Paper PP13

Is Complexity Worth It? A Case Study Comparing Multiple Imputation

Elnaz Ghadimi, Cytel, Toronto, Canada
David Bushnell, Cytel, Maryland, US
Khalif Halani, Cytel, Vancouver, Canada

ABSTRACT

Missing data is a pervasive challenge in clinical research and can introduce bias or reduce statistical power when not handled appropriately. Multiple imputation (MI) is widely used to address missingness, yet the practical benefits of more complex MI strategies over simpler approaches remain uncertain. This study compared seven methods for handling missing data within a repeated measures framework applied to a single clinical dataset, including variations of the Treatment Policy approach, Missing at Random (MAR) and Missing Not at Random (MNAR) assumptions, an observed case analysis, and a two-step MI procedure. Across all approaches, treatment effect estimates were directionally consistent, with no meaningful deviations from the primary analysis. The results suggest that more complex MI procedures do not materially alter substantive conclusions in this context. These findings support the use of simpler, transparent imputation strategies when the missing data mechanism is not strongly informative.

INTRODUCTION

Missing data is a common challenge in clinical research and can bias results or reduce statistical power when not handled appropriately. When missingness arises after an intercurrent event—as defined by the estimand framework—the chosen imputation approach must follow a prespecified strategy. Multiple imputation (MI) is widely used to address missingness, yet the practical advantages of more complex MI strategies over simpler approaches remain uncertain. When the proportion of missingness is moderate and the data structure is relatively straightforward, simpler methods may perform just as well as more elaborate ones. Missingness is typically classified into three mechanisms: Missing Completely at Random (MCAR), MAR, and MNAR[1].

This study compares several MI approaches within a repeated-measures framework applied to the same dataset to evaluate whether increasing methodological complexity meaningfully affects the estimated treatment effect.

METHODS

This study evaluates the effect of the investigational medications, administered as add-on therapy to a standard anti-condition medication regimen, in subjects with a cardiovascular condition. Participants were randomized in a 1:1:1 ratio to receive daily doses of Placebo, Investigational Product Arm 1 (Arm 1), or Investigational Product Arm 2 (Arm 2).

The primary endpoint is the change from baseline in the average cardiovascular measure at Month 12. The study includes one intermediate assessment at Month 5. A total of 298 subjects participated. Among them:

- 3 subjects (1%) had missing measurements only at Month 5
- 26 subjects (9%) had missing measurements only at Month 12
- 21 subjects (7%) had missing measurements at both Month 5 and Month 12

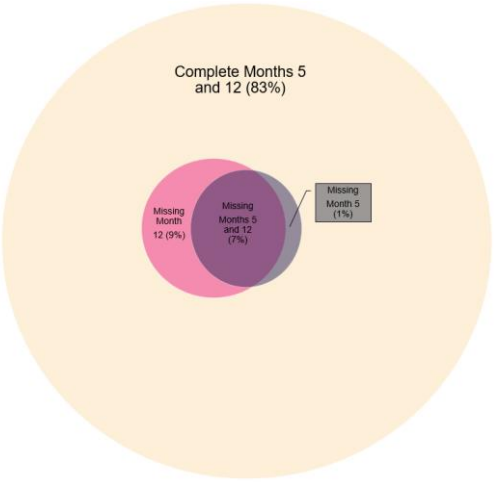

Figure 1 Number of subjects with and without missing data at Months 5 and 12

Among those who have a missing Month 12 measurement, almost 80% of them were in the Investigational Product Arms (Figure 2.)

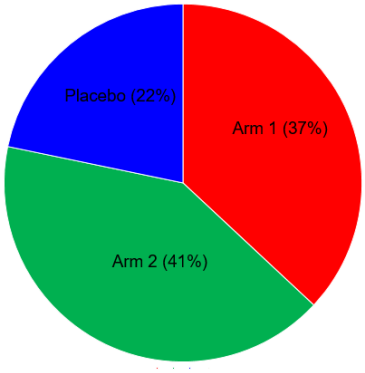

Figure 2 Distribution of subjects with missing measurement at Month 12 across treatment arms

To evaluate the treatment effect, the change from baseline in the cardiovascular measure at Month 12 was calculated for each subject. Missing Month 12 efficacy assessments were imputed using MI, applying one of the seven approaches described below.
The primary estimand for this study requires implementing MI with different imputation reference distributions depending on the type of intercurrent event: MAR for on-treatment incidental missing, retrieved dropout or MNAR for subjects who experienced an intercurrent event or discontinued, and a hypothetical worst-case imputation (worst 5%) for subjects who died.
Sensitivity analyses are used to evaluate the robustness of the assumptions underlying the MAR, retrieved-dropout, and MNAR imputations. While retrieved-dropout imputations are generally less conservative than MNAR, their usefulness depends on the proportion of subjects with retrieved data relative to the overall amount of missingness.
Seven approaches were applied to the dataset to handle missing data, each one using 100 multiply imputed datasets:

- Approach 1: Treatment Policy analysis, incorporating an interaction term for treatment-by-last-on-treatment visit.
- Approach 2: Treatment Policy analysis imputing missing values using only off-treatment retrieved dropouts within each treatment arm.
- Approach 3: Reference-Based Treatment Policy approach as outlined by Rogers et al. [2].
- Approach 4: MAR-based multiple imputation.
- Approach 5: NMAR-based multiple imputation.
- Approach 6: Observed-case Analysis of covariance (ANCOVA) with no imputation.
- Approach 7: Two-step MI strategy: intermediate missing data imputed using Markov Chain Monte Carlo (MCMC); patients with intercurrent events imputed using the placebo distribution; active-treatment patients without intercurrent events imputed within their own treatment distribution.

RESULTS

Across Approaches 1–3, all variations of the Treatment Policy framework produced directionally similar estimates with no meaningful shifts in treatment effect.
The MAR-based (Approach 4) and NMAR-based (Approach 5) sensitivity analyses also yielded results consistent with the primary analysis, suggesting limited sensitivity to departures from the MAR assumption.
Approach 6, the observed-case ANCOVA, supported the primary findings despite relying on stronger assumptions about the missing-data mechanism.
Approach 7, the two-step MI method, produced results aligned with the other approaches, further reinforcing the stability of the treatment effect across diverse imputation strategies.
Figure 3 presents the least-squares (LS) means for each approach across all treatment arms.

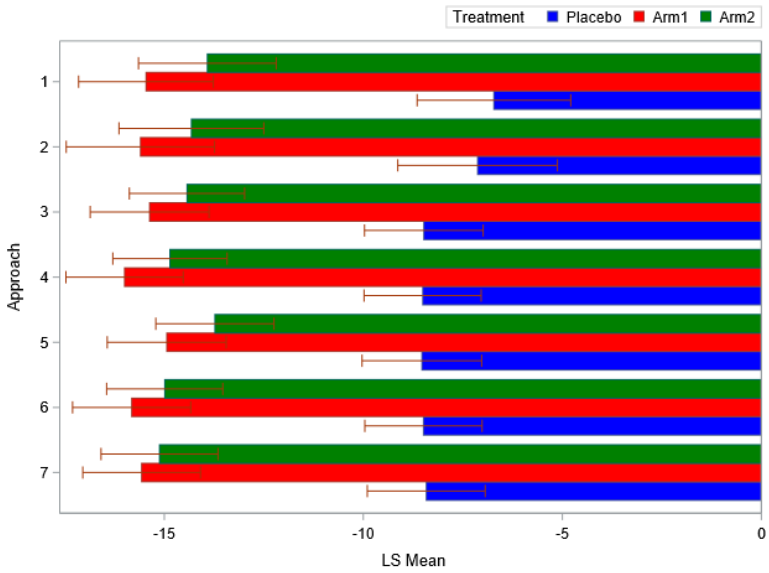

Figure 3 LS Mean (+/-SE) across all treatment arms

The LS mean for the placebo group is almost the same across Approaches 3 to 7, which indicates that the placebo response was stable and did not contribute to differences seen between approaches. When comparing against placebo (Figure 4), the LS mean differences did not statistically vary greatly.

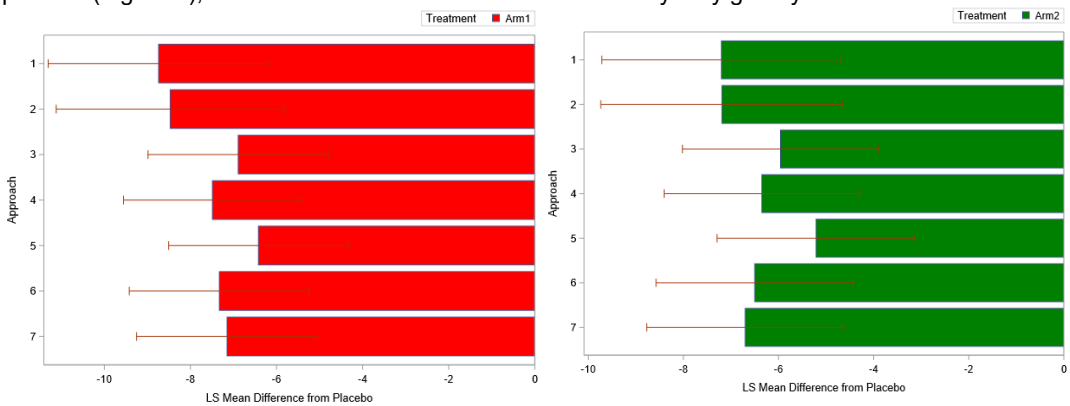

Figure 4 LS Mean difference from Placebo (+/-SE) for Arms 1 and 2

Overall, the results are consistent across methods. As illustrated in the figures, none of the approaches yield meaningfully different estimates, and no statistically significant differences are observed among them. This suggests that the, in this particular case, the choice of approach does not materially influence the conclusions drawn from the data.

CONCLUSION

Across all seven imputation approaches, the primary analysis produced consistent results, and the sensitivity analyses supported the plausibility of both the MAR and MNAR assumptions. Importantly, the more complex MI strategies did not materially change the overall conclusions.

These findings indicate that simpler and more transparent imputation methods—or even observed case analyses—may be adequate when the missing-data mechanism is not strongly informative, although the impact is now known prior to unblinding. The strong agreement across methods reinforces their suitability for routine analytical workflows and provides confidence in the robustness and stability of the study's conclusions.

Because all seven approaches produced similar results, it is reasonable to question whether such a large set of methods is always necessary. Running multiple MI strategies can be time-consuming and computationally demanding, and using so many approaches may also require additional explanation to reviewers, especially when the rationale for including all seven is not immediately clear. Although each method is well described in the literature, applying them all together can make the analysis harder to communicate and may create the impression of unnecessary complexity. These results highlight the value of considering whether a smaller, more focused set of approaches could provide the same insights with greater efficiency while still supporting a thorough and credible assessment of missing-data assumptions.

REFERENCES

1 Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med. 1991 Apr;10(4):585-98. doi: 10.1002/sim.4780100410. PMID: 2057657.
2 https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data#dia-missing-data Jonathan Bartlett, James Roger, James Carpenter

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:
Author Name: Elnaz Ghadimi
Company: Cytel
Address: 675 Massachusetts Ave, Cambridge, MA 02139 USA
Email: Elnaz.ghadimi@cytel.com