Performance of selected imputation techniques for missing variances in meta-analysis

NRN Idris\textsuperscript{1,4}, MH Abdullah\textsuperscript{2}, SMTolos\textsuperscript{3}

\textsuperscript{1,2,3}Department of Computational and Theoretical Sciences, Faculty of Science, International Islamic University Malaysia, Bandar Indera Mahkota, 25200 Kuantan, Pahang, MALAYSIA
Email: ruzni@iium.edu.my

Abstract. A common method of handling the problem of missing variances in meta-analysis of continuous response is through imputation. However, the performance of imputation techniques may be influenced by the type of model utilised. In this article, we examine through a simulation study the effects of the techniques of imputation of the missing SDs and type of models used on the overall meta-analysis estimates. The results suggest that imputation should be adopted to estimate the overall effect size, irrespective of the model used. However, the accuracy of the estimates of the corresponding standard error (SE) is influenced by the imputation techniques. For estimates based on the fixed effects model, mean imputation provides better estimates than multiple imputations, while those based on the random effects model respond more robustly to the type of imputation techniques. The results showed that although imputation is good in reducing the bias in point estimates, it is more likely to produce coverage probability which is higher than the nominal value.

1. Introduction
A meta-analysis is a statistical technique for integrating quantitative results of the same research question from several sources. Theoretically, combining the results from multiple trials should enhance the precision and accuracy of any pooled results. In practice however, there are a number of potential problems that may affect the validity of such results. One widely debated controversy related to meta-analysis concerns the choice between a fixed and random effects model for providing an overall estimate of the effect size [1]. When the differences in the effect sizes across the studies is due only to sampling error, they are considered homogeneous. This source of variation can be accommodated in meta-analysis by using the fixed effect model. However, if the variability in the effect size estimates exceeds those from sampling error alone, then a random effects model would be more appropriate as it takes into account the unexplained heterogeneity attributed to systematic differences between studies [2].

In meta-analysis of continuous data, variability measures, particularly the standard deviations (SDs), are often not reported in the published report of the trials [3]. A popular

\textsuperscript{4} Any correspondences should be addressed to ruzni@iium.edu.my
approach in handling this problem is through imputation of the missing SDs [4]. Earlier studies which examined the effects of imputing the missing SDs on the overall meta-analysis estimates [5]-[7] concluded that imputation recovers most of the lost information in the estimate of effect size and the corresponding SE. These studies did not examine the effects of different techniques of imputation used to estimate the missing variances.

This article investigates empirically, the effects of imputing the missing SDs and the choice of meta-analysis model on the overall meta-analysis estimates. Meta-analysis estimates were based on the fixed and random effect models obtained from three sets of simulated data, namely, (1) complete data, where all studies are assumed to report the SDs; (2) incomplete data, where studies with missing SDs were excluded; and (3) the imputed data, where the missing SDs were imputed, and the studies with imputed SDs were included in the analysis.

The imputation methods considered in this study are the weighted mean imputation and the multiple imputations (MI). These techniques of imputation are the most common techniques in meta-analyses, particularly in psychology and social science [3]. The mean imputation has some popularity among the analysts due to its simplicity and ease of understanding and interpretation. This technique is recommended if less than 10% of the data is missing [6] and known to perform better if the data is approximately normal, and the missing values are missing completely at random (MCAR) [8]. The MI is a popular choice of imputation technique, and the main advantage is that it allows for the computation of the uncertainty due to imputation, in addition to those due to sampling error [9]. In this study we assumed that the study level SDs is missing completely at random (MCAR).

2. Material and Method

2.1. Simulation of data

The data for each meta-analysis was simulated at individual levels and summarized into study or aggregate levels. The following model was used to generate the individual level data,

\[ y_{ij} = \beta_{0i} + \beta_{1i} T_{ij} + \epsilon_{ij} \]

where \( y_{ij} \) is the response of the \( j \)th patient in the \( i \)th study, \( \beta_{0i} \) is the random study effect, \( T_{ij} \) represents the dummy covariates for treatment, which takes two values (namely, 0 for the control and 1 for the treatment arm), \( \beta_{1i} \) is the random treatment effect which is assumed to vary across the studies, but would take a fixed value under the fixed effect model, and \( \epsilon_{ij} \) are the random error terms. \( y_{ij}, \beta_{0i}, \beta_{1i} \) and \( \epsilon_{ij} \) are assumed to be independent and normally distributed with \( \epsilon_{ij} \sim N(0, \sigma_{\epsilon}^2) ; \sigma_{\epsilon}^2 = 1 ; \beta_{0i} \sim N(\beta_0, \sigma_{\epsilon_0}^2) ; \beta_0 = 0, \sigma_{\epsilon_0}^2 = 1 ; \beta_{1i} \sim N(\beta_1, \sigma_{\epsilon_1}^2) ; \beta_1 = 1, \sigma_{\epsilon_1}^2 = 1 \). Additionally, the number of patients in each study for each \( N \) are assumed to be equal, i.e., \( n_i = n \) for \( i = 1, 2, \ldots, N \), with an equal number of patients undergoing the two treatments, i.e. \( n_{0i} = n_{1i} = \frac{1}{2} n \). Total responses in each simulation run was \( \Sigma n_{ij} \). The parameters varied across the simulation are the number of studies in each meta-analysis (\( N \)) at 10, 20, 30; the number of patients within each study in a meta-analysis (\( n \)) at 20, 60, 100; and the percentage of studies with missing SDs (\( x \% \)) at 10, 30, 50. The combinations of the number of studies, \( N \), the sample size, \( n \), and the percentage of missing studies, \( x \% \), generated 27 meta-analyses. Each of these combinations was replicated 500 times.
2.2. Creation of missing SD’s
To create the SDs missing completely at random, $x\% (x = 10, 30, 50)$ studies were selected at random from $N$ studies, and excluded from the data.

2.3. Imputation techniques
Mean imputation: The missing variances were replaced by the weighted mean of the available variances, where the weight is the number of patients in each study ($n$).

Multiple imputations: The missing variance was replaced by a randomly selected value from the available variances and estimates based on this data were recorded. This process was repeated 250 times in each of the 500 simulation runs, and the mean and standard deviations of the estimates were computed. This allows the uncertainty induced by the imputation to be incorporated into the overall estimate of the variance of the estimate. [8].

2.4. Measures of performance
The performance of the two imputation techniques were evaluated using the percentage relative bias (PRB) for the point estimates and the coverage probability for the interval estimates. The PRB is computed between the estimates from complete data (all studies reported the SDs) and the estimates from the imputed data and the incomplete data. To estimate the coverage probability, a random interval was computed based on estimates from the imputed and incomplete data in each simulation run for each meta-analysis. The coverage probability for $\hat{\theta}_{1,\text{imputed}}$ is estimated as the proportion of the random intervals, out of 500 replications, which contains $\hat{\theta}_{1,\text{imputed}}$. The nominal level used was 0.95.

3. Results

3.1. Fixed effects model
The percentage relative bias (PRB) for the estimates based on the fixed effects model is tabulated in Table 1. Clearly the PRB in the SE of the estimates were much higher compared to those of the effect size. Furthermore, the PRB was generally smaller when the missing SDs is imputed, compared to the approach of excluding the studies with missing SDs. The trends of the PRB in the effect estimates for the different values of $x\%$ are illustrated in Figure 1. There was not much difference in the magnitude of the PRB when the missing SDs were imputed using the two techniques of imputations. The mean imputation performed only slightly better than the MI. On the other hand, the PRB in the SE of the estimates (Figure 2) were much higher when the studies with missing SDs were excluded compared to when the missing SDs were imputed ($> 300\%$). Furthermore, unlike the trends for the effect estimates, there were significant differences in the PRB for the two imputation techniques. The mean imputation performed far better than MI, particularly for the larger percentages of missing SDs ($> 30\%$). Additionally, the effects of $x$ on PRB in SE was very notable, i.e. the bias increases with $x\%$.

Table 1. Fixed effects model: Percentage relative bias in effect size and the SE of the effect size (n=60)

| x % | N | % Relative bias of effect size | % Relative bias of SE of effect size |
|-----|---|-------------------------------|-----------------------------------|
|     |   | Not imputed | MI | Mean imp | Not imputed | MI | Mean imputed |
| 10  | 10 | 0.16         | 0.02 | 0.05      | -5.5       | -2.40 | -0.09       |
Figure 1. Fixed effects model : percentage relative bias in effect size ($n = 60$).
Legend: thin dotted-line - studies with missing SDs omitted ; thick break-line - mean imputation ; thick solid-line - multiple imputation.

Figure 2. Fixed effect model : percentage relative bias in the SE of the effect size ($n = 60$).
Legend: thin dotted-line - studies with missing SDs omitted ; thick break-line - mean imputation ; thick solid-line - multiple imputation.
Figure 3. The coverage probability for the true effect sizes based on fixed and random effects model.

The estimated coverage probability for $\hat{\theta}_{1,\text{re}}$ is illustrated in Figure 3. The coverage probability from both the mean imputed and MI data were above the nominal value for all the meta-analyses, while in incomplete data, the coverage were mostly well below the nominal level.

3.2. Random effects model

The PRB for the estimates based on the random effect model are presented in Table 2. Consistent with the results for the fixed effect model, the PRB were much smaller in the effect sizes compared to those in the SE of the effect size. Additionally the PRBs were much larger if the studies with missing SDs were excluded. However there were no notable differences in the magnitude of the PRB in the effect size using either the mean imputation or the MI (Figure 4).

As in the case for the fixed effect model, for estimates of effect size, the percentage of missing SDs and the sample sizes did not have much effect on the magnitude of the relative bias when imputation was used. By contrast, when studies with missing SDs were excluded, the PRB in the SE of the effect size increased significantly with increasing x%. In this case, the PRB increased up to 40% when half of the studies were excluded. Imputing the missing SDs appears to recover most of the information, as illustrated in Figure 5. The relative biases were all very close to zero for both techniques of imputation. This was different from
the results obtained from estimates based on the fixed effect model, where mean imputation seemed to be more superior in recovering the information on the SE of the estimates compared to the MI imputation. Additionally, when the SDs were imputed, the study sizes $n$ and the percentage of missing SDs, x %, had little effect on the bias in the SE.

Table 2. Random effect model: percentage relative bias in the effect size and the SE of the effect size

| X % | N | % Relative bias of effect size | % Relative bias of SE of effect size |
|-----|---|-------------------------------|-------------------------------------|
|     | Not imputed | MI Mean imp | Not imputed | MI Mean imp |
| 10  | -0.19 0.012 0.007 | -4.9 0.12 0.01 |
| 20  | 0.30 -0.003 0.003 | -5.4 0.01 0.01 |
| 30  | -0.29 0.0001 0.0002 | -5.4 0.03 0.02 |

| 30  | 10 -0.13 -0.022 | -0.007 -17.5 0.01 0.09 |
| 20  | 0.54 -0.008 0.001 | -19.0 0.01 0.02 |
| 30  | -0.83 -0.0001 | -0.004 -19.1 0.02 0.02 |

| 50  | 10 -0.25 0.082 0.002 | -35.8 0.03 0.40 |
| 20  | 0.60 0.005 -0.005 | -39.6 0.04 0.12 |
| 30  | -0.46 0.002 | -0.006 -39.7 0.09 0.08 |

Figure 4. Random effects model: percentage relative bias in effect size ($n = 60$). Legend: thick break-line - mean imputation; thick solid-line - multiple imputation.
Figure 5. Random effects model: percentage relative bias in SE of the effect size (n=60).

Legend: thin dotted-line - studies with missing SDs omitted; thick break-line - mean imputation; thick solid-line - multiple imputation.

4. Discussion

This paper investigates the influence of the imputation techniques and the choice of a meta-analysis model on the estimates of meta-analysis parameters—namely, the treatment effect size and the corresponding SE—in continuous response data with missing SDs in some of the studies. Comparisons in PRBs and the coverage probability were made for estimates based on the random and fixed effect model for the data imputed using the mean imputation and MI. The main conclusions had drawn from this project support many of the findings in previous literature. We have additionally illustrated that irrespective of whether the estimates of overall effect size is based on the fixed or random effect model, imputation is a good approach in handling the problem of missing study-level SDs. The PRB produced using this approach is relatively small compared to the studies with missing SDs. In fact, it was shown that the expected bias in this case will tend to be zero if the SDs across the studies is assumed to be completely homogeneous [10]. Imputation is therefore always recommended to estimate the SE of the effect size in data with missing SDs, as other serious bias may be introduced. The results show that if the random effect model is used to estimate the SE of the estimate for data with some missing SDs, both the MI and mean imputation will give equally good estimates (no difference in PRB; \( p < 0.337 \)). On the other hand, if the estimate of the SE is based on the fixed effect model, the techniques of imputation adopted will have some impact on the PRB introduced into the estimate of the SE. In this particular study, it was observed that if the fixed effect model is used, then mean imputation is expected to produce smaller PRB compared to those using the MI (difference; \( p < 0.001 \)). These results are clearly illustrated in Figure 2.

The results also showed that coverage probability for estimates based on excluded studies are mostly well below the nominal level despite a large percentage of overestimation in the SE. On the other hand, the coverage is mostly above nominal level for the data with imputed SDs due to overestimation of the effect sizes. This suggests that the
bias in the effects size has more influence on the coverage probability compared to the bias in the SE of the estimates. For estimates based on the fixed effect model, high coverage level using both the MI and mean imputed data suggests that the intervals are too wide. Despite a small bias in the effect sizes (from 0.01% – 1.7 %), the intervals are wider due to overestimation in the SE effect sizes. In the incomplete data, low coverage is mainly due to underestimation in the effect sizes. Although the SEs in this data are overestimated and much larger than those in imputed data, their influences are relatively small compared to those of the effect sizes. Similarly, for the random effect model, the coverage probability is consistently above the nominal value for both methods of imputation. This is expected, as the intervals are generally wider for the estimates using the random effect model due to incorporation of the between-study variances. For MI, the intervals are even wider, with the inclusion of additional uncertainty due to imputation in the estimate of the variances. Generally the coverage is more affected by the bias in effect size itself than the precision. We found no superior performance using the two techniques of imputation for the two models with respect to coverage probability. Both models generated wider intervals than the nominal level. Thus imputation in general may produce intervals which are on average too wide. Nonetheless, this approach is better than exclusion of studies with missing SDs, which produces coverage that is too low.

When selecting the imputation technique to compute missing SDs, an analyst should also consider the type of meta-analysis model that the estimate is based on in order to minimise the bias. The random effect model responds more robustly to the type of imputations used. Additionally, in the random effect model, both the MI and mean imputation produced smaller PRB compared to those from the fixed effect model.

Although the random effects model appears to be a safer choice, there are some concerns regarding its general application in practice, such as the assumptions of normally distributed random effects that pose problems with both validity and in our ability to check the validity for meta-analyses based on small number of studies [11].

In this study, the simulation of data was carried out under the assumption that the SDs is homogeneous across the studies. Further work should consider this limitation and that the extension of analyses is also possible on the effects of other imputation techniques, such as imputation by regression or the hot deck. These analyses are not intended to provide a specific guide for the model and imputation techniques to be utilized, but to investigate their influences on the estimates of meta-analysis parameters. However, we may suggest that when there are missing SDs, an analyst is therefore advised to exercise caution in choosing the technique of imputation which is best suited for the type of meta-analysis model used. This study shows that if the fixed effect model is used, then mean imputation is recommended. The random effect model is more vigorous and either the mean or MI technique is suitable.

References

[1] Brockwell S.E. and Gordon I.R., 2001. “A comparison of statistical method for meta-analysis,” Stat Med, 20, 825-840.
[2] Whitehead A., 2002. Meta-analysis of controlled clinical trials, London : John Wiley.
[3] Rothstein H.R., Sutton A.J., Borenstein M., 2005. Publication bias in meta-analysis; prevention, assessment and adjustments. London : John Wiley & Sons.
[4] Wiebe N., Vandermeer B., Platt R.W., Klassen T.P., Moher D., Barrowman N.J., 2006. “A systematic review identifies a lack of standardization in methods for handling missing variance data,” J. Clin. Epidemiol, 59, 342-353.

[5] Furukawa TA., Barbui C., Cipriani A., Brambilla P., Watanabe N., 2006. “Imputing missing standard deviations in meta analyses can provide accurate results,” J. Clin. Epidemiol, 59 7-10

[6] Idris, N.R.N., Robertson C., 2009. “The effects of imputing the missing standard deviations on the standard error of the meta-analysis estimates,” Commun Stat – Simul C, 38, 513-526

[7] Thiessen, P.H., Barrowman, N., Garg, A.X., 2007. “Imputing variance estimates do not alter the conclusions of a meta-analysis with continuous outcomes: a case study of changes in renal function after living kidney donation,” J. Clin. Epidemiol. 60: 3, 228-240.

[8] Little, J.A. and Rubin, D.B., 1987. Statistical analysis with missing data. London:John Wiley.

[9] Marshall, A., Altman, D.G., Holder, R.L., 2010. “Comparison of imputation methods for handling missing covariate data when fitting a Cox proportional hazards model: a resampling study,” BMC Medical Research Methodology, 10 112.

[10] Idris N.R., 2006. Estimating meta-analysis parameters in non-standard data. PhD Thesis, University of Strathclyde.

[11] Sutton, A.J., Abrams, K.R., Jones, D.R., Sheldon, T.A., Song, F., 2000. “Missing data,” In : Methods for meta-analysis in medical research, New York : John Wiley, 199-204.