPIOS 1990	120	114	107	144.73	23.77	97	144.88	23.96	107	80.98	11.83	97	83.37	13.5
Limburg 1990	12	14	10	141	26.65	6	150.83	24.58	10	86.1	18.53	6	84.17	13.57
Norris 1994	96	93	83	133.62	16.3	75	149.87	25.1	83	71.41	10.33	75	83.29	12.42
Bogousslavsky 1990	24	28	24	134	24	28	141	16	24	81	11	28	80	11
Kaste 1994/120 mg	176	174	160	146.6	25.81	163	148.5	25.81	160	84.3	14.84	161	89.2	14.24
Lowe 1993	56	56	54	145.19	24.65	54	141.02	22.39	54	85.11	13.29	54	81.3	11.33
Paci 1989/120 mg	19	22	19	136.7	17.4	22	145.7	19.7	19	82.8	10	22	85.2	9.8
Squire 1996	75	72	68	145.72	22.65	55	147.85	21.82	68	81.35	12.66	55	82.53	11.26
VENUS 1995	225	229	215	152.69	25.31	213	152.16	26.5	215	84.16	12.81	213	85.81	13.54
Lees 1995	30	30	27	147.48	28.27	25	139.52	23.82	27	79.63	17.24	25	76.92	15.79
IMAGES Pilot	26	25	26	147.35	17.54	25	155.94	27.31	26	72.94	10.12	25	82.12	10.9
Muir 1995	19	6	15	150.93	20.33	4	159.25	22.02	15	82.2	15.49	4	89.75	15.11
Strand 1984	13	13	13	168.46	35.73	12	150.42	18.4	13	84.62	15.87	12	80.83	10.19
PRISTINE	313	307	310	155.7	24.45	307	154.33	24.95	310	87.8	12.55	307	85.96	12.64
Steiner 1986	55	45	44	139.89	20.01	37	147.3	26.21	44	82.73	10.37	37	86.89	15.56
Herrschaft 1988	24	20	23	176.6	19.6	17	174.2	20.2	23	94.5	12.4	17	95.2	10.2
Huczynski 1988	15	15	15	140.4	27.84	15	146.33	26.94	15	87	11.19	15	86.47	17.31

Table 3: Effect size of acute stroke (Geeganage & Bath, 2010).

Study	Mean difference in SBP	Mean difference in DBP	$Var(MD_{SBP})$	$Var(MD_{DBP})$	$Cov(MD_{DBP})$
Barer 1988 atenolol	-2.47	-3.44	87.99	27.85	34.814090
Barer 1988 propanolol	1.61	-0.34	128.50	42.48	52.313831
Barer 1988/50 mg	-8.16	-6.44	17.50	6.89	7.754026
Barer 1988/80 mg	-3.17	-3.41	16.49	6.83	7.518547
ASCLEPIOS 1990	-0.15	-2.39	11.20	3.187	4.233494
Limburg 1990	-9.83	1.93	171.72	65.03	74.400131
Norris 1994	-16.25	-11.88	11.60	3.34	4.391508
Bogousslavsky 1990	-7	1	33.14	9.36	12.272857
Kaste 1994/120 mg	-1.9	-4.9	8.25	2.64	3.300570
Lowe 1993	4.17	3.81	20.54	5.65	7.642719
Paci 1989/120 mg	-9	-2.4	33.58	9.63	12.732678
Squire 1996	-2.13	-1.18	16.20	4.66	6.165673
VENUS 1995	0.53	-1.65	6.28	1.62	2.266717
Lees 1995	7.96	2.71	52.30	20.98	23.497898
IMAGES Pilot	-8.59	-9.18	41.67	8.69	13.301330
Muir 1995	-8.32	-7.55	148.77	73.07	73.964011
Strand 1984	18.04	3.79	126.42	28.03	42.062353
PRISTINE	1.37	1.84	3.96	1.03	1.432132
Steiner 1986	-7.41	-4.16	27.67	8.99	11.174236
Herrschaft 1988	2.4	-0.7	40.71	12.81	16.107739
Huczynski 1988	-5.93	0.53	100.06	28.32	36.818754

SBP = systolic blood pressure, DBP = diastolic blood pressure, MD = mean

Run	D_{11}^2	D_{22}^2	ρ_D	ρ	$E(\hat{\mu}_1)$				$E(\hat{\mu}_2)$				$SE(\hat{\mu}_1)$				$SE(\hat{\mu}_2)$			
					UFMA	URMA	BPMA	BBMA	UFMA	URMA	BPMA	BBMA	UFMA	URMA	BPMA	BBMA	UFMA	URMA	BPMA	BBMA
1	0	0	0	0	-0.0015	-0.0020	-0.0025	-0.0025	-0.0009	-0.0008	-0.0016	-0.0014	0.0804	0.1113	0.1091	0.1077	0.0790	0.0885	0.0834	0.0836
2	0	0.024	0	0	0.0004	0.0010	0.0016	0.0014	0.0002	0.0003	0.0003	0.0004	0.0816	0.0940	0.0947	0.0930	0.0783	0.0826	0.0803	0.0803
3	0	0.168	0	0	-0.0007	0.0007	0.0003	0.0003	-0.0007	-0.0002	-0.0005	-0.0005	0.0816	0.0959	0.0961	0.0944	0.0782	0.0835	0.0808	0.0807
4	0.024	0	0	0	-0.0007	0.0014	0.0010	0.0012	0.0009	0.0011	0.0008	0.0010	0.0803	0.1090	0.1068	0.1053	0.0796	0.0890	0.0842	0.0842
5	0.024	0.024	0	0	-0.0022	-0.0015	-0.0037	-0.0037	-0.0041	-0.0006	-0.0025	-0.0023	0.0809	0.1252	0.1181	0.1174	0.0782	0.1037	0.0960	0.0961
6	0.024	0.168	0	0	-0.0002	0.0004	-0.0005	-0.0004	-0.0030	-0.0024	-0.0034	-0.0031	0.0804	0.1096	0.1073	0.1058	0.07866	0.0887	0.0837	0.0838
7	0.168	0	0	0	0.0088	0.00163	0.0021	0.0021	0.0004	0.0006	0.0012	0.0011	0.0797	0.1635	0.1638	0.1628	0.0792	0.0836	0.0820	0.0821
8	0.0168	0.024	0	0	-0.0001	0.0002	0.0009	0.0008	-0.0003	-0.0003	-0.0003	-0.0004	0.0803	0.1631	0.1633	0.1624	0.0802	0.0851	0.0831	0.0833
9	0.168	0.168	0	0	0.0154	0.0118	0.0118	0.0116	0.0036	0.0036	0.0036	0.0036	0.0794	0.1620	0.1621	0.1611	0.0797	0.0843	0.08250	0.0826
10	0.024	0.024	0.7	0.7	0.0110	0.0078	0.0074	0.0074	0.0043	0.0028	0.0040	0.0037	0.0820	0.1132	0.1087	0.1074	0.0780	0.0958	0.0898	0.0898
11	0.024	0.168	0.7	0.7	0.0073	0.0083	0.0086	0.0082	0.0001	0.0022	0.0020	0.0020	0.1005	0.1268	0.1238	0.1229	0.0788	0.0985	0.0946	0.0947
12	0.168	0.024	0.7	0.7	-0.0032	-0.0034	-0.0024	-0.0028	-0.0023	-0.0019	-0.0023	-0.0022	0.0796	0.1649	0.1599	0.1595	0.0797	0.1092	0.1027	0.1031
13	0.168	0.168	0.7	0.7	-0.0033	-0.0001	0.0001	0.0001	-0.0012	-0.0023	-0.0023	0.0023	0.0808	0.1732	0.1621	0.1621	0.0793	0.1501	0.1384	0.1387
14	0.024	0.024	0.95	0.95	0.0001	0.0026	0.0021	0.0018	-0.0030	-0.0028	-0.0035	-0.0036	0.0844	0.1284	0.1224	0.1213	0.1095	0.1095	0.1032	0.1032
15	0.024	0.168	0.95	0.95	-0.0084	-0.0069	-0.0074	-0.0074	-0.0034	0.0006	0.0016	0.0010	0.1126	0.1894	0.1855	0.1844	0.0808	0.1616	0.1562	0.1565
16	0.168	0.024	0.95	0.95	-0.0040	-0.0023	-0.0024	-0.0025	-0.0016	0.0016	0.0005	0.0006	0.0795	0.1734	0.1667	0.16639	0.0800	0.1306	0.1220	0.1226
17	0.168	0.168	0.95	0.95	-0.0114	0.0055	0.0060	0.0060	-0.0048	0.0053	0.0061	0.0061	0.0843	0.2180	0.2121	0.2111	0.0753	0.2018	0.1961	0.1958
18	0.024	0.024	0.7	0	0.0037	0.0039	0.0044	0.0043	0.0018	0.0020	0.0019	0.0018	0.0823	0.1119	0.1064	0.1053	0.0764	0.0932	0.0874	0.0874
19	0.024	0.168	0.7	0	-0.0015	-0.0014	-0.0020	-0.0018	0.0010	0.0007	0.0001	0.0002	0.1277	0.12770	0.1247	0.1238	0.0789	0.0985	0.0948	0.0950
20	0.168	0.024	0.7	0	-0.0026	-0.0003	-0.0020	-0.0020	0.0806	-0.0038	-0.0040	-0.0040	0.0795	0.1683	0.1626	0.1623	0.0515	0.1096	0.1028	0.1032
21	0.168	0.168	0.7	0	0.0026	-0.0010	-0.0020	-0.0018	0.0006	-0.0041	-0.0028	-0.0030	0.0808	0.1716	0.1618	0.1616	0.0793	0.1491	0.1382	0.1385
22	0.024	0.024	0.95	0	-0.0032	-0.0028	-0.0046	-0.0041	-0.0044	-0.0032	-0.0038	-0.0037	0.0827	0.1139	0.1101	0.1087	0.0771	0.0950	0.0901	0.0900
23	0.024	0.168	0.95	0	0.0080	0.0083	0.0075	0.0078	0.0074	0.0065	0.0069	0.0068	0.1127	0.1404	0.1364	0.1355	0.0788	0.1034	0.0989	0.0990
24	0.168	0.024	0.95	0	0.0006	-0.0009	-0.0002	-0.0004	-0.0060	-0.0061	-0.0048	-0.0050	0.0797	0.1665	0.1612	0.1607	0.0798	0.1187	0.1112	0.1115
25	0.168	0.168	0.95	0	-0.0023	-0.0008	-0.0012	-0.0015	-0.0042	-0.0033	-0.0034	-0.0037	0.0849	0.1732	0.1635	0.1628	0.0751	0.1549	0.1470	0.1468

Table 4: Treatement effect estimates and standard errors for simulations

					$MSE(\hat{\mu_1})$				$MSE(\hat{\mu_2})$				$Coverage(\hat{\mu_1})$				$Coverage(\hat{\mu_2})$			
Run	D_{11}^2	D_{22}^2	ρ_D	ρ	UFMA	URMA	BPMA	BBMA	UFMA	URMA	BPMA	BBMA	UFMA	URMA	BPMA	BBMA	UFMA	URMA	BPMA	BBMA
1	0	0	0	0	0.0186	0.0257	0.0245	0.0241	0.0096	0.0122	0.0110	0.0110	84.5	92.9	93.6	93.2	84.5	92.9	98.6	98.9
2	0	0.024	0	0	0.0141	0.0178	0.0179	0.0173	0.0083	0.0093	0.0088	0.0088	95.2	96.3	96.0	95.8	95.2	96.3	99.5	99.4
3	0	0.168	0	0	0.0134	0.0175	0.0179	0.0172	0.0082	0.0095	0.0092	0.0091	96	97.0	96.9	96.9	96	97.0	98.8	98.9
4	0.024	0	0	0	0.0188	0.0250	0.0239	0.0235	0.0096	0.0120	0.0111	0.0110	84	92.4	92.6	92.1	84	92.4	98.2	98.0
5	0.024	0.024	0	0	0.0238	0.0324	0.0298	0.0295	0.0154	0.0198	0.0184	0.0183	77.2	92.5	91.2	90.9	77.2	92.5	94.7	94.7
6	0.024	0.168	0	0	0.0190	0.0252	0.0242	0.0237	0.0098	0.0124	0.0112	0.0112	83.9	92.9	93.1	92.5	83.9	92.9	97.9	97.9
7	0.168	0	0	0	0.0517	0.0588	0.0576	0.0574	0.0088	0.0099	0.0096	0.0096	52.9	89.5	90.4	89.8	52.9	89.5	99.2	99.2
8	0.0168	0.024	0	0	0.0533	0.0596	0.0586	0.0585	0.0091	0.0103	0.0099	0.0099	54.1	90.1	90.7	90.6	54.1	90.1	99.3	99.4
9	0.168	0.168	0	0	0.0496	0.0565	0.0551	0.0550	0.0093	0.0104	0.0010	0.0010	53.9	91.8	92.3	92.0	53.9	91.8	99.0	99.0
10	0.024	0.024	0.7	0.7	0.0200	0.0264	0.0252	0.0247	0.0132	0.0162	0.0152	0.0151	84.7	94.0	92.8	92.5	84.7	94.0	96.1	96.3
11	0.024	0.168	0.7	0.7	0.0258	0.0322	0.0327	0.0319	0.0188	0.0191	0.0191	0.0190	88.6	94.6	92.8	93.1	88.6	94.6	92.0	92.4
12	0.168	0.024	0.7	0.7	0.0475	0.0595	0.0561	0.0561	0.0190	0.0234	0.0219	0.0219	55.7	92.2	92.1	91.8	55.7	92.2	93.9	94.2
13	0.168	0.168	0.7	0.7	0.0536	0.0642	0.06012	0.0598	0.0371	0.0464	0.0436	0.0434	52.8	91.9	89.3	89.8	52.8	91.9	88.9	89.2
14	0.024	0.024	0.95	0.95	0.0254	0.0336	0.0324	0.0318	0.0188	0.0236	0.02260	0.0224	77.8	93.1	91.8	91.8	77.8	93.1	93.0	93.1
15	0.024	0.168	0.95	0.95	0.0588	0.0752	0.0769	0.0757	0.0591	0.0549	0.0546	0.0546	67.9	92.3	90.9	90.9	67.9	92.3	89.9	89.9
16	0.168	0.024	0.95	0.95	0.0598	0.0697	0.0666	0.0666	0.0322	0.0374	0.0349	0.0350	49.9	88.7	87.6	87.0	49.9	88.7	89.6	90.2
17	0.168	0.168	0.95	0.95	0.0978	0.0970	0.0965	0.0956	0.0847	0.0844	0.0848	0.0843	41.6	92.9	90.9	91.0	41.6	92.9	91.2	90.9
18	0.024	0.024	0.7	0	0.0192	0.0258	0.0246	0.0240	0.0122	0.0152	0.0143	0.0141	85	94.2	92.3	92.7	85	94.2	96.6	97.0
19	0.024	0.168	0.7	0	0.0255	0.0318	0.0320	0.0314	0.0178	0.0171	0.0179	0.01783	89.9	95.5	94.4	94.2	89.9	95.5	94.5	94.5
20	0.168	0.024	0.7	0	0.0515	0.0621	0.0590	0.0588	0.0201	0.0246	0.0232	0.0232	54.7	90.1	90.0	89.8	54.7	90.1	91.5	91.7
21	0.168	0.168	0.7	0	0.0530	0.0643	0.0602	0.0599	0.0363	0.0454	0.0431	0.0430	53.7	89.7	88.4	88.4	53.7	89.7	89.2	89.4
22	0.024	0.024	0.95	0	0.0199	0.0266	0.0258	0.0251	0.0132	0.0160	0.0152	0.0151	84.4	93.6	92.6	92.5	84.4	93.6	95.6	95.7
23	0.024	0.168	0.95	0	0.0300	0.0373	0.0377	0.0368	0.0217	0.0207	0.0203	0.0203	90.6	95.5	94.2	94.4	90.6	95.5	91.2	91.5
24	0.168	0.024	0.95	0	0.0508	0.0606	0.0575	0.0573	0.0230	0.0288	0.0271	0.0271	53.8	90.1	90.1	89.9	53.8	90.1	91.8	91.9
25	0.168	0.168	0.95	0	0.0531	0.0601	0.0577	0.0571	0.0426	0.0483	0.0469	0.0465	58.2	93.6	90.6	90.2	58.2	93.6	90.5	90.9

Table 5: MSE and coverage probabilities for simulations