A COMPARISON OF MISSING DATA METHODS IN TYPE I DIABETES CLINICAL TRIALS

Jessica Lavery[1], Meagan Spychala[1], Mario Ehlers[2], Carla Greenbaum[3] and Lynette Keyes-Elstein[1]

[1] Rho, [2] Immune Tolerance Network, [3] Benaroya Research Institute

Abstract

How missing data is handled in a clinical trial has the potential to affect the results of any study. This research uses simulations to compare various methods for accounting for missing data within the context of type 1 diabetes trials in which the primary endpoint is a difference in C-peptide levels at 12 months. Methods used to handle that missing data have the potential to drastically over- or under-estimate the true difference in C-peptide levels between groups. The primary ramification of choosing a method that under-estimates the effect size and is too conservative is that the true treatment difference may not be detected; thus, we may not be able to conclude that the treatment is better than the placebo. If the method over-estimates the difference between the placebo and treatment groups, we could conclude that a treatment is helpful in a case where it truly is not. This simulation study selected 99 subjects in a 2:1 ratio 5,000 times in order to examine which missing data methods over- and under-estimate the true effect size in Type-I diabetes clinical trials. Considering both the average accuracy of methods to measure effect sizes, and also the variability associated with a method's estimation allowed us to assess its true value in a clinical trial setting. We will discuss these results as well as identify the method that most consistently detects the true treatment effect while retaining desired statistical power. Missing data methods addressed are as follows: single imputation methods (such as LOCF), multiple imputation, and mixed effects models.

Missing Data Considerations

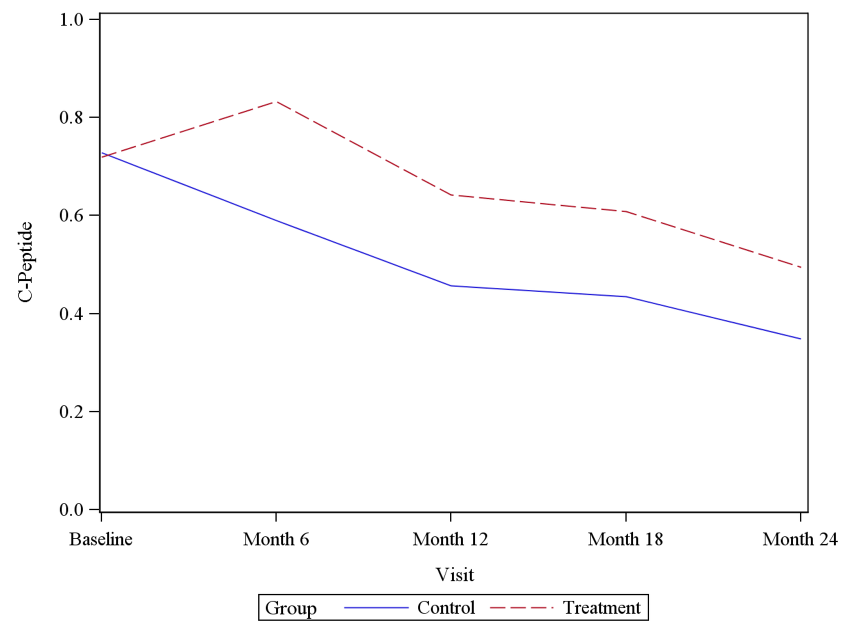
Pattern of Missingness
- Percent Missing
  - 7.5% and 15%
- Missingness Pattern
  - 80% Monotone Missing
  - 20% Non-Monotone Missing

Missing Mechanism
- Missing Completely at Random (MCAR)
- Missing at Random (MAR)
  - Group
  - Age
  - Baseline C-Peptide
- Missing Not at Random (MNAR)
  - Month 12 C-Peptide

Rate of Missingness
- Rate of Missingness
  - Equal dropout rates
  - Dropout twice as likely in treatment arm
  - Dropout twice as likely in control arm

Missing Data Methods
- Complete Case
- Single-Value Imputation
  - LOCF
  - Regression-Based Method
- Mixed Model
  - Single Slope
  - At Primary Endpoint
  - Spline
- Multiple Imputation

Study Design



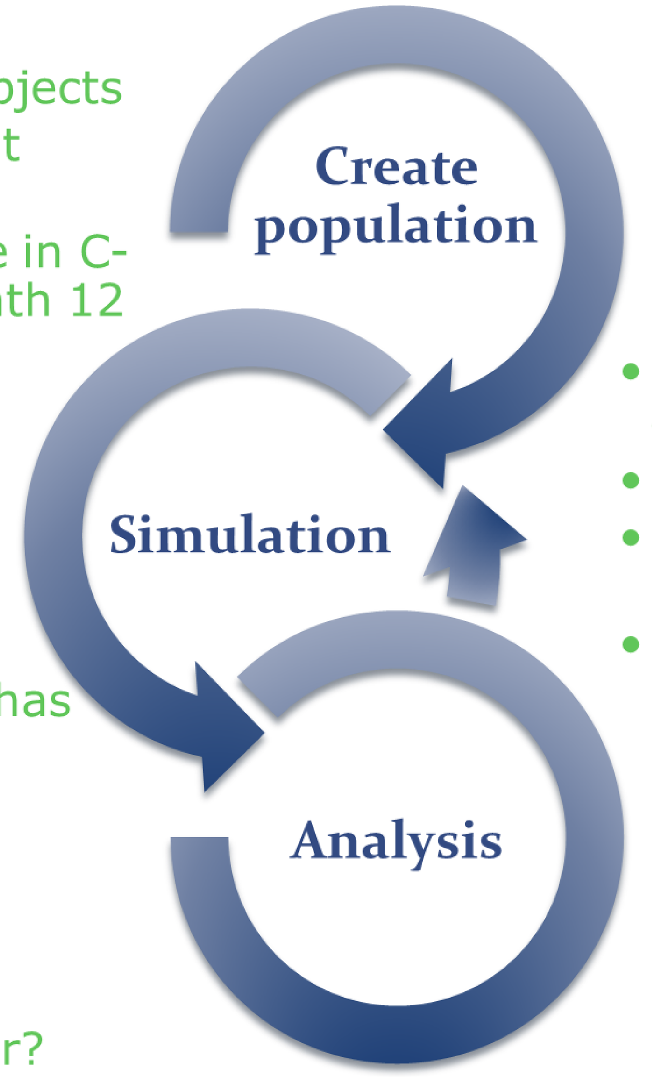
C-peptide: a measure of insulin production

Goal: Lessen the decline in insulin production over time

Outcome: Difference in change from baseline to 12 Months in C-peptide between the treatment and control groups

Methods

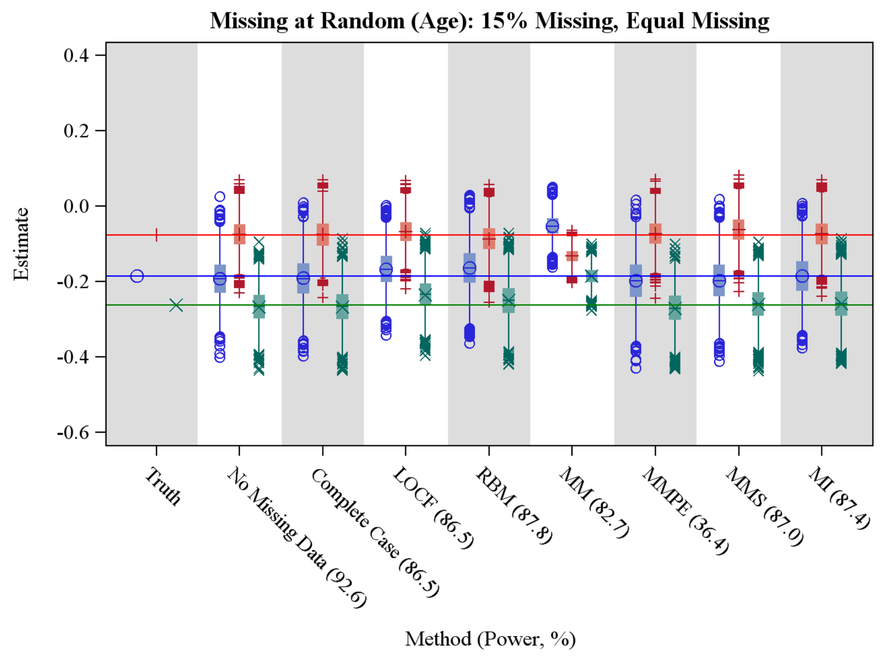
- 177 control subjects
- 1000 treatment subjects
- 40% difference in C-peptide at month 12

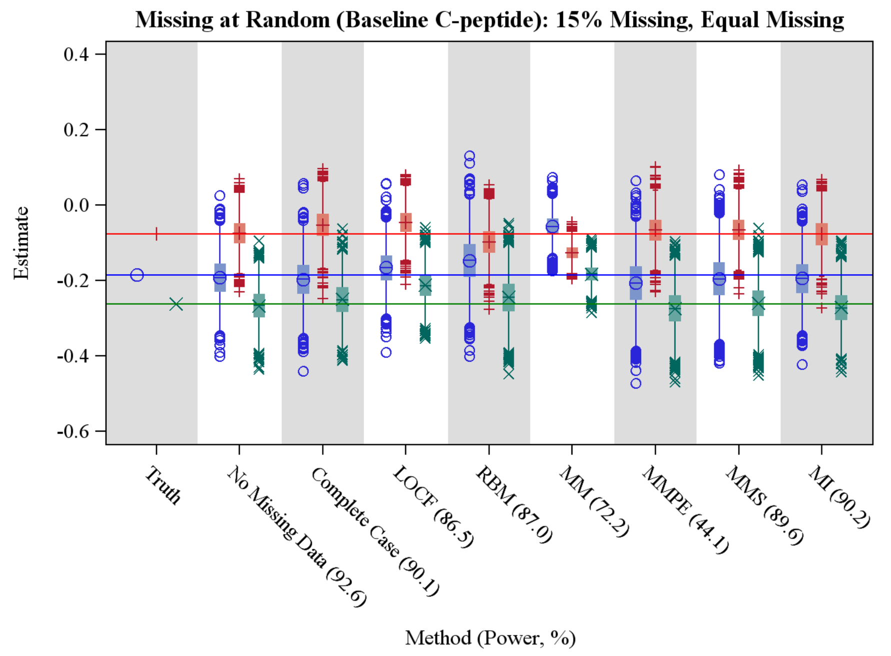


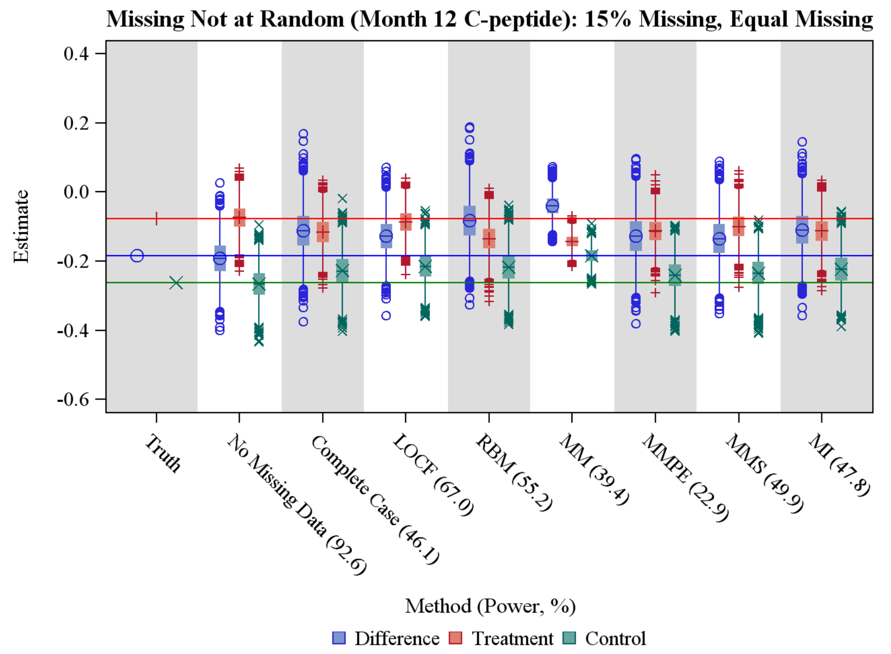
- Select 99 subjects
  - 2 treatment: 1 control
- Set data to missing
- Apply missing data method
- Test for treatment effect

- Which method has the smallest variability in its estimate?
- Which method retains the appropriate statistical power?

Results







LOCF = Last Observation Carried Forward
RBM = Regression-Based Method
MM = Mixed Model
MMPE = Mixed Model at Primary Endpoint
MMS = Mixed Model Spline
MI = Multiple Imputation

Discussion

- Power, effect size and type-I error rates were considered to compare performance of methods
- Multiple imputation and the spline mixed model performed consistently well across most scenarios considered
- Complete case performed well under MCAR and MAR, but only when data are equally missing in each arm
  - When more data is missing in a given arm, the performance of complete case declines drastically
- Regardless of the method used, for more data missing in the control arm, conclusions were drastically affected in the presence of missing data and the variability of each method was greatly increased.
- Under MNAR all methods perform poorly
  - Considered treatment subjects with higher C-peptide to have increased missingness and control subjects with lower C-peptide to have increased percentages of missingness
  - Expected result: Analysis methods considered are based on MCAR and MAR assumptions
    - Other methods not considered such as pattern mixture models or selection models may be more appropriate for the MNAR scenario
  - Highlights that precautions must be taken to avoid missing data
- The performance of each method is dependent on the mechanism and pattern of missing data; thorough work should be done to try to identify this pattern before applying a method
- These results are specific to this particular scenario; the performance of these methods on

Recommendations

Avoid missing data as much as possible!

If data appear to be MAR, MI and the spline mixed model perform well most consistently.

Conduct sensitivity analyses that assume MNAR (pattern mixture models, selection models)

Summary

In a true clinical trial setting, we are unable to identify the mechanism of missingness, so recommendations are made based on the pattern of missingness and the performance of each method across the mechanisms of missing data. The spline mixed model and multiple imputation performed well most consistently under MCAR and MAR. The performance of all methods under MNAR greatly declines, but the spline mixed model and multiple imputation are still the most reasonable.

The most drastic results were seen with more data missing in the control arm, indicating that great precaution needs to be taken when a large percentage of data are missing (these scenarios had 20% of the control subjects with missing data).

This project has been funded with federal funds from the National Institutes of Health, under contracts for the Statistical and Clinical Coordinating Center for Autoimmune Disease Clinical Trials (HHSN272200900057C) and the Statistical and Data Coordinating Center for the Immune Tolerance Network (HHSN272200800029C).

Immune Tolerance Network

Rho Giving flight to research