



On comparative performance of multiple imputation methods for moderate to large proportions of missing data in clinical trials: a simulation study

Sukhdev Mishra^{1*} and Diwakar Khare²

*Correspondence: Mishra.Sukhdev@gmail.com



CrossMark

← Click for updates

¹Division of Bio-Statistics, National Institute of Occupational Health, Meghani Nagar, Ahmedabad, India.

²Department of Statistics, Institute of Social Sciences, Dr. B. R. Ambedkar University, Agra, India.

Abstract

Background: Longitudinal clinical trial has measurements at successive occasions, and unavailability of patient at a scheduled visit causes missingness in expected full sequence of the measurements. Missing data is a major concern during the conduct of a clinical trial. It has been noted that missing data are not handled properly during final analysis which may considerably bias the results of analysis, reduce the power of the study and lead to invalid conclusions. A promising approach to handle this problem is to impute the missing values.

Methods: Multiple imputation (MI) methods provide a useful strategy to deal with the data sets with missing values, where missing values are filled in by estimate and the resulting data sets are analyzed by complete data methods. Statistical methods to address missingness have been actively pursued in recent years. This paper has attempted to provide a description of missing data mechanism and various imputation techniques for missing data analysis in longitudinal clinical trials. Further, the appropriateness of multiple imputation methods has been discussed under moderate to large proportion of missingness in a simulated clinical trial data, by comparing the various performance measures derived through intensive simulation procedure.

Results: For moderate proportion (~20 & 30%) of missingness MI-regression method scored minimum bias and MSE with increase in the sample size. However, other methods did not improve much despite increased sample size. For large proportion (50%) of missing data, MI-regression and MI-propensity score methods were close in performance but MI-regression method performed significantly well with increased number of subjects in the dataset.

Conclusions: Present investigation showed that MI-regression method is most appropriate for the analysis of data in presence of missingness with discussed sample size and missingness mechanism. Overall, the study findings will help researchers having limited knowledge of statistical methodology to choose a multiple imputation method accordingly, so that achieved estimates will be more precised.

Keywords: Missing data, missing mechanism, longitudinal data, multiple imputation

Introduction

Missing data is a major concern during conduct of a clinical trial. Dropout from the study before protocol specified completion time has been one of the major reason of the missingness. Withdrawal reason may be related to adverse event, lack of drug effectiveness, unpleasant study procedures *etc.* There could be other reasons which are not associated with the study, e.g., moving away from study place, external engagements

of patients. Missingness is more prevalent in clinical trials where a slow acting or intolerable treatment drug is being investigated [1]. Unfortunately missing data are not handled properly during final analysis which considerably bias the results of the analysis, reduce the power of the study and lead to invalid conclusions. Even the valid statistical analysis of a longitudinal study may not be of help if missingness is related to the outcome measure [2]. One obvious strategy to address

missing data during analyses is to include only completers, i.e., the subjects whose complete data are available on variable of analysis interest. However, interpretation from such research may be biased. An attractive approach to handle this problem is to impute the missing values. Imputation methods provide a useful strategy to deal with data sets with missing values, where missing values are filled in by imputed values and the resulting each completed data set is analyzed by standard methods. Numerous methods for handling missing data have been proposed, examined and implemented [3]. In fact, there are so many methods that choosing a suitable method can be difficult. In the presence of missingness, there is no universally accepted best strategy to analyze the data. Hence, it is quintessential to study missingness mechanism first, thereafter, the strategy for analysis should be decided. A description of mechanism for missingness is given in the next section. Simplest imputation technique involves replacing of value with observed data e.g., baseline value, average of variable and last value carried forward [4]. Multiple imputation is considered one of the most useful technique to deal with missing data and it is readily available in modern statistical softwares [5,6]. Bias in the results may vary based on the proportion of missingness in the dataset, hence, the results of imputation methods are of interests. This article reviews theory of missing data mechanisms and various multiple imputation methods applicable to missing data analysis in longitudinal clinical trial data and discusses the results based on findings of multiple imputation procedures applied to a simulated diabetic clinical trial data.

Missing data mechanism

Longitudinal clinical trial has measurements at successive occasions and unavailability of patient causes missingness in expected full sequence of measurements. The missing data may affect statistical estimates profoundly during analysis. Hence, it is important to assess the impact of missing data on subsequent statistical inference [7]. The probable impact of missing data on statistical analysis is understood by underlying mechanism that causes missing data. Nomenclature for missing data mechanism is defined as follows by Rubin [8]. Data are considered missing completely at random (MCAR), if the probability that an observation is missing does not depend on either the observed or unobserved value in the dataset. Data are missing at random (MAR), if the probability that an observation is missing depends on observed values in dataset but not on unobserved values. Data are missing not at random (MNAR), if the missingness probability depends on unobserved or missing values in the dataset.

The following examples illustrate the mechanism of missing data that may arise in certain situation. For example, if a patient shifted to another city which is located at large from clinical investigation site, so the patient dropouts from the study. Since dropout is no way related to the outcome of interest, the missingness will classify under MCAR mechanism.

In another situation, if a patient's condition is not improving on investigative medicine and patient decides to withdraw from the study; we note that this dropout is related to the outcome and related to the observed data till now, so the missingness mechanism in this case is MAR [2]. If a patient lost to follow up because after the last visit to the hospital the patient's condition got worsened at home and he sought medical treatment somewhere else. As this dropout is related to the outcome of interest but previously observed data do not explain this dropout and missingness is dependent on unobserved data, hence, the missingness is due to MNAR.

Most of the times data are not MCAR, as missingness is related to the outcome of interest. In well designed studies, such as clinical trials, the MAR assumption is much more prevalent than the MCAR assumption, because the observed data explain much of the missingness in most situations [9,10].

Methods

Multiple imputation

Imputation refers to fill in missing values of a data set with their plausible values. These plausible values are sampled from the predictive distribution of missing data. The selection procedure of plausible values is what differentiates the various imputation methods. The objective of any imputation method is to produce a complete dataset which can then be analyzed using standard statistical methods. Imputation methods have been used frequently in the last few years to deal with missing data in the clinical trials. Imputation methods that have been put under single and multiple imputation methods. Single imputation refers to imputing one plausible value for each missing value for a particular variable in a data set and then performing analysis as if all data were originally observed. The idea of multiple imputation was proposed by Rubin [11]. In multiple imputation (MI), instead of replacing a single value for each missing observation, we substitute it with multiple ($m > 1$) plausible values to reflect the uncertainty about the right value to impute. Rubin (1987) showed that efficiency of an estimate based on imputations can be approximated by following expression:

$$(1 + \gamma / m)^{-1}$$

where γ is the fraction of missing information for variable being estimated.

It is important to decide how many imputed data sets should be created. Rubin also stated that many may be surprised by the claim that only 3-10 imputations may be needed [11]. Unless missing information is unusually high there is no practical benefit of using more than five imputations [12]. The multiple imputations is motivated by Bayesian framework. To illustrate, let Q denote the scalar quantity to be estimated such as a mean, regression coefficient etc. Let Y_{miss} denote the missing data and Y_{obs} denote the observed data. Thus, the complete data set is (Y_{obs}, Y_{miss}) . Let Y follows the distribution $p(Y|\theta)$, where θ is the collection of all the

parameters of the model.

Let $\hat{Q} = \hat{Q}(Y_{obs}, Y_{miss})$ denote the imputed data estimate of Q and $\text{var}(\hat{Q}) = \text{var}(\hat{Q}(Y_{obs}, Y_{miss}))$ denote the estimated

variance of Q. It involves two steps for imputing Y_{miss} . First, we simulate a parameter value from the observed data posterior $p(\theta | Y_{obs})$. Second, we simulate a missing data vector from the conditional posterior distribution $p(Y_{miss} | Y_{obs}, \theta)$ using the value q generated in the first step. So, m values are imputed for each data set with missing values and $m > 1$ independent simulated imputed data sets $(Y_{obs}, \hat{Y}_{miss}^{(1)}), (Y_{obs}, \hat{Y}_{miss}^{(2)}), \dots, (Y_{obs}, \hat{Y}_{miss}^{(m)})$ are produced.

Now statistical analysis is carried out on each imputed data sets separately, using standard procedures for complete data. The analysis enables us to calculate each of the imputed-data estimates $\hat{Q}^{(t)} = \hat{Q}(Y_{obs}, \hat{Y}_{miss}^{(t)})$ along with their estimated variance

$$\hat{\text{var}}(\hat{Q}^{(t)}) = \hat{\text{var}}[\hat{Q}(Y_{obs}, \hat{Y}_{miss}^{(t)})]$$

Once the analysis is completed for each imputed data set, next step is to combine these results using imputed-data estimates. As per Rubin [11], the point estimate for Q is simply the average:

$$\hat{Q}_{MI} = \frac{1}{m} \sum_{t=1}^m \hat{Q}^{(t)}$$

The total variance of \hat{Q}_{MI} consists of two parts: the “between-imputation” variance and the “within-imputation” variance. The *between-imputation* and *within-imputation* variance are given as per below, respectively.

$$B = \frac{1}{m-1} \sum_{t=1}^m (\hat{Q}^{(t)} - \hat{Q}_{MI})^2$$

$$W = \frac{1}{m} \sum_{t=1}^m \hat{\text{var}}(\hat{Q}^{(t)})$$

Where $\hat{\text{var}}(\hat{Q}^{(t)})$ is an estimated variance on imputed data set $t, t = 1, 2, \dots, m$.

The estimated total variance is as follows:

$$T = (1 + \frac{1}{m})B + W$$

From the above, it is evident that multiple imputation techniques are advantageous over other methods, since statistical estimates after analysis of imputed datasets reflect the uncertainty or variability due to imputation [13]. Further, we have explored the performance of multiple imputation methods under varying sample size and proportion of missingness. The important multiple imputation methods are considered that can be implemented through MI procedure in SAS software: regression method, markov chain monte carlo(MCMC) and propensity score method. We also considered widely used

single imputation method called last observation carried forward (LOCF), where last observed value is carried forward for missing values [14].

Data simulation and procedure implementation

To assess the performance of various imputation methods we have considered a simulated diabetic clinical trial data under MAR dropout mechanism. Since, observed data in clinical trials explain much of information, the MAR assumption is relevant in many scenarios [2,9,15]. Simulated variable is Glycosylated Hemoglobin (HbA1c), which is an indicator of blood glucose level used in diabetes management. The data sets are simulated to resemble the time behavior of HbA1c in the diabetes clinical trial. Simulation was done using statistical analysis system SAS 9.1.3. If intervention is a slow acting treatment and patient dropouts with a perception of non-effectiveness of drug, then observed data have information for missingness, the mechanism can be assumed as missing at random. We assume that under MAR mechanism, missingness can appear at any visit to any subject but with different probability. At a visit, the probability of missingness is 3 times in the subjects where the measurement of HbA1c falls down the mean value of the visit compared to the subjects whose measurements are equal or higher than visit mean value. For example, the mean HbA1c at visit 3 is 9. Therefore, Probability [The observation is missing at visit 3|Observed value < 9] = 3 * Probability [The observation is missing at visit 3|Observed value ≥ 9].

In order to check the performance of various multiple imputation methods in different scenarios, we created missingness in the complete datasets from moderate (~20%, 30 %) to high proportion (~50%) for varied sample sizes (n=50, 75, 120). If missingness occurs at any post baseline visit for a subject, then it was assumed that all the successive post baseline visits have missing observations. Next, to impute the missingness through multiple imputation methods, procedure *proc mi* of SAS software was employed. For each imputation procedure a SAS macro was developed to impute missing values in data set. In order of computational steps, first a complete data set was generated through simulation procedure and then mean and standard deviation were calculated. Thereafter, varied proportion of missingness was created in the dataset under MAR mechanism for assumed proportion of missingness. Further, this missing data set was imputed using various studied imputation methods and descriptive statistics were calculated for the last visit of patient, further called as estimated mean and standard deviation. Incomplete datasets were imputed five times [12]. Statistical measures to assess the performance of methods were bias (absolute) and mean square error (MSE) of estimated mean and standard deviation, as defined below:

$$\text{Bias}(\hat{\theta}) = E |\hat{\theta} - \theta| \quad \text{and} \\ \text{MSE}(\hat{\theta}) = \text{Var}(\hat{\theta}) + (\text{Bias}(\hat{\theta}, \theta))^2$$

Further, we constructed 90% Confidence Interval (CI) for bias and mean square error using the range of 5th and 95th percentile.

Results

Table 1 summarizes the results based on 1000 simulations. Findings are interpreted in light of existing bias, MSE of estimated mean and standard deviation along with given 90% confidence interval. Inferences are drawn on the basis of performance of multiple imputation methods for moderate to high proportion of missingness in data. Datasets with 20% to 30% incompleteness are called moderate missingness datasets, while those of consisting half (50%) of incompleteness are assumed as large proportion missingness datasets.

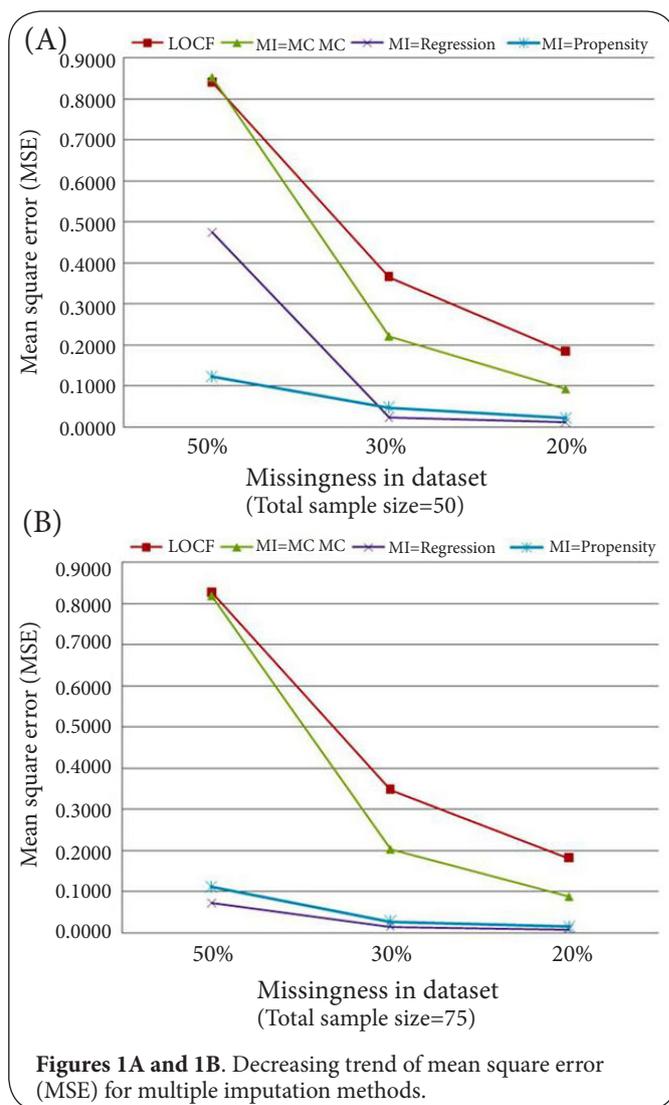
Moderate proportion of missingness

We observed that under moderate missingness MI-regression method has outperformed among all tested methods. For 20% missingness in data, it scored the least bias [0.0553(90% CI :0.0047, 0.1361)] as well as MSE [0.0048(90% CI :0.0000, 0.0185)] for mean when sample size of dataset increased to 120 patients. Similarly, when 30% of data was missing, the MI-regression showed improvement in bias and MSE of estimates with an increase in sample size. In other two multiple imputation methods, MI-propensity score method has scored less bias and MSE than MCMC method for both the moderate proportions of missingness (20% and 30%). Thus, in this scenario MCMC method has shown the lowest performance despite the increase in sample size. Now, when we check on LOCF method: a single imputation method, we observe that it has much higher bias and MSE for mean than multiple imputation methods, indicating that MI methods are better choice for discussed situation. It has been observed that much prevalent LOCF method far behind than multiple imputation methods. LOCF is widely used methods in clinical trial data analysis, but the study results made it very clear that their performance is the poorest even in the case of small missingness (20%).

Figures 1A and **1B** shows the performance of imputation methods, mean square error (MSE) of the MI-regression and propensity score methods can be observed decreasing with the lessening of missingness percentage in data.

Large proportion of missingness

In a situation when half of the data is missing in study and sample size is as small as 50 patients, result showed that MI-propensity is better than other two MI methods, but the difference with MI-regression methods was not much. MI-propensity method scored the lowest bias [0.2810(90% CI :0.0182, 0.6889)] and MSE [0.1219(90% CI : 0.0003, 0.4746)] for mean, while MI-regression method recorded the bias and MSE for mean as [0.3001(90% CI :0.0208, 0.7564)] and [0.4743(90% CI :0.0004, 0.5722)]. With the increase in sample size to 120 patients, MI-regression method had the lowest Bias [0.1649 (90% CI :0.0168, 0.3882)] and MSE [0.0417(90% CI :0.0003, 0.1507)] among other two MI methods. MI-propensity method



also improved the performance with the increase in sample size, however the performance of MI-MCMC method did not improve much despite the increase in sample size in case of large proportion of missingness in data. Again, MI-regression method performed well in the presence of large proportion of missingness in dataset.

It is also observed last observation carried forward method remains at the bottom in terms of comparative measures when missingness proportion is large.

Figure 2 shows the performance of imputation methods for large sample size of patients, MSE of the MI-regression and propensity score methods can be observed decreasing with the lessening of missingness percentage in data.

Discussion

Missing data invariably occur during conduct of a clinical trial and considered as a major concern for longitudinal studies. Unfortunately, the occurrence of missingness is unavoidable.

Table 1. Bias and MSE for estimated Mean and Standard deviation for varying proportion of missingness and sample sizes.

Proportion Missingness	Sample Size	Estimate	Single imputation LOCF method			Multiple imputation MCMC method			Multiple imputation regression method			Multiple imputation propensity score method		
			Bias (90% CI)	MSE (90% CI)	Bias (90% CI)	MSE (90% CI)	Bias (90% CI)	MSE (90% CI)	Bias (90% CI)	MSE (90% CI)	Bias (90% CI)	MSE (90% CI)	Bias (90% CI)	MSE (90% CI)
20%	50	Mean	0.4130 (0.2321, 0.6101)	0.1831 (0.0538, 0.3723)	0.2638 (0.0388, 0.5293)	0.0924 (0.0015, 0.2802)	0.0815 (0.0054, 0.2108)	0.0108 (0.0000, 0.0444)	0.1165 (0.0066, 0.2832)	0.0224 (0.0000, 0.0802)	0.0815 (0.0054, 0.2108)	0.0108 (0.0000, 0.0444)	0.1165 (0.0066, 0.2832)	0.0224 (0.0000, 0.0802)
		Std dev	0.0840 (0.0064, 0.2093)	0.0112 (0.0000, 0.0438)	0.1009 (0.0071, 0.2650)	0.0171 (0.0001, 0.0702)	0.0889 (0.0064, 0.2269)	0.0129 (0.0000, 0.0515)	0.1966 (0.0255, 0.4174)	0.0532 (0.0006, 0.1742)	0.0889 (0.0064, 0.2269)	0.0129 (0.0000, 0.0515)	0.1966 (0.0255, 0.4174)	0.0532 (0.0006, 0.1742)
		Mean	0.4157 (0.2638, 0.5674)	0.1809 (0.0696, 0.3220)	0.2641 (0.0532, 0.5101)	0.0875 (0.0028, 0.2602)	0.0704 (0.0045, 0.1784)	0.0081 (0.0000, 0.0318)	0.0948 (0.0064, 0.2442)	0.0148 (0.0000, 0.0596)	0.0704 (0.0045, 0.1784)	0.0081 (0.0000, 0.0318)	0.0948 (0.0064, 0.2442)	0.0148 (0.0000, 0.0596)
		Std dev	0.0696 (0.0063, 0.1745)	0.0079 (0.0000, 0.0305)	0.0782 (0.0075, 0.2007)	0.0098 (0.0001, 0.0403)	0.0794 (0.0058, 0.1961)	0.0099 (0.0000, 0.0385)	0.1814 (0.0289, 0.3502)	0.0424 (0.0008, 0.1227)	0.0794 (0.0058, 0.1961)	0.0099 (0.0000, 0.0385)	0.1814 (0.0289, 0.3502)	0.0424 (0.0008, 0.1227)
		Mean	0.4148 (0.3027, 0.5386)	0.1773 (0.0916, 0.2901)	0.2597 (0.0908, 0.4327)	0.0785 (0.0083, 0.1872)	0.0553 (0.0047, 0.1361)	0.0048 (0.0000, 0.0185)	0.0743 (0.0059, 0.1866)	0.0087 (0.0000, 0.0348)	0.0553 (0.0047, 0.1361)	0.0048 (0.0000, 0.0185)	0.0743 (0.0059, 0.1866)	0.0087 (0.0000, 0.0348)
		Std dev	0.0563 (0.0051, 0.1410)	0.0050 (0.0000, 0.0199)	0.0647 (0.0060, 0.1683)	0.0068 (0.0000, 0.0283)	0.0726 (0.0074, 0.1729)	0.0079 (0.0001, 0.0299)	0.1665 (0.0401, 0.3157)	0.0347 (0.0016, 0.0997)	0.0726 (0.0074, 0.1729)	0.0079 (0.0001, 0.0299)	0.1665 (0.0401, 0.3157)	0.0347 (0.0016, 0.0997)
	75	Mean	0.5909 (0.3825, 0.7974)	0.3652 (0.1463, 0.6358)	0.4231 (0.0960, 0.7751)	0.2212 (0.0092, 0.6007)	0.1147 (0.0073, 0.2938)	0.0223 (0.0001, 0.0863)	0.1685 (0.0106, 0.4328)	0.0475 (0.0001, 0.1873)	0.1147 (0.0073, 0.2938)	0.0223 (0.0001, 0.0863)	0.1685 (0.0106, 0.4328)	0.0475 (0.0001, 0.1873)
		Std dev	0.1003 (0.0052, 0.2605)	0.0165 (0.0000, 0.0679)	0.1399 (0.0129, 0.3636)	0.0316 (0.0002, 0.1322)	0.1259 (0.0096, 0.3258)	0.0255 (0.0001, 0.1061)	0.2500 (0.0304, 0.5292)	0.0858 (0.0009, 0.2801)	0.1259 (0.0096, 0.3258)	0.0255 (0.0001, 0.1061)	0.2500 (0.0304, 0.5292)	0.0858 (0.0009, 0.2801)
		Mean	0.5804 (0.4141, 0.7599)	0.3474 (0.1715, 0.5774)	0.4122 (0.1246, 0.7370)	0.2026 (0.0155, 0.5432)	0.0955 (0.0073, 0.2368)	0.0148 (0.0001, 0.0561)	0.1289 (0.0094, 0.3147)	0.0271 (0.0001, 0.0990)	0.0955 (0.0073, 0.2368)	0.0148 (0.0001, 0.0561)	0.1289 (0.0094, 0.3147)	0.0271 (0.0001, 0.0990)
		Std dev	0.0869 (0.0073, 0.2131)	0.0118 (0.0001, 0.0454)	0.1163 (0.0090, 0.2933)	0.0220 (0.0001, 0.0860)	0.1126 (0.0096, 0.2715)	0.0195 (0.0001, 0.0737)	0.2328 (0.0357, 0.4633)	0.0719 (0.0013, 0.2147)	0.1126 (0.0096, 0.2715)	0.0195 (0.0001, 0.0737)	0.2328 (0.0357, 0.4633)	0.0719 (0.0013, 0.2147)
		Mean	0.5833 (0.4479, 0.7157)	0.3472 (0.2006, 0.5123)	0.4095 (0.1790, 0.6448)	0.1875 (0.0320, 0.4158)	0.0781 (0.0070, 0.1922)	0.0095 (0.0000, 0.0370)	0.1030 (0.0100, 0.2589)	0.0164 (0.0001, 0.0670)	0.0781 (0.0070, 0.1922)	0.0095 (0.0000, 0.0370)	0.1030 (0.0100, 0.2589)	0.0164 (0.0001, 0.0670)
		Std dev	0.0733 (0.0073, 0.1769)	0.0081 (0.0001, 0.0313)	0.0933 (0.0062, 0.2431)	0.0139 (0.0000, 0.0591)	0.1050 (0.0118, 0.2312)	0.0156 (0.0001, 0.0535)	0.2146 (0.0409, 0.4067)	0.0579 (0.0017, 0.1654)	0.1050 (0.0118, 0.2312)	0.0156 (0.0001, 0.0535)	0.2146 (0.0409, 0.4067)	0.0579 (0.0017, 0.1654)
50%	50	Mean	0.9058 (0.6816, 1.1471)	0.8401 (0.4646, 1.3159)	0.8548 (0.2980, 1.4540)	0.8518 (0.0888, 2.1141)	0.3001 (0.0208, 0.7564)	0.4743 (0.0004, 0.5722)	0.2810 (0.0182, 0.6889)	0.1219 (0.0003, 0.4746)	0.3001 (0.0208, 0.7564)	0.4743 (0.0004, 0.5722)	0.2810 (0.0182, 0.6889)	0.1219 (0.0003, 0.4746)
		Std dev	0.1325 (0.0104, 0.3131)	0.0268 (0.0001, 0.0980)	0.3120 (0.0239, 0.7339)	0.1464 (0.0006, 0.5386)	0.2957 (0.0242, 0.6567)	0.1631 (0.0002, 0.6491)	0.3333 (0.0338, 0.7580)	0.1606 (0.0011, 0.5745)	0.1325 (0.0104, 0.3131)	0.0268 (0.0001, 0.0980)	0.3120 (0.0239, 0.7339)	0.1464 (0.0006, 0.5386)
		Mean	0.9012 (0.7167, 1.0871)	0.8256 (0.5136, 1.1817)	0.8610 (0.3887, 1.3452)	0.8174 (0.1511, 1.8095)	0.2124 (0.0175, 0.5172)	0.0726 (0.0003, 0.2675)	0.2628 (0.0217, 0.6422)	0.1112 (0.0005, 0.4124)	0.8256 (0.5136, 1.1817)	0.8174 (0.1511, 1.8095)	0.2124 (0.0175, 0.5172)	0.0726 (0.0003, 0.2675)
		Std dev	0.1189 (0.0125, 0.2682)	0.0214 (0.0002, 0.0720)	0.2799 (0.0259, 0.6286)	0.1164 (0.0007, 0.3952)	0.2248 (0.0188, 0.5445)	0.0778 (0.0004, 0.2965)	0.3253 (0.0247, 0.7417)	0.1555 (0.0006, 0.5500)	0.1189 (0.0125, 0.2682)	0.0214 (0.0002, 0.0720)	0.2799 (0.0259, 0.6286)	0.1164 (0.0007, 0.3952)
		Mean	0.8956 (0.7565, 1.0362)	0.8096 (0.5722, 1.0736)	0.8606 (0.5175, 1.2047)	0.7856 (0.2679, 1.4514)	0.1649 (0.0168, 0.3882)	0.0417 (0.0003, 0.1507)	0.2274 (0.0200, 0.5705)	0.0814 (0.0004, 0.3255)	0.8956 (0.7565, 1.0362)	0.8096 (0.5722, 1.0736)	0.8606 (0.5175, 1.2047)	0.7856 (0.2679, 1.4514)
		Std dev	0.1037 (0.0094, 0.2342)	0.0157 (0.0001, 0.0548)	0.2413 (0.0266, 0.5457)	0.0834 (0.0007, 0.2978)	0.1993 (0.0194, 0.4376)	0.0568 (0.0004, 0.1915)	0.3060 (0.0304, 0.6508)	0.1298 (0.0009, 0.4235)	0.1037 (0.0094, 0.2342)	0.0157 (0.0001, 0.0548)	0.2413 (0.0266, 0.5457)	0.0834 (0.0007, 0.2978)

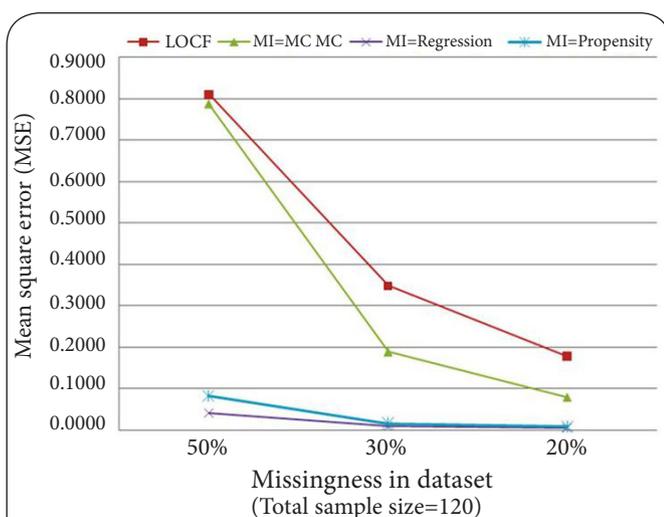


Figure 2. Performance of multiple imputation method along with LOCF; MI-regression achieved minimum MSE, while LOCF showed no improvement despite increased sample size.

able despite the carefulness in study design, conduct and preventive strategies. Due to complexity of methods while dealing with missing data analysis, researchers exclude the subjects with partial information missing on response variables [16]. Exclusion of subjects from a planned sample size directly impacts the power of study [17], especially when numbers of the subjects in study are already less or including more patients in study has significant financial implications. Multiple imputation methods have drawn researcher's attention in last decade due to their computational elegance and applicability across the wide spectrum of incomplete data problems [18]. Advancement of computers and availability of statistical softwares have given us the opportunity to apply complex and time taking analysis techniques efficiently and easily. However, the pace of development in methodology is not reflected in real data applications by its end users [18]. In the present work we have explored the efficiency and appropriateness of various multiple imputation methods under different possible circumstances, using the software Statistical Analysis System (SAS) Version 9.1.3. We evaluated the imputation methods for different sample sizes and varying missingness from moderate to large proportion. Based on the results and recommendations, the use of appropriate MI method is expected in analysis and reporting of longitudinal clinical studies.

In order to make an informed decision in presence of missing data, it is necessary to analyze the data with the available statistical framework to deal with missingness in clinical trial studies; this in return will add to the generalizability of the research findings. Due to easiness of procedure, single imputation methods have been the default choice by researchers, but the efficiency of this method has been under question [14]. In a longitudinal clinical trial where the treatment effect improves

with time, LOCF method has been a conservative approach to measure the change over time. Simulation result suggests that LOCF method has the maximum bias and mean square error. The multiple imputation method offers an opportunity to make the best out of collected data from an experiment on human being. Though missing data in clinical trials has received a growing amount of attention, there are still some key misunderstandings regarding selection of imputation methods. Based on the performance of methods, researchers can choose the most appropriate MI method from the Table 1. MI-regression method has performed better than other two MI methods. Hence, it can be applicable for various situations. MI-propensity method has also performed well but remains at the second position, notably the method is comparable for large proportion missingness in small sample size studies. Markov chain Monte Carlo (MCMC) method is far behind than other two MI methods, so method does not offer much applicability for the discussed situation. However, in general, all multiple imputation methods have performed better than single imputation method. It is also to be understood that performance of MI methods improves significantly with the increase of sample size.

Conclusions

In present investigation, we have explored the efficiency and appropriateness of various imputation methods under small size data sets with varying size of missingness. The results of the study indicate that MI-regression method is most appropriate method for the analysis of data in presence of missingness with considered sample size and missingness mechanism. It was also demonstrated that large proportion of missing data in study affect the performance of imputation technique. Study findings may have important implications, particularly for a limited sample size longitudinal study where varying proportions is encountered. As evident from the findings, single imputation may not be a suitable approach for missing data analysis; the reason is as simple as it does not incorporate the uncertainty of missingness in imputed value. The present study findings will help researchers to choose an imputation method according to their requirement and the achieved estimates will be more precise. Simultaneously, we also encourage investigators and readers to keep themselves abreast with new development of the multiple imputation methods for effective application of findings.

Limitations of the study

We have considered the data under MAR assumption on a single outcome of interest with specified mean and covariance structure in a diabetic clinical trial, but in practice the assumption of specified mean and covariance structure of missingness may differ. So, caution is needed while using the results from this simulation study when assumptions of the study differ significantly. Our study considerations for assumed sample sizes is based on ease of computational

procedure and commonly analyzed number of subjects for a treatment arm; to translate the findings in other therapeutic areas, a study with larger sample size with different missingness proportions can be taken. Further exploration with new known complex scenarios would also be beneficial.

List of abbreviations

MI: Multiple Imputation
MCAR: Missing completely at random
MAR: Missing at random
MNAR: Missing not at random

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Authors' contributions	SM	DK
Research concept and design	✓	✓
Collection and/or assembly of data	✓	--
Data analysis and interpretation	✓	--
Writing the article	✓	✓
Critical revision of the article	✓	--
Final approval of article	✓	--
Statistical analysis	✓	✓

Publication history

Editor: Feng Gao, Washington University School of Medicine, USA.
EIC: Jimmy Efrid, East Carolina University, USA.
Received: 16-Aug-2014 Final Revised: 29-Oct-2014
Accepted: 08-Nov-2014 Published: 14-Nov-2014

References

1. Myers WR. **Handling missing data in clinical trials: an overview.** *Drug Information Journal.* 2000; **34**:525-533. | [Article](#)
2. Mallinckrodt CH, Sanger TM, Dube S, DeBrotta DJ, Molenberghs G, Carroll RJ, Potter WZ and Tollefson GD. **Assessing and interpreting treatment effects in longitudinal clinical trials with missing data.** *Biol Psychiatry.* 2003; **53**:754-60. | [Article](#) | [PubMed](#)
3. Molenberghs G and Kenward MG. **Missing data in clinical studies.** Hoboken, New Jersey. John Wiley and Sons, 2007. | [Book](#)
4. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM and Carpenter JR. **Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.** *BMJ.* 2009; **338**:b2393. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
5. Royston P. **Multiple imputation of missing values: update of ice.** *Stata Journal.* 2005; **5**:527-536. | [Pdf](#)
6. SAS Institute. **The MI Procedure.** 1999. | [Pdf](#)
7. Molenberghs G, Thijs H, Jansen I, Beunckens C, Kenward MG, Mallinckrodt C and Carroll RJ. **Analyzing incomplete longitudinal clinical trial data.** *Biostatistics.* 2004; **5**:445-64. | [Article](#) | [PubMed](#)
8. Rubin DB. **Inference and missing data.** *Biometrika.* 1976; **63**:581-592. | [Pdf](#)
9. Little RJA and Rubin DB. **Statistical analysis with missing data.** New York, John Wiley & Sons Inc. 1987. | [Book](#)
10. Verbeke G and Kenward MG. **Linear mixed models for longitudinal data.** New York, Springer. 2000. | [Book](#)
11. Rubin DB. **Multiple imputation for nonresponse in surveys.** New York, Wiley. 1987. | [Book](#)

12. Schafer JL. **Multiple imputation: a primer.** *Stat Methods Med Res.* 1999; **8**:3-15. | [Article](#) | [PubMed](#)
13. Meng XL. **Multiple-imputation inferences with uncongenial sources of input.** *Statistical Science.* 1994; **9**:538-558. | [Article](#)
14. Kenward MG and Molenberghs G. **Last observation carried forward: a crystal ball?** *J Biopharm Stat.* 2009; **19**:872-88. | [Article](#) | [PubMed](#)
15. Mallinckrodt CH, Clark WS and David SR. **Accounting for dropout bias using mixed-effects models.** *J Biopharm Stat.* 2001; **11**:9-21. | [Article](#) | [PubMed](#)
16. Pigott T. **A review of methods for missing data.** *Educational Research and Evaluation.* 2001; **7**:353-383. | [Pdf](#)
17. Kang H. **The prevention and handling of the missing data.** *Korean J Anesthesiol.* 2013; **64**:402-6. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
18. Yucel RM. **State of the Multiple Imputation Software.** *J Stat Softw.* 2011; **45**:1-7. | [PubMed Abstract](#) | [PubMed Full Text](#)

Citation:

Mishra S and Khare D. **On comparative performance of multiple imputation methods for moderate to large proportions of missing data in clinical trials: a simulation study.** *J Med Stat Inform.* 2014; **2**:9.
<http://dx.doi.org/10.7243/2053-7662-2-9>