An Empirical Study Into The Effect Of The Choice OfMeta Analysis Model On The Overall Estimates For Continuous Data With Missing Standard Deviations

Nik Ruzni Nik Idris , Noraida Saidin

Department of Computational And Theoretical Science,

International Islamic University Malaysia

Kuantan , Pahang, Malaysia

ruzni@iiu.edu.my

noraida@iiu.edu.my

Abstract— **The choice between the fixed and random effects model for providing an overall meta analysis estimate in continuous data may affect the accuracy of these estimates. For studies with complete information, the Cochrane's Q-test could provide some guide on the choice, although the power of this test is quite low. If the study- level standard deviations (SDs) are not completely reported or "missing", selection of meta analysis model should be done with more caution. Many studies suggest that imputation is a good way of recovering the lost information in the effect size estimate and the corresponding standard error. In this article, we compare empirically, the effects of imputation of the missing SDs on the overall meta analysis estimates based on both the fixed and random effect model. The results suggest imputation is recommended to estimate the overall effect size. However, to estimate its corresponding standard error (SE), imputation is recommended for the estimates based on the random effect model. If the fixed effect model is used, imputation may lead to bias estimates of the SE.**

Keywords-Fixed effect model; Random effect model; imputation

I. INTRODUCTION

A meta analysis is a statistical techniques 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 concerned the choice between the fixed and random effects model for providing an overall estimate of the effect size [1]. When the difference in the effect sizes across the studies is due only to sampling error, they are considered homogeneous, and 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, and the variation may be attributed to systematic differences between studies, a random effects model which takes into account the unexplained heterogeneity as the variation of individual study effects around a population average effect would be more appropriate.

The most common method of assessing the presence of heterogeneity is to carry out a simple χ^2-test based on the magnitude of the Cochrane Q-statistics defined by

$$Q = \frac{\sum w_i(y_i - \hat{\theta})^2}{\sum w_i}$$

which is a weighted sum of squares of the deviations of the individual study estimates y_i from the overall estimate $\hat{\theta}$. The weight w_i is the inverse of the variance of study i $w_i = \frac{1}{V(y_i)}$. Large values of Q suggest high variation in the effects across studies, thus the random effects model would be preferable over the fixed effects model. However, it was suggested that the power of this test is quite low especially in the case of sparse data [2]. Thus decisions based solely on this test are not recommended. Other factors such as the clinical aspect of a study should be taken into consideration.

Another common problem with meta analysis and systematic reviews is that when variability measures, particularly the standard deviations (SDs), are not reported in the published report of the trials. A popular approach in handling this problem is through imputation of the missing SDs [3]. Earlier studies which examined the effects of imputing the missing SDs on the overall meta analysis estimates [4,5,6] concludes that imputation recovers most of the lost information in the estimate of effect size and the corresponding SE. These studies however did not look at the effect of the model used to estimate the effect size and the corresponding SE.

In this article, we investigate and compare, empirically, the effects imputing the missing SDs and the choice of meta analysis model on the overall meta analysis estimates. We used meta analysis estimates based on the fixed and random effect models obtained from three sets of simulated data, namely, (1) complete data – where the all studies are assume to report the SDs (2) the incomplete data - where studies with missing SDs were excluded (3) the imputed data - where the missing SDs were imputed, and the studies with imputed SDs are included in the analysis. The effect

estimates and their corresponding SE from (2) and (3) were compared to those based on (1). The imputation methods considered in this study is the weighted mean imputation and the multiple imputation (MI) and the SDs are assume to be missing completely at random (MCAR).

II. META ANALYSIS MODEL

The fixed effect model assumes that the true effect is homogeneous across studies and thus variation in the observed values is assumed to be due to random error. Suppose y_i is the estimated effect size for the *ith* study from a collection of N studies. A general fixed effect model is given by

$$y_i = \theta + \epsilon_i$$

where $\epsilon_i \sim N(0, \sigma_i^2)$, $i = 1, 2, \ldots, N$ are the random deviations from the true effect size, θ, which are assumed to be independent with mean zero and variance σ^2 while θ, the true effect, is the same for all studies. Thus the study specific variance is $V(y_i) = \frac{\sigma_i^2}{n_i}$. The overall estimate of the effect is $\hat{\theta}$ and the corresponding variance of the estimate $V(\hat{\theta})$ are

$$\begin{aligned} \hat{\theta} &= \frac{\sum w_i y_i}{\sum w_i} \\ V(\hat{\theta}) &= \frac{1}{\sum w_i} \end{aligned}$$

where $w_i = \frac{1}{V(y_i)} = \frac{n_i}{\sigma_i^2}$. The overall estimate of effect and its confidence interval therefore are specific to the trials included in the meta analysis and cannot be generalized to a larger population.

In contrast to the fixed effect model, the random effects model assumes that the treatment in each trial is itself a realization of random variable, which is usually assumed to be normally distributed. The random effect model incorporates a between-study random term, ν_i into the model where ν_i is assumed to be normally distributed with mean 0 and variance σ_ν^2 and is independent of the error terms ϵ_i. This gives

$$\begin{aligned} y_i^* &= \theta_i + \epsilon_i \\ \theta_i &= \theta + \nu_i \end{aligned}$$

where $\epsilon_i \sim N(0, \sigma_i^2)$ and $\nu_i \sim N(0, \sigma_\nu^2)$, $i = 1, 2, \ldots, N$. The variance of the study specific estimate in this case is

$$V(y_i^*) = \frac{\sigma_i^2}{n_i} + \sigma_\nu^2$$

and the overall estimate of the effect and the corresponding variance of the estimate for the random effect model are

$$\begin{aligned} \hat{\theta}^* &= \frac{\sum w_i^* y_i^*}{\sum w_i^*} \\ V(\hat{\theta}^*) &= \frac{1}{\sum w_i^*} \end{aligned}$$

where the weight is given by $w_i^* = [\frac{\sigma_i^2}{n_i} + \sigma_\nu^2]^{-1}$.

Consequently, the standard error of each trial estimate is increased due to the addition of this between-trial variation. By allowing the between-study variability to be accounted for in the overall estimate and in its standard error, the random effect model produces results which can be considered to be more generalizable and realistic in estimating treatment effect for a future study.

III. METHOD

A. *Simulation of Data*

The analysis was carried out using simulation study. Individual patient level data were simulated using the following model

$$y_{ij} = \beta_{0i} + \beta_{1i} x_{ij} + \varepsilon_{ij}$$

where β_{0i} is the random study effect, x_{ij} represents the dummy covariates for treatment which takes two values, namely, 0 for the control and 1 for the treatment arm, and β_{1i} is the random treatment effect, and ε_{ij} are the random error terms. y_{ij}, β_{0i}, β_{1i} and ε_{ij} are assumed to be independent and normally distributed with $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$; $\sigma_\varepsilon^2 = 1$; $\beta_{0i} \sim N(\beta_0, \sigma_{v0}^2)$; $\beta_0 = 0, \sigma_{v0}^2 = 1$; $\beta_{1i} \sim N(\beta_1, \sigma_{v1}^2)$; $\beta_1 = 1, \sigma_{v1}^2 = 1$.

Additionally, number of patients for in each study for each N are assumed to be equal, i.e. $n_i = n$ for i = 1,2,...,N, with equal number of patients undergoing the two treatments, i.e $n_{0i} = n_{1i} = 1/2\, n$. The parameters that were varied cross the simulation are the number of studies in each meta-analysis (N) at 10, 20, 30 ; the number of patients in each meta analysis (n) at 20, 60, 100 ; and the percentage of studies with missing SDs (x%) at 10, 30, 50. The assigned values for N and n are chosen as we believe these are the likely range of values in real data. A total responses in each simulation run was Σn_{ij}.

Each of the 27 combinations of the number of studies N, the sample size n and the percentage of missing studies x %, were repeated 500 times. The mean effect estimate and the mean SE over the 500 simulations were computed. For Multiple imputation, the computation of the SE of the estimates takes into account the variation due to imputation as described by Robertson et al [7].

B. *Creation Of Missing SDs*

To create the SDs missing completely at random, x %

studies were selected at random from N studies, and excluded from the data.

C. Imputation techniques

Mean imputation : The missing variances were replaced by the weighted mean of the available variances.

Multiple Imputation : Each missing variance was replaced by a randomly selected value from the available variances and estimates based on this data recorded. This process was repeated 500 times for each missing value, which resulted in a set of 500 estimated effect sizes and the corresponding variances. The overall estimates of effect size were computed by taking the mean of the 500 effect sizes, while its overall variance were computed by taking the mean of 500 variances attributed by the sampling variability, plus the uncertainty due to the imputation.

D. Performance Measures

The performance of the two imputation techniques for each model were compared using the percentage relative bias (PRB) between the estimates which are based on all studies with no missing SDs, and the corresponding estimates using studies with imputed SDs. The PRB for the effect size based on Mean Imputation , for instance will be computed as follows

$$PRB = \frac{|\bar{X}_{all} - \bar{X}_{mean}|}{\bar{X}_{all}} \times 100\%$$

where $\bar{X}_{all} = \sum_{i=1}^{N} X_{all}/N$, X_{all} is the estimate of effect size based on all studies, $\bar{X}_{mean} = \sum_{i=1}^{N} X_{mean}/N$ X_{mean} is the effect size based on studies which includes the SDs imputed using the Mean imputation, and N is the number of simulations. Similar procedure is used for the computation of PRB in SE of the effect size estimate.

IV. RESULT

A. Fixed Effect Model

The percentage relative bias (PRB) for the estimates based on the fixed effect model are tabulated in Table 1. Clearly the PRB in the SE of the estimates is much higher compared to those of the effect size. Furthermore, the PRB are generally smaller when the missing SDs are 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 is not much differences in the magnitude of the PRB when the missing SDs are imputed using the two techniques of imputations. The mean imputation performs only slightly better than the MI. (Mean : 0.005 % – 0.07 % ; MI : 0.01 % - 0.3%). The PRB in the SE of the estimates (Figure 2) are much more higher when the studies with missing SDs are excluded compared to when the missing SDs are imputed (>300 %). Furthermore, unlike the trends for the effect estimates, there is significant differences in the PRB for the two imputation techniques. The mean imputation performs far better than MI, particularly for the larger percentages of missing SDs (> 30 %). Additionally, the percentage of studies with missing SDs, x %, appear to have substantial impact on the PRB in SE, namely, the bias increases with x %. The trend is observed when the missing SDs are imputed as well as when studies are omitted.

B. Random Effect Models

The PRB for the estimates that are based on random effect model are presented in Table 2. Consistent with the results for the fixed effect model, the PRB are much smaller in the effect size compared to the biases in the SE of the effect size. Additionally the PRB are much larger if the studies with missing SDs are excluded, compared to those when the missing SDs are imputed. There is no notable differences in the magnitude of the PRB in the effect size using either the mean imputation or the MI (Figure 3). As in the case for fixed effect model, the percentage of missing SDs and the sample sizes do not have much effect on the magnitude of the relative bias when imputation is used. In contrast, when studies with missing SDs are excluded, the PRB in the SE of the effect size increases significantly with increasing x %. In this case, the PRB increases up to 40% when half of the studies are excluded. However, imputation of the missing SDs appear to recover most of the information, as illustrated in Figure 4, where the PRB are all very close to zero, for both techniques of imputation. This is different from the results obtained from estimates based on the fixed effect model, where mean imputation seem to be more superior in recovering the information on the SE of the estimates.compared to the MI imputation.

Table 1 : ***FIXED EFFECT 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		
		OMIT	MI	MEAN IMP	OMIT	MI	MEAN IMP
10	10	0.16	0.02	0.05	-5.5	-2.40	-0.09
	20	-0.45	0.02	0.01	-5.4	-2.64	-0.11
	30	-0.21	0.03	-0.01	-5.4	-2.70	0.03
30	10	-0.14	0.03	-0.17	-19.7	-6.01	-0.21
	20	-0.01	-0.07	0.14	-19.5	-7.77	-0.20
	30	0.62	0.02	-0.06	-19.5	-7.90	-0.23
50	10	0.57	0.34	-0.01	-41.5	-10.5	-0.36
	20	-0.83	-0.06	0.02	-41.5	-12.3	-0.39
	30	-0.82	0.08	0.06	-41.3	-12.6	-0.27

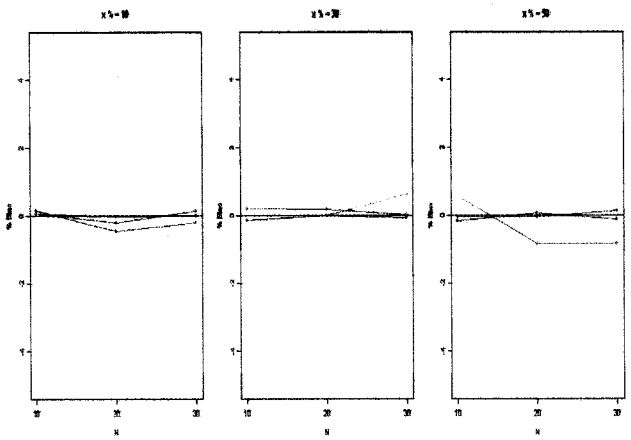

Figure 1 : Fixed Effect Model : *Percentage Relative bias in Effect size.*

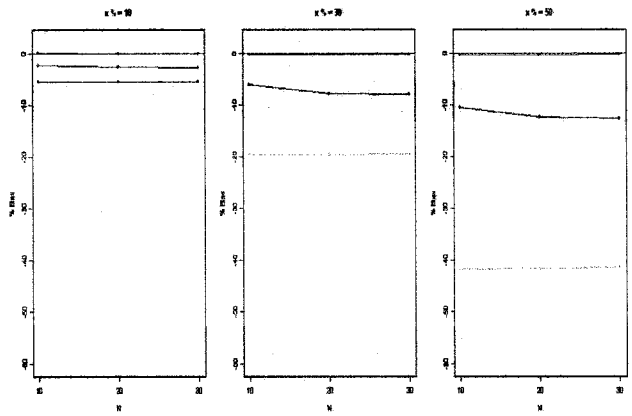

Figure 2 : Fixed Effect Model : *Percentage Relative bias in the SE of the Effect size*

Table 2 : *RANDOM EFFECT MODEL : Percentage Relative Bias in the 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		
		OMIT	MI	MEAN IMP	OMIT	MI	MEAN IMP
10	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	0.4
	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

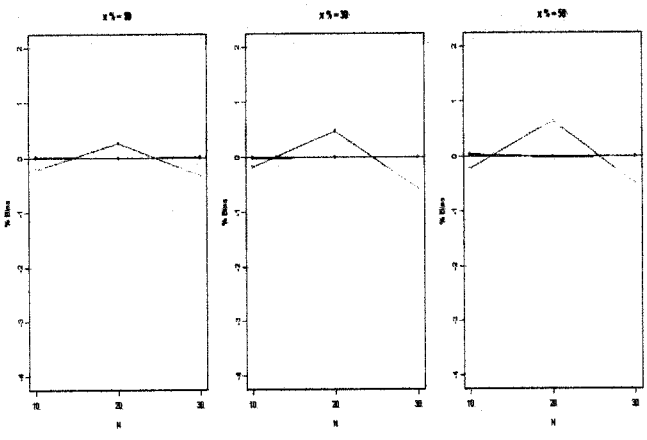

Figure 3 : Random Effect Model : *Percentage Relative bias in Effect size.*

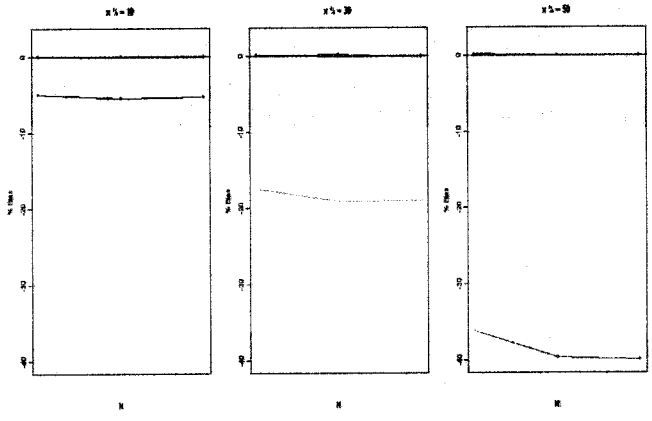

Figure 4 : Random Effect Model : *Percentage Relative bias in SE of the Effect size.*

NOTE : In all the figures, Red : Studies with missing SDs are omitted ; Green : Mean Imputation ; Blue : Multiple Imputation

V. CONCLUSION

In this paper we investigated the influence of the choice of meta analysis model and the imputation techniques on the estimates of meta analysis parameters, namely , the treatment effect size and the corresponding SE, in continuous data with missing SDs in some of the studies included in the meta analysis. Comparisons of PRB were made between the estimates based on the random effect model and the fixed effect model. The main conclusions drawn from this project support many of the findings from the previous literature [4,5,6]. We have illustrated that 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 much smaller compared to excluding the studies with missing

SDs. In fact, the expected bias in this case will tend to zero if the SDs across the studies are assumed to be completely homogeneous [8]. Nonetheless, while imputation of the missing SDs is recommended to estimate the effect size, it is not a must as the bias introduced are generally not very substantial particularly when the percentage of missing SDs is below 10%.

In contrast, imputation is always recommended to estimate the SE of the effect size in data with missing SDs as otherwise serious bias may be introduced [4]. The results show that if the random effect model is used to estimate the SE of the estimate in the data where there is some missing SDs, both the non-parametric MI and mean imputation will give equally good estimates (*no difference in estimates based on ; $p < 0.337$*). 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 is 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 5.

Therefore, in deciding the imputation technique to employ for the missing SDs, an analyst should look at the type of meta analysis model that the estimate is based on, in order to minimise the bias in the estimate of the SE. We also noted that the random effect model is more robust to the type of imputations used. Additionally, in random effect model, both the MI and mean imputation produces very small PRB in the SE compared to those from the fixed effect model.

Although the random effects model appear to be a safer choice, there are some concerns regarding its general application in practice such as the assumptions of normally distributed random effects or between study errors poses problems in both its validity, and in our ability to check that validity for meta analyses based on small number of studies. In presence of imputed SDs, the estimates would be affected and would depend largely on how close the values being replaced are to the true values. Hence both the technique of imputation and the meta analysis model utilised should be carefully considered as both have some influence over the accuracy of the estimates.

Although this paper considers only two techniques of non-parametric imputation, the principle generally applies to other type of imputation techniques. The analysts are therefore advised to exercise caution in choosing the techniques of imputation which is best suited for the type of meta analysis model used. Although the simulation of data was carried out under the assumption that the SDs are more or less homogeneous across the studies, the results are expected to be quite robust to minor departure from this assumptions. However further work should consider the effect of relaxing this assumption. Extension of analyses are also possible on the effects of other imputation techniques.

These analyses is not intended to provide specific guide for the model and techniques to be utilised but to investigate the influence of the model and imputation techniques on the possible bias in the estimates of meta analysis parameters. However we may suggest that when there is missing SDs, an analyst should look at the choice of model used before deciding on the technique of imputation, and our recommendation is that if the FE is used, used mean imputation. If RE model is used, then either one of the technique is good. However for this case, the MI is recommended as it takes into account the variations due to imputation

ACKNOWLEDGMENT

This work was supported by the IIUM Research Endowment Fund, grant no.EDW B0802-106, through the Department of CTS, Kulliyyah of Science.

REFERENCES

[1] Brockwell SE and Gordon IR (2001). A comparison of statistical method for meta-analysis. Statistics In Medicine ; 20: 825-840.

[2] Whitehead A (2002). Meta-analysis of controlled clinical trials. London:John Wiley . 215 - 237.

[3] Wiebe N, Vandermeer B, Platt RW, Klassen TP, Moher D, Barrowman NJ (2006).
A systematic review identifies a lack of standardization in methods for handling missing
variance data. Journal of Clinical Epidemiology ; 59: 342-353.

[4] Idris, NRN., Robertson C, (2009). The effects of imputing the missing standard deviations on the standard error of the meta analysis estimates. Communication in statistics – Simulation and Computation ; 38:513-526

[5] Furukawa TA., Barbui C, Cipriani A, Brambilla P, Watanabe N (2006). Imputing missing
standard deviations in meta analyses can provide accurate results. Journal of Clinical Epidemiology ; 59: 7-10

[6] 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

[7] Robertson C., Idris, NRN, Boyle P (2004). Beyond classical meta analysis: can inadequately reported studies be included ? Drug Discovery Today ; 9: 924-931.

[8] Idris NRN (2006). Estimating meta analysis parameters in non-standard data. PhD Thesis. University of Strathclyde.

31/3/2010