PharmaSUG 2023 - Paper SA-269

Statistical Considerations and Methods for Handling Missing Outcome Data During the Era of COVID-19

Xi Qian, Chengfei Lu, BioPier, Inc.

ABSTRACT

Missing data occurs when the values of variables of interest are not collected for all subjects at all scheduled visits in clinical trials. It is usually caused by subject dropout and/or loss to follow-up. Although missing data is a common problem for any clinical trials, COVID-19 pandemic has been presenting unprecedented challenges for many aspects of clinical trial studies, particularly for missing data analysis. Statistical methods for dealing with missing data are dependent on the assumptions regarding the data missing mechanisms and patterns. In this paper, we first introduce data missing mechanisms and patterns and summarize commonly used statistical methods for handling missing data. We then focus on multiple imputation for continuous outcome data and discuss issues in its practical implementation, including developing imputation model, how to handle data with monotone or non-monotone missing pattern, the number of imputed data sets need to be created, and how to examine the robustness of missing data assumption with sensitivity analysis. We illustrate the application of multiple imputation through analysis of quality-of-life score data on lung cancer patients in a hypothetical superiority trial.

INTRODUCTION

Clinical trials are crucial in determining the efficacy and safety of new treatments, drugs, and medical interventions. However, a common challenge faced by researchers conducting clinical trials is the problem of missing data. Missing data occurs when participants drop out of the study or are lost to follow-up, leading to incomplete or insufficient data for analysis. This can compromise the reliability and validity of the trial results, introduce bias, as well as reduce the statistical power of the study.

While missing data is a challenge for any clinical trial, the COVID-19 pandemic has been presenting unprecedented difficulties for trials conducted during this time. Regulatory agencies such as the FDA, EMA, and MHRA have recognized the impact of COVID-19 on clinical trials and issued guidance on trial conduct during the pandemic [1-3]. These guidelines recommend that data collection should prioritize the safety of both participants and investigators, and that remote data collection methods should be used when feasible. Remote methods of data collection may include virtual visits, telemedicine, and electronic self-reporting. Despite the availability of remote data collection methods, difficulties in adhering to protocol-defined follow-up have arisen during the pandemic. For example, infected participants may be unable to attend follow-up visits due to quarantine or hospitalization or may find remote intervention or assessments less engaging. Another challenge related to missing data during the pandemic is the early termination of clinical trials due to safety concerns or logistical challenges posed by the pandemic. These factors have contributed to an increased rate of missing data, particularly outcome data, in clinical trials during the pandemic. This can have significant implications for the validity and generalizability of the trial results, as well as for the development of new treatments and interventions.

This paper aims to provide an overview of commonly used statistical methods to address missing data in clinical trials, with a particular focus on multiple imputation (MI) for continuous outcome data. We begin by introducing data missing mechanisms, data missing patterns, and statistical methods for missing data. We then discuss practical issues related to implementing MI, such as developing the imputation model, handling data with monotone or non-monotone missing patterns, the number of imputed data sets required, and examining the robustness of missing data assumptions using sensitivity analysis. To illustrate the application of MI, we analyze the quality-of-life score data from a hypothetical superiority trial involving lung cancer patients.

MISSING DATA MECHANISM

To address missing data in clinical studies, it is important to understand the conditions under which data may be missing. Rubin proposed a framework for handling missing data and described three different

mechanisms by which data can be missing: Missing Completely at Random (MCAR), Missing at Random (MAR), and Missing Not at Random (MNAR) [4]. Common examples of the three missing data mechanisms are summarized in Table 1.

Missing Data Mechanisms	Examples From Clinical trial
MCAR	• Random failure of the laboratory equipment • Study participants move and are unable to attend study visits
MAR	• Subjects drop out based on recorded adverse events or other known characteristics. • Missingness due to study design such as participation termination because of violation of study protocol
MNAR	• Subjects drop out based on the unobserved response (e.g., subjects not responding to treatment or having unrecorded adverse events are more likely to drop out)

Table 1. Common Examples in Clinical Trials of the Three Missing Data Mechanisms (adapted from [5]).

MCAR

MCAR is a type of missing data where the probability of data being missing is completely random and unrelated to any observed or unobserved variables. In other words, the chance of missing data is the same for subjects in different treatment groups and those who have different baseline characteristics or treatment response. Examples of MCAR include missing data due to subject migration (study participants move and are unable to complete visits), administrative censoring (follow-up is terminated because the study has been ended), and random failure of laboratory instrument. In this scenario, the average effect of the treatment is the same in those with and without missing data, and therefore, the missing data is unlikely to introduce bias. However, MCAR may be accompanied by a loss of power in the test of significance (i.e., higher p-values), as well as a loss of precision in the estimation of the treatment effect (i.e., larger standard errors and wider confidence intervals).

MAR

MAR is a type of missing data where the probability of data being missing is related to observed variables, but not to unobserved variables. If in a clinical trial is a study investigating the effect of a new treatment on hypertension, where male participants is more likely to drop out compared to female subjects, the missing data mechanism is MAR. Other examples of MAR are exemplified by missing data caused by dropouts based on known baseline characteristics, and recorded side effects or lack of efficacy. In this case, the missing data may introduce some bias since the reason for the missingness is related to the observed variable. Yet assuming that data missingness was random after accounting for related variables, analyzing the available data may extend our knowledge about observed data to the missing data.

MNAR

MNAR, also sometimes labeled non-ignorable missing data or informative missing data, is a type of missing data where the probability of data being missing is related to unobserved variables. This can occur when (1) the missing value itself influences the probability of missingness, or (2) some unmeasured factor is linked to both the value of the missing variable and the probability of missingness. An example of MNAR in a clinical trial is a study evaluating the effectiveness of a new drug for treating depression, where participants who experienced adverse effects from the drug were more likely to drop out of the trial. In this case, the missing data is likely to introduce bias since the reason for the missingness is related to the unobserved variable (treatment efficacy), which is also associated with the trial outcome (effectiveness of the new drug).

Differentiating between MCAR, MAR, and MNAR missing data can be challenging, and it often requires us to make assumptions based on prior knowledge. To help assess if an MCAR assumption holds, we can 1) compare observed baselined characteristics and other endpoints related the main outcome of

interest between subjects with and without missing data, and 2) assess if there is a difference between treatment groups on the distribution of missing data. If there are group differences on baseline characteristics/surrogate endpoints or the randomized treatment do affect completion status, the missing data mechanism is most likely not MCAR. However, discerning MAR vs MNAR is typically not possible, as MNAR are characterized by systematic differences in unmeasured variables. Often, the best the investigators can do is to test how sensitive our results are to different missing data assumptions.

MISSING DATA PATTERN

Another way to categorize missing data is based on the data structure or pattern, which has practical implications for planning the analytical strategy to address missing data, such as the imputation approach. The missing data pattern is often depicted visually as an array, with observations as rows and variables as columns (see Figure 1). There are two missing patterns: arbitrary (Figure 1, Left) and monotone (Figure 1, Right) patterns. In the arbitrary pattern, the missing data are interspersed among full data values, while in monotone patterns, the missing data are at the end, from left to right, and there are no gaps between the missing data and full data. The arbitrary missing pattern is the most common pattern, while a monotonic missing data pattern may result from loss to follow-up, where subjects drop out at a particular visit. A monotone missing pattern is easier to work with and more flexible, but this pattern often suggests a MNAR mechanism.

	Variable		
Obs	V1	V2	V3
1			?
2	?	?	
3	?		
4		?	
5	?		?

	Variable		
Obs	V1	V2	V3
1			
2			
3			?
4		?	?
5	?	?	?

Figure 1. Missing Data Pattern. Left, arbitrary pattern; Right, monotone pattern.

STATISTICAL METHODS FOR HANDLING MISSING DATA

As mentioned above, failure to properly address the missing data mechanism during the analysis can result in erroneous conclusions. However, the problem is that there is no definitive way to determine which mechanism has led to the data missingness. Therefore, a single technique for analysis is not sufficient. Rather, a series of missing data analytic methods are usually employed to examine missing data bias and the robustness of the trial results. This strategy is also advocated in FDA's Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic [2]. The utility of each method under the different missing data mechanisms is summarized in Table 2 [5].

One common approach for dealing with missing data was to exclude subjects who had missing data on any necessary variables, leaving only those with complete data (known as "complete case" analysis). This approach may lead to biased estimated statistics and regression coefficients and reduce the sample size and power to detect a meaningful treatment effect. Given these drawbacks, complete case analysis approach may be used in exploratory studies, particularly in the early stage of drug development, however, it is not used be the primary analysis in confirmatory trials. Instead, it may be used as a secondary supportive analysis or sensitivity analysis to demonstrate the robustness of conclusions.

Single imputation replaces missing values with plausible values. Commonly used single imputation include Last Observation Carried Forward (LOCF), Baseline Observation Carried Forward (BOCF), Worst Observation Carried Forward (WOCF) and Mean Substitution. Other single imputation approach includes simple and conditional mean imputation. Single mean imputation replaces the values by the mean for that variable while in conditional mean imputation, the missing outcomes are imputed by regressing the outcome on other observed variables in the completers. These approaches do not fully capture the uncertainty that this value is correct and often underestimate the variability of treatment effect and false

positive results. The ICH E9 guidelines for missing data state that single imputation methods like “last observation carried forward” should not be used as the primary analytic approach unless the assumptions that underlie these methods are scientifically justified [6].

Missing Data Mechanisms	Statistical Methods
MCAR	• Unbiased effects and standard errors: Likelihood Based, MI, Inverse Probability Weighting, Complete Case • Unbiased effects: Single Mean Imputation, Conditional Mean Imputation • Unacceptable: LOCF, WOCF
MAR	• Unbiased effects and standard errors: Likelihood Based, MI, Inverse Probability Weighting • Unbiased effects: Conditional Mean Imputation • Unacceptable: LOCF, WOCF, Simple Mean Imputation
MNAR	• Sensitivity Analysis (Pattern Mixture Models and Tipping Point Analysis) to evaluate the robustness of the results to the deviations from the MAR assumption.

Table 2. A Summary of Statistical Methods for Three Types of Missing Data Mechanisms (adapted from [5]).

Acceptable approaches under MAR include likelihood-based analysis, inverse probability weighting, and MI. Among them, MI has gained enormous popularity in the past decades in trial analysis. Conditional mean imputation treats the observed and imputed values the same despite uncertainty in the imputed values and therefore consistently underestimates the variability of treatment effect. The method of MI corrects for this by generating multiple completed data sets, each with the observed values and different plausible values imputed for the missing observations. After creating the multiple complete data sets, each is analyzed using usual methods (e.g., ANCOVA, MMRM), and the results are then combined across the analyses. Under the MAR assumption, MI produces unbiased estimates of the intervention effect and correct p-values. In addition, MI along with pattern mixture model and tipping point analysis is often used as a MNAR sensitivity analysis to assess how severe departure from MAR must be to overturn conclusions.

MI METHOD

The theoretical development of MI methods for missing data is rooted in the Bayesian framework for statistical inference, which is beyond the scope of this paper. For detailed statistical considerations, please refer to Berglund and Heeringa’s book [7]. In this paper, we focus on the practical implementation of MI. Implementing MI requires three steps, as depicted in Figure 2.

MI STEPS

The imputation step (Step 1) involves specifying the variables and distributional assumptions of the imputation model and using specific MI algorithms to generate imputations for the missing values. The PROC MI procedure provides a range of options for imputation algorithms, which depend on the pattern of missing data (monotone vs. arbitrary) and the type of outcomes (continuous, binary, nominal, ordinal) that need to be imputed. Additionally, assumptions about the multivariate distribution of the variables in the imputation model must be taken into account. The imputation method available in PROC MI is summarized in Table 3. The output of this step is a completed data set for each of m=1,…,M repetitions of the imputation process. These completed data sets are stored as a stacked for, with the _IMPUTATION_ variable added to differentiate between the imputed data sets.

Step 2 involves analyzing the data using standard SAS statistical procedures such as PROC FREQ, PROC MIXED, and PROC GENMOD, in the same manner as we analyze non-imputed data. However, for MI imputed data sets, this analysis needs to be repeated for each data set separately. This can be accomplished by adding a BY statement to the relevant procedure with the _IMPUTATION_ variable, which allows for separate analyses of each imputed data set. An important task of this step is to specify

an appropriate output of the estimated statistics and their standard errors from each repetition of the analysis. The output data set of estimated statistics and standard errors is the required input for Step 3.

In Step 3, the results obtained in Step 2 are combined. This is achieved using the PROC MIANALYZE procedure, which takes as input the results of the M separate analyses and combines the results from each MI repetition and provides valid statistical inferences. Regardless of the method used to analyze the data in Step 2, PROC MIANALYZE combines the information to obtain a single result.

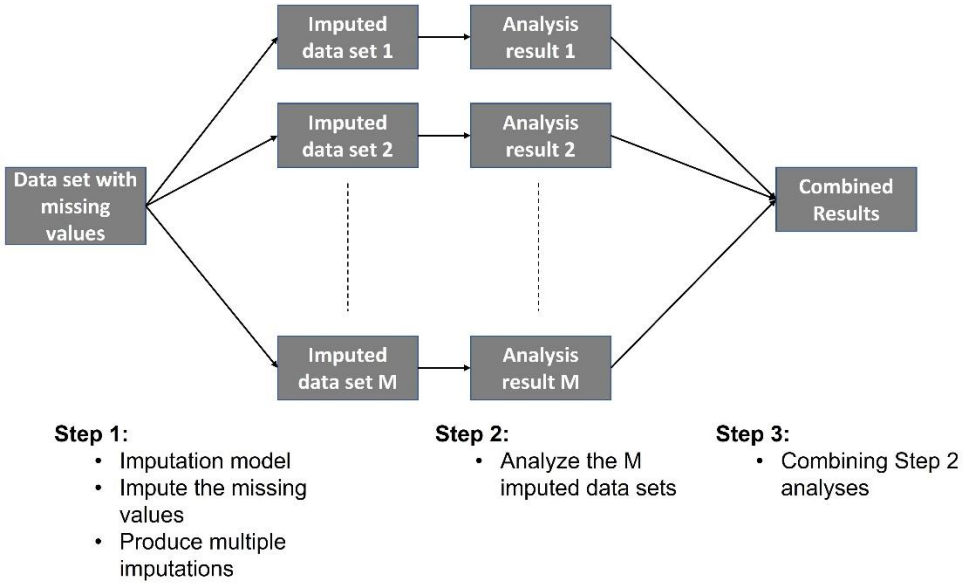

Figure 2. Three Steps in the MI Approach.

Pattern of Missingness	Type of Imputed Variable	Type of Covariates	Available Methods
Monotone	Continuous	Arbitrary	Monotone regression
			Monotone predicted mean matching
			Monotone propensity score
Monotone	Classification (ordinal)	Arbitrary	Monotone logistic regression
Monotone	Classification (nominal)	Arbitrary	Monotone discriminant function
			Monotone logistic regression
Arbitrary	Continuous	Continuous	MCMC full-data imputation
			MCMC monotone-data imputation
Arbitrary	Continuous	Arbitrary	FCS regression
			FCS predicted mean matching
Arbitrary	Classification (ordinal)	Arbitrary	FCS logistic regression
Arbitrary	Classification (nominal)	Arbitrary	FCS discriminant function
			FCS logistic regression

Table 3. Imputation Methods in PROC MI (adapted from [8])

HOW MANY IMPUTATIONS ARE NEEDED IN MI?

Determining the optimal number of imputed data sets in MI is an important question. Early recommendations suggested that 3 to 5 imputed data sets were sufficient, as long as the amount of missing information was not high. However, other experts suggested that 5 to 10 imputations were often necessary for adequate precision. These early recommendations were primarily based on the accuracy of estimated regression coefficients compared to those obtained from an infinite number of imputed data sets. However, in addition to accurate regression coefficients, analysts also require accurate standard errors for confidence intervals and significance tests. Therefore, it is ideal to select a number of imputations that yields stable estimates of both regression coefficients and standard errors, such that the pooled results do not vary substantially across repeated applications of MI (i.e., if the entire process was repeated with M new imputed data sets, one would obtain estimates similar to those obtained using the initial M imputed data sets). In practice, M=50 is usually proposed in the statistical analysis plan for trials.

APPLICATIONS OF MI IN CLINICAL TRIAL

IMPLEMENTATION OF MI USING SAS UNDER MONOTONE MISSING DATA PATTERN AND MAR

In this example, we employ the data from a hypothetical clinical trial in lung patients comparing Chemotherapy + Additive (experimental) vs. Chemotherapy alone (control). One of the secondary outcomes in this trial is to compare the change in Overall Health Score (OHS) from baseline at week 52, although the score is also measured at week 24. The higher the OHS score, the better the overall quality of health. The mock shell is shown in Figure 3 and a snapshot of ADaM data set (ADQLQ) used for analysis is shown Figure 4. Only key variables for subjects, 01002-001and 46011-003, are displayed. The TRTPN variable is coded as: 1, Chemotherapy + Additive; 2, Chemotherapy alone. There are four randomization strata in this trial.

Table x.x.x.x
Analysis of Change from Baseline in Overall Health Status Scores at Week 52
ITT Analysis Set

	Chemo + Addtive (N=xx)		Chemo (N=xx)	
Study Visit	Observed Value	Change From Baseline	Observed Value	Change From Baseline
Baseline				
n	xx		xx	
Mean (SD)	xx.x (xx.xx)		xx.x (xx.xx)	
Median	xx.x		xx.x	
Minimum, Maximum	xx.x, xx.x		xx.x, xx.x	
Week 52				
n	xx	xx	xx	xx
Mean (SD)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)
Median	xx.x	xx.x	xx.x	xx.x
Minimum, Maximum	xx.x, xx.x	xx.x, xx.x	xx.x, xx.x	xx.x, xx.x
Week 52				
LS Mean (95 % CI) [a]		xx.x (xx.x, xx.x)		xx.x (xx.x, xx.x)
Difference in LS Mean (SE), Chemo + Addtive vs. Chemo [a]		xx.x (xx.xx)		
95% CI of the Difference in LS Mean Chemo + Addtive vs. Chemo [a]		(xx.x, xx.x)		
p-value, Chemo + Addtive vs. Chemo [a]		0.xxx		

Figure 3. Example Table Shell of Analysis of Change from Baseline in Overall Health Scores at Week 52.

SUBJID	TRTPN	STRATUM	AVISIT	AVAL
01002-001	2	STRAUM2	BASELINE	55.12850587
01002-001	2	STRAUM2	WEEK 24	57.54135646
01002-001	2	STRAUM2	WEEK 52	60.19275881
46011-003	1	STRAUM1	BASELINE	50.59689156

Figure 4. A Snapshot of ADaM Data Set, ADQLQ, used for Analysis.

Imputation Step

As shown in Figure 4, the ADQLQ data set is created in the form of basic data structure (BDS). However, PROC MI procedure requires a horizontal data structure, e.g., one record per subject. Therefore, we need to first transpose the BDS data set into a horizontally structured data set. The following code serves the purpose here:

```
proc transpose data=adqlq out=ohs1;
    by subjid trtpn stratum;
    id avisit;
    var aval;
run;
```

Figure 5 below shows the data structure after transposing for the two same subjects. After transposing, the variables BASELINE, WEEK_24 and WEEK_52 contain the AVAL values from baseline, week 24 and 52, respectively.

SUBJID	TRTPN	STRATUM	BASELINE	WEEK_24	WEEK_52
01002-001	2	STRAUM2	55.12850587	57.54135646	60.19275881
46011-003	1	STRAUM1	50.59689156		

Figure 5. A Snapshot of Transposed ADQLQ, used for Analysis.

Before applying the MI procedure, the missing data pattern needs to be checked, which can help determine the correct imputation model to be used. In the following code, the CLASS statement is used to identify categorical variables, the FCS (or MONOTONE) statement is specified as a CLASS statement is used here, and the VAR statement identifies all the variables of interest. Please note, the NIMPUTE is set to 0 as no imputations are carried out here. From Figure 6, we notice that 73.06% of records are complete cases (without any missing data for all variables checked), 20.41% of the records miss for week 52's score, and 6.53% of the records miss for the scores for both week 24 and 52.

```
proc mi data=ohs1 nimpute=0;
    class subjid trtpn stratum;
    ods select misspattern;
    fcs;
    var subjid trtpn stratum baseline week_24 week_52;
run;
```

SUBJID	TRTPN	STRATUM	BASELINE	WEEK_24	WEEK_52	Freq	Percent
X	X	X	X	X	X	179	73.06
X	X	X	X	X	.	50	20.41
X	X	X	X	.	.	16	6.53

Figure 6. Missing Data Pattern of Transposed ADQLQ. The 'X' indicates non-missing data while the '.' indicates missing data.

Since the missing pattern follows a monotone structure, the MONOTONE REG method is applied. In PROC MI code below, the SEED option specifies the random number seed to reproduce the results. There are three complete data sets imputed (NIMPUTE=3), which are then output to the miout1 data set (OUT=miout1). Note in the MONOTONE REGRESSION statements: we impute week 24 score before week 52 score. The VAR statement is needed, due to the monotone missing data pattern, BASELINE needs to be listed before WEEK_24, which needs to be listed before WEEK_52. A snapshot of resulting output for subjects, 01002-001and 46011-003, is displayed in Figure 7. The _IMPUTATION_ variable indicates 3 repetitions of MI data sets.

```
proc mi data=ohs1 seed=113241 out=miout1 nimpute=3;
    class trtpn stratum;
    monotone regression (week_24=baseline trtpn stratum/details);
    monotone regression (week_52=baseline week_24 trtpn stratum/details);
    var trtpn stratum baseline week_24 week_52;
run;
```

IMPUTATION	SUBJID	TRTPN	STRATUM	BASELINE	WEEK_24	WEEK_52
1	01002-001	2	STRAUM2	55.12850587	57.54135646	60.19275881
1	46011-003	1	STRAUM1	50.59689156	57.384924519	154.6867436
2	01002-001	2	STRAUM2	55.12850587	57.54135646	60.19275881
2	46011-003	1	STRAUM1	50.59689156	61.105864304	61.663247455
3	01002-001	2	STRAUM2	55.12850587	57.54135646	60.19275881
3	46011-003	1	STRAUM1	50.59689156	61.849064976	43.08323034

Figure 7. A Snapshot of Multiple Imputed Data Set.

Analysis Step

After performing imputation on the data, the next step is to analyze the results and generate the estimates for each of the imputed data sets. To do this, the multiple imputed data set is transposed back to its original ADaM BDS structure, along with the required variables such as SUBJID, TRTPN, STRATUM, BASE, BASE, and AVISIT. The changes from baseline at week 24 and week 52 are then calculated. This process results in a data set with imputations, comprising 3 repetitions of the data, indexed by the column _IMPUTATION_. A sample BDS structure data set with imputations is shown below (Figure 8).

```
  proc transpose data=miout1 (drop=_name_)
            out=miout2 (rename=(name_=avisit baseline=base col1=aval));
      by _imputation_ subjid trtpn stratum baseline;
run;
```

IMPUTATION	SUBJID	TRTPN	STRATUM	BASE	AVISIT	AVAL	CHG
1	01002-001	2	STRAUM2	55.12850587	WEEK_24	57.54135646	2.413
1	01002-001	2	STRAUM2	55.12850587	WEEK_52	60.19275881	5.064
1	46011-003	1	STRAUM1	50.59689156	WEEK_24	57.384924519	6.788
1	46011-003	1	STRAUM1	50.59689156	WEEK_52	154.6867436	104.090
2	01002-001	2	STRAUM2	55.12850587	WEEK_24	57.54135646	2.413
2	01002-001	2	STRAUM2	55.12850587	WEEK_52	60.19275881	5.064
2	46011-003	1	STRAUM1	50.59689156	WEEK_24	61.105864304	10.509
2	46011-003	1	STRAUM1	50.59689156	WEEK_52	61.663247455	11.066
3	01002-001	2	STRAUM2	55.12850587	WEEK_24	57.54135646	2.413
3	01002-001	2	STRAUM2	55.12850587	WEEK_52	60.19275881	5.064
3	46011-003	1	STRAUM1	50.59689156	WEEK_24	61.849064976	11.252
3	46011-003	1	STRAUM1	50.59689156	WEEK_52	43.08323034	-7.514

Figure 8. A Snapshot of Transposed MI Data Set.

Next, mixed-effects model for repeated measures (MMRM) analysis is used to analyze each of three MI data sets. In the model, change from baseline in OHS score is used as the outcome, and treatment group, visit, the interaction between treatment group and visit, and the stratification factors are included as covariates. A compound symmetry (CS) covariance matrix is used. The REF='2' option specifies the control group, Chemotherapy only, as reference group per mock shell. The ODS OUTPUT DIFFS= statement outputs the estimate differences from the LSMEANS statement. This contains an OHS score comparison for each treatment combination within each analysis visit (Figure 9). The BY _IMPUTATION_ statement allows independent analyses for each imputation. These results will later be pooled to obtain overall estimates for each visit.

```
ods output diffs=lsmd1 (where=(avisit=_avisit));
proc mixed data= miout2;
    by _imputation_;
    class trtpn(ref='2') avisit subjid stratum;
    model chg = trtpn avisit trtpn*avisit stratum/s;
    repeated avisit/subject=subjid type=cs;
    lsmeans avisit*trtpn/cl diff;
run;
```

IMPUTATION	EFFECT	TRTPN	AVISIT	_TRTPN	_AVISIT	ESTIMATE	STDERR	DF	TVALUE	PROBT	ALPHA	LOWER	UPPER
1	TRTPN*AVISIT	1	WEEK_24	2	WEEK_24	8.0424	5.5624	243	1.45	0.1495	0.05	-2.9144	18.9991
2	TRTPN*AVISIT	1	WEEK_24	2	WEEK_24	7.9632	5.4235	243	1.47	0.1433	0.05	-2.7199	18.6462
3	TRTPN*AVISIT	1	WEEK_24	2	WEEK_24	8.1177	5.7480	243	1.41	0.1592	0.05	-3.2046	19.4400
1	TRTPN*AVISIT	1	WEEK_52	2	WEEK_52	19.1941	5.5624	243	3.45	0.0007	0.05	8.2373	30.1508
2	TRTPN*AVISIT	1	WEEK_52	2	WEEK_52	20.1567	5.4235	243	3.72	0.0003	0.05	9.4736	30.8397
3	TRTPN*AVISIT	1	WEEK_52	2	WEEK_52	16.6715	5.7480	243	2.90	0.0041	0.05	5.3492	27.9938

Figure 9. The OHS Score Comparison for Each Treatment Combination, Each Analysis Visit, and Each Imputed Data Set.

Results Pooling Step

To produce pooled statistical inferences, PROC MIANALYZE is applied to analyze lsmd1 data set. In order to produce pooled results by analysis visit, the data needs to be sorted by AVISIT before PROC MIANALYZE implementation. The following sample code and output (Figure 10) show the pooled statistical inferences. The results indicate that the average change from baseline in OHS score at week 52 in the Chemotherapy + Additive group is 18.67 [95% CI, (5.95, 6.90)] higher than the Chemotherapy Only group, with a p-value=0.002.

```
ods output parameterestimates=difbyvis;
proc mianalyze data=lsmd1;
    by avisit;
    modeleffects estimate;
    stderr stderr;
run;
```

AVISIT	NIMPUTE	PARM	ESTIMATE	STDERR	LCLMEAN	UCLMEAN	DF	PROBT
WEEK_24	3	estimate	8.041073	5.580280	-2.8961	18.97822	3.06E7	0.1496
WEEK_52	3	estimate	18.674085	5.954055	6.8986	30.44958	134.73	0.0021

Figure 10. Pooled Results by PROC MIANALYZE.

IMPLEMENTATION OF MI USING SAS UNDER ARBITRARY MISSING DATA PATTERN AND MAR

What if the missing data pattern is arbitrary in the lung cancer example (Figure 11)? For data sets with arbitrary missing patterns, there are two methods that can be used to impute missing values. The first method is a Markov chain Monte Carlo (MCMC) method that assumes multivariate normality, and the second is a fully conditional specification (FCS) method that assumes a joint distribution exists for all variables. The MCMC method can be used for continuous variables, either to impute all the missing values or to impute enough missing values to create a monotone missing pattern. This enables the use of other imputation methods such as the regression method. A different set of covariates can also be specified for each imputed variable.

SUBJID	TRTPN	STRATUM	BASELINE	WEEK_24	WEEK_52	Freq	Percent
X	X	X	X	X	X	174	71.02
X	X	X	X	X	.	34	13.88
X	X	X	X	.	X	9	3.67
X	X	X	X	.	.	13	5.31
X	X	X	.	X	X	9	3.67
X	X	X	.	X	.	6	2.45

Figure 11. Arbitrary Missing Data Pattern of Transposed ADQLQ. The 'X' indicates non-missing data while the '.' indicates missing data.

In case of arbitrary missing data pattern, the following code can be applied to impute the few missing values causing the non-monotone pattern to make the data pattern monotone before the MONOTONE REGRESSION implementation. The VAR statement lists the variables in the order of the desired monotone missing data pattern (i.e., for this example, if a subject is missing baseline score, they are missing week 24 and 52 in order to maintain the monotone missing pattern; if a person is missing week 24 score, then they are missing week 52 in order to maintain the monotone missing pattern).

```
proc mi data=ghs2 seed=113241 out=miout nimpute=3;
    mcmc impute=monotone;
    var baseline week_24 week_52;
run;
```

In the above code, only continuous variables are listed, as the MCMC approach requires both the outcome variable and covariates to be continuous variables. To impute missing values for both continuous and classification variables in data sets with arbitrary missing patterns, FCS methods can be used. The regression and predictive mean matching methods can be used to impute missing values for continuous variables, while the logistic regression method is used for classification variables with a

binary, nominal, or ordinal response, and the discriminant function method is used for classification variables with a binary or nominal response. The following code can be used to replace the PROC MI procedure in the previous example, if the FCS method is decided to be used.

```
proc mi data=ghs2 seed=113241 out=miout nimpute=3;
    fcs regression (week_52=baseline week_24 trtpn stratum/details);
    var trtpn stratum baseline baseline week_24 week_52;
run;
```

TIPPING POINT ANALYSIS – MI FOR SENSITIVITY ANALYSIS UNDER MNAR

The MI approaches presented above assumes that missing data follows MAR mechanism, and such an assumption often makes sense for the primary analysis. However, as discussed above, the observed data can never tell whether this assumption is correct. Therefore, to assess robustness, sensitivity analyses are recommended in the analysis of clinical trials. The tipping point approach has emerged as a popular tool for conducting sensitivity analysis when the missing data is MNAR. This approach involves a progressive stress-testing procedure to gauge the magnitude of departures from the MAR assumption that would be required to overturn the conclusions drawn from the primary analysis. Specifically, the robustness of the results is assessed by determining whether implausible deviations from MAR would be necessary to shift the results from being statistically significant (p-value≤0.05) to statistically insignificant (p-value >0.05). By demonstrating that the results are robust to deviations from the MAR assumption, greater confidence can be placed in the conclusions drawn from statistical methods that assume MAR. It is worth noting that the tipping point approach is not intended for use as a primary analysis method but is designed exclusively for sensitivity analysis purposes.

The tipping point approach can be viewed as a specific application of MI with pattern-mixture model approach. To implement the tipping point analysis, the following steps are taken, with the first three steps following the standard MI method: 1) create M multiple imputed data sets using selected MI algorithm; 2) analyze the M multiple imputed data sets by using standard SAS procedures; 3) pooling the results from the M complete data sets for statistical inference; 4) repeat step 1 to generate multiple imputed data sets with a specified shift parameter that adjusts the imputed values only for observations in the treatment group (i.e., force the imputed values from experimental group to be ‘bad’); 5) repeat step 2 for the M imputed data sets with the shift parameter applied; 6) repeat step 3 to obtain the p-value and check whether it is still ≤0.05; 7) repeat steps 4-6 with a more stringent shift parameter until the p-value becomes >0.05.

In the following code, the MNAR statement is added in to the PROC MI code from the previous section of IMPLEMENTATION OF MI USING SAS UNDER MONOTONE MISSING DATA PATTERN AND MAR. This statement tells PROC MI that, after the MI imputation of missing week 24 and 52 scores, do as follows but only for the Chemotherapy + Additive group [ADJUSTOBS=(TRTPN=’1’)]: multiply the imputed value by 1 (SCALE=1) and then lower the imputed week 24 value by 2.5 (SHIFT=2.5), and multiply the imputed value by 1 (SCALE=1) and then lower the imputed week 24 value by 3 (SHIFT=3). In other words, we request the PROC MI to worsen the imputed OHS scores for the experimental arm by lowering them 2.5 units for week 24 score and by 3 units for week 52 score. A series of shift values, usually determined by clinicians, are tested. If the treatment effect is qualitatively maintained for the range of shift values that are considered to be clinically plausible, then the findings are considered to be robust.

```
proc mi data=ghs2 seed=113241 out=miout nimpute=3 noprint;
    class trtpn stratum;
    monotone regression (week_24=baseline trtpn stratum/details);
    monotone regression (week_52=baseline week_24 trtpn stratum/details);
    mnar adjust (week_24 / scale=1 shift= -2.5 adjustobs=(trtpn='1'))
         adjust (week_52 / scale=1 shift= -3 adjustobs=(trtpn='1'));
    svar trtpn stratum baseline week_24 week_52;
run;
```

CONCLUSION

To summarize, MI is a method that replaces missing values with plausible values, providing a means to

address the inherent uncertainty in imputed values. Traditional approaches, such as complete case analysis, mean imputation, and single imputation, may lead to biased estimates of standard errors, incorrect tests of statistical significance, and other inaccuracies. As we have discussed, the COVID-19 outbreak continues to have major impact on planned and ongoing clinical trials, leading to exacerbated missing data issues. As most data that are missing due to pandemic reasons may be argued to be MCAR or MAR, especially if missingness is due to structural reasons, it may be plausible to use standard MI procedure as primary analysis. However, the results from primary analysis should be carefully examined sensitivity analyses under alternative assumptions for the missing data distribution to assess the robustness of conclusions.

REFERENCES

[1] U.S. Department of Health and Human Services Food and Drug Administration . FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic. 2020.

[2] European Medicines Agency . Implications of coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials. 2020.

[3] Medicines and Healthcare products Regulatory Agency . Managing clinical trials during Coronavirus (COVID-19) 2020.

[4] Little, R. J., & Rubin, D. B. (2019). Statistical analysis with missing data (Vol. 793). John Wiley & Sons.

[5] 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(3), 343.

[6] US Food and Drug Administration. (2018). ICH E9 'Statistical Principles for Clinical Trials.'. US Food and Drug Administration. https://www. fda. gov/regulatory-information/search-fda-guidance-documents/e9-statistical-principles-clinical-trials (accessed Jun. 10, 2022).

[7] Berglund, P., & Heeringa, S. G. (2014). Multiple imputation of missing data using SAS. SAS Institute.

[8] SAS Institute Inc. 2018. SAS/STAT® 15.1 User's Guide: High-Performance Procedures. Cary, NC: SAS Institute Inc.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Xi Qian
BioPier, Inc.
xqian@biopier.com

Chengfei Lu
BioPier, Inc.
clu@biopier.com