Global Sensitivity Analysis of Randomized Trials with Non-Monotone Missing Binary Outcomes:

Application to Studies of Substance Use Disorders

Daniel Scharfstein (PI)
Johns Hopkins University
dscharf@jhu.edu

June 8, 2020

Co-Investigators

- Aimee Campbell (Columbia)
- Edward Nunes (Columbia)
- Abigail Matthews (EMMES)
- Aidan McDermott (Johns Hopkins)
- Chenguang Wang (Johns Hopkins)
- Jon Steingrimmson (Brown)

Aims

1. Develop and evaluate a sensitivity analysis methodology for the analysis of randomized clinical trials with repeatedly measured binary outcomes and non-monotone missing data.
2. Develop open-source, user-friendly software.
3. Conduct sensitivity analysis of 23 CTN-sponsored trials with public-use datasets available on the NIDA Data Share website.
4. Link the results to study characteristics in order to identify patterns.
5. Disseminate the methodology and software to researchers interested in substance use disorder clinical trials.

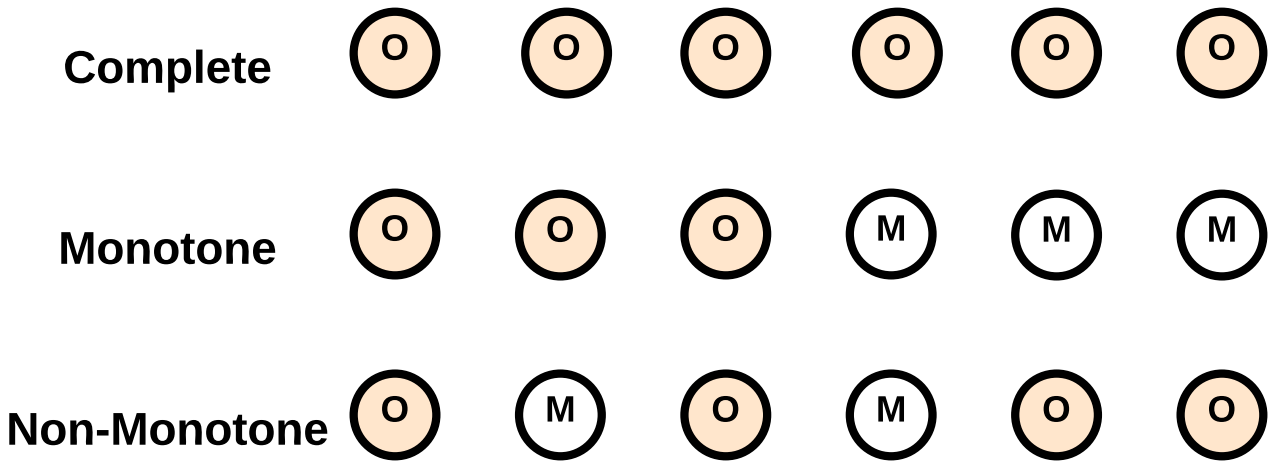
Sensitivity Analysis

- Missing outcome data threaten the validity of randomized clinical trials because inference about treatment effects then necessarily relies on untestable assumptions, which wrongly stated can lead to incorrect conclusions.
- The National Research Council (NRC) in its report entitled "The Prevention Treatment of Missing Data in Clinical Trials" recommended that evaluating the sensitivity of trial results to assumptions about the missing data mechanism should be a mandatory component of reporting.

Sensitivity Analysis

- Chapter 5 of the NRC Report presents an approach whereby one posits a broad class of untestable missing data assumptions that is
  1. indexed by sensitivity analysis parameters,
  2. anchored around a plausible benchmark assumption (sensitivity parameters equal to a reference value), and
  3. sensitivity analysis parameters further from the reference value represent larger deviations from the benchmark assumption.
- The goal of this "global" sensitivity analysis approach is to determine how much deviation from a benchmark assumption is required in order for inferences to change.
- If the deviation is judged to be sufficiently far from the benchmark assumption, then greater credibility is lent to the benchmark analysis; if not, the benchmark analysis can be considered to be fragile.

Missing Data



Non-Monotone Missing Data

- Positing plausible assumptions and specifying flexible models for studies with non-monotone missing data is challenging because of the potentially large number of missingness patterns (as many as $2^K - 1$ patterns, where K is the number of post-baseline assessments).
- Ibrahim and Molenberghs (2009) indicate that "[s]uch data present a considerable modeling challenge for the statistician".
- The NRC report highlighted the need for development and application of "novel, appropriate methods of model specification and sensitivity analyses to handle non-monotone missing data patterns".

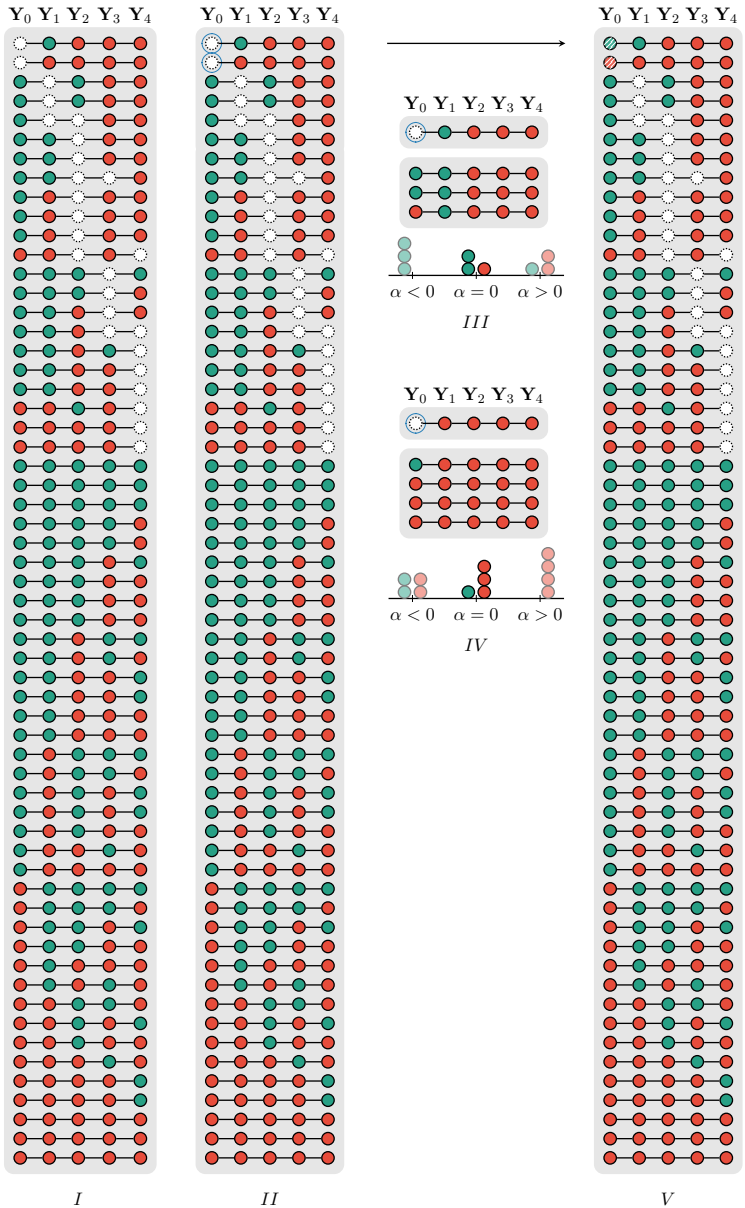
Robins (1997); Sadinle and Reiter (2017b)

- For individuals who share the same outcomes (observed or not) prior to a scheduled visit and the same observed data after the visit, the distribution of the outcome for those missing the visit is the same as the distribution of the outcome for those attending the visit.
- No global sensitivity analysis procedure was developed.

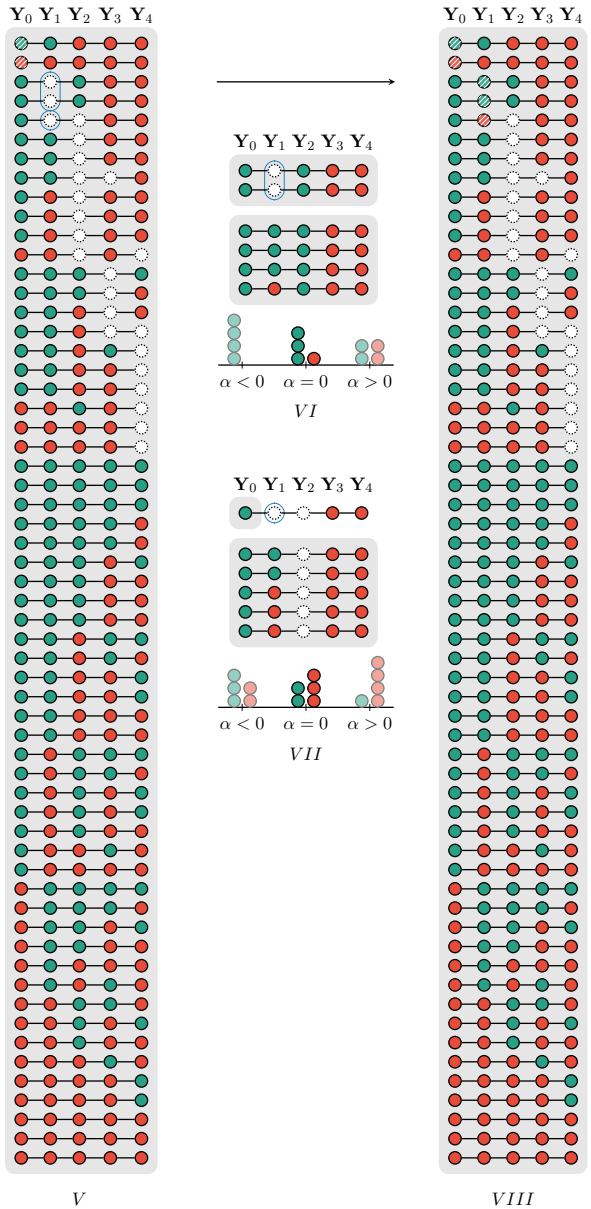
In Other Words ...

- Imagine the stratum of people who share the same outcomes prior to visit k (*observed or not*) and who share the same *observed* outcomes after visit k.
  - Sub-stratum A: people who show up at visit k
  - Sub-stratum B: people who do not show up at visit k
- Probability of outcome at visit k is the same for those in sub-stratum A and those in sub-stratum B.

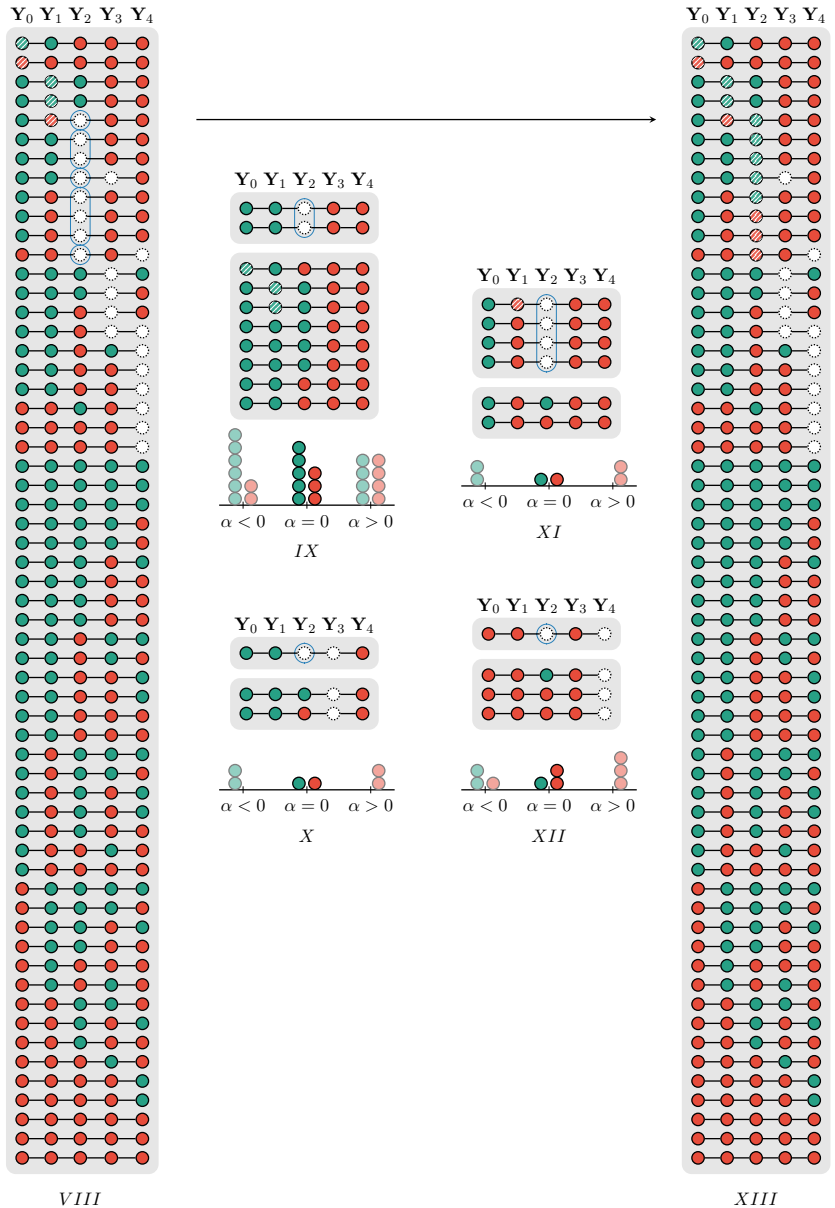
How It Works


$\mathbf{Y}_0$ $\mathbf{Y}_1$ $\mathbf{Y}_2$ $\mathbf{Y}_3$ $\mathbf{Y}_4$
$\alpha < 0$ $\alpha = 0$ $\alpha > 0$
I
II
III
IV
V

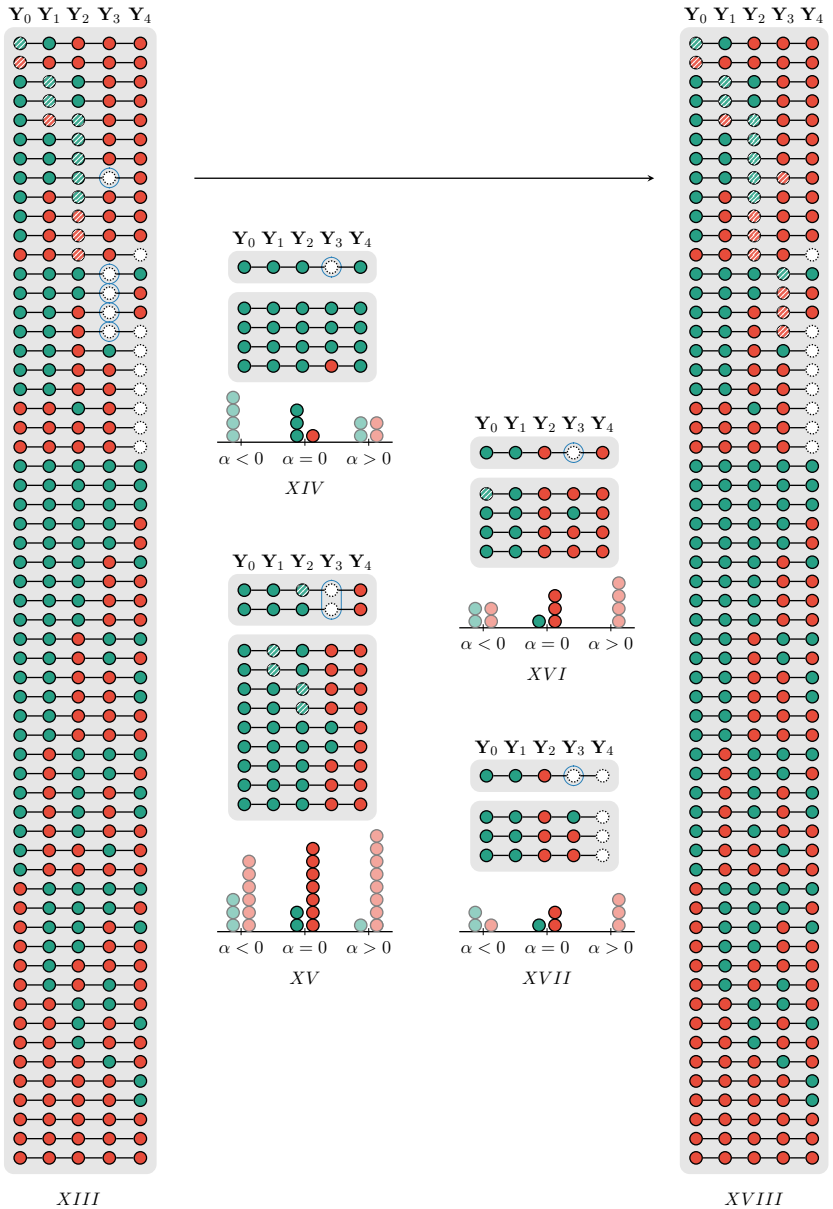
How It Works


$\mathbf{Y}_0$ $\mathbf{Y}_1$ $\mathbf{Y}_2$ $\mathbf{Y}_3$ $\mathbf{Y}_4$
$\alpha < 0$ $\alpha = 0$ $\alpha > 0$
V
VI
VII
VIII

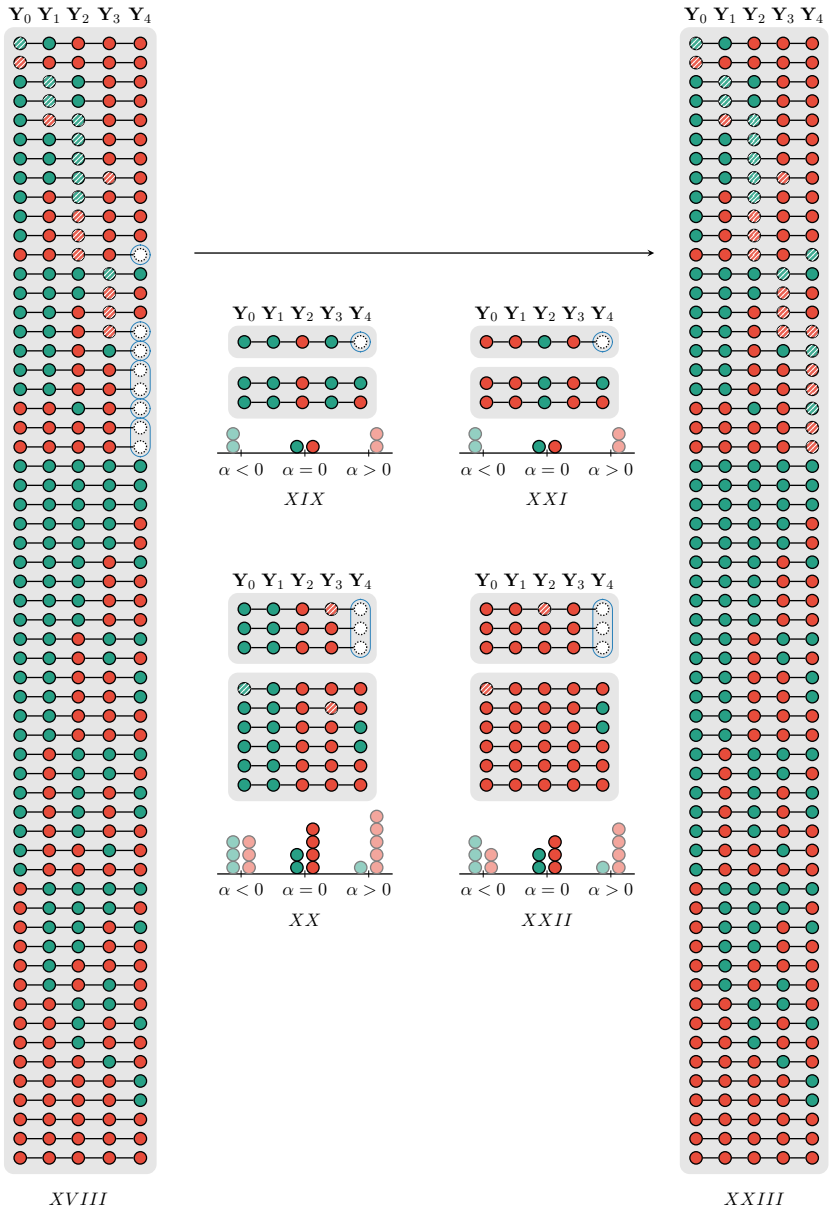
How It Works


$\mathbf{Y}_0\ \mathbf{Y}_1\ \mathbf{Y}_2\ \mathbf{Y}_3\ \mathbf{Y}_4$
$\alpha < 0$ $\alpha = 0$ $\alpha > 0$
VIII
IX
X
XI
XII
XIII

How It Works


$\mathbf{Y}_0$ $\mathbf{Y}_1$ $\mathbf{Y}_2$ $\mathbf{Y}_3$ $\mathbf{Y}_4$
$\alpha < 0$ $\alpha = 0$ $\alpha > 0$
XIII
XIV
XV
XVI
XVII
XVIII

How It Works


$\mathbf{Y}_0$ $\mathbf{Y}_1$ $\mathbf{Y}_2$ $\mathbf{Y}_3$ $\mathbf{Y}_4$
$\alpha < 0$ $\alpha = 0$ $\alpha > 0$
XVIII
XIX
XX
XXI
XXII
XXIII

How It Works

- Fit a model for the observed data using random forests, a "machine-learning" algorithm
- We use the estimated distribution of the observed data *plus* missing data assumptions to estimate the complete data quantities of interest.

Case Study: CTN -0044

- Two-arm randomized trial designed to evaluate a new approach to reducing substance use among patients entering outpatient addiction treatment.
- Treatment-as-usual (TAU) vs. treatment-as-usual plus a computerized therapeutic education system and contingent incentives (TAU+).
- TAU: individual and group counseling.
- TAU+: substituted 2 hours of usual care per week with computer-interactive modules covering skills for achieving and maintaining abstinence and prize-based motivational incentives contingent on abstinence and treatment adherence.
- Urine samples scheduled to be collected twice weekly.
- Outcome: number of negative urine samples during first 6 weeks.

Case Study: CTN -0044

- Among the 252 individuals randomized to TAU
  - 42 (16.7%) had a complete record of urine samples
  - 11 (4.4%) had no urine samples
  - 28 (11.1%) had at least one urine sample and a monotone missing data pattern
  - 171 (67.9%) individuals had an intermittent missing data pattern
- Among the 255 individuals randomized to TAU+
  - 81 (31.8%) had a complete record of urine samples,
  - 3 (1.2%) had no urine samples
  - 18 (7.1%) had at least one urine sample and a monotone missing data pattern
  - 153 (60.0%) individuals had an intermittent missing data pattern

Case Study: CTN -0044

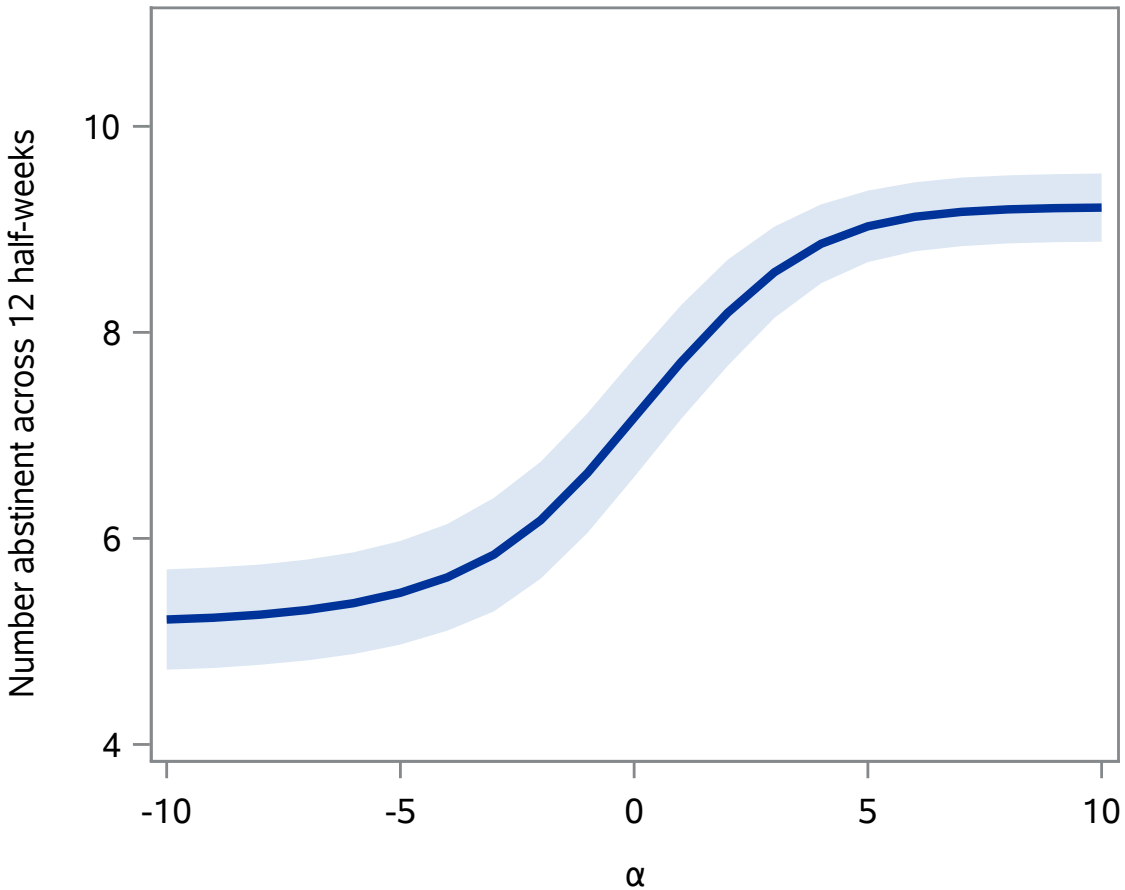
We first used the random forest algorithm to estimate the distribution of the observed data. We used 1000 trees.

- To evaluate the model fit, we compared empirical and model-based estimates of the joint distribution of the observed data at all 66 pairs of time points.
- For each pair, the joint distribution is represented by the cell probabilities of a three by three table.
- For each table, we computed the maximum of the absolute differences between the empirical and model-based estimates of the cell probabilities.
- The largest of these maximums over the 66 tables was 1.82%.
- In contrast, the largest of the maximums based on a first-order Markov model was 12.98%.

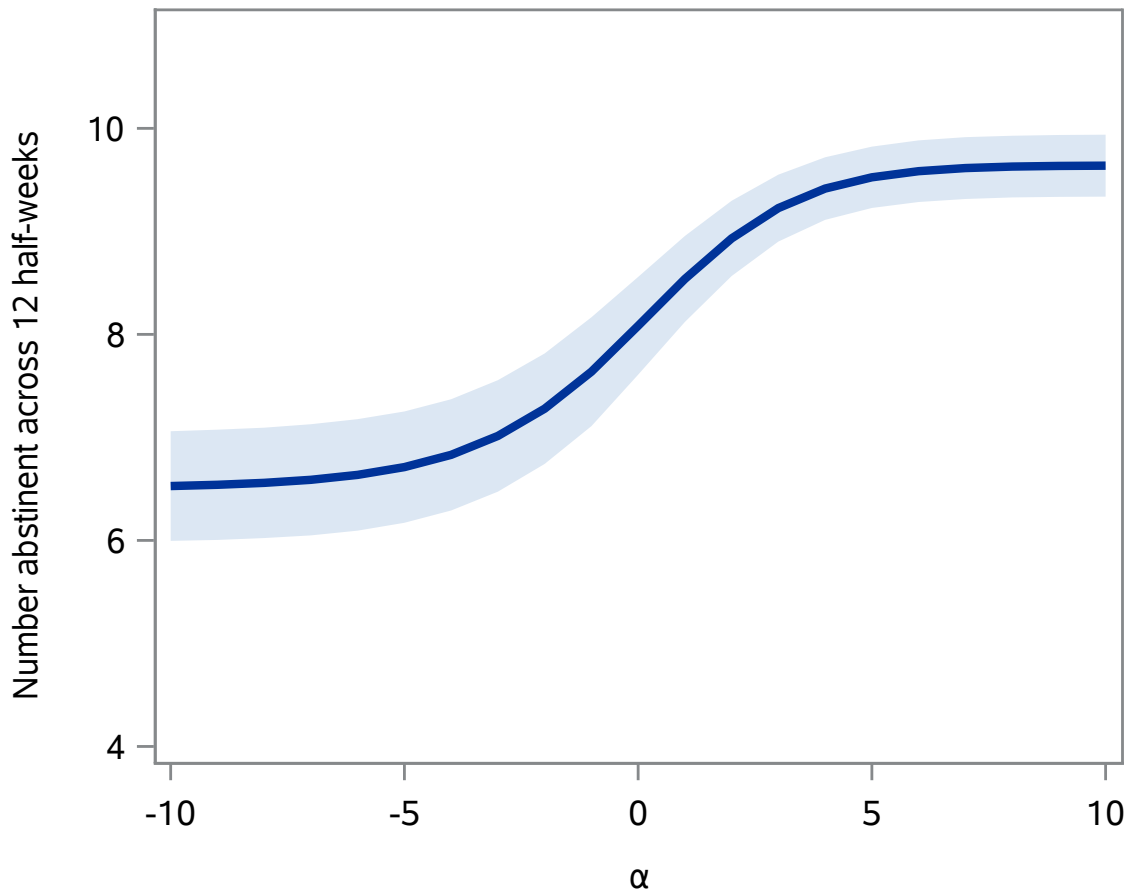
Case Study: CTN -0044

Assumption	TAU	TAU+	Difference
MCAR	7.86 (7.25, 8.47)	8.83 (8.28, 9.38)	0.97 (0.17, 1.76).
Missing=Non-Abstinent	5.14 (4.60, 5.69)	6.48 (5.90, 7.06)	1.34 (0.58, 2.10)
Missing=Abstinent	9.27 (8.87, 9.67)	9.64 (9.24, 10.04)	0.37 (-0.18, 0.92)
Benchmark	7.17 (6.60, 7.75)	8.08 (7.61, 8.56)	0.91 (0.06, 1.76)

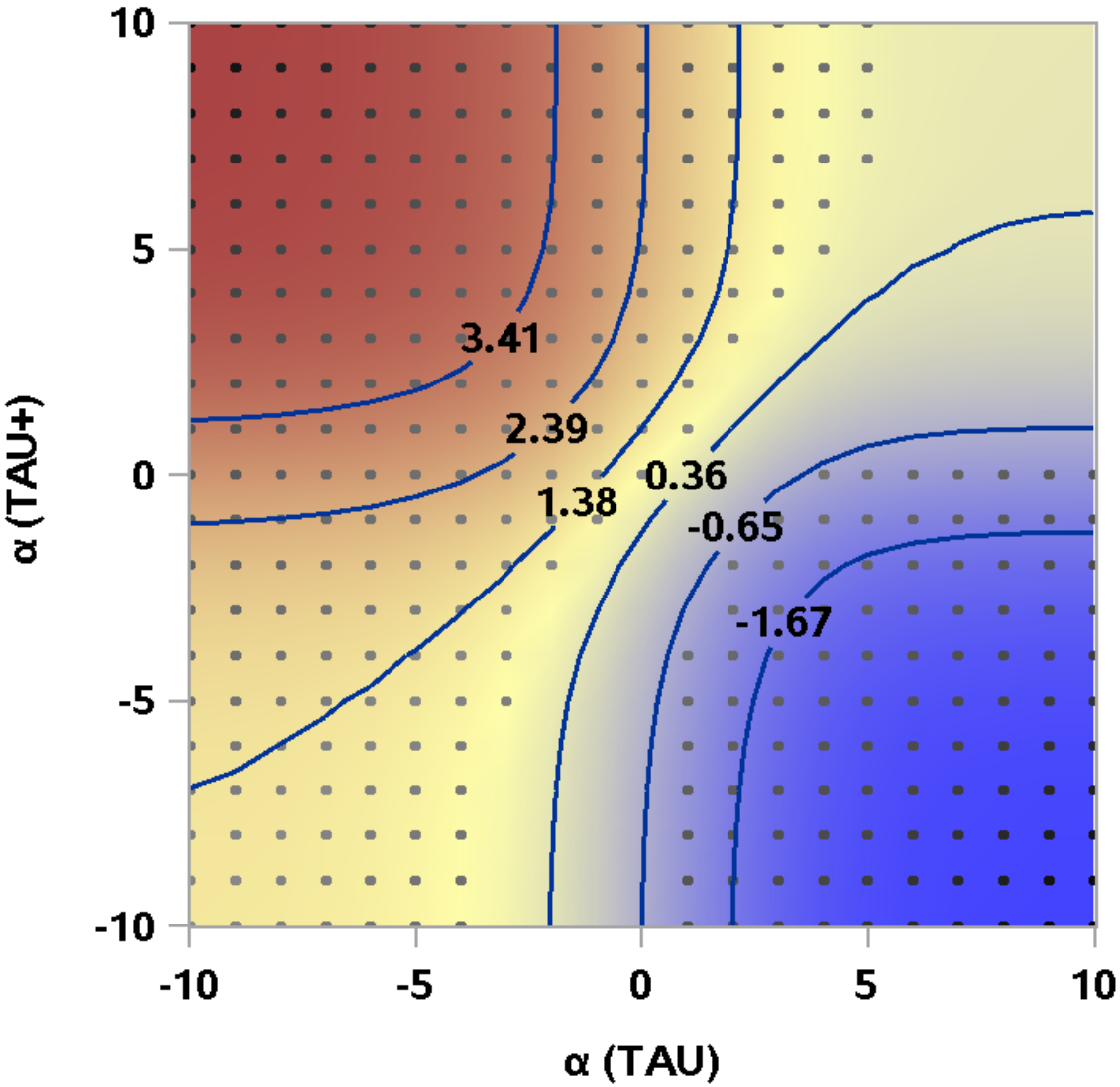
Case Study: CTN -0044



Case Study: CTN -0044



Case Study: CTN -0044



Discussion

- Software has been developed - `salbm`.
- The package can be installed in `R` from Github by

```
install_github("olssol/salbm")
```

- We have re-analyzed 21 CTN trials.
- Computationally infeasible when $K > 15$.
- Reduce dimension by introducing additional conditional independence restrictions.
- Cannot be easily extended to handle continuous outcomes.