Other statistical software for continuous longitudinal endpoints:

SAS macros for multiple imputation

James Roger
London School of Hygiene & Tropical Medicine, London.

Joint EFSPI and BBS virtual event, 8 December 2022.
Addressing intercurrent events: Treatment policy and hypothetical strategies.

Setting the scene

- Same background as previous slides by Marcel Wolbers & Alessandro Noci.

I will concentrate on:

- The history before using "jacknife with conditional mean imputation" was proposed.
  An open discussion is needed on suitability of classic type 1 error as measure of "confidence".

- A series of SAS macros to carry out these Multiple Imputations available from the DIA working Party TAB on the LSHTM Missing Data web site.
  https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data

Emergence of Reference-based imputation 1

200x Importance of Missing not at Random (MNAR) in data from
      clinical trials.
      As evidenced by EMA Guideline on Missing Data in Confirmatory
      Clinical Trials (Agreed by Working Party, April 2010)

2008 Carpenter & Kenward: Monologue; Missing data in clinical trials — a
      practical guide.
      Sensitivity analysis to MAR using Delta.

Emergence of Reference-based imputation 1

2008 Roger(PSI presentation):
      Sensitivity analysis based on "ITT-like imputation".
      Start of Reference-based imputation using MI as a Sensitivity analysis.

   ▶ The %MIWithD SAS macro (Multiple Imputation to handle
     withdrawals).
   ▶ Based on MCMC statement in proc MI.

2012 The GSK 5 macros for Reference based Imputation (PSI conference).
      Based on very early version of proc MCMC.

   ▶ A flexible system for Reference-based imputation.
   ▶ GSK willing to make code openly available and support it.
     [James Roger, Carly Donnovan & Tom Drury]

Early references

2013 Carpenter, Roger & Kenward: Analysis of longitudinal trials with
protocol deviation: a framework for relevant, accessible assumptions,
and inference via multiple imputation.

Jump-to-Reference (J2R), Copy-increment-from-Reference (CIR) etc.
along with a marginal RM model (unstructured covariance by treatment
arm).

2014 Ayele, Lipkovich, Molenberghs, Mallinckrodt: A
multiple-imputation-based approach to sensitivity analyses and
effectiveness assessments in longitudinal clinical trials.

Placebo Multiple Imputation (pMI) using MONOTONE statement in Proc
MI for conditional RM model.
Named Copy-Reference (CR) by Carpenter et al based on marginal
model.

Two algorithms for MI with continuous outcome measured repeatedly

- Stepwise
  Visit-by-Visit, each univariate Outcome is regressed conditional upon previous data.

- Marginal Multivariate
  A full multivariate Normal Repeated Measures model is fitted. Allows more complex models.

Conditional (Stepwise) algorithm

- The MONOTONE statement in proc MI is a classic example.

- If the baseline covariates remain completely unchanged from regression to regression this is equivalent to the marginal RM model (unstructured covariance) with these baseline covariates crossed with time.

- Regressing on previous residuals rather than previous observed values gives correct marginal model, as in J2R and CIR, despite covariates changing with time.

- The %MIStep and %MIStepWrap SAS macros allow sequential fitting based on previous residuals.

%MIStep and %MIStepWrap SAS macros and Off-treatment modelling

- Written to confirm the equivalence of regressing on residuals to the marginal model.

- Generates imputed data set for reference-based, or off-treatment models.

- Roger 2017 PSI conference poster, "Joint modelling of On-treatment and Off-treatment data".

- Polverejan & Dragalin (2019) use the macro to impute using off-treatment data.

Markov Chain Monte Carlo (MCMC) and Multiple Imputation

- When fitting a Bayesian model, random missing outcomes can be treated as if they are Bayesian parameters.

- When the imputation model matches the estimation model, a sample from the posterior can be used for imputation.

- For J2R, CIR etc. (but not CR) or off-treatment models, imputed values can be generated driectly.

- For reference-based imputation fix the design matrix for the missed values (e.g. switch to reference).

- For Off-Treatment data simply build the required design matrix for all subject/visit combinations.

Conjugate prior based macros

- With a single covariance matrix within each subject, conjugate priors can be used for all parameters allowing direct sampling of linear model, covariance and missing-value parameters.

- The %mymcmc and %RMConj and %RMCONJPlus SAS macros take advantage of this delivering independent imputations with minimal thinning using proc MCMC.

- A second imputation stage in (RMConjPlus) handles separate imputation models such as CR and also Carpenter et al separate covariances for each arm.

- The %RefBTVT macro builds the required design matrix for Reference based imputation.

- The %mymcmc macro allows extension to models such as the Diggle-Kenward MNAR model.

Other SAS MCMC based routes

- The SAS BGLIMM procedure fits Bayesian generalized linear models and handles missing data automatically.

- The %BGI SAS macro extracts sampled missed values and builds a multiple imputed data set.

- Useful for non-Normal data.

Analyzing imputed data set

- The %MIAnalyze SAS macro fits a univariate linear model to a multiple imputed data set and combines results using Rubin's rules like proc MI.

- BY= allows for analysis at repeated visits (not MMRM).

- Adds an estimate of Monte Carlo error for each treatment difference.

- Not needed for GSK 5 macros as a complete system.

Now . . .

- Availability of formally validated Stata and R code for RBI is timely.

- My personal choice for research . . .
  The %RMConjPlus SAS macro as it is fast and flexible.

References

Ayele, B. T., Lipkovich, I., Molenberghs, G., & Mallinckrodt, C. H. (2014). A multiple-imputation-based approach to sensitivity analyses and effectiveness assessments in longitudinal clinical trials. Journal of Biopharmaceutical Statistics, 24(2), 211 228. https://doi.org/10.1080/10543406.2013.859148

Carpenter JR, Roger JH & Kenward MG. (2013) Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. J Biopharm Stat. 23(6):1352-1371.

Elena Polverejan & Vladimir Dragalin (2019)(: Aligning Treatment Policy Estimands and EstimatorsA Simulation Study in Alzheimers Disease, Statistics in Biopharmaceutical Research, DOI: 10.1080/19466315.2019.1689845)

Ian White, Royes Joseph & Nicky Best (2020) A causal modelling framework for reference-based imputation and tipping point analysis in clinical trials with quantitative outcome, Journal of Biopharmaceutical Statistics, 30:2, 334-350, DOI: 10.1080/10543406.2019.1684308

The supporting slides.

. . . and the future perhaps . . .

- Rather than use multiple imputation, fit joint Bayesian modelling of the outcome and the missing-data/withdrawal process and predict estimand value based on sample from posterior. Repeat over draws from posterior and summarize.

- Easy with lots of computer power, a suitable model and some well defined priors.

ITT and collection of off-treatment data

- Leads into importance of Estimand and ICH E9 (R1).

- The term "sensitivity analysis" changes to mean within the same estimand unlike in the EMA guidance.

- Reference-based MI becomes more than simply about sensitivity.

- Start to collect off-treatment data.
  Model off-treatment data to impute under "Treatment-policy" (ITT).

- Both require a linear model for mean and a multivariate Normal model for residuals.
  Often about half off-treatment data is "missing" so perhaps impute.

The legal stuff

- Apart from GSK 5 macros all were written to research and demonstrate methodology. Extensively tested but not formally validated.

- Open source licence, inbuilt parameter checking, diagnostics and ongoing support from James Roger
  james.roger@lshtm.ac.uk OR james@livedata.co.uk .

- All macros available at the DIA working Party TAB on the LSHTM Missing Data web site.
  https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data

- Also here, . . .
  - ▶ "Sequential imputation with tipping point and delta adjustment" based on proc MI and MIANALZE.
  - ▶ Imputation for recurrent event and also time-to-event data.

What is now available

- MIWithD
  RBI: Still used by some. Based on MCMC statment in proc MI.

- GSK 5 macros
  RBI: proc MCMC based. Slow but allows several methods based on a single Bayesian fit. System includes summary of Univariate analysis. No Treatment*Covariate interactions.

- MIStep
  RBI & OT: Sequential algorithm. Monotone missing only. Uses DATA step only.

- mymcmc
  Most flexible Bayesian conjugate prior fitting macro.

- RMConj and RMConjPlus
  RBI & OT: General conjugate RM model with ability to add separate imputation stage using same model but differeng covariate values for subject*visit. RefBTVT macro builds RBI model.

- MIAnalyze
  Bundles univariate analysis with summary using Rubin's rules.

- BGI
  Allows extraction of imputed data sets from proc BGLIMM output.