Missing data treatments in intervention studies: What was, what is, and what should be

Charlie Rioux[1] and Todd D. Little[1,2]

Abstract

Missing data are ubiquitous in studies examining preventive interventions. This missing data need to be handled appropriately for data analyses to yield unbiased results. After a brief discussion of missing data mechanisms, inappropriate missing data treatments and appropriate missing data treatments, we review the current state of missing data treatments in intervention studies as well as how they have evolved over the years. Although missing data treatments have improved over the years, antiquated missing data treatments associated with biased results are still prevalent. Furthermore, many studies do not appropriately report their rates of missing data and missing data treatments. Using appropriate missing data treatments is elemental to accurately identify effective preventive interventions and properly inform practice and policy.

Keywords

Missing data, attrition, dropout, multiple imputation, full information maximum likelihood, interventions, clinical trials

Intervention studies are an important component of prevention science. Although study designs can (and should) aim to limit incomplete assessments and participant dropout (de Leeuw, 2001; Dziura, Post, Zhao, Fu, & Peduzzi, 2013; R. J. Little et al., 2012; Wisniewski, Leon, Otto, & Trivedi, 2006), missing data remain ubiquitous in intervention studies, and some level of missing data is typically unavoidable in most studies. Missing data are an important problem since if not handled properly, missing data can compromise conclusions regarding the effectiveness of an intervention and the generalizability of findings. Indeed, participants with complete data may not be a random subsample of the original sample and may thus not be representative of the population of interest. Furthermore, in the specific case of randomized controlled trials, while the groups may have been equivalent within the original sample, they may not be equivalent nor fully generalizable when only considering the sample with complete data. Moreover, a properly planned quantitative intervention study may have planned sample size based on power analyses, but without proper missing data treatments, the sample size may be reduced by missing data, which decreases the power of the study to detect intervention effects. Proper missing data treatments are key as some problematic missing data treatments can introduce bias in the findings even when they recover the original sample size. Despite the increasing literature on proper missing data handling, reviews in the last decade have consistently shown that problematic missing data treatments remain commonly used, with deletion techniques being the most frequent missing data treatment in counseling psychology (Schlomer, Bauman, & Card, 2010), pediatric psychology (T. D. Little, Jorgensen, Lang, & Moore, 2014), epidemiology (Eekhout, de Boer, Twisk, de Vet, & Heymans, 2012), and prevention science (Lang & Little, 2018). However, the state of missing data handling has not been reviewed in intervention studies specifically. Some missing data treatments are unique to intervention studies (e.g., worst-case imputation, see below). Furthermore, methodological

recommendations and reporting standards specific to interventions studies (e.g., Schulz, Altman, Moher, & the CONSORT Group, 2010) may influence missing data handling and reporting in that field. An accurate representation of the state of missing data reporting and handling in intervention studies can help inform these reporting guidelines and support a stronger methodology for future intervention research. Accordingly, the present article aims to foster a better understanding of how missing data is typically handled in intervention studies and how missing data treatments may be improved. First, missing data theory is summarized, followed by the review of both problematic and appropriate missing data treatments for intervention studies (for thorough discussions of missing data theory and treatments, readers can refer to Enders (2010), Graham (2012), and R. J. A. Little and Rubin (2014)). Second, we review missing data treatments in intervention studies, looking at how they have evolved over the years, what the current state of missing data handling is, and what recommendations can be made to improve missing data treatments based on this current state.

Missing Data Mechanisms

Missing data can occur under three basic mechanisms, which have different implications in terms of preventing missingness and their impact on data analyses (Enders, 2010; Seaman, Galati, Jackson, & Carlin, 2013). These three mechanisms are respectively known as

---

[1] Texas Tech University, USA
[2] North-West University, South Africa

Corresponding author:
Charlie Rioux, Department of Educational Psychology and Leadership, Texas Tech University, Box 41071, Lubbock TX, 9409-1071, USA.
Email: charlie.rioux@ttu.edu

missing completely at random (MCAR), *missing at random* (MAR), and *missing not at random* (MNAR). MCAR is a truly random process. If the missing data are MCAR, then the reason for a missing data point is unrelated to either the observed data or the unobserved data. For example, a questionnaire may be lost in the mail or a participant may have missed an assessment because of bad weather. If all missing data are MCAR, then the sample of participants without missing data would be a random subsample of the original sample. Thus, an analysis with only complete cases would provide unbiased results, but standard errors (SEs) would be larger, and power would be lower than with the original sample. With appropriate missing data treatments, results would remain unbiased, and most of the power would be recovered. Unfortunately, MCAR is the least common missing data mechanism in practice (unless a planned missing data design is implemented—see, e.g., T. D. Little and Rhemtulla (2013); Rhemtulla and Hancock (2016)). MCAR is the assumed mechanism when using traditional missing data treatments such as listwise or pair-wise deletion.

MAR is a predictable process. Thus, if the missing data are MAR, the missing data points are predicted from other variables in the observed data. For example, if the odds of dropping out are higher for participants with a lower socioeconomic status (SES), then the missingness related to attrition would be predicted by SES. MAR is the most common assumption of modern missing data treatments. If the MAR assumption is met, modern missing data treatments can provide unbiased results and recover power. Importantly, to meet the MAR assumption when implementing these missing data treatments, the variables that predict missing data points must be included as *auxiliary variables* (i.e., variables that are not examined for the objectives of the study, but that are included as predictors of missingness for the missing data treatments).

Finally, MNAR is an unpredictable process. If the missing data are MNAR, then the missing data points are predicted by unobserved variables, which can be the missing value itself, or an unobserved covariate. For example, if participants with a higher income tended to skip the question on income, and there is no other observed indicators associated with income, then there would be no way to predict the missing data points on income with observed variables. In the MNAR case, there is no way to recover the missing information, even with modern missing data treatments.

Problematic Missing Data Treatments

Deletion Techniques

Deletion techniques, as their name implies, involve deleting cases with missing data. In listwise deletion, also called complete-case analysis, any case with incomplete data is deleted. In pair-wise deletion, also called available-case analysis, cases are deleted on a variable-by-variable analysis basis; that is, for each analysis, those cases with incomplete data are deleted on a pair-wise basis (e.g., for the correlation between X_1 and Y, responses with missing data on X_1 and/or Y are deleted, respectively, while for the correlation between X_2 and Y, responses with missing data on X_2 and/or Y are deleted, respectively). Generally, the advantage of deletion techniques would be that they are convenient and the default of many statistical packages. Among the two deletion techniques, listwise deletion has the advantage of having the same sample of participants for all analyses, while pair-wise deletion has the advantage of retaining more power for the analyses (Enders, 2010).

However, for both techniques, the disadvantages are usually far greater than the advantages.

Deletion techniques make the assumption that the participants with complete data are a random subsample of the original sample. Accordingly, deletion techniques require MCAR data (Enders, 2010). If the MCAR assumption is not met, using deletion techniques is associated with biased parameter estimates (Brockmeier, Kromrey, & Hogarty, 2003; Eekhout et al., 2014; Marshall, Altman, & Holder, 2010; Nicholson, Deboeck, & Howard, 2017; Tang, Song, Belin, & Unutzer, 2005). Furthermore, no matter the missing data mechanism, deletion techniques will reduce sample size and thus reduce power. In the specific case of intervention studies, deletion techniques are also problematic for intent-to-treat analyses, which should include all participants regardless of anything that happens after randomization, including withdrawal (Gupta, 2011). In addition to these main problems, there are additional issues specific to pair-wise deletion. Indeed, the inconsistency in degrees of freedom for the different sufficient statistics can be problematic and has been associated with the occurrence of non-positive definite covariance matrices and biased SEs (Enders, 2010; Marsh, 1998). Despite the fact that these issues have been known for decades and deletion techniques have long been considered as being among the worst options for missing data treatments (Wilkinson & Task Force on Statistical Inference, 1999), previous reviews have consistently shown that the use of deletion techniques has remained prevalent (Lang & Little, 2018; T. D. Little et al., 2014; Peugh & Enders, 2004; Schlomer et al., 2010).

Regression-Based Single Imputation

Single imputation includes several techniques that consist of substituting missing values with a predicted value, resulting in one complete data set. With regression-based single imputation, missing values are filled in by predicting them from a regression equation in which the incomplete variable is the dependent variable. Deterministic regression imputation, also called conditional mean substitution, will only fill in values using the regression equation, while stochastic regression imputation will also add a normally distributed residual term for each predicted score. With deterministic regression imputation, missing values are imputed to fall directly on the regression line, without the variability that would be expected in complete data. This method underestimates variances and overestimates the magnitude of linear associations with the imputed variables, even when the missing data are MCAR (Brockmeier et al., 2003; Marshall et al., 2010; Musil, Warner, Yobas, & Jones, 2002; Olinsky, Chen, & Harlow, 2003). By adding residuals, stochastic regression imputation restores the lost variability, which will produce unbiased parameter estimates under the MAR and MCAR mechanisms. It will, however, still underestimate SEs, thus increasing Type I error rates, because the uncertainty of the imputation model is not taken into consideration (Brockmeier et al., 2003; Eekhout et al., 2014; Musil et al., 2002).

Mean Substitution

Mean substitution, also called arithmetic mean imputation and unconditional mean imputation, is a single imputation technique where missing values are replaced with the mean of the available cases. Mean substitution reduces the variability of the data and substitutes missing values with scores that are uncorrelated with

other variables in the analytic model. In turn, the magnitude of standard deviations, variances, correlations and covariances is reduced (Enders, 2010). Accordingly, mean substitution is associated with biased results under all missing data mechanisms, including MCAR (Brockmeier et al., 2003; Brown, 1994; Eekhout et al., 2014; Farhangfar, Kurgan, & Dy, 2008; Musil et al., 2002; Olinsky et al., 2003).

Last Observation Carried Forward

Last observation carried forward (LOCF) is a single imputation technique specific to longitudinal designs, which includes longitudinal intervention studies and clinical trials. When a participant drops out of the study or misses an assessment, the missing repeated measures variables are substituted with the last observed value. LOCF makes the assumptions that scores do not change when they are missing. This assumption, however, is almost always implausible and can lead to biased results. Although it is often thought that LOCF is a conservative method that would yield a smaller treatment effect than if there were no missing data, this depends on the study because LOCF can also lead to a liberal bias that exaggerates treatment effects (Cook, Zeng, & Yi, 2004; Lachin, 2016; Lane, 2008; Mallinckrodt, Clark, & David, 2001; Saha & Jones, 2009; Siddiqui, Hung, & O'Neill, 2009; Tang et al., 2005; Wood, White, Hillsdon, & Carpenter, 2005). This bias can inflate Type I and Type II error rates (Mallinckrodt et al., 2001; Saha & Jones, 2016; Siddiqui et al., 2009) and occurs even when the data is MCAR (Molenberghs et al., 2004). LOCF is one of the most egregious methods of treating missing data.

Worst-Case Imputation

Worst-case imputation is a single imputation technique specific to intervention studies where missing outcome values are imputed as the worst-case value. For example, a study testing an intervention to prevent relapse in ex-smokers may impute missing follow-up values as having relapsed. Worst-case imputation makes the assumption that dropouts have poor outcomes, which is usually more specifically based on the assumption that participants avoid follow-up because of this poor outcome (e.g., ex-smokers drop out of the study because they relapsed). Although this assumption may be true for some, or even a majority of participants depending on the study (Hajek & West, 2010), it is unlikely to hold for all dropouts and missed assessments. Indeed, many factors are associated with attrition and missed assessments in intervention studies, including adverse life events, interference with work, schedule conflicts between the research team and participant, relocation, and measurement burden (Janson, Alioto, & Boushey, 2001; R. J. Little et al., 2012). Furthermore, if several consecutive assessments are missing, worst-case imputation makes the assumptions that scores do not change (like LOCF). Because of these unrealistic assumptions, worst-case imputation is clearly unrealistic and has been shown to be associated with biased results (Blankers et al., 2016; Hedeker, Mermelstein, & Demirtas, 2007; Smolkowski, Danaher, Seeley, Kosty, & Severson, 2010; Wood et al., 2005). This bias may lead to a smaller or higher treatment effect as it tends to favor the group with the lowest proportion of missingness (Blankers et al., 2016). Like LOCF, this method is egregious with regard to the validity of any statistical inferences.

Appropriate Missing Data Treatments

Full Information Maximum Likelihood

Full information maximum likelihood (FIML), also referred to as direct maximum likelihood, is a model-based approached that deals with the missing data by estimating parameters and SEs in a single step. FIML is an extension of maximum likelihood estimation that is robust to missing data by using case-wise log-likelihoods that consider the observed variables for each case. Accordingly, FIML can only be used for data analyses for which maximum likelihood estimation is applicable. FIML has been shown to be valid and yield unbiased estimates (Cham, Reshetnyak, Rosenfeld, & Breitbart, 2017; Enders, 2001; Enders & Bandalos, 2001; Larsen, 2011). For parameter estimates to be unbiased under the MAR mechanism, however, variables that predict missingness and are not part of the main analytical model must be included as auxiliary variables (Collins, Schafer, & Kam, 2001; Graham, 2003; Howard, Rhemtulla, & Little, 2015).

Multiple Imputation

Like single imputation, multiple imputation involves substituting missing values with a predicted value, but this is done multiple times, resulting in multiple complete data sets. Analytic results after multiple imputation take into account both the variance within each imputed data set and the variance between imputed data sets. Thus, unlike stochastic regression imputation, multiple imputation provides appropriate SEs by taking into account the uncertainty of the imputation model.

Some researchers have been reluctant to use this method as they believe it "makes up" data. Multiple imputation, however, does not aim to make inferences regarding the individual scores that were imputed, but rather to correct bias introduced by missing data in the analytical model, and thus make accurate inferences. The distributions or associations in the observed data are used to estimate the missing values, and there are two main methods for generating the imputations. The first approach, joint modeling, fills in missing values simultaneously for all incomplete variables using a multivariate distribution while the second approach, fully conditional specification (also called multiple imputation by chained equations (MICE)), imputes variables one at a time with a series of conditional models (Enders, 2017; Kenward & Carpenter, 2007; Li & Stuart, 2019). While joint modeling has better theoretical underpinnings and calculations per iterations are less intensive, fully conditional specification has the advantage of being more flexible in creating multivariate models (Murray, 2018; van Buuren, 2018). Accordingly, fully conditional specification can maintain unique features of the data, such as skip patterns, bracketed responses, and bounds. While joint modeling and fully conditional specification were found to perform similarly for continuous and binary variables, fully conditional specification outperforms joint modeling for categorical variables (Kropko, Goodrich, Gelman, & Hill, 2014; Lee & Carlin, 2010).

Regardless of the approach, multiple imputation is done in two steps. First, in the imputation step, m values are estimated for each missing data point and used to create m imputed data sets. Second, in the analysis step, the analysis model is run on each of the m imputed data sets and the m sets of estimates are aggregated using Rubin's Rules (Rubin, 1987). For multiple imputation to yield unbiased parameter estimates under the MAR mechanism, auxiliary

variables must be included in the imputation step (Collins et al., 2001). Details on the implementation of multiple imputation can be found in the works of Carpenter and Kenward (2013) and van Buuren (2018).

Censoring in Time-to-Event Analyses

Censoring is specific to time-to-event analyses (i.e., survival analyses), which examine time until an event of interest occurs as their outcome. Censoring is used in time-to-event analyses when there is information about a participant not experiencing the event, but no information about when they might experience the event (Singh & Mukhopadhyay, 2011). This may occur when a participant does not experience the event by the end of the study, but also if a participant is lost to follow-up, in which case random censoring may be used by censoring participants at their missed assessments. Random censoring makes the assumption that cases with and without missing data have the same chance of experiencing the event, thus usually requiring a MCAR mechanism. Still, censoring can be used for a MAR mechanism. Indeed, if the probability of being censored is dependent on covariates, Inverse Probability of Censoring Weights can be used, which will weigh participants by the inverse of their probability of not dropping out given the included covariates (van der Laan & Robins, 2003; Willems, 2014). Multiple imputation may also be used for time-to-event analyses, in which case imputation in the wide data format is recommended since imputation in the tall data format was found to be biased and inefficient (Young & Johnson, 2015). Furthermore, imputation models for time-to-event analyses should include the dependent variable and the Nelson-Aalen estimator of the cumulative hazard to the survival time (White & Royston, 2009).

Review of Missing Data Treatments in Intervention Studies

We reviewed missing data treatments in intervention studies by searching the keyword prevent* in APA PsycNET® and restricting the search to journal articles with the methodology *clinical trials* or *treatment outcome*. The search was done on all articles up to January 30, 2019, and yielded 7941 results. Ten articles per year were selected randomly to be reviewed. When an article was not about an intervention or the full text could not be obtained, another one was selected randomly within the same year as replacement. When there were fewer than 10 articles for a given year, they were all included. This process yielded a total sample of 291 studies (1966–1989, $n = 13$; 1990–1994, $n = 29$; 1995–1999, $n = 49$; 2000–2004, $n = 50$; 2005–2009, $n = 50$; 2010–2014 $n = 50$; 2015–2019 $n = 50$). All studies were coded for missing data treatments (i.e., methods used to treat the missing data). Studies from 2015 to 2019 were also coded for missing data reporting (i.e., acknowledgement of missing data, nonresponse rates reported, explicit mention, and testing of missing data mechanisms).

The prevalence of missing data treatments across years was examined, excluding studies that did not have any missing data or did not provide information on their missing data treatments. Missing data treatments included deletion techniques, LOCF, worst-case imputation, single imputation (regression-based—also included studies that report using imputation without further details), multiple imputation, and model-based approaches (i.e., FIML, which comprised 93% of included studies, while 7% used censoring in time-to-event analyses). Two missing data treatments were each used in only one study and were thus not included in the review below; namely, mean substitution (1 study, 1990) and using the individual's mean from previous observations to replace missing values (1 study, 2004).

What was

Results showed historical patterns in the treatment of missing data (see Figure 1). Deletion was the most prevalent missing data treatment across years, except for recent years when model-based approaches became more prevalent (note that if studies that do not provide information on their missing data treatments are considered as having used deletion, it would be the most prevalent across all years). Still, the use of deletion techniques decreased over time, from 89% of studies before 1990 to 38% of studies in 2015–2019. The second most prevalent missing data treatment until 2010 was LOCF, for which the use remained relatively stable except for an increase in the years 2000–2004. In the studies reviewed, worst-case imputation began in the early 2000s, increasing thereafter, but decreasing in recent years. This method was mostly used in studies examining smoking. Single imputation had a very low prevalence throughout all years.

The modern recommended missing data treatments increased in prevalence over the years. Model-based approaches were first used several years before multiple imputation (1995 vs. 2005 in this review). The use of model-based approaches was relatively stable from 1995 to 2014, but increased substantially in recent years. This recent increase is most likely due to the recent increase in the availability and use of structural equation modeling programs and the ease of implementation of FIML in these programs (note that all studies reviewed that used model-based approaches in 2015–2019 used FIML, with no time-to-event study using random censoring). After its first occurrence in this review in 2005, the use of multiple imputation increased steadily.

What is

Studies from 2015 to 2019 were reviewed in more detail regarding their missing data reporting; 96% of studies mentioned missing data. Among those studies, 79% reported the rate of missing data, which among those that did have missing data ranged from 1% to 28%. Furthermore, among studies that mentioned missing data, 15% did not report their missing data treatment, 38% used FIML, 31% used deletion, 19% used multiple imputation, 15% used LOCF, 2% used worst-case imputation, and 4% had no missing data (note that some studies used multiple missing data treatments, thus the count exceeds 100%). Among studies that reported a missing data treatment, only 12% ($n = 5$) mentioned missing data mechanisms. Note that all five studies mentioning missing data mechanisms were from 2018 and 2019. One of these studies assumed a MNAR mechanism; one assumed a MAR mechanism but did not mention auxiliary variables; one assumed a MAR mechanism and mentioned the auxiliary variables included in the imputation model, but did not mention testing whether these auxiliary variables predicted missing data points; one study conducted Little's MCAR test (R. J. A. Little, 1988) but did not mention or test the MAR mechanism, even for variables for which the test showed data was not MCAR; and one study did both Little's MCAR test
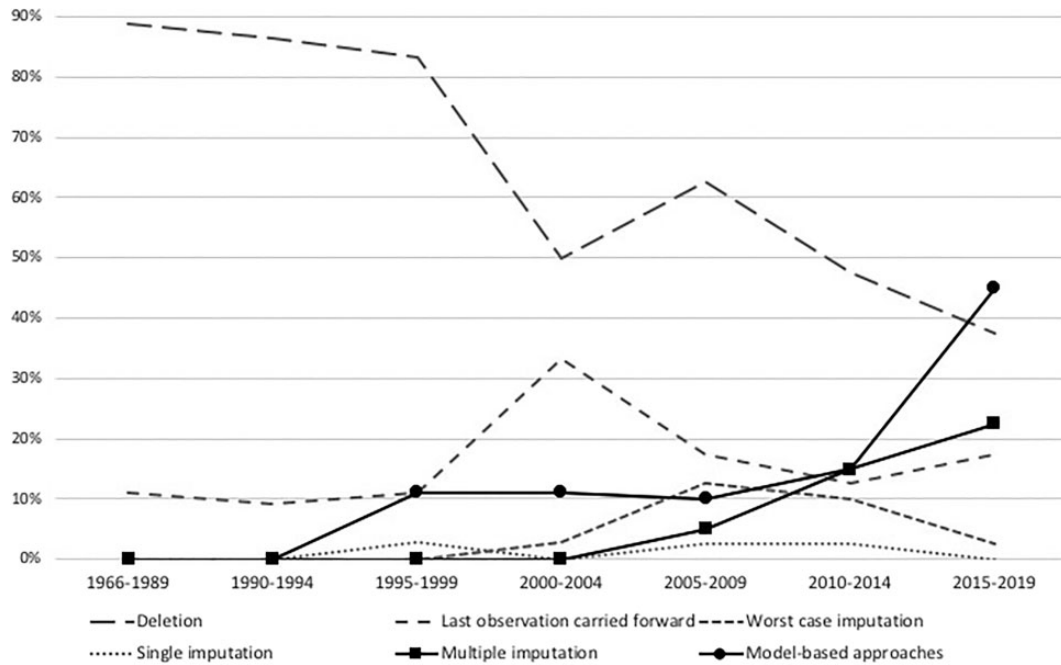
Figure 1. Prevalence of Missing Data Treatments Over the Years. Prevalence Based on Studies With Missing Data That Reported Their Missing Data treatment.
Note: 1966–1989, *n* = 9; 1990–1994, *n* = 22; 1995–1999, *n* = 36; 2000–2004, *n* = 36; 2005–2009, *n* = 40; 2010–2014 *n* = 39; 2015–2019 *n* = 41.

and *t*-tests to identify variables associated with the MAR mechanism.

The missing data treatments used in these intervention studies are relatively similar to rates found in recent prevention science articles (Lang & Little, 2018), although LOCF was used considerably more in the intervention studies than in the general prevention science literature (15% vs. 1%). Rates for recommended missing data treatments were in a similar range in both reviews, although they were slightly higher in the intervention studies (multiple imputation 19% vs. 15%; FIML 38% vs. 32%)[1]. Furthermore, missing data were mentioned and nonresponse rates were reported slightly more frequently in the intervention literature (missing data mentioned 96% vs. 84%; nonresponse rates provided 79% vs. 73%), which may be due in part to reporting guidelines, such as the CONSORT statement (Schulz et al., 2010), that are followed by many journals and require nonresponse rates to be reported in clinical trials.

What should be

Although this historical review showed that missing data treatments in intervention studies have improved over the years, the more detailed review of studies from recent years showed that there is still much room for improvement. Although the majority of studies included basic missing data reporting, many studies did not report their missing data rates and missing data treatments, which should be reported in all published studies. Multiple imputation and FIML have become more prevalent, but there are still many studies using antiquated techniques, mostly deletion and LOCF. As previously discussed, deletion techniques lead to biased results unless all missing data are MCAR (in addition to reducing power), and LOCF is associated with biased results under all missingness mechanisms, even MCAR. Using appropriate missing data treatments is

imperative to obtain unbiased results and make proper inferences regarding the effect of an intervention.

Still, the majority of studies implementing multiple imputation and FIML did not explicitly support the assumptions of these missing data treatments. As previously discussed, multiple imputation and FIML provide unbiased results under the MAR mechanism only if the auxiliary variables are included in the missing data treatment. The majority of studies implementing these missing data treatments did not mention missing data mechanisms or auxiliary variables. Our review did suggest that error of omission is currently improving since studies mentioning missing data mechanisms were the most recent ones. Among studies that did mention missing data mechanisms, however, few provided details on testing the mechanisms and on the inclusion of auxiliary variables. Potential correlates of missing data should be examined (see Nicholson et al., 2017) and included as auxiliary variables, and these details should be reported when publishing studies. With the vast majority of journals now accepting online supplementary materials, there are no space restrictions to properly reporting these analyses. Reporting practices for imputation are readily available (see, e.g., Manly & Wells, 2015; Sterne et al., 2009). Articles should at a minimum report the rates, reasons, and mechanisms for missing data; state the missing data handling procedure; and describe all decisions made when handling the missing data. A few aspects that should be considered for proper missing data handling but are less often discussed are reviewed next.

The level of aggregation of the data for multi-item scales before handling missing data should be taken into consideration and reported. If multiple imputation is used, item-level imputation is associated with higher precision and power than imputing the composite score (Eekhout et al., 2014; Gottschall, West, & Enders, 2012). For a data set with many scales, where there may be too many items to impute, parcel summary scores have been shown to be a reliable option (Eekhout, de Vet, de Boer, Twisk, & Heymans,

2018). With this method, the items from one scale are imputed using the average of the available items on other scales as auxiliary variables. Another option when the number of variables is high compared to the number of observations is the use of principal component auxiliary variables (Howard et al., 2015), which can be implemented using the R Package PcAux Beta (Lang, Little, & PcAux Team, 2018). With this method, principal component scores that capture information from the items are extracted and act as auxiliary variables in the imputation model instead of the original items. If FIML is used to handle missing data, entering scales with missingness without further information for missing data handling was shown to provide highly biased parameters, but this bias was eliminated if the individual items were included as auxiliary variables. As too many auxiliary variables can lead to convergence issues, using an average of complete items with the individual incomplete items as auxiliary variables is also a valid appropriate method that is equivalent to item-level imputation in terms of bias and power (Mazza, Enders, & Ruehlman, 2015).

The compatibility of the imputation model and the analysis model should be ensured and explicitly reported, especially if any transformed variables are included in the analysis model. Indeed, if transformed variables, such as power terms and interaction terms, are included in the analysis model, they should be included in the imputation model as well for unbiased results (von Hippel, 2009). Failing to do so biases associations with these transformed variables toward zero. In line with the requirement for compatibility, multilevel data should also be taken into consideration in the imputation phase in addition to the main analyses (Drechsler, 2015).

Researchers may also perform checks of their imputation models to examine how the main analysis results may be affected by modeling decisions in the imputation phase. Any imputation checking efforts and their results should be reported. These checks may notably include examining the distribution of imputed values, graphical comparisons between the observed and imputed data, examining the regression goodness-of-fit of the imputation model, assessing the predictive validity of the imputation model, and posterior predictive checking (see Nguyen, Carlin, & Lee, 2017).

For optimal missing data handling, missing data should be an important consideration in the design stage of a study. Notably, potential predictors of missingness, and thus of the MAR mechanism, should be anticipated from previous studies with similar designs, populations, and/or measures, and these predictors should be included in the questionnaire protocol. Furthermore, even though missing data can be handled in the analysis stage, precautions should be implemented to prevent missed assessments and dropout and thus keep the missingness rate at a minimum (de Leeuw, 2001; Dziura et al., 2013; R. J. Little et al., 2012; Wisniewski et al., 2006). One possibility in the design stage is also to implement a *planned missing data design* (T. D. Little & Rhemtulla, 2013; Rhemtulla & Hancock, 2016), which consists in randomly assigning participants to having missing items (i.e., multiform design), missing measurement occasions (i.e., wave-missing design), or missing measures (i.e., two-method design). Because these missing data are introduced randomly in the research design by the researcher, they are MCAR and thus by using a modern missing data treatment, the lost power is recovered without any added bias. These designs have many advantages, notably reducing data collection costs and participant burden. The reduction in participant fatigue is associated with more valid results, stronger effects, and reduced unplanned missingness (Harel, Stratton, & Aseltine, 2015).

Conclusion

Missing data treatments in intervention studies have improved in recent years, but many studies still fail to properly report missing data rates and handling. Furthermore, inappropriate missing data treatments that have been unequivocally demonstrated to be associated with biased results are still prevalent. Missing data analysis is an integral part of proper statistical inference and should be given as much thought as substantive design and analyses. Research designs should anticipate missing data and modern missing data treatments should be used to obtain the most accurate results when faced with missing data. Preventive interventions are a key component of prevention science and must be conducted with methodological rigor to yield accurate results that will properly guide practice and policy (T. D. Little, 2015). Proper missing data treatments are an essential component of correctly quantifying the effect of preventive interventions and properly informing clinicians and other stakeholders.

ORCID iD

Charlie Rioux https://orcid.org/0000-0002-7943-0710
Todd D. Little https://orcid.org/0000-0002-4146-4712

Note

1. Percentage of studies using the missing data treatments is based on studies that mentioned missing data.

References

Blankers, M., Smit, E. S., van der Pol, P., de Vries, H., Hoving, C., & van Laar, M. (2016). The missing=smoking assumption: A fallacy in internet-based smoking cessation trials? *Nicotine & Tobacco Research*, 18, 25–33. doi:10.1093/ntr/ntv055

Brockmeier, L. L., Kromrey, J. D., & Hogarty, K. Y. (2003). Nonrandomly missing data in multiple regression analysis: An empirical comparison of ten missing data treatments. *Multiple Linear Regression Viewpoints*, 29, 8–29.

Brown, R. L. (1994). Efficacy of the indirect approach for estimating structural equation models with missing data: A comparison of five methods. *Structural Equation Modeling-a Multidisciplinary Journal*, 1, 287–316. doi:10.1080/10705519409539983

Carpenter, J. R., & Kenward, M. G. (2013). *Multiple imputation and its application*. Chichester, UK: Wiley.

Cham, H., Reshetnyak, E., Rosenfeld, B., & Breitbart, W. (2017). Full information maximum likelihood estimation for latent variable interactions with incomplete indicators. *Multivariate Behavioral Research*, 52, 12–30. doi:10.1080/00273171.2016.1245600

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6, 330–351. doi:10.1037// 1082-989x.6.4.330

Cook, R. J., Zeng, L. L., & Yi, G. Y. (2004). Marginal analysis of incomplete longitudinal binary data: A cautionary note on LOCF imputation. *Biometrics*, 60, 820–828. doi:10.1111/j.0006-341X. 2004.00234.x

de Leeuw, E. D. (2001). Reducing missing data in surveys: An overview of methods. *Quality & Quantity*, 35, 147–160. doi:10.1023/a:1010395805406

Drechsler, J. (2015). Multiple imputation of multilevel missing data-rigor versus simplicity. *Journal of Educational and Behavioral Statistics*, 40, 69–95. doi:10.3102/1076998614563393

Dziura, J. D., Post, L. A., Zhao, Q., Fu, Z., & Peduzzi, P. (2013). Strategies for dealing with missing data in clinical trials: From design to analysis. *The Yale Journal of Biology and Medicine*, 86, 343.

Eekhout, I., de Boer, M. R., Twisk, J. W. R., de Vet, H. C. W., & Heymans, M. W. (2012). Missing data a systematic review of how they are reported and handled. *Epidemiology*, 23, 729–732. doi:10.1097/EDE.0b013e3182576cdb

Eekhout, I., de Vet, H. C. W., de Boer, M. R., Twisk, J. W. R., & Heymans, M. W. (2018). Passive imputation and parcel summaries are both valid to handle missing items in studies with many multi-item scales. *Statistical Methods in Medical Research*, 27, 1128–1140. doi:10.1177/0962280216654511

Eekhout, I., de Vet, H. C. W., Twisk, J. W. R., Brand, J. P. L., de Boer, M. R., & Heymans, M. W. (2014). Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. *Journal of Clinical Epidemiology*, 67, 335–342. doi:10.1016/j.jclinepi.2013.09.009

Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement*, 61, 713–740. doi:10.1177/00131640121971482

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

Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour Research and Therapy*, 98, 4–18. doi:10.1016/j.brat.2016.11.008

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling-a Multidisciplinary Journal*, 8, 430–457. doi:10.1207/s15328007sem0803_5

Farhangfar, A., Kurgan, L., & Dy, J. (2008). Impact of imputation of missing values on classification error for discrete data. *Pattern Recognition*, 41, 3692–3705. doi:10.1016/j.patcog.2008.05.019

Gottschall, A. C., West, S. G., & Enders, C. K. (2012). A comparison of item-level and scale-level multiple imputation for questionnaire batteries. *Multivariate Behavioral Research*, 47, 1–25. doi:10.1080/00273171.2012.640589

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling*, 10, 80–100. doi:10.1207/s15328007sem1001_4

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

Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research*, 2, 109–112. doi:10.4103/2229-3485.83221

Hajek, P., & West, R. (2010). Commentary on Smolkowski, et al. (2010): Why is it important to assume that non-responders in tobacco cessation trials have relapsed? *Addiction*, 105, 1016–1017. doi:10.1111/j.1360-0443.2010.02964.x

Harel, O., Stratton, J., & Aseltine, R. (2015). Designed missingness to better estimate efficacy of behavioral studies-application to suicide prevention trials. *Journal of Medical Statistics and Informatics*, 3, 2. doi:10.7243/2053-7662-3-2

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2007). Analysis of binary outcomes with missing data: Missing = smoking, last observation carried forward, and a little multiple imputation. *Addiction*, 102, 1564–1573. doi:10.1111/j.1360-0443.2007.01946.x

Howard, W. J., Rhemtulla, M., & Little, T. D. (2015). Using principal components as auxiliary variables in missing data estimation. *Multivariate Behavioral Research*, 50, 285–299. doi:10.1080/00273171.2014.999267

Janson, S. L., Alioto, M. E., & Boushey, H. A. (2001). Attrition and retention of ethnically diverse subjects in a multicenter randomized controlled research trial. *Controlled Clinical Trials*, 22, 236S–243S. doi:10.1016/s0197-2456(01)00171-4

Kenward, M. G., & Carpenter, J. (2007). Multiple imputation: Current perspectives. *Statistical Methods in Medical Research*, 16, 199–218. doi:10.1177/0962280206075304

Kropko, J., Goodrich, B., Gelman, A., & Hill, J. (2014). Multiple imputation for continuous and categorical data: Comparing joint multivariate normal and conditional approaches. *Political Analysis*, 22, 497–519. doi:10.1093/pan/mpu007

Lachin, J. M. (2016). Fallacies of last observation carried forward analyses. *Clinical Trials*, 13, 161–168. doi:10.1177/1740774515602688

Lane, P. (2008). Handling drop-out in longitudinal clinical trials: A comparison of the LOCF and MMRM approaches. *Pharmaceutical Statistics*, 7, 93–106. doi:10.1002/pst.267

Lang, K. M., & Little, T. D. (2018). Principled missing data treatments. *Prevention Science*, 19, 284–294. doi:10.1007/s11121-016-0644-5

Lang, K. M., Little, T. D., & PcAux Team, P. D. (2018). *PcAux: Automatically eextract auxiliary features for simple, principled missing data analysis* [R package]. Retrieved from https://github.com/PcAux-Package/PcAux/

Larsen, R. (2011). Missing data imputation versus full information maximum likelihood with second-level dependencies. *Structural Equation Modeling-a Multidisciplinary Journal*, 18, 649–662. doi:10.1080/10705511.2011.607721

Lee, K. J., & Carlin, J. B. (2010). Multiple imputation for missing data: Fully conditional specification versus multivariate normal imputation. *American Journal of Epidemiology*, 171, 624–632. doi:10.1093/aje/kwp425

Li, P., & Stuart, E. A. (2019). Best (but oft-forgotten) practices: Missing data methods in randomized controlled nutrition trials. *The American Journal of Clinical Nutrition*, 109, 504–508. doi:10.1093/ajcn/nqy271

Little, R. J., D'Agostino, R., Cohen, M. L., Dickersin, K., Emerson, S. S., Farrar, J. T., . . . Stern, H. (2012). The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine*, 367, 1355–1360. doi:10.1056/NEJMsr1203730

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83, 1198–1202. doi:10.2307/2290157

Little, R. J. A., & Rubin, D. B. (2014). *Statistical analysis with missing data*. Hoboken, NJ: John Wiley & Sons.

Little, T. D. (2015). Methodological practice as matters of justice, justification, and the pursuit of verisimilitude. *Research in Human Development*, 12, 268–273. doi:10.1080/15427609.2015.1068044

Little, T. D., Jorgensen, T. D., Lang, K. M., & Moore, E. W. G. (2014). On the joys of missing data. *Journal of Pediatric Psychology*, 39, 151–162. doi:10.1093/jpepsy/jst048

Little, T. D., & Rhemtulla, M. (2013). Planned missing data designs for developmental researchers. *Child Development Perspectives*, 7, 199–204. doi:10.1111/cdep.12043

Mallinckrodt, C. H., Clark, W. S., & David, S. R. (2001). Type I error rates from mixed effects model repeated measures versus fixed effects anova with missing values imputed via last observation carried forward. *Drug Information Journal*, 35, 1215–1225. doi:10.1177/009286150103500418

Manly, C. A., & Wells, R. S. (2015). Reporting the use of multiple imputation for missing data in higher education research. *Journal Research in Higher Education*, 56, 397–409. doi:10.1007/s11162-014-9344-9

Marsh, H. W. (1998). Pairwise deletion for missing data in structural equation models: Nonpositive definite matrices, parameter estimates, goodness of fit, and adjusted sample sizes. *Structural Equation Modeling-a Multidisciplinary Journal*, 5, 22–36. doi:10.1080/10705519809540087

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, 10. doi:10.1186/1471-2288-10-112

Mazza, G. L., Enders, C. K., & Ruehlman, L. S. (2015). Addressing item-level missing data: A comparison of proration and full information maximum likelihood estimation. *Multivariate Behavioral Research*, 50, 504–519. doi:10.1080/00273171.2015.1068157

Molenberghs, G., Thijs, H., Jansen, I., Beunckens, C., Kenward, M. G., Mallinckrodt, C., & Carroll, R. J. (2004). Analyzing incomplete longitudinal clinical trial data. *Biostatistics*, 5, 445–464. doi:10.1093/biostatistics/kxh001

Murray, J. S. (2018). Multiple imputation: A review of practical and theoretical findings. *Statistical Science*, 33, 142–159. doi:10.1214/18-sts644

Musil, C. M., Warner, C. B., Yobas, P. K., & Jones, S. L. (2002). A comparison of imputation techniques for handling missing data. *Western Journal of Nursing Research*, 24, 815–829. doi:10.1177/019394502237390

Nguyen, C. D., Carlin, J. B., & Lee, K. J. (2017). Model checking in multiple imputation: An overview and case study. *Emerging Themes in Epidemiology*, 14, 12. doi:10.1186/s12982-017-0062-6

Nicholson, J. S., Deboeck, P. R., & Howard, W. (2017). Attrition in developmental psychology: A review of modern missing data reporting and practices. *International Journal of Behavioral Development*, 41, 143–153. doi:10.1177/0165025415618275

Olinsky, A., Chen, S., & Harlow, L. (2003). The comparative efficacy of imputation methods for missing data in structural equation modeling. *European Journal of Operational Research*, 151, 53–79. doi:10.1016/s0377-2217(02)00578-7

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, 74, 525–556. doi:10.3102/00346543074004525

Rhemtulla, M., & Hancock, G. R. (2016). Planned missing data designs in educational psychology research. *Educational Psychologist*, 51, 305–316. doi:10.1080/00461520.2016.1208094

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

Saha, C., & Jones, M. P. (2009). Bias in the last observation carried forward method under informative dropout. *Journal of Statistical Planning and Inference*, 139, 246–255. doi:10.1016/j.jspi.2008.04.017

Saha, C., & Jones, M. P. (2016). Type I and Type II error rates in the last observation carried forward method under informative dropout.

Journal of Applied Statistics, 43, 336–350. doi:10.1080/02664763.2015.1063112

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology*, 57, 1–10. doi:10.1037/a0018082

Schulz, K. F., Altman, D. G., & Moher, D., & the CONSORT Group. (2010). CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *Bmc Medicine*, 8, 9. doi:10.1186/1741-7015-8-18

Seaman, S., Galati, J., Jackson, D., & Carlin, J. (2013). What is meant by "missing at random"? *Statistical Science*, 28, 257–268. doi:10.1214/13-sts415

Siddiqui, O., Hung, H. M. J., & O'Neill, R. (2009). MMRM vs. LOCF: A comprehensive comparison based on simulation study and 25 NDA datasets. *Journal of Biopharmaceutical Statistics*, 19, 227–246. doi:10.1080/10543400802609797

Singh, R., & Mukhopadhyay, K. (2011). Survival analysis in clinical trials: Basics and must know areas. *Perspectives in Clinical Research*, 2, 145–148. doi:10.4103/2229-3485.86872

Smolkowski, K., Danaher, B. G., Seeley, J. R., Kosty, D. B., & Severson, H. H. (2010). Modeling missing binary outcome data in a successful web-based smokeless tobacco cessation program. *Addiction*, 105, 1005–1015. doi:10.1111/j.1360-0443.2009.02896.x

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., ... Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *British Medical Journal*, 339, 9. doi:10.1136/bmj.b2393

Tang, L. Q., Song, J. W., Belin, T. R., & Unutzer, J. (2005). A comparison of imputation methods in a longitudinal randomized clinical trial. *Statistics in Medicine*, 24, 2111–2128. doi:10.1002/sim.2099

van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). Boca Raton, FL: CRC Press.

van der Laan, M. J., & Robins, J. M. (2003). *Unified methods for censored longitudinal data and causality*. New York, NY: Springer.

von Hippel, P. T. (2009). How to impute interactions, squares, and other transformed variables. In Y. Xie (Ed.), *Sociological methodology* (Vol. 39, pp. 265–291). Malden, MA: Wiley-Blackwell.

White, I. R., & Royston, P. (2009). Imputing missing covariate values for the Cox model. *Statistics in Medicine*, 28, 1982–1998. doi:10.1002/sim.3618

Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals – Guidelines and explanations. *American Psychologist*, 54, 594–604. doi:10.1037//0003-066x.54.8.594

Willems, S. J. W. (2014). *Inverse probability censoring weights for routine outcome monitoring data*. Leiden, The Netherlands: Universiteit Leiden.

Wisniewski, S. R., Leon, A. C., Otto, M. W., & Trivedi, M. H. (2006). Prevention of missing data in clinical research studies. *Biological Psychiatry*, 59, 997–1000. doi:10.1016/j.biopsych.2006.01.017

Wood, A. M., White, I. R., Hillsdon, M., & Carpenter, J. (2005). Comparison of imputation and modelling methods in the analysis of a physical activity trial with missing outcomes. *International Journal of Epidemiology*, 34, 89–99.

Young, R., & Johnson, D. R. (2015). Handling missing values in longitudinal panel data with multiple imputation. *Journal of Marriage and Family*, 77, 277–294. doi:10.1111/jomf.12144