# Global Sensitivity Analysis of Randomized Trials with Missing Data

**Daniel Scharfstein**
Johns Hopkins University
dscharf@jhu.edu

May 17, 2015

# Missing Data Matters

- Missing outcome data are a widespread problem in randomized trials
- We reviewed all randomized trials reporting five major patient-reported outcomes published in five leading general medical journals between January 1, 2008 and March 14, 2015
  - 83.5% reported percentages greater than 10%,
  - 46.1% reported percentages greater than 20%
  - 23.1% reported percentages greater than 30%.

# Missing Data Matters

- While unbiased estimates of treatment effects can be obtained from trials with no missing data, this is no longer true when data are missing on some patients.
- The essential problem is that inference about treatment effects relies on *unverifiable* assumptions about the nature of the mechanism that generates the missing data.
- While we usually know the reasons for missing data, we do not know the distribution of outcomes for patients with missing data, how it compares to that of patients with observed data and whether differences in these distributions can be explained by the observed data.

# Robert Temple and Bob O'Neil (FDA)

- "*During almost 30 years of review experience, the issue of missing data in ... clinical trials has been a major concern because of the potential impact on the inferences that can be drawn .... when data are missing .... the analysis and interpretation of the study pose a challenge and the conclusions become more tenuous as the extent of 'missingness' increases.*"

# NRC Report and Sensitivity Analysis

- In 2010, the National Research Council (NRC) issued a reported entitled "The Prevention and Treatment of Missing Data in Clinical Trials."
- This report, commissioned by the FDA, provides 18 recommendations targeted at (1) trial design and conduct, (2) analysis and (3) directions for future research.
- Recommendation 15 states
    - *Sensitivity analyses should be part of the primary reporting of findings from clinical trials. Examining sensitivity to the assumptions about the missing data mechanism should be a mandatory component of reporting.*

- 1998 International Conference of Harmonization (ICH) Guidance document (E9) entitled "Statistical Principles in Clinical Trials" states: "*it is important to evaluate the robustness of the results to various limitations of the data, assumptions, and analytic approaches to data analysis*"

- European Medicines Agency 2009 draft "Guideline on Missing Data in Confirmatory Clinical Trials" states "*[i]n all submissions with non-negligible amounts of missing data sensitivity analyses should be presented as support to the main analysis.*"

# PCORI and Sensitivity Analysis

- In 2012, Li *et al.* issued the report "Minimal Standards in the Prevention and Handling of Missing Data in Observational and Experimental Patient Centered Outcomes Research"

- This report, commissioned by PCORI, provides 10 standards targeted at (1) design, (2) conduct, (3) analysis and (4) reporting.

- Standard 8 echoes the NRC report, stating
  - *Examining sensitivity to the assumptions about the missing data mechanism (i.e., sensitivity analysis) should be a mandatory component of the study protocol, analysis, and reporting.*

# Sensitivity Analysis

The set of possible assumptions about the missing data mechanism is very large and cannot be fully explored. There are different approaches to sensitivity analysis:

- Ad-hoc
- Local
- Global

# Ad-hoc Sensitivity Analysis

- Analyzing data using a few different analytic methods, such as last or baseline observation carried forward, complete or available-case analysis, mixed models or multiple imputation, and evaluate whether the resulting inferences are consistent.
- The problem with this approach is that the assumptions that underlie these methods are very strong and for many of these methods unreasonable.
- More importantly, just because the inferences are consistent does not mean that there are no other reasonable assumptions under which the inference about the treatment effect is different.

# Local Sensitivity Analysis

- Specify a reasonable benchmark assumption (e.g., missing at random) and evaluate the robustness of the results within a small neighborhood of this assumption.
- What if there are assumptions outside the local neighborhood which are plausible?

## Global Sensitivity Analysis

- Evaluate robustness of results across a much broader range of assumptions that include a reasonable benchmark assumption and a collection of additional assumptions that trend toward best and worst case assumptions.

- Emphasized in Chapter 5 of the NRC report.

- This approach is substantially more informative because it operates like "stress testing" in reliability engineering, where a product is systematically subjected to increasingly exaggerated forces/conditions in order to determine its breaking point.

# Global Sensitivity Analysis

- In the missing data setting, global sensitivity analysis allows one to see how far one needs to deviate from the benchmark assumption in order for inferences to change.
- "Tipping point" analysis (Yan, Lee and Li, 2009; Campbell, Pennello and Yue, 2011)
- If the assumptions under which the inferences change are 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.

# Global Sensitivity Analysis

- Restrict consideration to follow-up randomized study designs that prescribe that measurements of an outcome of interest are to be taken on each study participant at fixed time-points.
- Focus on monotone missing data pattern
- Consider the case where interest is focused on a comparison of treatment arm means at the last scheduled visit.

## Global Sensitivity Analysis

- The missingness mechanism is typically not under the control of the investigator
- Inference about the treatment arm means requires two types of assumptions:
  - (i) *unverifiable* assumptions about the distribution of outcomes among those with missing data and
  - (ii) additional testable assumptions that serve to increase the efficiency of estimation.

# Global Sensitivity Analysis

- Type (i) assumptions are necessary to identify the treatment-specific means.
- By *identification*, we mean that we can write it as a function that depends only on the distribution of the observed data.
- When a parameter is identified we can hope to estimate it as precisely as we desire with a sufficiently large sample size,
- In the absence of identification, statistical inference is fruitless as we would be unable to learn about the true parameter value even if the sample size were infinite.

# Global Sensitivity Analysis

- To address the identifiability issue, it is essential to conduct a sensitivity analysis, whereby the data analysis is repeated under different type (i) assumptions, so as to investigate the extent to which the conclusions of the trial are dependent on these subjective, unverifiable assumptions.
- The usefulness of a sensitivity analysis ultimately depends on the plausibility of the unverifiable assumptions.
- It is key that any sensitivity analysis methodology allow the formulation of these assumptions in a transparent and easy to communicate manner.

# Global Sensitivity Analysis
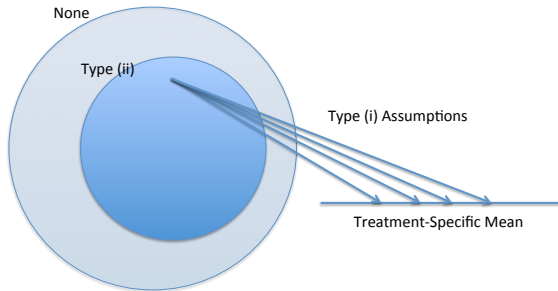
- There are an infinite number of ways of positing type (i) assumptions.
- Ultimately, however, these assumptions prescribe how missing outcomes should be "imputed."
- A reasonable way to posit these assumptions is to
    - stratify individuals with missing outcomes according to the data that we were able to collect on them and the occasions at which the data were collected
    - separately for each stratum, hypothesize a connection (or link) between the distribution of the missing outcome with the distribution of the outcome among those with the observed outcome and who share the same recorded data.

# Global Sensitivity Analysis

- Type (i) assumptions will not suffice when the repeated outcomes are continuous or categorical with many levels. This is because of *data sparsity*.

- For example, the stratum of people who share the same recorded data will typically be small. As a result, it is necessary to draw strength across strata by "smoothing."

- Without smoothing, the data analysis will rarely be informative because the uncertainty concerning the treatment arm means will often be too large to be of substantive use.

- As a result, it is necessary to impose type (ii) smoothing assumptions.

- Type (ii) assumptions should be scrutinized with standard model checking techniques.

Restrictions on Distribution of Observed Data

None

Type (ii)

Type (i) Assumptions

Treatment-Specific Mean

# Notation

- $K$ scheduled post-baseline assessments.
- There are $(K + 1)$ patterns representing each of the visits an individual might last be seen, i.e., $0, \ldots, K$.
- The $(K + 1)$st pattern represents individuals who complete the study.
- Let $Y_k$ be the outcome scheduled to be measured at visit $k$, with visit 0 denoting the baseline measure (assumed to be observed).
- Let $Y_k^- = (Y_0, \ldots, Y_k)$ and $Y_k^+ = (Y_{k+1}, \ldots, Y_K)$.

# Notation

- Let $R_k$ be the indicator of being on study at visit $k$
- $R_0 = 1$; $R_k = 1$ implies that $R_{k-1} = 1$.
- Let $C$ be the last visit that the patient is on-study.
- We focus inference separately for each treatment arm.
- The observed data for an individual is $O = (C, Y_C^-)$.
- We want to estimate $\mu^* = E[Y_K]$.

- Full Data: $(Y_0, Y_1, Y_2)$
- Observed Data: $O = (C, Y_C^-)$
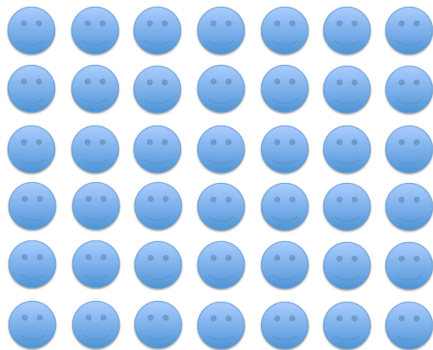- Estimate $\mu^* = E[Y_2]$

# Missing at Random (MAR)

- In this setting, MAR postulates

$$f(Y_1, Y_2 | R_1 = 0, Y_0) = f(Y_1, Y_2 | R_1 = 1, Y_0)$$
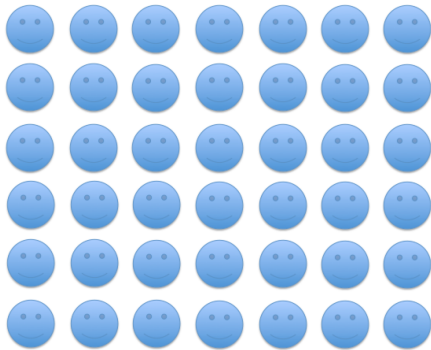$$f(Y_2 | R_2 = 0, R_1 = 1, Y_1^-) = f(Y_2 | R_2 = 1, Y_1^-)$$

or

$$P[R_1 = 0 | Y_2^-] = P[R_1 = 0 | Y_0]$$
$$P[R_2 = 0 | R_1 = 1, Y_2^-] = P[R_2 = 0 | R_1 = 1, Y_1^-]$$
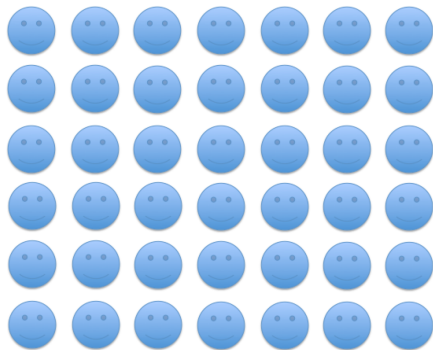
# Missing at Random (MAR)
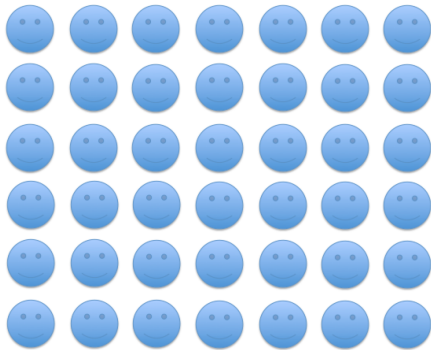
$R_1 = 1, Y_0 = y_0$

$R_1 = 0, Y_0 = y_0$

# Missing at Random (MAR)

$R_2 = 1, Y_1^- = y_1^-$



$R_2 = 0, R_1 = 1, Y_1^- = y_1^-$

# Missing at Random (MAR)

- MAR is a type (i) assumption. It is "unverifiable."
- For patients last seen at visit $k$, we cannot learn from the observed data about the conditional (on observed history) distribution of outcomes after visit $k$.
- For patients last seen at visit $k$, any assumption that we would make about the conditional (on observed history) distribution of the outcomes after visit $k$ will be unverifiable from the data available to us.
- For patients last seen at visit $k$, the assumption that the conditional (on observed history) distribution of outcomes after visit $k$ is the same as those who remain on-study after visit $k$ and have the same observed history is unverifiable.

# Missing at Random (MAR)

$$\mu^* = \int_{y_0} \int_{y_1} \int_{y_2} y_2 dF(y_2|R_2 = 1, Y_1 = y_1, Y_0 = y_0) dF(y_1|R_1 = 1, Y_0 = y_0) dF(y_0)$$

$$\mu^* = E\left[\frac{R_2 Y_2}{\prod_{k=1}^2 P[R_k = 1|R_{k-1} = 1, Y_{k-1}^-]}\right]$$

- $\mu^*$ is written as a function of the distribution of the observed data.

$$R_{k+1} \perp Y_k^+ \mid R_k = 1, Y_k^-$$

$$f(Y_k^+|R_{k+1} = 0, R_k = 1, Y_k^-) = f(Y_k^+|R_{k+1} = 1, Y_k^-)$$

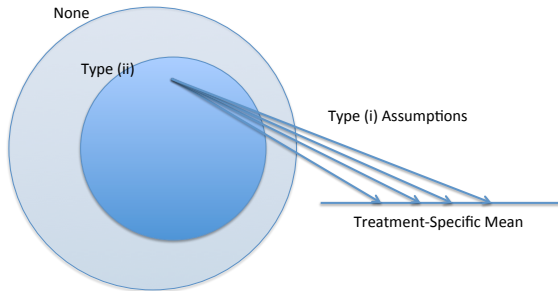$$P[R_{k+1} = 0 \mid R_k = 1, Y_K^-] = P[R_{k+1} = 0 \mid R_k = 1, Y_k^-]$$

$$\mu^* = \int_{y_0} \ldots \int_{y_K} y_K \prod_{k=K}^{1} dF(y_k|R_k = 1, Y_k^- = y_k^-)dF(y_0)$$

$$\mu^* = E\left[\frac{R_K Y_K}{\prod_{k=1}^{K} P[R_k = 1|R_{k-1} = 1, Y_{k-1}^-]}\right]$$

- $\mu^*$ is written as a function of the distribution of the observed data.

# Global Sensitivity Analysis



Restrictions on Distribution of Observed Data

None

Type (ii)

Type (i) Assumptions

Treatment-Specific Mean

The MAR assumption is not the only one that is (1)
unverifiable and (2) admits identification of $\mu^*$.

# Missing Not at Random (MNAR)

- Non-future Dependence

$$f(Y_2|R_1 = 0, Y_1^-) = f(Y_2|R_1 = 1, Y_1^-)$$

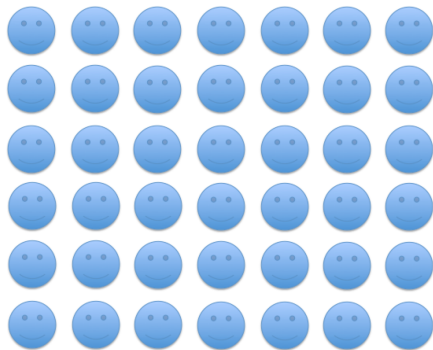$$R_1 \perp Y_2 \mid Y_1, Y_0$$

- Exponential Tilting

$$f(Y_1|R_1 = 0, Y_0) \propto \underbrace{f(Y_1|R_1 = 1, Y_0)}_{\text{Reference Distribution}} \exp\{\alpha r(Y_1)\}$$

$$f(Y_2|R_2 = 0, R_1 = 1, Y_1^-) \propto \underbrace{f(Y_2|R_2 = 1, Y_1^-)}_{\text{Reference Distribution}} \exp\{\alpha r(Y_2)\}$$
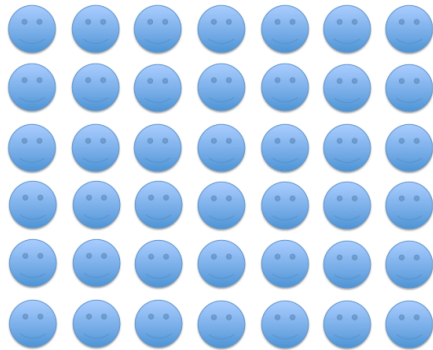
- $r(y)$ is a specified function; $\alpha$ is a sensitivity analysis parameter.
- $\alpha = 0$ is MAR.

$R_1 = 1, Y_1^- = y_1^-$

$R_1 = 0, Y_1^- = y_1^-$
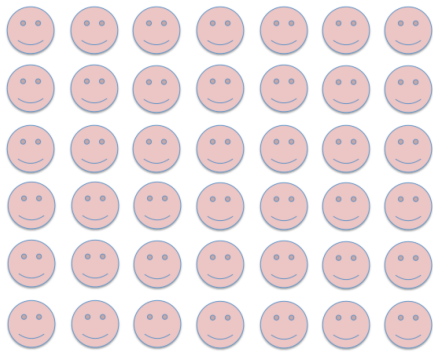
$R_1 = 1, Y_0 = y_0$

$R_1 = 0, Y_0 = y_0$

$R_1 = 1, Y_0 = y_0$

$R_1 = 0, Y_0 = y_0$

# Missing Not at Random (MNAR)
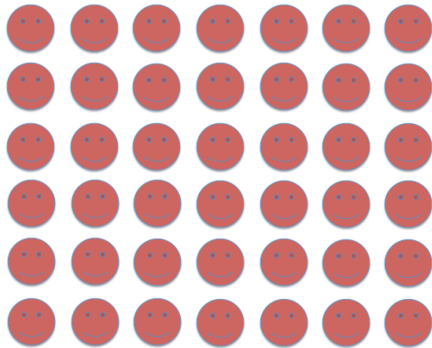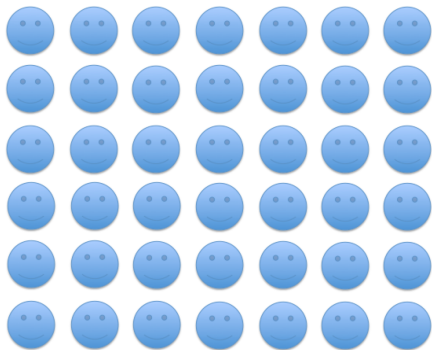
$R_2 = 1, Y_1^- = y_1^-$
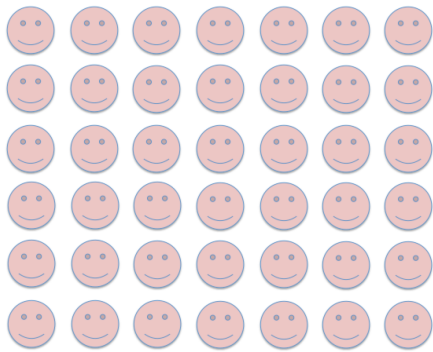
$R_2 = 0, R_1 = 1, Y_1^- = y_1^-$

# Missing Not at Random (MNAR)
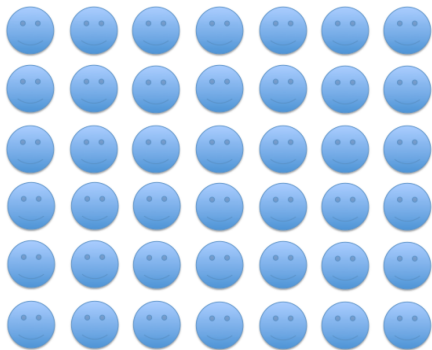
$R_2 = 1, Y_1^- = y_1^-$
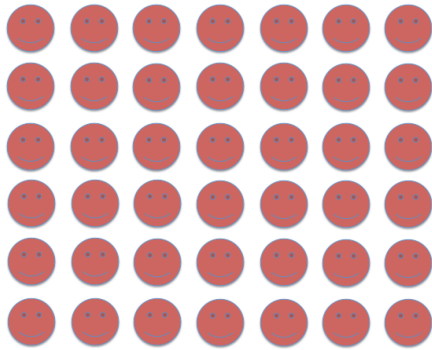
$R_2 = 0, R_1 = 1, Y_1^- = y_1^-$

# Missing Not at Random (MNAR)

$$\text{logit } P[R_1 = 0 | Y_2^-] = h_1(Y_0) + \alpha r(Y_1)$$
$$\text{logit } P[R_2 = 0 | R_1 = 1, Y_2^-] = h_2(Y_1^-) + \alpha r(Y_2)$$

where

$$
\begin{aligned}
h_1(Y_0) &= \text{logit } P[R_1 = 0 | Y_0] - \\
&\quad \log\{E[\exp\{\alpha r(Y_1)\} | R_1 = 1, Y_0]\} \\
h_2(Y_1^-) &= \text{logit } P[R_2 = 0 | R_1 = 1, Y_1^-] - \\
&\quad \log\{E[\exp\{\alpha r(Y_2)\} | R_2 = 1, Y_1^-]\}
\end{aligned}
$$

# Exponential Tilting Explained

$$f(Y|R = 0) \propto f(Y|R = 1)\exp\{\alpha r(Y)\}$$

- If $[Y|R = 1] \sim N(\mu, \sigma^2)$ and $r(Y) = Y$,
  $[Y|R = 0] \sim N(\mu + \alpha\sigma^2, \sigma^2)$
- If $[Y|R = 1] \sim Beta(a, b)$ and $r(Y) = \log(Y)$,
  $[Y|R = 0] \sim Beta(a + \alpha, b)$, $\alpha > -a$.
- If $[Y|R = 1] \sim Gamma(a, b)$ and $r(Y) = \log(Y)$,
  $[Y|R = 0] \sim Gamma(a + \alpha, b)$, $\alpha > -a$.
- If $[Y|R = 1] \sim Gamma(a, b)$ and $r(Y) = Y$,
  $[Y|R = 0] \sim Gamma(a, b - \alpha)$, $\alpha < b$.
- If $[Y|R = 1] \sim Bernoulli(p)$ and $r(Y) = Y$,
  $[Y|R = 0] \sim Bernoulli\left(\frac{p\exp(\alpha)}{p\exp(\alpha)+1-p}\right)$.

# Beta
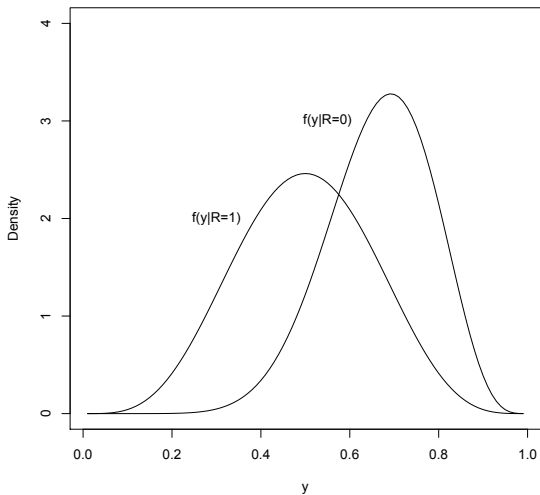
# Gamma
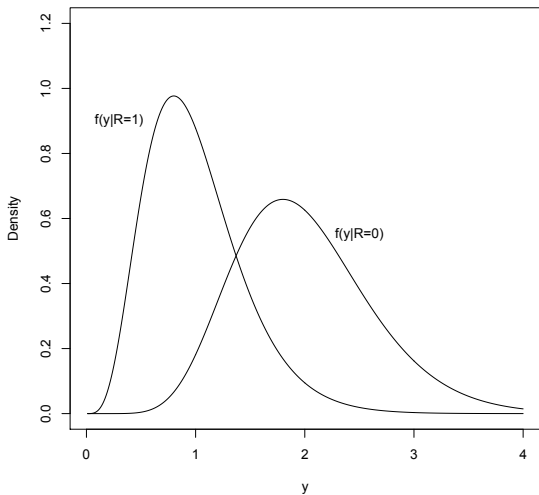
$$\mu^* = \int_{y_0} \int_{y_1} \int_{y_2} y_2 \left\{ dF(y_2|R_2 = 1, Y_1 = y_1, Y_0 = y_0) \frac{1}{1 + \exp\{h_2(y_1) + \alpha r(y_2)\}} + \right.$$
$$\frac{\exp(\alpha r(y_2)) dF(y_2|R_2 = 1, Y_1 = y_1, Y_0 = y_0)}{E[\exp(\alpha r(Y_2))|R_2 = 1, Y_1 = y_1, Y_0 = y_0]} \left. \frac{\exp\{h_2(y_1) + \alpha r(y_2)\}}{1 + \exp\{h_2(y_1) + \alpha r(y_2)\}} \right\} \times$$
$$\left\{ dF(y_1|R_1 = 1, Y_0 = y_0) \frac{1}{1 + \exp\{h_1(Y_0) + \alpha r(Y_1)\}} + \right.$$
$$\frac{\exp(\alpha r(y_1)) dF(y_1|R_1 = 1, Y_0 = y_0)}{E[\exp(\alpha r(Y_1))|R_1 = 1, Y_0 = y_0]} \left. \frac{\exp\{h_1(y_0) + \alpha r(y_1)\}}{1 + \exp\{h_1(y_0) + \alpha r(y_1)\}} \right\} dF(y_0)$$

$$\mu^* = E\left[ \frac{R_2 Y_2}{\prod_{k=1}^{2} (1 + \exp\{h_k(Y_{k-1}^-) + \alpha r(Y_k)\})^{-1}} \right]$$

- $\mu^*$ is written as a function of the distribution of the observed data (depending on $\alpha$).

Restrictions on Distribution of Observed Data

None

Type (ii)

Type (i) Assumptions

Treatment-Specific Mean

# Missing Not at Random (MNAR) - In General

- Non-future Dependence

$$f(Y_k^+ | R_k = 0, R_{k-1} = 1, Y_k^-) = f(Y_k^+ | R_k = 1, Y_k^-)$$

$$R_k \perp Y_k^+ \mid R_{k-1} = 1, Y_k^-$$

- Exponential Tilting

$$f(Y_k | R_k = 0, R_{k-1} = 1, Y_{k-1}^-) \propto f(Y_k | R_k = 1, Y_{k-1}^-) \exp\{\alpha r(Y_k)\}$$

- $r(y)$ is a specified function; $\alpha$ is a sensitivity analysis parameter.
- $\alpha = 0$ is MAR.

$$\text{logit } P[R_k = 0 | R_{k-1} = 1, Y_K^-] = h_k(Y_{k-1}^-) + \alpha r(Y_k)$$

where

$$
\begin{aligned}
h_k(Y_{k-1}^-) &= \text{logit } P[R_k = 0 | R_{k-1} = 1, Y_{k-1}^-] - \\
&\quad \log\{E[\exp\{\alpha r(Y_k)\} | R_k = 1, Y_{k-1}^-]\}
\end{aligned}
$$

# Missing Not at Random (MNAR) - In General

$$\mu^* = \int_{y_0} \cdots \int_{y_K} y_K \prod_{k=K}^{1} \left\{ dF(y_k | R_k = 1, Y_{k-1}^- = y_{k-1}^-) \frac{1}{1 + \exp\{h_k(y_{k-1}^-) + \alpha r(y_k)\}} + \right.$$
$$\left. \frac{\exp(\alpha r(y_k)) dF(y_k | R_k = 1, Y_{k-1}^- = y_{k-1}^-)}{E[\exp(\alpha r(Y_k)) | R_k = 1, Y_{k-1}^- = y_{k-1}^-]} \frac{\exp\{h_k(y_{k-1}^-) + \alpha r(y_k)\}}{1 + \exp\{h_k(y_{k-1}^-) + \alpha r(y_k)\}} \right\} dF(y_0)$$

$$\mu^* = E\left[ \frac{R_K Y_K}{\prod_{k=1}^{K}(1 + \exp\{h_k(Y_{k-1}^-) + \alpha r(Y_k)\}^{-1}} \right]$$

- $\mu^*$ is written as a function of the distribution of the observed data (depending on $\alpha$).

Restrictions on Distribution of Observed Data

None

Type (ii)

Type (i) Assumptions

Treatment-Specific Mean

# Inference
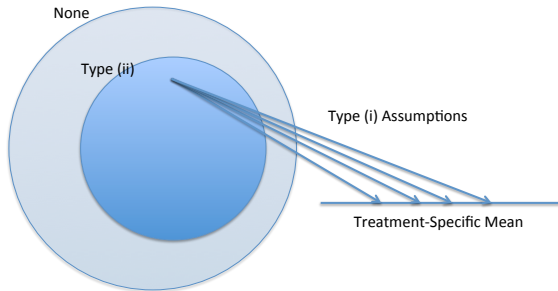
Need to estimate:

$$dF(Y_0)$$
$$dF(Y_k|R_k = 1, Y_{k-1}^-)$$
$$P[R_k = 0|R_{k-1} = 1, Y_{k-1}^-]$$

- Estimate $dF(Y_0)$ by its empirical distribution
- Can't estimate $dF(Y_k|R_k = 1, Y_{k-1}^-)$ and $P[R_k = 0|R_{k-1} = 1, Y_{k-1}^-]$ non-parametrically due to curse of dimensionality. Need Type (ii) assumptions.

First-order Markov

$$dF(Y_k|R_k = 1, Y_{k-1}^-) = dF(Y_k|R_k = 1, Y_{k-1}) \equiv dF_k(Y_k|Y_{k-1})$$

$$P[R_k = 0|R_{k-1} = 1, Y_{k-1}^-] = P[R_k = 0|R_{k-1} = 1, Y_{k-1}] \equiv H_k(Y_{k-1})$$

# Smoothing

$$\widehat{F}_k(y_k|y_{k-1}; \sigma_F) = \frac{\sum_{i=1}^{n} R_{k,i} I(Y_{k,i} \leq y_k) \phi\left(\frac{Y_{k-1,i} - y_{k-1}}{\sigma_F}\right)}{\sum_{i=1}^{n} R_{k,i} \phi\left(\frac{Y_{k-1,i} - y_{k-1}}{\sigma_F}\right)}$$

$$\widehat{H}_k(y_{k-1}; \sigma_H) = \frac{\sum_{i=1}^{n} R_{k-1,i}(1 - R_{k,i}) \phi\left(\frac{Y_{k-1,i} - y_{k-1}}{\sigma_H}\right)}{\sum_{i=1}^{n} R_{k-1,i} \phi\left(\frac{Y_{k-1,i} - y_{k-1}}{\sigma_H}\right)}$$

# Estimating of Smoothing Parameters

J-fold Cross-Validation

$$\widehat{\mathcal{L}}_F^{cv}(\sigma_F) = \frac{1}{J}\sum_{j=1}^{J}\frac{1}{n_j}\sum_{i\in V_j}\sum_{k=1}^{K}R_{k,i}\left[\frac{\sum_\ell R_{k,\ell}\{I(Y_{k,i}\le Y_{k,\ell}) - \hat{F}_k^{(j)}(Y_{k,\ell}\mid Y_{k-1,i};\sigma_F)\}^2}{\sum_\ell R_{k,\ell}}\right]$$

$$\widehat{\mathcal{L}}_H^{cv}(\sigma_H) = \frac{1}{J}\sum_{j=1}^{J}\frac{1}{n_j}\sum_{i\in V_j}\sum_{k=1}^{K}\frac{R_{k-1,i}[1-R_{k,i}-\hat{H}_k^{(j)}(Y_{k-1,i};\sigma_H)]^2\sum_\ell R_{k-1,\ell}(1-R_{k,\ell})}{\sum_\ell R_{k-1,\ell}}$$

- Minimize these weighted loss functions to find optimal $\sigma_F$ and $\sigma_H$, denoted by $\widehat{\sigma}_F$ and $\widehat{\sigma}_H$

# Estimation

- Estimate $F_k(Y_k|Y_{k-1})$ and $H_k(Y_{k-1})$ by $\widehat{F}_k(y_k|y_{k-1}; \widehat{\sigma}_F)$ and $\widehat{H}_k(y_{k-1}; \widehat{\sigma}_H)$; these estimators will not converge at $\sqrt{n}$ rates.

- Plug in these estimators into the $\mu^*$ formula

- This plug-in estimator can suffer from non-standard asymptotics.

- To correct this problem, we use a one-step estimator:

    plug-in $+$ average of estimated influence functions

- The influence function for a patient by $\psi(O; F, H)$. The estimated influence function is $\psi(O; \widehat{F}, \widehat{H})$.

# Uncertainty

- An influence function-based 95% confidence interval takes the form $\widehat{\mu} \pm 1.96 \widehat{se}(\widehat{\mu})$, where

$$\widehat{se}(\widehat{\mu}) = \sqrt{E_n[\psi(O; \widehat{F}, \widehat{H})^2]/n}$$

- In studentized bootstrap, the confidence interval takes the form $[\widehat{\mu} - t_{0.975}\widehat{se}(\widehat{\mu}), \widehat{\mu} - t_{0.025}\widehat{se}(\widehat{\mu})]$, where $t_q$ is the $q$th quantile of $\left\{ \frac{\widehat{\mu}^{(b)} - \widehat{\mu}}{\widehat{se}(\widehat{\mu}^{(b)})} : b = 1, \ldots, B \right\}$

# Uncertainty - Double Bootstrap

- For the $b$th bootstrapped dataset, $n$ observed patient records are repeatedly re-sampled with replacement to create $S$ new datasets.
- For each of these datasets the entire estimation procedure is executed to obtain parameter estimates $\{\widehat{\mu}^{(b,s)} : s = 1, \ldots, S\}$.
- Let $\tilde{t}_q^{(b)}$ to be $q$th quantile of $\left\{ \frac{\widehat{\mu}^{(b,s)} - \widehat{\mu}^{(b)}}{\widehat{se}(\widehat{\mu}^{(b,s)})} : s = 1, \ldots, S \right\}$
- Solve for $q$ such that

$$\left| \frac{1}{B} \sum_{b=1}^{B} I(\widehat{\mu} \in [\widehat{\mu}^{(b)} - \tilde{t}_{1-q}^{(b)} \widehat{se}(\widehat{\mu}^{(b)}), \widehat{\mu}^{(b)} - \tilde{t}_q^{(b)} \widehat{se}(\widehat{\mu}^{(b)})]) - 0.95 \right|$$

  is minimized; denote the solution by $q^*$.
- The 95% double bootstrap confidence interval takes the form $[\widehat{\mu} - t_{1-q^*} \widehat{se}(\widehat{\mu}), \widehat{\mu} - t_{q^*} \widehat{se}(\widehat{\mu})]$.

- The drawback of double bootstrap is that it is computationally intensive.
- Nankervis (2005) has an algorithm that incorporates stopping rules to avoid unnecessary computations.
- Another idea is to set $S = 1$ and define $\tilde{t}_q^{(b)} = \tilde{t}_q$ above to be $q$th quantile of $\left\{ \frac{\widehat{\mu}^{(b,1)} - \widehat{\mu}^{(b)}}{\widehat{se}(\widehat{\mu}^{(b,1)})} : b = 1, \ldots, B \right\}$.

# Bootstrap Bias Correction

$$\widehat{\mu}_{bc} = 2\widehat{\mu} - \frac{1}{B} \sum_{b=1}^{B} \widehat{\mu}^{(b)}$$

# Case Study: SCA-3004

- Randomized trial designed to evaluate the efficacy and safety of once-monthly, injectable paliperidone palmitate (PP1M) relative to placebo (PBO) in delaying the time to relapse in subjects with schizoaffective disorder.
- Open-label phase consisting of a flexible-dose, lead-in period and a fixed-dose, stabilization period.
- Stable subjects entered a 15-month relapse-prevention phase and were randomized to receive PP1M or placebo injections at baseline (Visit 0) and every 28 days (Visits 1-15).
- Additional clinic visit (Visit 16) scheduled for 28 days after the last scheduled injection.
- 170 and 164 subjects were randomized to the PBO and PP1M arms.

## Case Study: SCA-3004

- Research question: Are functional outcomes better in patients with schizoaffective disorder better maintained if they continue on treatment or are withdrawn from treatment and given placebo instead?
- An ideal study would follow all randomized subjects through Visit 16 while maintaining them on their randomized treatment and examine symptomatic and functional outcomes at that time point.
- Since clinical relapse can have a major negative impact, the study design required that patients who had signs of relapse were discontinued from the study.
- In addition, some patients discontinued due to adverse events, withdrew consent or were lost to follow-up.
- 38% and 60% of patients in the PBO and PP1M arms were followed through Visit 16 (p=0.0001).
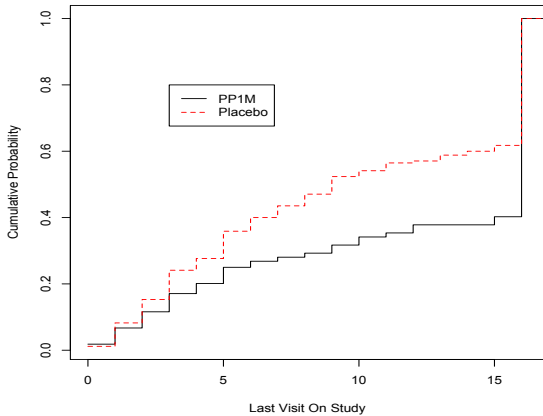
## Case Study: SCA-3004

- Focus: Patient function as measured by the Personal and Social Performance (PSP) scale.
- The PSP scale is scored from 1 to 100 with higher scores indicating better functioning based on evaluation of 4 domains (socially useful activities, personal/social relationships, self-care, and disturbing/aggressive behaviors).
- Estimate treatment-specific mean PSP at Visit 16 in the counterfactual world in which all patients who are followed to Visit 16.
- The mean PSP score among completers was 76.05 and 76.96 in the PBO and PP1M arms; the estimated difference is -0.91 (95%: -3.98:2.15).
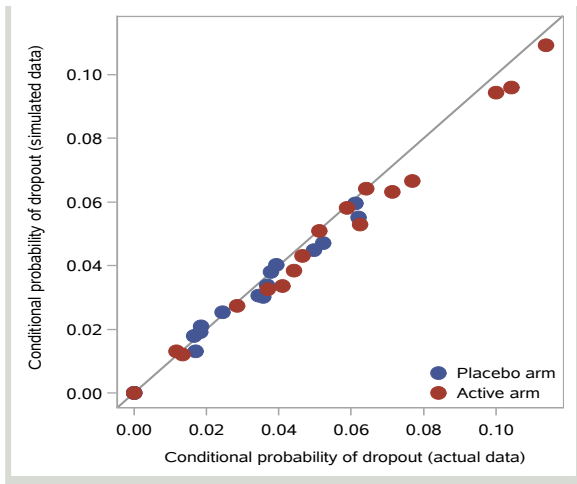
# Case Study: SCA-3004

| L | n | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|
| 0 | 3 | 70.7 | | | | | | | | | | | | | | | | |
| 1 | 8 | 67.6 | 65.3 | | | | | | | | | | | | | | | |
| 2 | 8 | 76.3 | 74.3 | 60.6 | | | | | | | | | | | | | | |
| 3 | 9 | 70.9 | 71.8 | 68.7 | 58.1 | | | | | | | | | | | | | |
| 4 | 5 | 75.2 | 75.2 | 75.6 | 67.6 | 64.6 | | | | | | | | | | | | |
| 5 | 8 | 74.8 | 77.3 | 75.1 | 76.4 | 78.9 | 74.9 | | | | | | | | | | | |
| 6 | 3 | 72.7 | 74.7 | 73.7 | 73.0 | 74.3 | 73.0 | 68.7 | | | | | | | | | | |
| 7 | 2 | 72.0 | 68.5 | 68.5 | 71.0 | 72.5 | 72.5 | 72.5 | 68.5 | | | | | | | | | |
| 8 | 2 | 80.5 | 79.5 | 74.0 | 73.0 | 71.5 | 72.0 | 72.5 | 71.5 | 63.5 | | | | | | | | |
| 9 | 4 | 69.8 | 69.0 | 70.3 | 71.8 | 73.3 | 72.8 | 71.8 | 73.5 | 70.5 | 59.8 | | | | | | | |
| 10 | 4 | 74.3 | 71.8 | 73.3 | 72.5 | 73.5 | 74.0 | 73.8 | 78.0 | 78.0 | 78.0 | 67.3 | | | | | | |
| 11 | 2 | 72.0 | 71.0 | 70.0 | 71.5 | 69.5 | 72.0 | 75.0 | 71.0 | 72.5 | 76.5 | 75.5 | 74.0 | | | | | |
| 12 | 4 | 76.5 | 78.0 | 72.8 | 74.5 | 74.0 | 74.0 | 74.5 | 77.5 | 76.8 | 76.3 | 75.5 | 78.3 | 72.0 | | | | |
| 15 | 4 | 69.8 | 70.8 | 70.0 | 69.8 | 70.8 | 72.8 | 71.5 | 72.0 | 68.0 | 67.3 | 67.0 | 68.3 | 68.0 | 66.0 | 67.0 | 70.3 | |
| 16 | 98 | 73.0 | 73.8 | 73.7 | 74.4 | 74.9 | 75.3 | 74.9 | 75.0 | 75.5 | 75.9 | 76.3 | 76.6 | 76.8 | 76.8 | 76.6 | 77.0 | 77.0 |
| T | 164 | 72.9 | 73.3 | 72.5 | 72.9 | 74.3 | 74.8 | 74.4 | 74.8 | 74.9 | 75.1 | 75.6 | 76.3 | 76.3 | 76.3 | 76.2 | 76.7 | 77.0 |

| L | n | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|
| 0 | 2 | 67.5 | | | | | | | | | | | | | | | | |
| 1 | 12 | 68.3 | 60.2 | | | | | | | | | | | | | | | |
| 2 | 12 | 67.3 | 66.0 | 57.4 | | | | | | | | | | | | | | |
| 3 | 15 | 67.2 | 67.7 | 68.1 | 60.1 | | | | | | | | | | | | | |
| 4 | 6 | 73.3 | 75.7 | 75.0 | 79.7 | 63.7 | | | | | | | | | | | | |
| 5 | 14 | 69.9 | 72.3 | 72.2 | 72.1 | 71.9 | 60.9 | | | | | | | | | | | |
| 6 | 7 | 70.9 | 71.6 | 69.4 | 68.6 | 70.0 | 70.7 | 65.7 | | | | | | | | | | |
| 7 | 6 | 69.7 | 71.5 | 70.8 | 68.8 | 69.8 | 71.5 | 72.0 | 59.7 | | | | | | | | | |
| 8 | 6 | 79.0 | 80.0 | 80.7 | 80.5 | 79.5 | 79.2 | 78.2 | 79.3 | 74.8 | | | | | | | | |
| 9 | 9 | 72.3 | 73.4 | 72.9 | 73.2 | 74.3 | 74.0 | 73.2 | 72.6 | 74.1 | 58.0 | | | | | | | |
| 10 | 3 | 73.3 | 75.0 | 75.7 | 75.7 | 80.0 | 79.7 | 80.0 | 72.7 | 75.7 | 76.3 | 52.3 | | | | | | |
| 11 | 4 | 72.5 | 71.0 | 71.0 | 68.5 | 70.0 | 68.8 | 70.5 | 72.5 | 70.3 | 67.8 | 64.3 | 60.8 | | | | | |
| 12 | 1 | 62.0 | 62.0 | 62.0 | 62.0 | 62.0 | 62.0 | 62.0 | 62.0 | 62.0 | 62.0 | 63.0 | 63.0 | 62.0 | | | | |
| 13 | 3 | 81.7 | 75.0 | 73.7 | 78.3 | 76.3 | 75.7 | 77.7 | 71.3 | 78.3 | 78.7 | 77.3 | 73.0 | 70.0 | 55.3 | | | |
| 14 | 2 | 77.0 | 79.5 | 74.5 | 76.5 | 80.0 | 74.0 | 81.0 | 81.0 | 82.5 | 77.0 | 81.5 | 75.5 | 74.5 | 75.0 | 65.0 | | |
| 15 | 3 | 65.7 | 65.7 | 65.3 | 66.0 | 66.3 | 66.0 | 67.0 | 68.0 | 67.3 | 67.3 | 68.7 | 70.0 | 68.7 | 68.7 | 67.3 | 65.3 | |
| 16 | 65 | 72.1 | 73.0 | 73.2 | 73.3 | 73.2 | 73.3 | 73.5 | 74.3 | 74.6 | 75.3 | 75.3 | 75.3 | 76.0 | 76.3 | 76.0 | 76.5 | 76.0 |
| T | 170 | 71.1 | 71.2 | 71.3 | 71.8 | 72.7 | 71.8 | 73.2 | 73.2 | 74.4 | 73.0 | 73.7 | 74.1 | 75.3 | 75.1 | 75.3 | 76.0 | 76.0 |
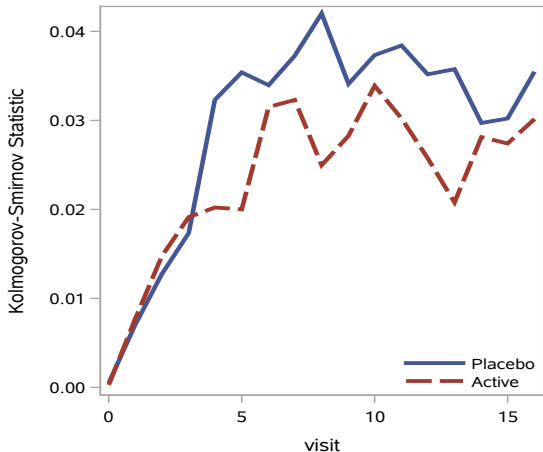
# Case Study: SCA-3004

- Under MAR (i.e., $\alpha = 0$), the estimated means of interest are 69.60 and 74.37 for the PBO and PP1M arms.
- The estimated treatment difference is $-4.77$ (95% CI: -10.89 to 0.09).
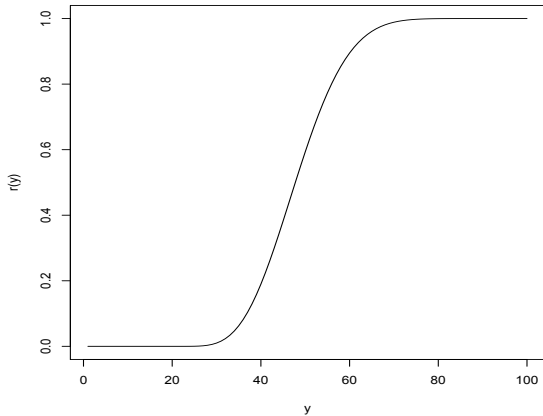
| $y_{k+1}$ | $y_{k+1}^*$ | Log Odds Ratio |
|-----------|-------------|----------------|
| 30 | 20 | $\alpha \times 0.01$ |
| 40 | 30 | $\alpha \times 0.18$ |
| 50 | 40 | $\alpha \times 0.40$ |
| 60 | 50 | $\alpha \times 0.30$ |
| 70 | 60 | $\alpha \times 0.09$ |
| 80 | 700 | $\alpha \times 0.01$ |
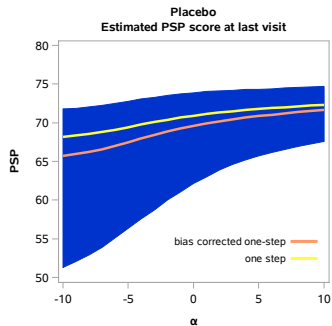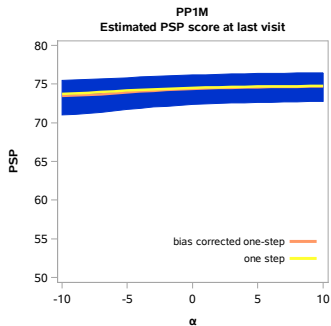
**Difference in PSP score at last visit**
**non-completers - completers**
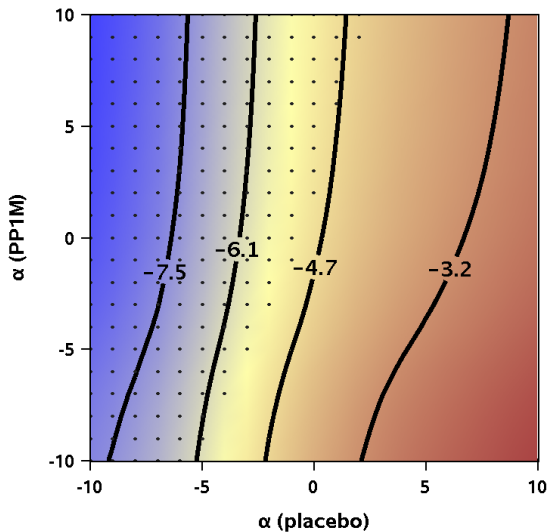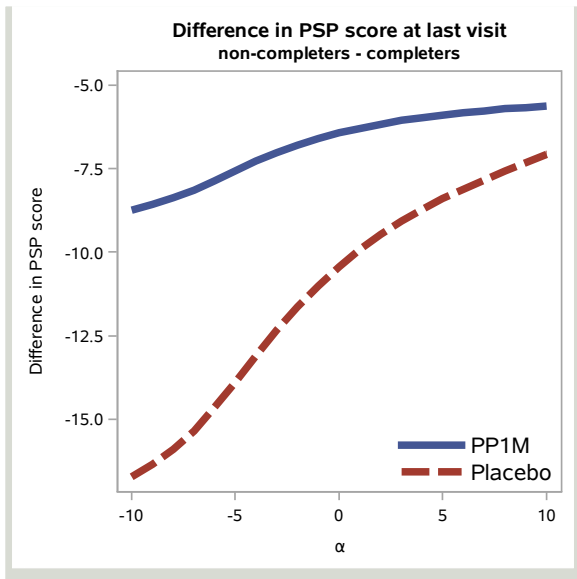
# Simulation Study

| | | PP1M | | | PBO | | |
|---|---|---|---|---|---|---|---|
| $\alpha$ | Estimator | $\mu^*$ | Bias | MSE | $\mu^*$ | Bias | MSE |
| -10 | $\mu(\widehat{P})$ | 73.64 | 0.43 | 1.41 | 69.06 | 2.04 | 7.47 |
| | $\widehat{\mu}$ | | 0.33 | 1.29 | | 1.53 | 6.47 |
| | $\widehat{\mu}_{bc}$ | | 0.07 | 2.28 | | 0.55 | 9.03 |
| -5 | $\mu(\widehat{P})$ | 74.25 | 0.29 | 1.17 | 70.23 | 1.55 | 5.12 |
| | $\widehat{\mu}$ | | 0.19 | 1.08 | | 1.13 | 4.54 |
| | $\widehat{\mu}_{bc}$ | | -0.00 | 1.98 | | 0.38 | 6.86 |
| -1 | $\mu(\widehat{P})$ | 74.59 | 0.20 | 1.04 | 71.47 | 0.94 | 3.05 |
| | $\widehat{\mu}$ | | 0.09 | 0.96 | | 0.59 | 2.84 |
| | $\widehat{\mu}_{bc}$ | | -0.07 | 1.82 | | 0.08 | 4.98 |
| 0 | $\mu(\widehat{P})$ | 74.63 | 0.19 | 1.03 | 71.70 | 0.82 | 2.75 |
| | $\widehat{\mu}$ | | 0.08 | 0.95 | | 0.50 | 2.61 |
| | $\widehat{\mu}_{bc}$ | | -0.07 | 1.82 | | 0.04 | 4.68 |
| 1 | $\mu(\widehat{P})$ | 74.67 | 0.18 | 1.01 | 71.90 | 0.72 | 2.52 |
| | $\widehat{\mu}$ | | 0.07 | 0.94 | | 0.43 | 2.42 |
| | $\widehat{\mu}_{bc}$ | | -0.07 | 1.79 | | 0.01 | 4.44 |
| 5 | $\mu(\widehat{P})$ | 74.77 | 0.16 | 0.99 | 72.41 | 0.48 | 2.04 |
| | $\widehat{\mu}$ | | 0.06 | 0.92 | | 0.27 | 2.03 |
| | $\widehat{\mu}_{bc}$ | | -0.07 | 1.75 | | -0.03 | 3.87 |
| 10 | $\mu(\widehat{P})$ | 74.84 | 0.15 | 0.97 | 72.74 | 0.34 | 1.80 |
| | $\widehat{\mu}$ | | 0.06 | 0.91 | | 0.20 | 1.82 |
| | $\widehat{\mu}_{bc}$ | | -0.06 | 1.73 | | -0.01 | 3.54 |

# Software - SAMON

www.missingdatamatters.org

- R version
- SAS version coming soon.

# Software - SAMON

- Requires that the outcome measure to be coded as a positive integer (1,2,3, . . . )
- Dataset should be constructed with one record per subject (i.e., wide format)
- Missing data should be represented with the numeric value -1
- The first time-point (i.e., baseline value) should not contain any missing values

# Software - SAMON

```
        V1  V2  V3  V4  V5  V6
 [1,]   90  87  86  93  72  87
 [2,]  112  -1  -1  -1  -1  -1
 [3,]   99  76  62  52  57  49
 [4,]   86  78  91 113  89  68
 [5,]   80  85  -1  -1  -1  -1
 [6,]   72  64  78 113  -1  -1
 [7,]   67  -1  -1  -1  -1  -1
 [8,]   96  -1  -1  -1  -1  -1
 [9,]   93  90  -1  -1  -1  -1
[10,]   78  70  53  85  -1  -1
[11,]   93  86  92  94  -1  -1
[12,]  111 112  95  -1  -1  -1
[13,]  107  79  90  -1  -1  -1
[14,]  132 113 114  -1  -1  -1
[15,]   97  78  60  55  73  72
[16,]  121  -1  -1  -1  -1  -1
```

# Software - SAMON

```
library(samon, lib.loc="../../samlib")

data("samonPANSS1")    # Treatment 1 data.

Results1a <- samon(
    mat             = samonPANSS1,
    Npart           = 10,

    InitialSigmaP   = 10.0,          # initial value
    HighSigmaP      = 50.0,          # high value for p

    InitialSigmaQ   = 8.0,           # initial value
    HighSigmaQ      = 50.0,          # high value for q

    LowAlpha        = -10,           # alphas
    HighAlpha       =  10,
    IncrementAlpha  =   1,

    lb              = 30,            # parameters to
    ub              = 210,           # cumulative
    zeta1           = 4.0,           # beta distribution
    zeta2           = 7.0,

    NBootstrapPairs = 500,           # number of bootstraps
    seed0           = 826847827,

    SAconvg         = 1E-5,          # stopping criteria
    FAconvg         = 1E-5,
    FRconvg         = 1E-5)

saveRDS(Results1a, "Results1a.rds") # save results for latter use
```