Handling Missing EQ-5D-5L data in Clinical Trials – A Simulation Study and Empirical Application

Juliana Schmidt[1], Ângela Jornada Ben[2,3], Johanna M. van Dongen[2], Wolfgang Greiner[1], Judith E. Bosmans[2]
[1] School of Public Health, Bielefeld University, Bielefeld, Germany
[2] Department of Health Sciences, Vrije Universiteit Amsterdam, the Netherlands
[3] Amsterdam Institute for Global Health & Development

Background and aim

- The EQ-5D-5L (5L) has a considerable risk of missing data, potentially leading to biased estimates when not handled adequately (1,2).
- Multiple imputation strategies (MI) and longitudinal linear mixed models (LLM) have shown to be promising (1,3).
- MI can be divided into two broad categories:
    - Joint Modelling (JM)
    - Fully Conditional Specification (FCS) (4).
- It remains unclear which MI approach best suits longitudinal 5L data and whether to impute missing 5L data at the response or index level.
- We explored the performance of various methodological strategies to handle 5L data in clinical trials using simulated and empirical data.
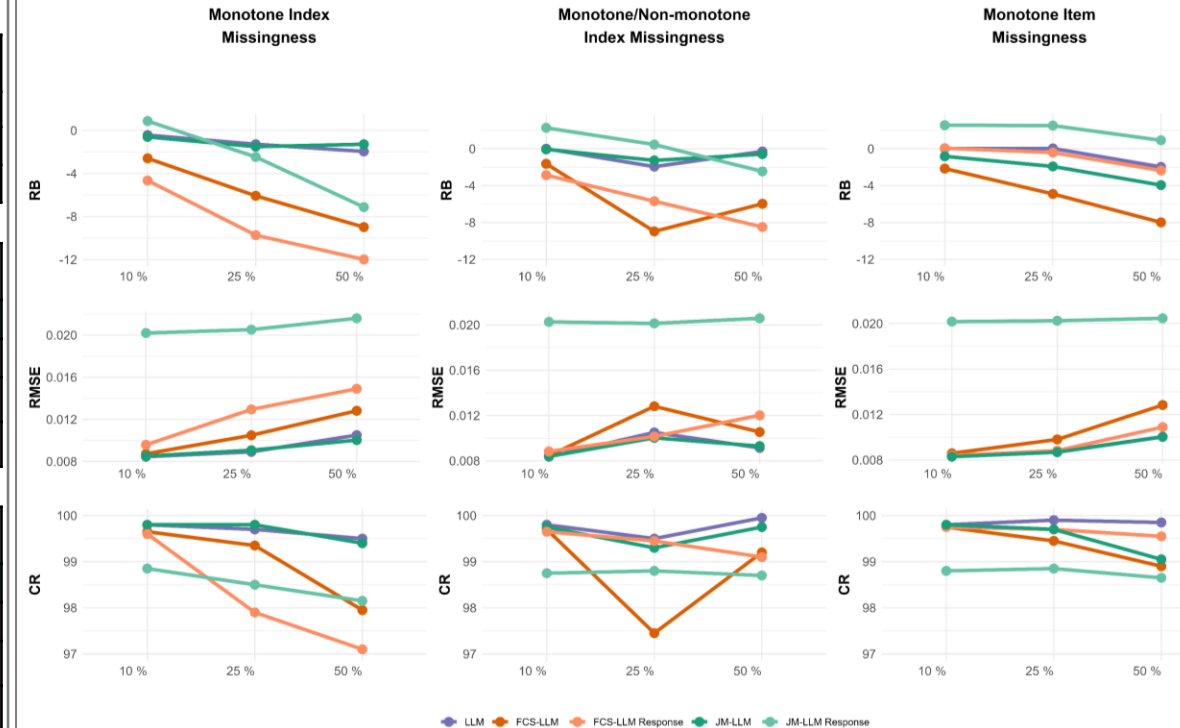
Methods

- 2,000 complete data sets were simulated based on empirical trial data, including 5L responses at baseline and two time points.
- Missing data was introduced
    - at rates of 10%, 25%, and 50%
    - under Missing At Random (MAR), Missing Completely At Random (MCAR), or Missing Not at Random (MNAR).
- We simulated different missingness patterns (Figure 1).
- Six strategies were evaluated (Table 1):
- Performance was assessed using empirical bias (EB), root-mean-square error (RMSE), and coverage rate (CR).
- The six strategies were applied to empirical trial data.

Table 1: Methodological strategies

1	Complete-case analysis with ordinary least squares (CCA-OLS)
2	LLM alone
3	FCS with LLM at index-level (FCS-LLM index)
4	FCS with LLM at response-level (FCS-LLM response)
5	JM with LLM at index-level (JM-LLM index)
6	JM with LLM at response-level (JM-LLM response)
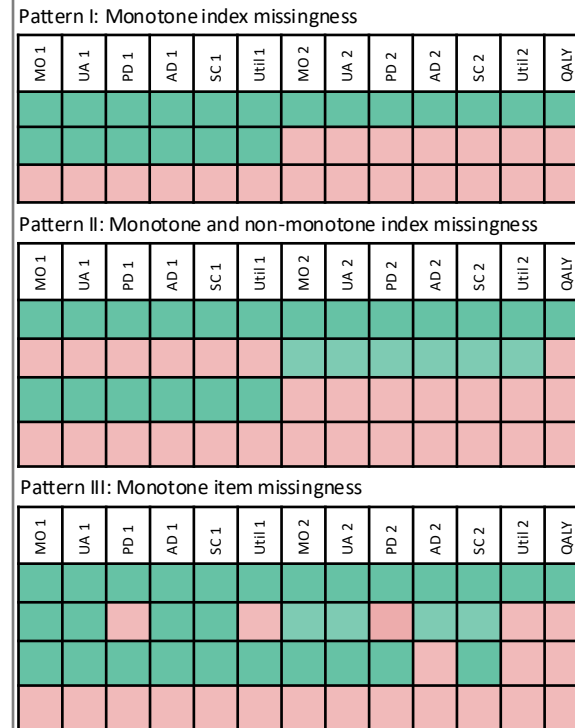
Figure 1: Simulated missingness pattern

Figure 2: Performance measure of the methodological strategies under the MAR assumption

Results

- CCA-OLS resulted in highly biased estimates.
- LLM alone had similar and oftentimes slightly smaller EBs and RMSEs than JM-LLM at 10%, 25%, and 50% missing data, irrespective of missingness pattern.
- FCS-LLM was outperformed by LLM alone and JM-LLM in all scenarios.
- Among MI strategies, JM-LLM had smaller EBs and RMSEs than FCS-LLM.
- In all scenarios, all methods were slightly overfitted.
- MI performed worse when imputing at the response-level rather than on the index-level.
- The MCAR results aligned with the MAR findings.
- The MNAR results exhibited larger EBs and lacked a clear pattern indicating a superior or inferior strategy (results not shown on the poster).
- In the empirical trial data, estimates from CCA-OLS deviated considerably from those of the other strategies, whereas estimates derived using LLM and MI at index-level were highly similar. This similarity was not seen for MI at the response-level.

Conclusion

- CCA-OLS yields biased estimates.
- LLM seems to perform slightly better than MI strategies for handling missing 5L data in trial-based evaluations, provided baseline data are complete.
- If baseline data is not complete, JM-LLM is preferred, among MI strategies.
- Index-level imputation is advised.

1 Rösel I, Serna-Higuita LM, Al Sayah F, Buchholz M, Buchholz I, Kohlmann T, et al. What difference does multiple imputation make in longitudinal modeling of EQ-5D-5L data? Empirical analyses of simulated and observed missing data patterns. Qual Life Res. 2022;31:1521–32. doi:10.1007/s11136-021-03037-3.
2 Grady KL, Jones PG, Cristian-Andrei A, Naftel DC, Myers S, Dew MA, et al. Causes and Consequences of Missing Health-Related Quality of Life Assessments in Patients Who Undergo Mechanical Circulatory Support Implantation: Insights From INTERMACS (Interagency Registry for Mechanically Assisted Circulatory Support). Circ Cardiovasc Qual Outcomes. 2017;10:e003268. doi:10.1161/CIRCOUTCOMES.116.003268
3 Ben ÂJ, van Dongen JM, Alili ME, Heymans MW, Twisk JWR, MacNeil-Vroomen JL, et al. The handling of missing data in trial-based economic evaluations: should data be multiply imputed prior to longitudinal linear mixed-model analyses? Eur J Health Econ 2022. doi:10.1007/s10198-022-01525-y
4 van Buuren S. Flexible imputation of missing data. Boca Raton: CRC Press; 2012.