Missing Data: The Importance and Impact of Missing Data from Clinical Research

Christine R. Padgett,[1] Clive E. Skilbeck[1,2] and Mathew James Summers[1,3]

[1] *School of Psychology, University of Tasmania, Australia*
[2] *Tasmanian Neurotrauma Register, Royal Hobart Hospital, Tasmania, Australia*
[3] *Wicking Dementia Research & Education Centre, University of Tasmania, Hobart, Australia*

There is compelling evidence that traditional methods used to address the detrimental impacts of missing data are inadequate. Despite this, researchers have been slow to utilise newer statistical approaches known to be more effective. The aim of the current article is to offer a conceptual explanation of the rationale for using newer missing data techniques, with a focus on multiple imputation (MI). To illustrate the relative efficacy of deletion, single imputation and multiple imputation techniques in the clinical setting, 20 cases were selected randomly from a population study investigating the cognitive sequelae of traumatic brain injury (TBI), and 8 out of 20 cases had scores on one variable deleted to simulate a missing data set. Comparing the parameter estimates obtained by each technique to the known parameters of the complete data set revealed that MI outperformed deletion and single imputation approaches. It is therefore recommended that more sophisticated techniques such as MI should be considered in clinical research.

Keywords: missing data, multiple imputation, traumatic brain injury

Missing data (often referred to as missingness) is commonplace in clinical research, where participants may miss appointments, refuse certain tasks, drop out of or die during the course of longitudinal studies. Statistical techniques do not enable the analysis of data in which there are missing cases, consequently different approaches for dealing with the issue of missing values have evolved. Traditionally, the most commonly used approaches for dealing with missing values are deletion or single imputation. Deletion techniques exclude the missing data from analyses, thus statistical analysis is conducted only on complete sets of observed scores. While this may appear to be the most pragmatic approach, ignoring missing data is known to be problematic and has been criticised by methodologists as being one of the most dubious approaches to analysis (Wilkinson & American Psychological Task Force on Statistical Inference, 1999). Imputation techniques involve substituting the missing value with an estimated value. Single imputation techniques provide estimates based on the observed scores of the variable for which the data is missing. The most commonly used single imputation techniques are mean imputation and regression imputation. (Stochastic regression can also be used; however, this is rarely employed in clinical research, and is not addressed in this paper.) Mean imputation involves substituting the mean of the observed scores of a given variable for the missing value. Alternatively, where it is possible to conduct a regression analysis, a missing value can be estimated based on a regression equation for the variable.

It has been argued for some time that deletion and single imputation techniques fail to adequately address the impact missing data may have on analysis, with the primary concern being that such

Address for correspondence: Christine Padgett, School of Psychology, University of Tasmania, Australia. E-mail: Christine.Padgett@utas.edu.au

https://doi.org/10.1017/BrImp.2014.2 Published online by Cambridge University Press

approaches do not account for the increased variance that may occur had the actual score(s) been obtained (Allison, 2002; Enders, 2010; Graham, 2012; Schafer & Graham, 2002; Schafer & Olsen, 1998). An additional consequence of deletion techniques is the reduction of statistical power that occurs from the reduction in sample size (Graham, 2009; McKnight, McKnight, Sidani, & Figueredo, 2007). Therefore missing data has the potential to introduce bias and reduce the integrity of results. With the emergence of increasingly powerful statistical software, it is increasingly easy to employ more sophisticated missing data techniques that better compensate for the problems associated with traditional missing data approaches. Two of the more sophisticated approaches are multiple imputation (MI) and maximum likelihood (ML) (Allison, 2002; Baraldi & Enders, 2010; Sinharay, Stern, & Russell, 2001). This report focuses on MI, as it has been suggested that this is the most generalisable and flexible missing data technique. MI is robust to breaches of the assumption of normality, is impacted less by the nature of the missing data than other approaches, and performs relatively well even in small samples (Graham, 2012; Graham & Schafer, 1999; Jelicic, Phelps, & Lerner, 2010; McKnight et al., 2007; Sinharay et al., 2001).

Despite repeated recommendations for employing newer missing data techniques such as MI, these approaches are rarely used in clinical/psychological research (Enders, 2010; Jelicic, Phelps, & Lerner, 2009; Roth, 1994; Schafer & Graham, 2002). This is possibly due to a lack of awareness of such approaches, and because the literature regarding these newer approaches can appear daunting for non-statisticians. Furthermore, the detrimental effect that missing data has on analysis is also often underestimated by researchers, thus methods that ameliorate the impact of missing data are not considered necessary. With this in mind, the aim of this report is to offer a conceptual demonstration of how deletion, single imputation and MI compare in estimating missing data when applied in clinical research, in order to demonstrate the impact missing data can have. First, to clarify the effect of missing data on analysis, a description of the nature of missing data, and the limitations of various approaches in addressing missing data, are provided.

Why Does Missing Data Matter?

In order to understand why missing data can bias results, it is necessary to briefly explain the nature of missingness itself. The underlying premise of missing data theory is that missingness is often not random in nature. Rubin (1976) proposed that there are three types of missing data. First, data can be missing completely at random (MCAR). MCAR occurs in instances where missing data is not related to the scores on the variable in question, and is also not related to scores on any other variables under analysis. The second category proposed by Rubin is missing at random (MAR). Missing data in this category is not related to scores on the variable with the missing data; however, there may be systematic missingness in relation to another variable also being measured. For example, if assessing cognitive function during the acute phase following traumatic brain injury (TBI), it may be that individuals who also experienced orthopaedic injuries when the TBI was sustained would be less likely to complete a design-learning task, due to restricted mobility. In this case, missingness is systematic on one variable (presence of orthopaedic injury/polytrauma) but not on the variable being measured (design-learning), and is therefore MAR. Finally, missing values can be said to be missing not at random (MNAR). This occurs when the missing values are directly related to values of the variable in question. For example, if only participants who are experiencing high levels of depression refuse to complete a questionnaire assessing depression, high scores will be systematically missing on the variable with the missing data.

Missing data therefore has the potential to influence analysis in a number of ways. Clearly, where there is a relatively large amount of missing data and only the complete cases are analysed, the sample size may be seriously reduced. This results in a loss of power, leading to increased risk of Type II errors, whereby meaningful differences are missed (McKnight et al., 2007). An equally pressing concern, long recognised by statisticians but often unrecognised by clinical researchers, is that traditional techniques (deletion, single imputation) are based on an assumption that missing data occurs randomly, or is MCAR. However, data is often not MCAR, especially in the clinical setting where often there are one or more variables that have the potential to influence which scores are missing. As such, missingness is typically either MAR and MNAR, and therefore bias can occur (Graham, 2009). A further limitation of both deletion and mean and regression imputation is that they only utilise the variance of the observed data set. These methods assume that missing data will fall within the variance range of the existing data set. This is concerning, as the missing data is likely to introduce greater variability and, as such, ignoring missing data may reduce variance. This attenuates differences where scores are being compared, increasing the likelihood of a Type I error

https://doi.org/10.1017/BrImp.2014.2 Published online by Cambridge University Press

occurring; identifying a significant effect when one is not present (Schafer & Olsen, 1998). Alternatively, where correlational analysis is conducted, if mean imputation is used the strength of association becomes weaker, whereas employing regression imputation will strengthen the association, again increasing the risk of misinterpretation of results (Baraldi & Enders, 2010; Schafer & Olsen, 1998). In short, traditional approaches to replacing missing values are based on the frequently erroneous assumption that data is MCAR; increasing the risk of bias and errors in interpretation. Furthermore, the use of traditional techniques such as deletion and single imputation may lead to loss of power, and/or increased risk of Type I or Type II errors.

A Conceptual Explanation of Multiple Imputation (MI)

A brief conceptual description of the processes involved in MI is provided below. For more detailed explanations see Enders (2010), Graham (2012), Schafer (1999) and/or Sinharay et al. (2001).

Three separate phases occur in MI; imputation, analysis and pooling. The imputation phase involves the creation of multiple data sets where missing scores are replaced by scores obtained from regression equations that are computed from the variable in question and other related variables, as selected by the researcher. Generally, all variables that will be used in the final analysis, and any variable expected to interact with the variables under investigation, should be included in the MI (Enders, 2010). Indeed, including additional auxiliary variables (those which may interact with the outcome variable, but may not be of relevance to the research question) is recommended and does not mean that such variables need to be used in later analysis (Enders, 2010). Several algorithms are available for use when conducting the imputation phase; however, data augmentation is the most commonly used as it assumes multivariate normality (Enders, 2010; Graham, 2009), and is the focus of the following description.

The imputation phase involves two steps; the imputation step (I-step) and posterior step (P-step). In the I-step one or more regression equations (based on observed scores of selected variables) are calculated to estimate missing values. To increase variability, an error term derived from the regression is added to each imputed score. From this set of scores, the P-step calculates new means and covariances, and again a residual error term is added. The data set obtained in the P-step can then be used to create another imputed set of scores. In order to ensure that each imputed data set is independent, a large number of iterations (I-steps and P-steps) are run before the next imputed data set is selected for use in the analysis phase. Because each data set is based on different regression equations, and generates a different set of parameter estimates, the amount of variance is increased. Thus, MI techniques better account for the impact of unknown variance in the missing data than if a single imputation technique were to be employed. Parenthetically, by comparing the variability of parameters between each of the data sets, the researcher can also estimate the relative impact of the missing data – greater variation between data sets indicating a greater impact of missingness (McKnight et al., 2007). It should also be noted that the number of imputed data sets needed will vary depending on the amount and nature of the missing data, and predicted effect size (Graham, Olchowski, & Gilreath, 2007; Graham & Schafer, 1999)

When the required number of imputed data sets have been created, the analysis phase is performed, in which the desired analysis is conducted on each data set separately. In other words, whatever analysis had been planned (regression, analysis of variance (ANOVA)) can be conducted separately on each of the multiple data sets. Finally, in the pooling phase, the results obtained for each imputed data set during the analysis phase are combined to obtain a single value reflective of all data sets. For example, means and standard errors could be calculated for each imputed data set during the analysis phase, and then combined to create a pooled mean and standard error, which can then be reported in the final analysis. The process of pooling varies in complexity, depending on the statistic being calculated. In some situations, such as pooling means, the formula for pooling results is the same as for any other data. However, in other situations the process is more complicated; for example, pooling standard errors employs a formula in which the variance between and within each imputed data set is taken into account (for further detail on method of reporting MI results and formulae for pooling, see Enders, 2010).

Despite the conceptual nature of the above description, it may appear that MI is a complicated process. However, these steps are automated in many statistical software packages, and therefore MI can be conducted with relative ease. Nevertheless, it should be noted that popular statistical programs vary in capacity to perform the various stages of MI. For example, while SPSS conducts the imputation and analysis phases automatically, it currently doesn't always pool test statistics (for example, if *F*-tests are performed). Where the test statistic is not calculated, it is possible to create syntax or calculate manually (Enders, 2010;

https://doi.org/10.1017/BrImp.2014.2 Published online by Cambridge University Press

TABLE 1
Injury and Demographic Characteristics of Participants ($n = 20$)

	Range	*M*	*SD*
Age at injury (years)	21–70	45.56	14.71
Post-traumatic amnesia (days)	0.03–9.00	1.11	2.15
Pre-morbid FSIQ*	75–121	104.35	11.27

*Premorbid full-scale IQ was estimated using the National Adult Reading Test (NART).

Howell, 2012; Schafer & Olsen, 1998), although clearly this is not ideal and it is hoped that forthcoming versions will soon be more user-friendly. SAS does pool test statistics routinely, and there are also a number of MI-specific programs, such as NORM, which also have better functionality.

It is also important to have some understanding of the processes involved in MI, as the researcher may need to modify the default settings in some software. Of particular importance are decisions regarding the number of iterations allowed to run between each data set selected for analysis, and the total number of imputed data sets that should be obtained. For example, in SPSS (version 21) the default number of iterations when choosing the custom setting is 10, whereas it is likely that many more iterations may be necessary, with current recommendations suggesting that 100–200 iterations are typically required (Baraldi & Enders, 2010; Graham & Schafer, 1999). Furthermore, in SPSS the default number of imputations is 5, and while some authors suggest that this is acceptable (Sinharay et al., 2001), others suggest that 10–20 imputations are required, although this is dependent on the amount of missing data and predicted effect size (Enders, 2010; Graham et al., 2007; Graham & Schafer, 1999).

The Current Study

Baraldi and Enders (2010) offer a coherent and insightful demonstration that clearly illustrates traditional and newer techniques and their effect. The current study replicates this approach, but rather than create a simulated data set, the data reported here has been extracted from an existing clinical research project investigating cognitive and psychosocial function following traumatic brain injury (TBI). In order to demonstrate the application of missing data approaches within the clinical research setting, a small number of complete cases were randomly selected from the data set. Parameter estimates (means and standard deviation/error) were calculated, and then some scores were excluded to create a hypothetical missing data set. The means and standard deviations obtained by deletion, mean imputation, regression imputation and MI were then compared to the known mean and standard deviation of the complete data set in order in order to determine the relative efficacy of each approach.

Description of Data Set

Data was obtained from the Neurotrauma Register; a population study investigating cognitive and psychosocial recovery following TBI (ethics approval obtained). Individuals with TBI presenting at the Royal Hobart Hospital, Tasmania from 2003 to 2007 were invited to participate in the Register. A total of 1227 (m = 798, f = 429) individuals consented and underwent a battery of psychosocial and neuropsychological assessments at admission, with follow-up assessments at 3, 6 and 12 months, and annually up to 5 years post injury. For the purpose of the current study, a random draw of 20 participants was selected (m = 11, f = 9), and scores obtained at the initial assessment were used. The injury and demographic variables of the sample are shown in Table 1.

Materials

Two cognitive measures were selected from the assessment battery. The Controlled Oral Word Association Task (COWAT; Benton & Hamsher, 1989) requires participants to list as many words as possible, starting with a particular letter, within 60 s. Three trials, using the letters 'F', 'A' and 'S' were used for all participants, with the total number of words across all three conditions being used in analysis. Coughlan's Information Processing Speed Task (IP Speed; Coughlan & Hollows, 1985) is a measure of processing speed, in which the participant identifies the second highest number in each row of five numbers over a set of 105 rows of numbers. The total number of lines in which the correct number is crossed out was the score used in the current analysis.

Procedure

From the 20 selected cases, participants who had attended between June 2005 and June 2006 had

https://doi.org/10.1017/BrImp.2014.2 Published online by Cambridge University Press

TABLE 2

Means and Standard Deviations Obtained Using Deletion, Mean and Regression Imputation Techniques for Hypothetical Missing Data Set, Compared to Complete Data Set

Complete (observed) data		Hypothetical missing COWAT data set		
IP Speed	COWAT	Deletion technique	Mean imputation	Regression imputation[a]
38.00	23.00	–	41.67	30.07
67.00	19.00	–	41.67	43.93
60.00	26.00	–	41.67	40.59
59.00	44.00	–	41.67	40.11
52.00	31.00	–	41.67	36.76
38.00	26.00	–	41.67	30.07
72.00	65.00	–	41.67	46.32
58.00	46.00	–	41.67	39.63
50.00	36.00	36.00	36.00	36.00
39.00	18.00	18.00	18.00	18.00
63.00	50.00	50.00	50.00	50.00
42.00	32.00	32.00	32.00	32.00
87.00	54.00	54.00	54.00	54.00
67.00	45.00	45.00	45.00	45.00
86.00	63.00	63.00	63.00	63.00
87.00	47.00	47.00	47.00	47.00
41.00	28.00	28.00	28.00	28.00
68.00	34.00	34.00	34.00	34.00
36.00	44.00	44.00	44.00	44.00
79.00	49.00	49.00	49.00	49.00
Mean	**39.00**	**41.67**	**41.67**	**40.37**
SD	**13.76**	**12.45**	**9.48**	**10.26**

[a] Regression imputation equation: $\hat{Y} = 11.989 + 0.478$ (IPSpeed).

their scores on the COWAT deleted, thereby simulating a data set in which some data was missing due to a hypothetical data entry error. As such, the simulated missing data set can be said to have data missing at random (MAR), as the likelihood of any given score being missing is not directly related to the variable being measured. Deletion, mean and regression imputation, and MI were then conducted.

Results

SPSS (version 21) was used for all analyses. Means and standard deviations were calculated for the complete set, and then COWAT scores for individuals assessed between June 2005 and June 2006 were excluded, resulting in a total of 8 out of the 20 scores being deleted. Standard approaches to missing data (deletion, and mean and regression imputation) were conducted, and the means and standard deviations obtained using each approach were compared to the known values of the full data set (Table 2). Scatterplots for the complete data set (Figure 1), the deletion data set (Figure 2), mean imputation (Figure 3) and regression imputation (Figure 4) are also provided to enhance illustration.

When the deletion technique was used, there was a slight increase in the mean and reduction in variance, as compared to the known parameters (Table 2). The mean obtained from the deletion data set was then used for the mean imputation approach. Unsurprisingly, this resulted in the mean remaining the same as for the deletion technique, with a further decrease in variance (Table 2). A bivariate regression analysis was conducted using the IP Speed scores and the COWAT scores from the hypothetical missing data set (Table 2), with the resulting regression equation being used to imputate the missing COWAT values (see footnote Table 2). The regression imputation mean was closer to the true mean than either the deletion or mean imputation technique. However, while the variance was slightly more accurate than variance obtained by

https://doi.org/10.1017/BrImp.2014.2 Published online by Cambridge University Press

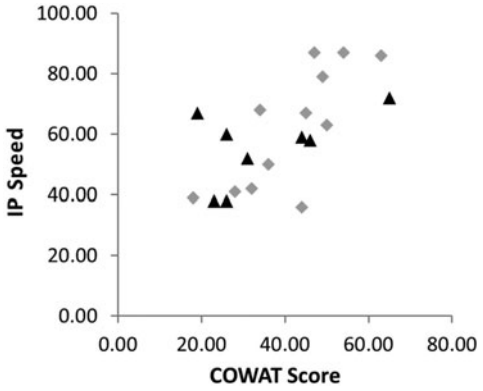


FIGURE 1

Scatterplot of observed data set. ▲ Denotes observed data point deleted to create missing data set.

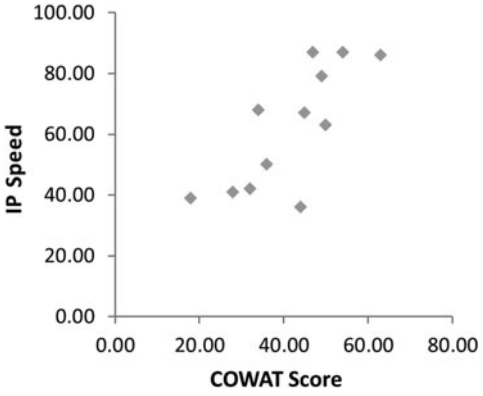


FIGURE 2

Scatterplot of the hypothetical data set cases with missing values removed using the deletion technique.

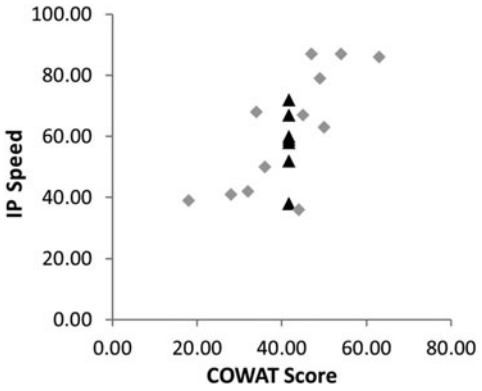


FIGURE 3

Scatterplot of the hypothetical data set missing values replaced by mean imputation. ▲ Denotes mean imputed data point.

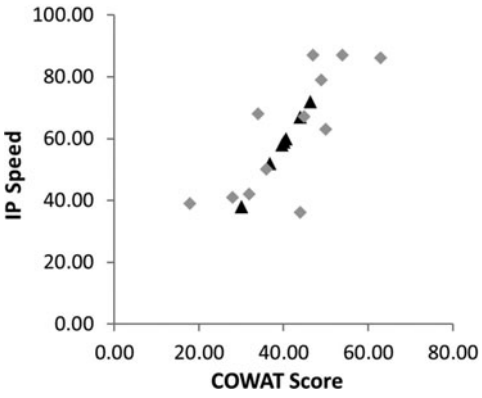


FIGURE 4

Scatterplot of the hypothetical data set missing values replaced by regression imputation. ▲ Denotes regression imputed data point.

mean imputation, it was not as close to the observed *SD* as when the deletion technique was employed.

MI was then conducted, using IP Speed and COWAT scores in the missing data set. Although it has been recommended that at least 20 imputations be used in most situations (Graham et al., 2007), for the purpose of demonstration a total of five imputations were performed (Table 3). In order to increase independence of each imputation, the number of iterations between each imputation was increased from the SPSS default setting of 10 to 200, consistent with Baraldi and Enders' demonstration (2010). As is evident in Table 3, each imputation generates different values for a given missing score, thus creating differing means and standard errors (*SE*). The mean and standard error are pooled for the final analysis. When comparing the pooled mean and standard error to the complete data set, it is apparent that MI has yielded both a mean and variance very similar to the known parameter estimates, and is closer on both measures than the deletion or single imputation techniques.

Discussion

In the current study, a random sample of 20 cases was drawn from a clinical research study to compare the means and variance obtained from a

https://doi.org/10.1017/BrImp.2014.2 Published online by Cambridge University Press

TABLE 3

Comparison of Means and Standard Deviations (*SD*s) for Complete Data Set, and Hypothetical Missing Data Sets Employing Multiple Imputation Techniques

Original data			Imputed COWAT scores					
IP Speed	COWAT	Missing data COWAT	MI Set 1	MI Set 2	MI Set 3	MI Set 4	MI Set 5	MI Pooled
38.00	23.00	–	10.31	33.95	22.94	23.09	50.17	
67.00	19.00	–	33.66	51.92	32.77	26.33	33.12	
60.00	26.00	–	43.81	42.86	38.59	36.45	49.10	
59.00	44.00	–	25.14	52.26	50.86	25.12	34.03	
52.00	31.00	–	35.16	34.57	35.31	0.04	47.08	
38.00	26.00	–	37.51	35.46	25.69	24.30	23.13	
72.00	65.00	–	48.43	32.17	17.29	60.61	27.59	
58.00	46.00	–	16.13	61.45	16.31	43.04	54.99	
50.00	36.00	36.00	36.00	36.00	36.00	36.00	36.00	
39.00	18.00	18.00	18.00	18.00	18.00	18.00	18.00	
63.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	
42.00	32.00	32.00	32.00	32.00	32.00	32.00	32.00	
87.00	54.00	54.00	54.00	54.00	54.00	54.00	54.00	
67.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	
86.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	
87.00	47.00	47.00	47.00	47.00	47.00	47.00	47.00	
41.00	28.00	28.00	28.00	28.00	28.00	28.00	28.00	
68.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	
36.00	44.00	44.00	44.00	44.00	44.00	44.00	44.00	
79.00	49.00	49.00	49.00	49.00	49.00	49.00	49.00	
Mean	**39.00**	41.67	37.51	42.23	36.98	36.95	40.96	**38.93**
SD*/*SE	**13.76**	12.45	13.46	11.58	13.23	15.45	11.93	**13.47**[a]

[a]As SPSS does not pool *SE* this was calculated manually using the formula provided in Baraldi & Enders (2010).

complete data set to an identical data set, except with missing values. As demonstrated, MI yielded parameter estimates closer to the known estimates than any of the traditional missing data approaches. Each of the deletion and single imputation procedures was limited in its capacity to reflect the complete sample parameters, yielding means and/or variance estimates that were markedly different to the original data set. In contrast, MI was found to provide both a mean and standard error that was close to identical to the known values, and therefore outperformed the deletion or single imputation approaches.

This finding provides similar results to Baraldi and Enders (2010), who used a hypothetical data set to compare missing data techniques. Indeed, when considering the traditional approaches, in both studies the regression imputation yielded the closest mean, and deletion yielded the closest standard deviation when compared to the complete data sets. An advantage of the current study was that the parameter estimates obtained by each missing data technique were able to be compared to the known parameters of a complete data set which has been obtained in the clinical setting, thus indicating that missing data techniques such as MI can be applied in this context.

While the above results indicate that MI is an effective method for addressing missing data, it must be acknowledged that the problems associated with missing data cannot be entirely ameliorated by any missing data techniques. MI and other new missing data procedures are useful for data that is MCAR or MAR; however, they are less reliable when data is MNAR. Unfortunately, although there are methods to determine if data is MCAR (see Enders, 2010), it is not possible to determine whether data is MAR and MNAR, and to date there is no effective technique for addressing data that is MNAR (Enders, 2010; Graham, 2012). Nevertheless, MI has been found to work relatively well in the MNAR setting (Enders, 2011), and may remain

https://doi.org/10.1017/BrImp.2014.2 Published online by Cambridge University Press

a better alternative than traditional approaches, or even techniques specifically designed to deal with data that is MNAR, as techniques designed for MNAR data rely on untestable and stringent assumptions (Enders, 2010; McKnight et al., 2007; Schafer & Graham, 2002). However, where there are strong violations of the MAR assumption, use of MI or other modern techniques may be questionable.

It should also be noted that even where missingness is MCAR or MAR, MI may not be the most appropriate approach. McKnight and colleagues (2007) state that small sample sizes may reduce the reliability of MI; however, others have suggested that this likely to be a reflection of the inherent limitations of small samples, rather than inefficiency of MI *per se* (Graham, 2009; Graham & Schafer, 1999). Using standard MI in some longitudinal designs and with multilevel data may also be problematic, requiring either the use of an MI algorithm other than data augmentation, or the use of a technique other than MI (Enders, 2010, 2011; Graham, 2012).

A further limitation is that statistical packages still vary in terms of ease of use, especially in relation to pooling parameters and test statistics. Given recent advances, and increased use of modern missing data approaches, it is likely that techniques such as MI will become better integrated into many software packages in the near future. Nevertheless, it is important to be aware of the capacity of a given program with regard to MI prior to commencing analysis.

An issue not often addressed in the literature to date is the proportion of data missing from a data set that can be reliably replaced when employing missing data techniques such as MI. McKnight et al. (2007) observe that any such guideline would be arbitrary, and would also depend on the nature of the research. However, it may be advantageous to clarify whether there is a proportion of missingness at which the efficacy of MI is reduced, and whether there are any considerations specific to applying MI in the clinical research setting, as have been observed for other statistical techniques (Proust-Lima, Dartigues, & Jacqmin-Gadda, 2011). Furthermore, given that clinical researchers may have little experience with such techniques, some provisional guidelines may foster confidence and encourage greater utilisation of these approaches.

It is hoped that by providing this comparison of the various missing data techniques, the benefit of utilising MI in the clinical setting is apparent. It should also be clear that regardless of whether missing data is actively dealt with, it is important to have a sound understanding of the nature and impact of missing data. Ideally, while modern missing data approaches can address missingness on a *post hoc* basis, it is suggested that consideration should be given to the possible factors that may contribute to missing data prior to commencement of research. Any factors that are likely to influence missingness can then be measured, allowing more rigorous analysis. That being said, Graham (2012) states that while it is useful to include and understand the key causes of missing data, it is not necessary to include all variables contributing to missing data for MI to be effective. Finally, whilst beyond the scope of this paper, it should be noted that understanding the nature of missing data can also allow more effective research, both in terms of analysis and in creating cost-effective studies by utilising a planned missingness design (for further explanation, see Enders, 2010 and Graham, 2012).

In conclusion, the nature and detrimental impact of missing data, and the associated limitations of traditional methods of dealing with missingness, are often not fully appreciated. The above findings indicate that modern missing data approaches, such as MI, outperform traditional approaches by reintroducing the variance that was likely to have occurred should missing scores have been obtained. This minimises the amount of bias caused by missing data, and increases the power to detect meaningful effects. MI and other approaches are now readily available in many statistical software programs, and are relatively fast and straightforward to use. While some additional effort may be required to learn new techniques such as MI, it is suggested that the long-term benefits make learning these techniques worthwhile, and it is hoped that approaches such as MI will be conducted routinely in the clinical research setting.

Financial Support

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Declaration of Interest

M.J.S. reports honoraria from Eli Lilly (Australia) Pty Ltd and grants from Novotech Pty Ltd, outside the submitted work.

References

Allison, P.D. (2002). *Missing data*. Thousand Oaks, CA: Sage.

Baraldi, A.N., & Enders, C.K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, *48*(1), 5–37.

Benton, A.L., & Hamsher, K.D. (1989). *Multilingual aphasia examination*. Iowa City: AJA Associates.

https://doi.org/10.1017/BrImp.2014.2 Published online by Cambridge University Press

Coughlan, A.K., & Hollows, S.E. (1985). *The Adult Memory & Information Processing Battery (AMIPB): manual*. England: St James' University Hospital, Leeds.

Enders, C.K. (2010). *Applied missing data analysis* (1st ed.). New York: The Guildford Press.

Enders, C.K. (2011). Analyzing longitudinal data with missing values. *Rehabilitation Psychology*, *56*(4), 267–288.

Graham, J.W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549–576.

Graham, J.W. (2012). *Missing data: analysis and design*. New York: Springer.

Graham, J.W., Olchowski, A.E., & Gilreath, T.D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, *8*(3), 206–213.

Graham, J.W., & Schafer, J.L. (1999). On the performance of multiple imputation for multivariate data with small sample size. In R.H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 1–29). California: Sage Publications.

Howell, D.C. (2012). Multiple imputation using SPSS. Retrieved from http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data/MissingData SPSS.html (accessed 11 March 2014).

Jelicic, H., Phelps, E., & Lerner, R.A. (2009). Use of missing data methods in longitudinal studies: The persistence of bad practices in developmental psychology. *Developmental Psychology*, *45*(4), 1195–1199.

Jelicic, H., Phelps, E., & Lerner, R.M. (2010). Why missing data matter in the longitudinal study of adolescent development: using the 4-H Study to understand the uses of different missing data methods. *Journal of Youth and Adolescence*, *39*(7), 816–835.

McKnight, P.E., McKnight, K.M., Sidani, S., & Figueredo, A.J. (2007). *Missing data: A gentle introduction*. New York: The Guildford Press.

Proust-Lima, C., Dartigues, J.F., & Jacqmin-Gadda, H. (2011). Misuse of the linear mixed model when evaluating risk factors of cognitive decline. *American Journal of Epidemiology*, *174*(9), 1077–1088.

Roth, P.L. (1994). Missing data – a conceptual review for applied psychologists. *Personnel Psychology*, *47*(3), 537–560.

Rubin, D.B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–590.

Schafer, J.L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, *8*(1), 3–15.

Schafer, J.L., & Graham, J.W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177.

Schafer, J.L., & Olsen, M.K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research*, *33*(4), 545–571.

Sinharay, S., Stern, H.S., & Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods*, *6*(4), 317–329.

Wilkinson, L., & American Psychological Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*(8), 594–604.

https://doi.org/10.1017/BrImp.2014.2 Published online by Cambridge University Press