Web-Based Supplementary Materials for "A matrix-based method of moments for fitting multivariate network meta-analysis models with multiple outcomes and random inconsistency effects" by Jackson, Bujkiewicz, Law, Riley and White.

July 10, 2017

1 Multivariate estimation

1.1 An important result

In order to evaluate the expectations required, we will need to be able to compute expressions of the form $\text{btr}(\mathbf{A}(\mathbf{M}\otimes\mathbf{\Sigma})\mathbf{B})$, where $\mathbf{A}$ and $\mathbf{B}$ are $np\times np$ matrices, $\mathbf{M}$ is an $n\times n$ matrix and $\mathbf{\Sigma}$ is a $p\times p$ matrix. We continue to use the notation $\mathbf{A}_{i,j}$ to denote the ith by jth block of $\mathbf{A}$, where these blocks are $p\times p$ matrices. For any three $np\times np$ matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, we have

$$(\mathbf{ACB})_{k,l}=\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{A}_{k,i}\mathbf{C}_{i,j}\mathbf{B}_{j,l}.$$

This is just the law of matrix multiplication applied to blocks. Then taking $\mathbf{C}=\mathbf{M}\otimes\mathbf{\Sigma}$ so that from the definition of the Kronecker product, $\mathbf{C}_{i,j}=m_{ij}\mathbf{\Sigma}$, we have

$$(\mathbf{A}(\mathbf{M}\otimes\mathbf{\Sigma})\mathbf{B})_{k,l}=\sum_{i=1}^{n}\sum_{j=1}^{n}m_{ij}\mathbf{A}_{k,i}\mathbf{\Sigma}\mathbf{B}_{j,l}.$$

To obtain the block trace, we sum the matrices along the main diagonal. Hence to obtain the block trace we take $l=k$ to obtain the matrices along the main diagonal and sum over k so obtain

$$\text{btr}(\mathbf{A}(\mathbf{M}\otimes\mathbf{\Sigma})\mathbf{B})=\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}m_{ij}\mathbf{A}_{k,i}\mathbf{\Sigma}\mathbf{B}_{j,k}. \quad (1)$$

The use of equation (1), with the appropriate matrices, almost immediately results in the expected values required in section 4.

1.2 The estimating equations

In this section we prove the results given in section 4 of the main paper. We do not redefine all quantities or give the size of all matrices and vectors, see the main paper for these details. As in the univariate approach of Jackson *et al.* (2016), we will base our estimation on the two quantities $\text{btr}(\mathbf{Q})$ and $\sum_{d=1}^{D}\text{btr}(\mathbf{Q}_d)$ where D is the number of different designs. We match these quantities to their expectations to estimate the unknown variance parameters. We therefore need to evaluate $\text{E}[\text{btr}(\mathbf{Q})]$ and $\text{E}[\text{btr}(\mathbf{Q}_d)]$.

1.2.1 Evaluating E[btr(Q)] and deriving the first estimating equation

As in Jackson *et al.* (2013), by direct calculation we have that $\mathbf{WHW}^{-1}=\mathbf{H}^T$ and $((\mathbf{I}_{np}-\mathbf{H})^T)^2=(\mathbf{I}_{np}-\mathbf{H})^T$; if $\mathbf{W}$ is not invertible because outcome data are missing then we can justify the use of the

identity $\mathbf{WHW}^{-1} = \mathbf{H}^T$ and the expectation that follows in the limit, where the precision p attributed to missing data tends towards zero from above, $p \to 0^+$ (Jackson *et al.*, 2013). Furthermore we can use the identity $\mathbf{W} = \mathbf{S}^{-1}$ in this limit. We also have that $\mathbf{Y} - \hat{\mathbf{Y}} = (\mathbf{I}_{np} - \mathbf{H})\mathbf{Y}$ and $\mathrm{E}[\mathbf{Y} - \hat{\mathbf{Y}}] = \mathbf{0}$. Hence from the definition of $\mathbf{Q}$ we have $\mathrm{E}[\mathbf{Q}] = \mathbf{W}\mathrm{Var}[\mathbf{Y} - \hat{\mathbf{Y}}]\mathbf{R}$. From these results, taking the variance of $\mathbf{Y}$ from model (3) of the main paper, we can evaluate

$$\mathrm{E}[\mathbf{Q}] = \mathbf{A}(\mathbf{M_1} \otimes \mathbf{\Sigma}_\beta + \mathbf{M_2} \otimes \mathbf{\Sigma}_\omega)\mathbf{B} + \mathbf{B},$$

where

$$\mathbf{A} = (\mathbf{I}_{np} - \mathbf{H})^T\mathbf{W},$$

and

$$\mathbf{B} = (\mathbf{I}_{np} - \mathbf{H})^T\mathbf{R}.$$

Here $\mathbf{A}$ and $\mathbf{B}$ are known $np \times np$ matrices. For estimation purposes we require $\mathrm{E}[\mathrm{btr}(\mathbf{Q})] = \mathrm{btr}(\mathrm{E}[\mathbf{Q}])$. We write $\mathbf{A}_{i,j}$ and $\mathbf{B}_{i,j}$ to mean the ith by jth blocks of $\mathbf{A}$ and $\mathbf{B}$ respectively, so that $\mathbf{A}_{i,j}$ and $\mathbf{B}_{i,j}$ are both $p \times p$ matrices. Then, using (1), we have

$$\mathrm{E}[\mathrm{btr}(\mathbf{Q})] = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} m_{1ij}\mathbf{A}_{k,i}\mathbf{\Sigma}_\beta\mathbf{B}_{j,k} +$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} m_{2ij}\mathbf{A}_{k,i}\mathbf{\Sigma}_\omega\mathbf{B}_{j,k} + \mathrm{btr}(\mathbf{B}).$$

1.2.2 Evaluating $\mathrm{E}[\mathbf{btr}(\mathbf{Q}_d)]$ and deriving the second estimating equation

Then we follow very similar, but much simpler, arguments as in the previous section to derive the result that we require. We define design specific hat matrices

$$\mathbf{H}_d = \mathbf{X}_d(\mathbf{X}_d^T\mathbf{W}_d\mathbf{X}_d)^{-1}\mathbf{X}_d^T\mathbf{W}_d, \tag{2}$$

and also design specific $pn_d \times pn_d$ $\mathbf{A}$ and $\mathbf{B}$ matrices

$$\mathbf{A}_d = (\mathbf{I}_{pn_d} - \mathbf{H}_d)^T\mathbf{W}_d,$$

and

$$\mathbf{B}_d = (\mathbf{I}_{pn_d} - \mathbf{H}_d)^T\mathbf{R}_d.$$

In equation (2) we take the matrix inverse to be the Moore-Penrose pseudoinverse. This is because, in the presence of missing outcome data, the design-specific regression corresponding to this hat matrix may not be identifiable (for example, if studies of a particular design do not provide data for one or more of the outcomes). In such instances this design may still provide information about some of the unknown between-study variance components and so it is not desirable to exclude the design from this part of the estimation procedure. By computing (2) using this pseudoinverse we obtain a suitable hat matrix (Searle, 1971; page 221, his equations 126 and 127). Furthermore all the necessary properties of the hat matrix are retained when using the pseudoinverse when computing (2) and we retain unbiased fitted values (Searle, 1971; page 181).

Following a simpler version of the arguments in the previous section and the main paper, taking the variance of $\mathbf{Y}_d$ from model (5) of the main paper, and upon applying the vec operator, we obtain

$$\mathrm{vec}(\mathrm{E}[\mathrm{btr}(\mathbf{Q}_d)]) = \mathbf{C}_d\mathrm{vec}(\mathbf{\Sigma}_\beta) + \mathbf{E}_d, \tag{3}$$

where

$$\mathbf{C}_d = \sum_{i=1}^{n_d}\sum_{j=1}^{n_d}\sum_{k=1}^{n_d} m_{1ij}^d\mathbf{B}_{d,j,k}^T \otimes \mathbf{A}_{d,k,i},$$

and

$$\mathbf{E}_d = \mathrm{vec}(\mathrm{btr}(\mathbf{B}_d)).$$

We then sum equation (3) across all designs in order to obtain

$$\mathrm{vec}\left(\mathrm{E}[\sum_{d=1}^{D}\mathrm{btr}(\mathbf{Q}_d)]\right) = \left(\sum_{d=1}^{D}\mathbf{C}_d\right)\mathrm{vec}(\mathbf{\Sigma}_\beta) + \sum_{d=1}^{D}\mathbf{E}_d. \tag{4}$$

1.3 Special cases of the estimation procedure (an extended version of section 4.5)

The proposed method reduces to two previous methods in special cases. If all studies are two arm studies (and so provide a single contrast) and consistency is assumed then the proposed method reduces to the matrix based method for multivariate meta-regression (Jackson *et al.*, 2013). This is because we then have $\mathbf{\Sigma}_\omega = \mathbf{0}$, so that the second triple sum in our expression for E[btr($\mathbf{Q}$)] is zero; furthermore the first triple summation in this expression can be reduced to a double summation, because $\mathbf{M_1}$ is an identity matrix for multivariate meta-regression (Jackson *et al.*, 2013; their equation A.1.).

Furthermore the proposed multivariate method also reduces to the univariate DerSimonian and Laird method for network meta-analysis (Jackson *et al.*, 2016) when $p = 1$. This is because, in one dimension, the $\mathbf{Q}$ matrices all reduce to the Q random scalars used in the estimation procedure suggested by Jackson *et al.* (2016). This can be shown by replacing the block trace operator with the more familiar trace of a matrix (btr is the trace when $p = 1$) in the definition of the $\mathbf{Q}$ matrices and using the identity tr($\mathbf{AB}$) = tr($\mathbf{BA}$). These two special cases are in turn generalisations of methods such as that proposed by DerSimonian and Laird (1986).

There is however one caveat when stating that the new multivariate method reduces to the univariate method proposed by (Jackson *et al.*, 2016) when $p = 1$. This is because the account of Jackson *et al.* (2016) does not mention the possibility of missing outcome data and so we have implicitly taken all data to be observed in the argument used in the previous paragraph.

2 Example of matrices $\mathbf{M_1}$ and $\mathbf{M_2}$

A referee suggested that we provide a concrete example of matrices $\mathbf{M_1}$ and $\mathbf{M_2}$, in order to clarify how they are computed. We take such an example from Law *et al.* (2016) which comprises thirteen studies with the following study designs: AB, BC, BC, BC, BC, BC, BD, BD, CD, CD, ABD, BCD, BCD. This is the same type of network as used in the simulation study below. The two matrices for this example are given explicitly below, where we can see that these matrices contain blocks that are comprised of blocks of $\mathbf{P}_{c_d}$, where in $\mathbf{M_1}$ the blocks are formed by studies and in $\mathbf{M_2}$ the blocks are formed by designs.

$$\mathbf{M_1} = \left(\begin{array}{cccccccccccccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 1
\end{array}\right)$$

$$
\mathbf{M_2} = \left(\begin{array}{c|ccccc|cc|cc|cc|cccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 1 & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2} & 1 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2} & 1 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 1
\end{array}\right)
$$

3 Simulation study

A simulation study was performed in order to investigate the use of the proposed estimation method. This simulation study was based on the second example of Jackson *et al.* (2016), in order to motivate it by a real example that we are already familiar with. This example involves 13 studies, four treatment groups and six designs (one AB study, five BC studies, two BD studies, two CD studies, one ABD study and two BCD studies). Bivariate ($p = 2$) datasets were simulated assuming 13 studies of these particular designs, where all within-study variances (the entries on the main diagonal of $\mathbf{S}$) were sampled with replacement from the within-study variances from the real example. Although the mean of these within-study variances is 0.39, the typical within-study variance proposed by Higgins and Thompson (2002), their equation (9), is 0.24. All within-study correlations between different contrasts involving the same outcome, or different outcomes involving the same contrast, were taken to be 0.5. Within-study correlations between different contrasts and outcomes were taken to be 0.25. Data were simulated from model (6) of the main paper throughout.

Twenty different sets of $\mathbf{\Sigma}_\beta$ and $\mathbf{\Sigma}_\omega$ were used; the basic parameters were all set to zero but this is immaterial because the estimation of the variance components is location-invariant and the point estimation of the means is just translated when using an alternative sets of treatment effects. One thousand simulated datasets were produced for each combination $\mathbf{\Sigma}_\beta$ and $\mathbf{\Sigma}_\omega$, so that 20,000 datasets were simulated in total. For fifteen of the simulation runs, both between-study variances (the main diagonal entries of $\mathbf{\Sigma}_\beta$) were set to 0.24, so that the extent of the between-study heterogeneity is comparable with the within-study variation; this is the case for the real data where the point estimates of the between-study variance are similar to the typical within-study variance (Jackson *et al.*, 2016). In these fifteen simulation runs the inconsistency variances (the main diagonal entries of $\mathbf{\Sigma}_\omega$) were set to either 0 or 0.12, to explore the cases where consistency assumptions are either true or violated, but where the departure from consistency is not very severe; since it is sometimes argued that considerable inconsistency should strongly discourage the use of models for network meta-analysis (Jackson *et al.*, 2016), we wished to investigate only relatively minor inconsistencies in the networks. In runs sixteen to nineteen we explore the use of two different between-study variances (0.12 and 0.6) and two different inconsistency variances (0.06 and 0.24). Finally in run twenty we explored the extreme case where the between-study variances are large (both 0.6), a very strong between-study covariance (0.59) and inconsistency variances of 0 (so that the consistency assumption is true). With very considerable and highly correlated between-study variance components, and no inconsistency, run twenty was performed in order to try to create the circumstances where borrowing of strength (Jackson *et al*, 2015a) is more likely to occur. We analysed all datasets using both the proposed multivariate method and the previous univariate method (Jackson *et al.*, 2016), where the univariate method was applied to both outcomes separately. We also applied both the univariate and multivariate methods under the assumption of consistency, in order to explore the implications of making this assumption.

To investigate a missing data scenario, for all 20,000 datasets we removed the estimated effect for the second outcome for four of the five *BC* studies. This was done because there is no direct comparison of treatments *A* and *C* in the simulated networks, because these include only the six designs in the real dataset. Hence the identification of the basic parameters δ_1^{AC} and δ_2^{AC} must rely greatly on the *BC* studies.

By removing four estimated effects in this way, we hoped to create a situation where borrowing of strength for δ_2^{AC} was likely in the incomplete datasets. We performed all univariate analyses using the univariate code provided previously by Jackson *et al.* (2016). Using univariate code reduces computation time when fitting 40,000 univariate network meta-analysis models.

The simulation study results are extensive and are shown in eight tables in these supplementary materials. Supplementary Tables 1 and 2 show that, for complete data and before truncation, both the univariate and multivariate estimates of the unknown variance components are unbiased. Truncation results in bias however. This is as expected for the unknown variance parameters, because truncation forces their estimates to be positive. However the results suggest that the multivariate approach may be helpful in reducing (but not removing) the upward bias in the truncated estimates of the inconsistency variances. Similar observations apply in the missing data scenario in Supplementary Tables 3 and 4. Supplementary Tables 5 and 6 show that the coverage probabilities of nominal 95% confidence intervals for the basic parameters are close to the correct level but also suggest that the multivariate approach may help to more accurately attain this level. This is as expected, because by using more data in the multivariate setting we can expect the asymptotic approximation of taking the variance components as known to be more accurate. Supplementary Table 7 shows that analyses under the consistency assumption fail to achieve the nominal significance level. This is even the case when the consistency assumption is true, because the extent of the between-study heterogeneity is quite large and the uncertainty in the between-study variance parameters is not taken into account. The extent to which the consistency analysis fails to achieve the nominal coverage probability probability is comparable to the univariate results in Jackson *et al.* (2016) when the inconsistency there is mild.

Finally, Supplementary Table 8 shows the ratio of empirical variances of the multivariate and univariate estimates of the basic parameters. For complete data these ratios are close to 100% so that, as expected, there is little or no borrowing of strength. Recalling that the incomplete data scenario was intended to allow borrowing of strength for δ_2^{AC}, in Supplementary Table 8 we can see some evidence of percentage efficiencies of slightly less than 100% for this parameter in runs where the correlation in the data is larger. For the final run, which we performed in order to try to create a situation where borrowing of strength may occur, we obtain a percentage efficiency of 88% for δ_2^{AC} in the missing data scenario, which is appreciably smaller than all other values in Supplementary Table 8 (the second smallest value is 95%). Hence we have achieved the most borrowing of strength exactly where we expected it. This corresponds to a borrowing of strength statistic (Jackson *et al*, 2015a) of 12%. This may appear small but it should be recalled that we only removed four univariate estimates to achieve this and furthermore that the within-study correlations are not large. For example, Jackson *et al.* (2015a) consider a bivariate meta-analysis where all within-study correlations, and the between-study correlation, are close to one, and further that one outcome is missing in 17 out of 31 studies (and is complete for the the other outcome). Even then, the borrowing of strength statistic is only around a half (53%; Jackson *et al* (2015a), their example 2) for the outcome with missing data. Our borrowing of strength statistic of 12% is therefore proof of concept that the multivariate approach can provide more accurate inference than the univariate approach, in situations that facilitate this.

In order to try to encourage a little more borrowing of strength, and also to test the numerical algorithms when not all design specific regressions are identifiable (see section 4.3.2), we repeated the simulation study for the final run where the second outcome was removed for all five BC studies; $\mathbf{H}_d$ is then not computable for the BC design using standard matrix inversion and it is necessary to use the Moore-Penrose pseudoinverse in this instance. As expected, the estimation performed very similarly to the missing data scenario described above, and the slightly larger borrowing of strength statistic of 13% was obtained for δ_2^{AC}.

To summarise, the simulation study suggests that the multivariate approach has three main advantages over the univariate approach proposed by Jackson *et al.* (2016): it can help reduce the upward bias of estimates of the inconsistency variance, it can help better attain the nominal coverage probability of confidence intervals and it can result in borrowing of strength. The better nominal coverage probability provided by the multivariate approach can be explained because multivariate analyses incorporate more information so that the large sample normal approximations are then more accurate.

Table 1: Average estimates of the entries of the between-study heterogeneity covariance matrix $\mathbf{\Sigma}_\beta$ for complete data. 1000 simulated datasets were produced for each run. The average estimates of the between-study variances from the proposed multivariate ('Multi') method are very similar to those from the univariate ('Uni') method. The proposed multivariate method estimates the between-study covariance; two 'untruncated' estimates of this covariance are obtained as explained in section X. The untruncated ('Untrunc.') estimates are within Monte Carlo error of the true parameter values but truncation ('Trunc') results in notable bias. **For the average untruncated estimates, the Monte Carlo standard errors range from 0.007 to 0.016.**

	$\mathbf{\Sigma}_\beta(11)$					$\mathbf{\Sigma}_\beta(22)$					$\mathbf{\Sigma}_\beta(12)$ and $\mathbf{\Sigma}_\beta(21)$			
		Untrunc.		Trunc.			Untrunc.		Trunc.			Untrunc.		Trunc.
Run	Truth	Uni	Multi	Uni	Multi	Truth	Uni	Multi	Uni	Multi	Truth	Multi	Multi	Multi
1	0.24	0.251	0.249	0.290	0.287	0.24	0.236	0.234	0.276	0.273	0	0.008	0.009	0.028
2	0.24	0.236	0.238	0.268	0.271	0.24	0.247	0.244	0.287	0.281	0	0.004	0.013	0.026
3	0.24	0.256	0.258	0.288	0.291	0.24	0.249	0.248	0.287	0.283	0	0.000	0.003	0.018
4	0.24	0.236	0.235	0.269	0.269	0.24	0.250	0.251	0.277	0.282	0	0.008	0.006	0.022
5	0.24	0.241	0.240	0.272	0.273	0.24	0.242	0.240	0.274	0.273	0	0.006	0.005	0.022
6	0.24	0.234	0.233	0.268	0.265	0.24	0.228	0.228	0.266	0.261	0.16	0.150	0.148	0.150
7	0.24	0.251	0.250	0.283	0.280	0.24	0.237	0.238	0.271	0.269	0.16	0.165	0.16	0.162
8	0.24	0.210	0.212	0.255	0.247	0.24	0.236	0.234	0.276	0.266	0.16	0.138	0.148	0.145
9	0.24	0.241	0.243	0.274	0.274	0.24	0.227	0.225	0.265	0.260	0.16	0.150	0.157	0.155
10	0.24	0.242	0.242	0.275	0.274	0.24	0.240	0.240	0.272	0.270	0.16	0.158	0.159	0.159
11	0.24	0.244	0.245	0.285	0.300	0.24	0.234	0.232	0.267	0.285	-0.16	-0.157	-0.152	-0.114
12	0.24	0.238	0.235	0.273	0.290	0.24	0.249	0.251	0.284	0.305	-0.16	-0.149	-0.157	-0.112
13	0.24	0.237	0.239	0.274	0.293	0.24	0.238	0.239	0.274	0.290	-0.16	-0.155	-0.154	-0.115
14	0.24	0.245	0.243	0.276	0.296	0.24	0.225	0.226	0.264	0.282	-0.16	-0.173	-0.177	-0.134
15	0.24	0.250	0.254	0.283	0.303	0.24	0.235	0.235	0.273	0.287	-0.16	-0.157	-0.148	-0.115
16	0.12	0.114	0.111	0.177	0.160	0.6	0.593	0.595	0.601	0.608	0.16	0.151	0.136	0.146
17	0.12	0.139	0.139	0.195	0.188	0.6	0.610	0.612	0.619	0.625	0.16	0.178	0.179	0.180
18	0.12	0.125	0.122	0.184	0.190	0.6	0.576	0.575	0.587	0.601	-0.16	-0.147	-0.151	-0.119
19	0.12	0.132	0.131	0.196	0.201	0.6	0.579	0.578	0.590	0.606	-0.16	-0.166	-0.168	-0.134
20	0.6	0.601	0.601	0.612	0.630	0.6	0.604	0.602	0.612	0.629	0.59	0.582	0.588	0.564

Table 2: Average estimates of the entries of the inconsistency covariance matrix $\mathbf{\Sigma}_\omega$ for complete data. 1000 simulated datasets were produced for each run. The average estimates of the inconsistency variances from the proposed multivariate ('Multi') method are very similar to those from the univariate ('Uni') method. The proposed multivariate method estimates the inconsistency covariance; two 'untruncated' estimates of this covariance are obtained as explained in section X. The untruncated ('Untrunc.') estimates are within Monte Carlo error of the true parameter values but truncation ('Trunc') results in notable bias. **For the average untruncated estimates, the Monte Carlo standard errors range from 0.007 to 0.017.**

	$\mathbf{\Sigma}_\omega(11)$					$\mathbf{\Sigma}_\omega(22)$					$\mathbf{\Sigma}_\omega(12)$ and $\mathbf{\Sigma}_\omega(21)$			
		Untrunc.		Trunc.			Untrunc.		Trunc.			Untrunc.		Trunc.
Run	Truth	Uni	Multi	Uni	Multi	Truth	Uni	Multi	Uni	Multi	Truth	Multi	Multi	Multi
1	0	-0.003	0.001	0.203	0.130	0	0.005	0.004	0.193	0.127	0	-0.014	-0.004	0.020
2	0.12	0.123	0.122	0.258	0.217	0	-0.010	-0.007	0.193	0.123	0	-0.005	-0.016	0.019
3	0.12	0.106	0.105	0.255	0.212	0.12	0.116	0.116	0.264	0.222	0	0.000	-0.004	0.021
4	0.12	0.115	0.115	0.253	0.213	0.12	0.100	0.100	0.252	0.207	0.08	0.070	0.068	0.077
5	0.12	0.114	0.115	0.247	0.220	0.12	0.097	0.097	0.236	0.205	-0.08	-0.074	-0.068	-0.030
6	0	0.003	0.004	0.200	0.123	0	0.005	0.003	0.193	0.120	0	0.001	0.003	0.050
7	0.12	0.120	0.120	0.257	0.212	0	0.010	0.012	0.194	0.128	0	0.009	0.004	0.054
8	0.12	0.133	0.133	0.261	0.220	0.12	0.131	0.132	0.259	0.219	0	0.017	0.013	0.054
9	0.12	0.119	0.118	0.255	0.210	0.12	0.137	0.140	0.259	0.225	0.08	0.088	0.080	0.114
10	0.12	0.120	0.121	0.269	0.231	0.12	0.122	0.120	0.255	0.223	-0.08	-0.075	-0.066	-0.005
11	0	-0.001	-0.002	0.199	0.133	0	0.007	0.010	0.197	0.138	0	0.005	-0.007	0.003
12	0.12	0.124	0.127	0.259	0.229	0	-0.003	-0.004	0.196	0.134	0	-0.006	0.005	0.005
13	0.12	0.132	0.133	0.268	0.240	0.12	0.133	0.127	0.263	0.232	0	-0.013	0.005	-0.002
14	0.12	0.114	0.115	0.254	0.223	0.12	0.122	0.123	0.252	0.229	0.08	0.088	0.089	0.068
15	0.12	0.097	0.093	0.248	0.209	0.12	0.125	0.127	0.254	0.228	-0.08	-0.070	-0.085	-0.055
16	0.24	0.249	0.248	0.307	0.307	0.06	0.047	0.043	0.356	0.229	0.08	0.086	0.090	0.113
17	0.24	0.218	0.222	0.295	0.300	0.06	0.060	0.061	0.377	0.257	-0.08	-0.087	-0.081	-0.014
18	0.24	0.249	0.253	0.319	0.325	0.06	0.059	0.058	0.345	0.245	0.08	0.068	0.080	0.062
19	0.24	0.235	0.238	0.303	0.314	0.06	0.079	0.077	0.367	0.268	-0.08	-0.086	-0.076	-0.053
20	0	-0.022	-0.023	0.339	0.177	0	-0.025	-0.024	0.336	0.178	0	-0.017	-0.021	0.110

Table 3: As Table 1 but after removing the second outcome from 4 of the 5 BC trials. **For the average untruncated estimates, the Monte Carlo standard errors range from 0.007 to 0.023**

	$\boldsymbol{\Sigma}_\beta(11)$					$\boldsymbol{\Sigma}_\beta(22)$					$\boldsymbol{\Sigma}_\beta(12)$ and $\boldsymbol{\Sigma}_\beta(21)$			
		Untrunc.		Trunc.			Untrunc.		Trunc.			Untrunc.		Trunc.
Run	Truth	Uni	Multi	Uni	Multi	Truth	Uni	Multi	Uni	Multi	Truth	Multi	Multi	Multi
1	0.24	0.251	0.250	0.290	0.303	0.24	0.219	0.216	0.320	0.299	0.00	-0.005	0.009	0.029
2	0.24	0.236	0.237	0.268	0.287	0.24	0.247	0.243	0.338	0.324	0.00	0.006	0.010	0.033
3	0.24	0.256	0.258	0.288	0.307	0.24	0.251	0.250	0.341	0.329	0.00	0.002	0.009	0.032
4	0.24	0.236	0.235	0.269	0.288	0.24	0.246	0.251	0.334	0.327	0.00	0.021	0.014	0.040
5	0.24	0.241	0.239	0.272	0.289	0.24	0.242	0.240	0.334	0.320	0.00	0.016	0.008	0.035
6	0.24	0.234	0.233	0.268	0.285	0.24	0.238	0.239	0.328	0.312	0.16	0.159	0.151	0.153
7	0.24	0.251	0.251	0.283	0.300	0.24	0.242	0.240	0.326	0.311	0.16	0.157	0.166	0.157
8	0.24	0.210	0.210	0.255	0.263	0.24	0.229	0.231	0.324	0.303	0.16	0.147	0.146	0.147
9	0.24	0.241	0.242	0.274	0.291	0.24	0.218	0.217	0.316	0.293	0.16	0.146	0.157	0.153
10	0.24	0.242	0.243	0.275	0.295	0.24	0.241	0.245	0.331	0.320	0.16	0.170	0.157	0.163
11	0.24	0.244	0.243	0.285	0.315	0.24	0.239	0.240	0.331	0.338	-0.16	-0.148	-0.154	-0.097
12	0.24	0.238	0.235	0.273	0.309	0.24	0.250	0.251	0.338	0.349	-0.16	-0.150	-0.165	-0.103
13	0.24	0.237	0.238	0.274	0.311	0.24	0.229	0.228	0.334	0.335	-0.16	-0.177	-0.170	-0.113
14	0.24	0.245	0.245	0.276	0.319	0.24	0.222	0.217	0.319	0.326	-0.16	-0.189	-0.181	-0.123
15	0.24	0.250	0.250	0.283	0.317	0.24	0.225	0.229	0.326	0.331	-0.16	-0.146	-0.145	-0.089
16	0.12	0.114	0.114	0.177	0.189	0.60	0.600	0.598	0.642	0.642	0.16	0.139	0.145	0.145
17	0.12	0.139	0.139	0.195	0.212	0.60	0.602	0.600	0.650	0.648	0.16	0.170	0.177	0.177
18	0.12	0.125	0.124	0.184	0.211	0.60	0.590	0.590	0.638	0.649	-0.16	-0.155	-0.153	-0.109
19	0.12	0.132	0.132	0.196	0.223	0.60	0.596	0.599	0.638	0.659	-0.16	-0.146	-0.152	-0.103
20	0.60	0.601	0.601	0.612	0.655	0.60	0.599	0.599	0.642	0.661	0.59	0.588	0.583	0.545

Table 4: As Table 2 but after removing the second outcome from 4 of the 5 BC trials. **For the average untruncated estimates, the Monte Carlo standard errors range from 0.008 to 0.023**

	$\boldsymbol{\Sigma}_\omega(11)$					$\boldsymbol{\Sigma}_\omega(22)$					$\boldsymbol{\Sigma}_\omega(12)$ and $\boldsymbol{\Sigma}_\omega(21)$			
		Untrunc.		Trunc.			Untrunc.		Trunc.			Untrunc.		Trunc.
Run	Truth	Uni	Multi	Uni	Multi	Truth	Uni	Multi	Uni	Multi	Truth	Multi	Multi	Multi
1	0	-0.003	0.000	0.203	0.152	0	0.021	0.021	0.281	0.193	0	-0.001	-0.001	0.026
2	0.12	0.123	0.123	0.258	0.240	0	-0.011	-0.006	0.284	0.186	0	-0.002	-0.012	0.026
3	0.12	0.106	0.105	0.255	0.235	0.12	0.115	0.112	0.357	0.283	0	-0.010	-0.012	0.024
4	0.12	0.115	0.115	0.253	0.240	0.12	0.106	0.101	0.351	0.275	0.08	0.065	0.064	0.078
5	0.12	0.114	0.115	0.247	0.243	0.12	0.093	0.097	0.329	0.270	-0.08	-0.082	-0.078	-0.029
6	0	0.003	0.003	0.200	0.143	0	-0.003	-0.005	0.287	0.174	0	-0.001	0.000	0.054
7	0.12	0.120	0.120	0.257	0.231	0	0.003	0.009	0.289	0.186	0	0.020	0.004	0.063
8	0.12	0.133	0.135	0.261	0.239	0.12	0.140	0.137	0.337	0.273	0	0.010	0.018	0.061
9	0.12	0.119	0.119	0.255	0.229	0.12	0.144	0.146	0.338	0.275	0.08	0.088	0.084	0.116
10	0.12	0.120	0.119	0.269	0.250	0.12	0.123	0.117	0.355	0.284	-0.08	-0.084	-0.070	0.005
11	0	-0.001	0.000	0.199	0.160	0	0.002	0.002	0.300	0.207	0	-0.003	-0.005	0.004
12	0.12	0.124	0.128	0.259	0.258	0	0.000	0.001	0.291	0.210	0	0.001	0.016	0.010
13	0.12	0.132	0.132	0.268	0.263	0.12	0.131	0.126	0.369	0.306	0	0.006	0.015	0.009
14	0.12	0.114	0.113	0.254	0.249	0.12	0.129	0.137	0.349	0.307	0.08	0.104	0.089	0.071
15	0.12	0.097	0.096	0.248	0.240	0.12	0.135	0.133	0.345	0.304	-0.08	-0.083	-0.093	-0.054
16	0.24	0.249	0.246	0.307	0.328	0.06	0.041	0.041	0.480	0.301	0.08	0.089	0.076	0.113
17	0.24	0.218	0.223	0.295	0.322	0.06	0.051	0.052	0.527	0.337	-0.08	-0.097	-0.083	-0.003
18	0.24	0.249	0.251	0.319	0.346	0.06	0.050	0.050	0.483	0.323	0.08	0.072	0.077	0.059
19	0.24	0.235	0.238	0.303	0.342	0.06	0.058	0.050	0.514	0.345	-0.08	-0.111	-0.085	-0.054
20	0	-0.022	-0.022	0.339	0.212	0	-0.014	-0.015	0.478	0.264	0	-0.015	-0.012	0.125

Table 5: Coverage probabilities of approximate 95% confidence intervals for complete data, using both the proposed multivariate and univariate methods. AJ(x) denotes the coverage probability of confidence intervals for the average treatment effect of J relative to A, for outcome x.

Run	Univariate						Multivariate					
	AB(1)	AC(1)	AD(1)	AB(2)	AC(2)	AD(2)	AB(1)	AC(1)	AD(1)	AB(2)	AC(2)	AD(2)
1	0.944	0.955	0.950	0.939	0.940	0.928	0.957	0.960	0.960	0.952	0.950	0.946
2	0.949	0.941	0.938	0.957	0.953	0.951	0.957	0.952	0.946	0.962	0.962	0.954
3	0.939	0.944	0.948	0.939	0.933	0.945	0.950	0.958	0.957	0.947	0.947	0.950
4	0.934	0.944	0.943	0.932	0.932	0.925	0.947	0.957	0.947	0.934	0.940	0.943
5	0.936	0.951	0.947	0.944	0.936	0.934	0.946	0.958	0.954	0.956	0.946	0.947
6	0.955	0.957	0.952	0.944	0.953	0.950	0.965	0.961	0.960	0.947	0.966	0.959
7	0.935	0.915	0.929	0.950	0.953	0.945	0.948	0.926	0.939	0.957	0.965	0.957
8	0.940	0.936	0.936	0.933	0.927	0.938	0.948	0.943	0.942	0.944	0.940	0.942
9	0.941	0.939	0.937	0.943	0.948	0.942	0.952	0.948	0.942	0.950	0.957	0.944
10	0.947	0.939	0.944	0.939	0.938	0.934	0.959	0.952	0.953	0.939	0.946	0.947
11	0.955	0.955	0.947	0.951	0.955	0.952	0.963	0.976	0.971	0.963	0.968	0.966
12	0.935	0.938	0.937	0.963	0.967	0.954	0.950	0.956	0.948	0.975	0.972	0.969
13	0.946	0.945	0.949	0.952	0.951	0.959	0.959	0.959	0.960	0.960	0.966	0.967
14	0.939	0.946	0.929	0.939	0.940	0.941	0.953	0.962	0.944	0.954	0.960	0.952
15	0.940	0.940	0.933	0.934	0.936	0.936	0.949	0.952	0.959	0.949	0.950	0.957
16	0.935	0.921	0.931	0.948	0.948	0.946	0.945	0.938	0.942	0.953	0.954	0.949
17	0.949	0.934	0.947	0.951	0.949	0.954	0.952	0.951	0.962	0.953	0.952	0.956
18	0.941	0.937	0.948	0.940	0.946	0.935	0.952	0.944	0.962	0.951	0.956	0.946
19	0.929	0.927	0.926	0.937	0.943	0.937	0.951	0.947	0.944	0.948	0.953	0.948
20	0.951	0.952	0.948	0.944	0.935	0.940	0.952	0.960	0.952	0.953	0.946	0.950

Table 6: As Table 5 but after removing the second outcome from 4 of the 5 BC trials.

Run	Univariate						Multivariate					
	AB(1)	AC(1)	AD(1)	AB(2)	AC(2)	AD(2)	AB(1)	AC(1)	AD(1)	AB(2)	AC(2)	AD(2)
1	0.944	0.955	0.950	0.945	0.941	0.935	0.961	0.966	0.968	0.957	0.963	0.956
2	0.949	0.941	0.938	0.960	0.957	0.957	0.961	0.956	0.956	0.971	0.971	0.967
3	0.939	0.944	0.948	0.946	0.941	0.947	0.958	0.962	0.962	0.954	0.951	0.959
4	0.934	0.944	0.943	0.942	0.943	0.933	0.954	0.962	0.956	0.953	0.950	0.944
5	0.936	0.951	0.947	0.955	0.943	0.939	0.951	0.959	0.956	0.968	0.956	0.953
6	0.955	0.957	0.952	0.949	0.959	0.949	0.965	0.968	0.968	0.961	0.963	0.968
7	0.935	0.915	0.929	0.946	0.962	0.953	0.959	0.939	0.948	0.962	0.968	0.953
8	0.940	0.936	0.936	0.939	0.936	0.933	0.950	0.951	0.947	0.953	0.949	0.953
9	0.941	0.939	0.937	0.940	0.952	0.946	0.957	0.955	0.947	0.958	0.966	0.961
10	0.947	0.939	0.944	0.943	0.941	0.946	0.957	0.959	0.957	0.944	0.946	0.955
11	0.955	0.955	0.947	0.962	0.961	0.962	0.974	0.971	0.970	0.971	0.975	0.976
12	0.935	0.938	0.937	0.962	0.962	0.961	0.956	0.960	0.956	0.976	0.976	0.972
13	0.946	0.945	0.949	0.951	0.951	0.954	0.964	0.962	0.962	0.967	0.964	0.967
14	0.939	0.946	0.929	0.950	0.959	0.941	0.956	0.962	0.943	0.964	0.971	0.966
15	0.940	0.940	0.933	0.940	0.946	0.935	0.955	0.956	0.962	0.957	0.962	0.958
16	0.935	0.921	0.931	0.941	0.940	0.944	0.955	0.949	0.949	0.953	0.954	0.956
17	0.949	0.934	0.947	0.941	0.940	0.944	0.957	0.963	0.965	0.950	0.956	0.954
18	0.941	0.937	0.948	0.942	0.940	0.947	0.956	0.953	0.965	0.953	0.956	0.958
19	0.929	0.927	0.926	0.942	0.949	0.937	0.953	0.956	0.953	0.956	0.963	0.955
20	0.951	0.952	0.948	0.949	0.936	0.941	0.959	0.962	0.954	0.956	0.959	0.952

Table 7: As Table 5 but making the consistency assumption.

Run	Univariate						Multivariate					
	AB(1)	AC(1)	AD(1)	AB(2)	AC(2)	AD(2)	AB(1)	AC(1)	AD(1)	AB(2)	AC(2)	AD(2)
1	0.930	0.932	0.933	0.923	0.908	0.905	0.930	0.939	0.938	0.919	0.917	0.905
2	0.927	0.911	0.912	0.926	0.924	0.927	0.927	0.909	0.917	0.928	0.927	0.930
3	0.916	0.912	0.920	0.912	0.900	0.915	0.917	0.919	0.925	0.911	0.900	0.913
4	0.904	0.907	0.910	0.905	0.890	0.898	0.906	0.909	0.912	0.908	0.889	0.897
5	0.906	0.913	0.926	0.911	0.903	0.902	0.906	0.913	0.931	0.918	0.903	0.900
6	0.926	0.928	0.930	0.925	0.928	0.931	0.933	0.922	0.933	0.918	0.930	0.928
7	0.916	0.895	0.907	0.922	0.923	0.922	0.918	0.895	0.913	0.917	0.933	0.921
8	0.909	0.901	0.906	0.907	0.891	0.912	0.907	0.892	0.910	0.909	0.896	0.916
9	0.919	0.918	0.914	0.919	0.913	0.913	0.922	0.922	0.908	0.924	0.921	0.912
10	0.931	0.912	0.920	0.909	0.897	0.909	0.924	0.909	0.919	0.908	0.899	0.906
11	0.936	0.924	0.931	0.935	0.930	0.938	0.942	0.942	0.937	0.940	0.937	0.941
12	0.910	0.895	0.909	0.944	0.940	0.930	0.917	0.907	0.918	0.952	0.945	0.943
13	0.916	0.916	0.926	0.934	0.915	0.937	0.920	0.921	0.930	0.931	0.921	0.933
14	0.918	0.913	0.903	0.901	0.912	0.912	0.921	0.912	0.906	0.917	0.914	0.915
15	0.918	0.905	0.897	0.905	0.895	0.910	0.916	0.917	0.923	0.925	0.907	0.927
16	0.905	0.877	0.899	0.926	0.911	0.929	0.901	0.880	0.899	0.924	0.909	0.921
17	0.919	0.892	0.913	0.935	0.929	0.937	0.914	0.890	0.913	0.930	0.920	0.928
18	0.901	0.881	0.906	0.920	0.910	0.916	0.910	0.884	0.912	0.921	0.912	0.916
19	0.896	0.872	0.889	0.914	0.919	0.911	0.902	0.876	0.898	0.923	0.920	0.918
20	0.922	0.920	0.922	0.920	0.916	0.919	0.927	0.929	0.920	0.925	0.919	0.928

Table 8: Percentage efficiency of the univariate method (compared to the proposed multivariate method). The percentages shown are the ratios of the empirical variances of the multivariate estimates and the empirical variances of the univariate estimates. Percentages of less than 100 indicate that the univariate method is not as efficient as the multivariate method. Results are shown for both complete and incomplete data, where for incomplete data we removed the second outcome from 4 of the 5 BC trials.

Run	Complete data						Incomplete data					
	AB(1)	AC(1)	AD(1)	AB(2)	AC(2)	AD(2)	AB(1)	AC(1)	AD(1)	AB(2)	AC(2)	AD(2)
1	100	100	99	100	100	100	100	99	99	100	98	99
2	99	99	99	100	101	100	99	100	100	100	100	99
3	99	99	99	99	100	100	99	99	100	100	100	100
4	100	99	100	100	101	101	100	99	101	100	103	101
5	99	101	100	99	99	98	99	100	101	99	101	99
6	99	99	99	99	97	97	99	99	99	99	95	98
7	98	98	98	98	100	99	98	98	98	99	96	98
8	99	100	99	99	99	99	99	100	99	98	97	99
9	98	99	100	97	98	98	99	100	100	97	96	98
10	98	99	99	100	100	100	98	99	99	100	98	99
11	99	100	100	100	100	100	99	101	100	100	102	101
12	99	99	99	99	100	99	100	99	99	100	100	98
13	100	100	100	100	100	100	100	100	100	100	101	100
14	99	99	99	100	100	100	100	100	100	99	101	101
15	100	99	99	99	98	99	99	99	99	99	103	101
16	98	99	99	99	100	100	99	100	100	100	99	100
17	101	101	101	100	100	100	101	100	100	100	99	99
18	100	101	100	101	100	100	100	100	101	101	102	100
19	100	100	100	100	100	100	100	100	101	100	102	100
20	97	98	96	96	96	97	97	98	97	96	88	96

4 Obtaining the within-study covariance structure for the RRMS example

Denote data in a three-arm study i with treatments A, B, C with three outcomes $1, 2, 3$ by $Y_{di} = (y_{1AB}, y_{2AB}, y_{3AB}, y_{1AC}, y_{2AC}, y_{3AC})^T$, where y_{jbk} represent estimates of difference in treatment effect of treatment k vs treatment b on outcome j. Then the within-study covariance matrix has the following form (study index i have been dropped in the matrix elements):

$$\mathbf{S}_{di} = \begin{bmatrix} \sigma^2_{1AB} & \sigma_{1AB}\sigma_{2AB}\rho_{w12} & \sigma_{1AB}\sigma_{3AB}\rho_{w13} & \psi_{14} & \psi_{15} & \psi_{16} \\ \sigma_{1AB}\sigma_{2AB}\rho_{w12} & \sigma^2_{2AB} & \sigma_{2AB}\sigma_{3AB}\rho_{w23} & \psi_{24} & \psi_{25} & \psi_{26} \\ \sigma_{1AB}\sigma_{3AB}\rho_{w13} & \sigma_{2AB}\sigma_{3AB}\rho_{w23} & \sigma^2_{3AB} & \psi_{34} & \psi_{35} & \psi_{36} \\ \psi_{41} & \psi_{42} & \psi_{43} & \sigma^2_{1AC} & \sigma_{1AC}\sigma_{2AC}\rho_{w12} & \sigma_{1AC}\sigma_{3AC}\rho_{w13} \\ \psi_{51} & \psi_{52} & \psi_{53} & \sigma_{1AC}\sigma_{2AC}\rho_{w12} & \sigma^2_{2AC} & \sigma_{2AC}\sigma_{3AC}\rho_{w23} \\ \psi_{61} & \psi_{62} & \psi_{63} & \sigma_{1AC}\sigma_{3AC}\rho_{w13} & \sigma_{2AC}\sigma_{3AC}\rho_{w23} & \sigma^2_{3AC} \end{bmatrix}$$

This within-study covariance matrix comprises of 4 blocks: two covariance matrices within each treatment contrast (one for B vs A and one for C vs A) and two covariance matrices between treatment arms.

The σ_{jbk} are the standard errors of the estimates y_{jbk}. The correlations ρ_{wjl} are the within-study correlations between treatment effect difference on outcome j and treatment effect difference on outcome l.

The covariances ψ_{qr} are present for studies with multiple arms (more than two arms) resulting in multiple treatment contrasts (in this example two contrasts; B vs A and C vs A). They form two sub-blocks with covariances on the diagonal (of the sub-block) $\psi_{qr} = var(y_{j(q)A})$ (the variance of the treatment effect in the control arm on outcome $j(q) = j(r)$) and off the diagonal $\psi_{qr} = \psi_{rq} = \rho^{\star}_{j(q),j(r)}$ (where $j(q)$ is the outcome corresponding to $Y_{di}[q]$ and $j(r)$ outcome corresponding to $Y_{di}[r]$). For example,

$$\begin{aligned} S_{di}[1,4] &= \psi_{14} = cov(y_{1AB}, y_{1AC}) = cov(y_{1B} - y_{1A}, y_{1C} - y_{1A}) \\ &= cov(y_{1B}, y_{1C}) - cov(y_{1B}, y_{1A}) - cov(y_{1A}, y_{1C}) + cov(y_{1A}, y_{1A}) = var(y_{1A}). \end{aligned}$$

$$\begin{aligned} S_{di}[2,4] &= \psi_{24} = cov(y_{2AB}, y_{1AC}) = cov(y_{2B} - y_{2A}, y_{1C} - y_{1A}) \\ &= cov(y_{2B}, y_{1C}) - cov(y_{2B}, y_{1A}) - cov(y_{2A}, y_{1C}) + cov(y_{2A}, y_{1A}) = \rho^{\star}_{1,2}\sqrt{var(y_{1A}) * var(y_{2A})}. \end{aligned}$$

Here y_{jb} represents the treatment effect in arm b on outcome j and $\rho^{\star}_{j(q),j(r)}$ is the correlation between treatment effects on outcome $j(q)$ and $j(r)$ in arm b (here outcomes 1 and 2 in arm A).
Similarly

$$\begin{aligned} S_{di}[2,5] &= \psi_{25} = cov(y_{2AB}, y_{2AC}) = cov(y_{2B} - y_{2A}, y_{2C} - y_{2A}) \\ &= cov(y_{2B}, y_{2C}) - cov(y_{2B}, y_{2A}) - cov(y_{2A}, y_{2C}) + cov(y_{2A}, y_{2A}) = var(y_{2A}). \end{aligned}$$

and

$$\begin{aligned} S_{di}[2,6] &= \psi_{26} = cov(y_{2AB}, y_{3AC}) = cov(y_{2B} - y_{2A}, y_{3C} - y_{3A}) \\ &= cov(y_{2B}, y_{3C}) - cov(y_{2B}, y_{3A}) - cov(y_{2A}, y_{3C}) + cov(y_{2A}, y_{3A}) = \rho^{\star}_{2,3}\sqrt{var(y_{2A}) * var(y_{3A})}. \end{aligned}$$

5 Data for the RRMS example

The data for the example in relapsing remitting multiple sclerosis (RRMS) are listed in Supplementary Table 9. In this table when we write, for example, 'IFNbeta-1b vs PBO', we mean that PBO is the reference group and IFNbeta-1b is the treatment that we compare this to when computing the treatment effect. In the notation of the paper, this means that we write B versus A to mean an 'AB trial where A is the reference group.

As explained in the previous section, the correlations include those between the differences in treatment effects (scale of the data) ρ_{wjl} and also those between treatment effects in reference treatment arm $\rho^{\star}_{jl}$ (that are necessary for a complete specification of the within-study matrices for the multi-arm studies). The within-study correlations ρ_{wjl} and $\rho^{\star}_{jl}$ are assumed the same across studies (and treatments). They are listed in Supplementary Table 10. The covariance matrix also contains standard errors (or variances) of the average effects in control arms for each study – those are listed in Supplementary Table 11.

Table 9: RRMS data.

Study	contrast	follow-up (months)	number of patients	annualized relapse rate log ARR (SE)	disability progression log OR (SE)	number of MRI patients	MRI lession log RR (SE)
IFNB SG (1)	IFNbeta-1b vs PBO	24	186	-0.08 (0.10)	0.00 (0.35)	186	-0.99 (0.35)
IFNB SG (2)	IFNbeta-1b vs PBO	24	186	-0.42 (0.11)	-0.44 (0.36)	186	-0.89 (0.44)
Johnson	GA vs PBO	24	251	-0.34 (0.10)	-0.17 (0.29)		
Jacobs/Simon	IFNbeta-1a vs PBO	24	172	-0.39 (0.13)	-0.63 (0.35)	158	-0.40 (0.16)
PRISMS (1)	IFNbeta-1a vs PBO	24	282	-0.34 (0.08)	-0.32 (0.27)	282	-1.11 (0.13)
PRISMS (2)	IFNbeta-1a vs PBO	24	278	-0.39 (0.09)	-0.46 (0.27)	278	-1.51 (0.15)
Durelli	IFNbeta-1b vs IFNbeta-1a	24	188	-0.34 (0.14)	-1.05 (0.37)		
Mikol	IFNbeta-1a vs GA	24	764	0.03 (0.14)	0.33 (0.24)	335	-0.31 (0.23)
O'Connor (1)	IFNbeta-1b vs GA	24	1121	0.06 (0.07)	0.06 (0.19)		
O'Connor (2)	IFNbeta-1b vs GA	24	1123	-0.03 (0.08)	0.12 (0.19)		
FREEDOMS 1	Fin 0.5 mg vs PBO	24	639	-0.80 (0.10)	-0.39 (0.17)	589	-1.35 (0.17)
FREEDOMS 1	Fin 1.25 mgvs PBO	24	643	-0.92 (0.10)	-0.47 (0.17)	557	-1.35 (0.14)
FREEDOMS 2	Fin 0.5 mg vs PBO	24	248	-0.64 (0.10)	-0.19 (0.17)	389	-1.35 (0.22)
FREEDOMS 2	Fin 1.25 vs PBO	24	249	-0.69 (0.10)	-0.39 (0.17)	371	-1.71 (0.24)
TRANSFORMS	Fin 0.5 mg vs IFNbeta-1a	12	644	-0.73 (0.15)	-0.31 (0.27)	644	-0.43 (0.15)
TRANSFORMS	Fin 1.25 vs IFNbeta-1a	12	636	-0.49 (0.14)	-0.18 (0.26)	636	-0.54 (0.14)

PBO – placebo, IFNbeta-1a(-1b) – interferon beta -a (-1b), GA – glatiramer, Fin – fingolimod
ARR – annualized rate ratio, RR – rate ratio, OR – odds ratio
Fingolimod trials are included as three-arm studies. Other three-arm studies (IFNB SG, PRISM and by O'Connor) have been included as separate two-arm trials (with reduced number of participants in "each" control arm by half). This was done to help the network structure – to ensure some repetition of the contrasts. Also differences in doses of interferon beta(1a and 1b) were ignored to simplify the geometry of the network.

The procedures for obtaining all data elements; the summary measures of treatment effects y_{jbk} on appropriate scales, with corresponding variances $var(y_{jbk})$ and $var(y_{jb})$ and the correlations between them, ρ_{wjl} and $\rho^{\star}_{jk}$ (listed in Tables 9–11) are described in detail in Section 2.2 and Appendix A of manuscript by Bujkiewicz *et al.* (2016).

Table 10: Within-study correlations.

Correlations			
ρ_{wjl}			
	y_{1bk}	y_{2bk}	y_{3bk}
y_{1bk}	1.00	0.25	0.09
y_{2bk}		1.00	0.09
y_{3bk}			1.00
$\rho^{\star}_{jk}$			
	y_{1b}	y_{2b}	y_{3b}
y_{1b}	1.00	0.4	0.15
y_{2b}		1.00	0.17
y_{3b}			1.00

Table 11: Standard errors of treatment effects in control arm

	relapse	disability	MRI
author	SE(log AR)	SE(log odds)	SE (log rate)
IFNB SG (1)	0.08	0.28	0.27
IFNB SG (2)	0.08	0.28	0.27
Johnson	0.07	0.20	
Jacobs/Simon	0.08	0.23	0.10
PRISMS (1)	0.06	0.22	0.08
PRISMS (2)	0.06	0.21	0.08
Durelli	0.09	0.23	
Mikol	0.10	0.18	0.14
O'Connor (1)	0.06	0.17	
O'Connor (2)	0.06	0.17	
FREEDOMS1	0.08	0.16	0.09
FREEDOMS1	0.08	0.16	0.09
FREEDOMS2	0.08	0.17	0.14
FREEDOMS2	0.08	0.16	0.14
TRANSFORMS	0.12	0.25	0.15
TRANSFORMS	0.12	0.25	0.15

AR – annualized rate