



PHUSE 2026

Estimand-Aligned Multiple Imputation Strategies in CNS Trials with Complex Intercurrent Events

Ian Lees, Veera Thota

Outline of presentation

1. What is Multiple Imputation (MI)
2. Intercurrent Events (ICEs)
3. Use Case –Study Outline and Primary Endpoint
4. MI Approach & Examples
5. Programming Challenges & Solutions
6. Conclusions

SIDENOTE:

- We focus on the use of MI with ICEs in central nervous system (CNS) **clinical trials** (interventional studies)

Abbreviations

- **ICE:** Intercurrent Event
- **MAR:** Missing at Random
- **MCMC:** Markov Chain Monte Carlo
- **MI:** Missing Imputation
- **MMRM:** Mixed Model for Repeated Models

What is MI

- MI fills in missing values by generating plausible numbers derived from distributions of and relationships among observed variables in the data set
- MI inference involves three distinct phases
 - The missing data are filled in m times to generate m complete data sets
 - The m complete data sets are analyzed by using standard procedures
 - The results from the m complete data sets are combined for the inference

ICEs

- An ICE is defined as an event occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest.
- Examples of ICEs include (but are not limited to):
 - Discontinued treatment
 - Withdrew consent
 - Had an adverse event
 - Lost to follow up

Use Case –Study Outline and Primary Endpoint

- This study was a Randomized, Placebo-Controlled, Double-Blind Study in Subjects with Mild to Moderate Alzheimer's Disease
- Primary Endpoint was to evaluate the efficacy of treatment in subjects not on background acetylcholinesteraseinhibitors (AChEIs) through the Global Statistical Test (GST)
- GST was defined as the composite of cognition and function, calculated as the average of the two change from baseline z-scores between the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog₁₁) and the Alzheimer's Disease Cooperative Study – Activities of Daily Living, 23-item version (ADCS-ADL23) score

Use Case – ICE-1 and ICE-2

- **Primary Analysis**

- ICE-1: Discontinuation of treatment due to (1) an AE or (2) lack of efficacy. This event (1) was defined as selection of “Any Adverse Event” as a reason for study discontinuation/withdrawal, even if other reasons are selected. The (2) lack of efficacy was determined via the protocol deviation report/listing.
- ICE-2: Discontinuations due to other reasons other than (1) an AE or (2) lack of efficacy.

- **Sensitivity Analysis**

- ICE-1: This ICE was handled using a “control-based” hypothetical strategy assuming that patients would have switched to the control treatment (i.e., placebo) starting from the occurrence of the ICE to the end of the 26-week treatment period.
- ICE-2: This ICE was handled using a treatment policy strategy implying that all outcomes following this ICE will be included in the analysis regardless of the actual treatment, with a possibly different implementation of the imputation model depending on the availability/sparsity of the observed outcomes following the ICE.

MI Approach

1. Creation of ICE flags
2. Imputed non-monotone missing data unrelated to ICEs (imputed using MCMC to create monotone missing)
3. Post-ICE1 endpoint values were set to missing for ICE-1 patients though data had been collected
4. ICE-1 (MI using control-based strategy)
5. ICE-2 (MI using treatment policy strategy)
6. Combined all imputed datasets into a final analysis dataset
7. Performed the final analysis and generated results

ICE flags definitions:

1. ICE1FL are those subjects who discontinued due to either Adverse event or Lack of Efficacy.
2. ICE2FL are those subjects who discontinued due to reasons other than mentioned in the ICE1FL.

Example Code to Create Monotone Data Set

Steps:

- Prepare analysis dataset with required population and covariate flags
- Restructure data to support visit-level imputation
- Apply MCMC-based multiple imputation as seen in code below
- Convert intermittent (non-monotone) missingness into a monotone pattern

```
PROC MI data = mcmc_in out = mcmc_out nimpute = 1000 seed = 2024;  
MCMC impute=monotone;  
BY <Treatment> <Additional Variables>;  
VAR <Other Covariates>;  
RUN;
```

- The default impute value is 5 which is generally insufficient for proper MI interference. In practice and expert consensus recommends more imputations (e.g., 20–100 or more) to ensure stable inference

Example Code to Create Control-Based Data Set Per ICE

Steps:

- Read in data set with correct ICE-1 flags isolated
- Prepare analysis dataset with ICE-1 subjects identified
- Ensure monotone missing data structure
- Apply monotone regression MI under a controlled MNAR framework using Pattern-Mixture Model (PMM) as shown in code below
- PMMs provide flexible framework for sensitivity analysis which formulate assumptions regarding missing data in a transparent and clinically interpretable manner
- Estimate imputation models using reference-arm observations

```
proc mi data=ice1_in out=ice1_out nimpute=1 seed=2024;  
  by _Imputation_;  
  class <Any Categorical Covariates>;  
  monotone reg(/details);  
  var <Covariates>;  
  mnar model<Visits to be analyzed in MMRM>/ modelobs= (<Treatment>='1');  
run;
```

Example Code to Create Treatment Policy Strategy Per ICE

Steps:

- Prepare analysis dataset with ICE-2 subjects identified
- Ensure monotone missing data structure
- Apply monotone regression imputation within each imputation replicate

```
PROC MI data=DATA_ICE2_IN out=DATA_ICE2_OUT nimpute=1 seed=2024;  
  CLASS <Any Categorical Covariates>;  
  MONOTONE REG;  
  VAR <Covariates>;  
  BY IMPUTATION_;  
RUN;
```

Example Code – MMRM for Primary Endpoint

Steps:

- Prepare analysis-ready datasets from imputed outputs.
- Apply MMRM separately within each imputation replicate.
- Estimate treatment effects and visit-specific contrasts.
- Repeat analysis for each ICE-specific dataset

```
proc mixed data = mixed_in;  
  by _imputation_ ;  
  class <Any Categorical Covariates>;  
  model <Endpoint> = <Covariates> / ddfm = kr;  
  repeated <Visit>/ type=un subject= <Subject ID>;  
  lsmeans <Treatment>*<Visit> / pdiff cl alpha=.05;  
  lsmestimate <Treatment>*<Visit>  
    "Drug_vs_Placebo_Week_X" [x, x x] [x, x x],  
    "Drug_vs_Placebo_Week_X" [x, x x] [x, x x],  
    "Drug_vs_Placebo_Week_X" [x, x x] [x, x x],  
    "Drug_vs_Placebo_Week_X" [x, x x] [x, x x],  
    "Drug_vs_Placebo_Week_X" [x, x x] [x, x x]  
  / cl;  
  ods output lsmeans=lsmeans lsmestimates=lsmestimates;  
run;
```

Example Code - Final Results for Primary Endpoint

Steps:

- Prepare parameter estimates from each imputed dataset.
- Combine estimates and standard errors across imputations.
- Derive overall treatment effect estimates, confidence intervals, and p-values.

```
proc mianalyze parms=lsmeans;  
  by <Visit> <Treatment>;  
  modeleffects <Treatment>*<Visit>;  
  ods output ParameterEstimates=lsmeans_params;  
run;
```

Programming Challenges & Solutions

- **MI non-convergence** → simplified covariates in MCMC step
- **Complex ICE logic** → modular datasets by ICE type
- **Traceability** → explicit ADaM derivations aligned with SAP
- **Validation** → distribution checks of observed vs imputed values

Conclusions

- ICEs are becoming more prevalent in clinical trials
- Need to take time and careful consideration when setting up the MI procedures
- Running Proc MI is not the end of the story
- Review datasets and outputs throughout to make sure MI worked properly

References

Benton, A. L., Hamsher, K., Rey, G. L., & Sivan, A. B. (1994).
Multilingual Aphasia Examination (3rd ed.). Iowa City, IA: AJA Associates.

Cui, L., Hung, H.M., Wang, S.J. (1999).
Modification of sample size in group sequential clinical trials. Biometrics. 55, 853-857.

Cummings J.L., Mega M., Gray K., Rosenberg-Thompson S., Carusi D.A., Gornbein J. (1994).
The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 44: 2308-2314

Folstein M.F., Folstein S.E., McHugh P.R. (1975).
“Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. J Psychiatric Research 12 (3), 189-198.

Galasko D., Bennett D., Sano M., Ernesto C., Thomas R., Grundman M., et al. (1997).
An inventory to assess activities of daily living for clinical trials in Alzheimer’s disease. The Alzheimer’s disease Cooperative Study. Alzheimer’s Disease and Associated Disorders, 11(S2), S33-S39.

Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. (1982).
A new clinical scale for the staging of dementia. Br J Psychiatry 140, 566-572.

Kordzakhia, G., Dmitrienko, A., Ishida, E. (2018).
Mixture-based gatekeeping procedures in adaptive clinical trials. Journal of Biopharmaceutical Statistics. 28, 129-145.

McKhann, G.M., et al. (2011).
The diagnosis of dementia due to Alzheimer’s disease: Recommendations from the National Institute on Aging – Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. Alzheimer’s & Dementia: The Journal of the Alzheimer’s Association 7, 263-269.

O’Brien, P.C. (1984).
Procedures for comparing samples with multiple endpoints. Biometrics. 40, 1079-1087.

References

Rosen WG, Mohs RC, Davis KL (1984).

A new rating Scale for Alzheimer's disease. Am J Psychiatry 141:1356-1364.

Schneider LS, Olin JT, Doody RS, Clark CM, Morris JC, Resiberg B, et al. (1997)

Validity and reliability of the Alzheimer's Disease Cooperative Study - Clinical Global Impression of Change. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord. 11(2):S22-S32.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006).

Compendium of neuropsychological tests: Administration, norms, and commentary (3rd ed., p. 502). Oxford: Oxford University Press.

The EuroQol Group (1990).

"EuroQol - a new facility for the measurement of health-related quality of life". Health Policy (Amsterdam, Netherlands). 16 (3):199-208.

Wassmer, G., Brannath, W. (2016).

Group Sequential and Confirmatory Adaptive Designs in Clinical Trials. New York: Springer.

Wimo A, Wetterholm AL, Mastey V, Winblad B. (1998).

Evaluation of the resource utilization and caregiver time in Anti-dementia drug trials – a quantitative battery. In: Wimo A, Karlsson G, Jönsson B, Winblad B (eds). The Health Economics of Dementia, Wiley's, London, UK.

Zarit, SH, Reeve, KE, Back-Peterson, J. (1980).

Relatives of the impaired elderly: correlates of feelings of burden. The Gerontologist 20:649-655.

SAS/Stat(R) 9.3 user's guide. (2011, July 15). https://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#statug_mi_sect017.htm

Li, P., Stuart, E. A., & Allison, D. B. (2015, November 10). *Multiple imputation: A flexible tool for handling missing data*. JAMA. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4638176/>

Team, S. C. (2026, January 15). *A guide to estimands in clinical trials*. Quanticate. <https://www.quanticate.com/blog/estimands-clinical-trials>

1 Ich E9(R1) estimands and sensitivity analysis in clinical trials. (n.d.). [https://database.ich.org/sites/default/files/E9\(R1\)%20Training%20Material%20-%20PDF_0.pdf](https://database.ich.org/sites/default/files/E9(R1)%20Training%20Material%20-%20PDF_0.pdf)



PHUSE 2026

Thank you

Ian Lees: ilees@mmsholdings.com

Veera Thota: vthota@mmsholdings.com