

Global Sensitivity Analysis of Randomized Trials with Missing Data

FDA 2015 Shortcourse

Daniel Scharfstein

Aidan McDermott

Chenguang Wang

Johns Hopkins University

dscharf@jhu.edu

November 30, 2015

Agenda

- 9-10:30: Methods for Studies with Monotone Missing Data (DS)
- 10:30-11: Break
- 11-12: Software for Studies with Monotone Missing Data (AM)
- 12-1: Lunch
- 1-2:30: Methods and Software for Studies with Death and Intermittent Missing Data (CW)
- 2:30-3 Break
- 3-4 Open Discussion

Funding Acknowledgments

- FDA
- PCORI

Case Study: Chronic Schizophrenia

- Major breakthroughs have been made in the treatment of patients with psychotic symptoms.
- However, side effects associated with typical and atypical neuroleptics have limited their usefulness.
- RIS-INT-3 (Marder and Meibach, 1994, Chouinard *et al.*, 1993) was a multi-center study designed to assess the effectiveness and adverse experiences of four fixed doses of risperidone compared to haliperidol and placebo in the treatment of chronic schizophrenia.

- At selection, patients were required to have a PANSS (Positive and Negative Syndrome Scale) score between 60 and 120.
- Prior to randomization, there was a single-blind, one-week washout phase during which all anti-psychotic medications were to be discontinued.
- If acute psychotic symptoms occurred, patients were randomized to a double-blind treatment phase, scheduled to last 8 weeks.
- Patients were randomized to one of 6 treatment groups: risperidone 2, 6, 10 or 16 mg, haliperidol 20 mg, or placebo.
- Dose titration occurred during the first week of the double-blind phase.

- Patients scheduled for 5 post-baseline assessments at weeks 1,2,4,6, and 8 of the double-blind phase.
- Primary efficacy variable: PANSS score
- 521 patients randomized to receive placebo ($n = 88$), haliperidol 20 mg ($n = 87$), risperidone 2mg ($n = 87$), risperidone 6mg ($n = 86$), risperidone 10 mg ($n = 86$), or risperidone 16 mg ($n = 87$).

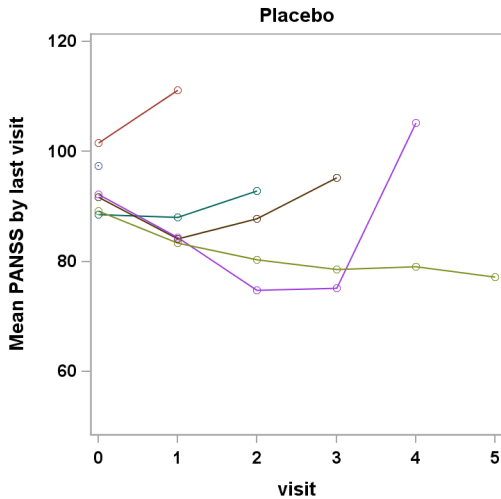
Premature Withdrawal

- Only 49% of patients completed the 8 week treatment period.
- The most common reason for discontinuation was “insufficient response.”
- Other main reasons included: adverse events, uncooperativeness, and withdrawal of consent.

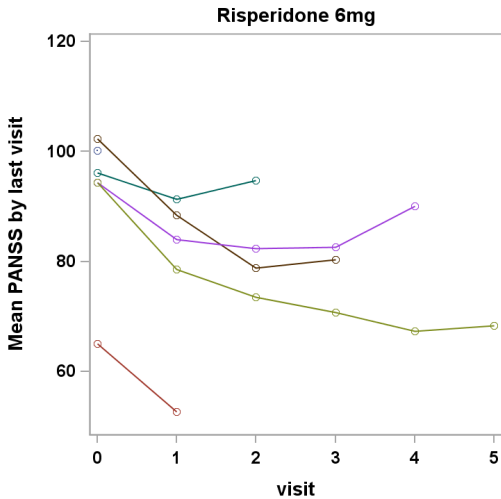
Premature Withdrawal

| | Placebo (n = 88) | | Haliperidol (n = 87) | | Risp 2mg (n = 87) | | Risp 6mg (n = 86) | | Risp 10mg (n = 86) | | Risp 16 mg (n = 87) | |
|------------------|---------------------|-----|-------------------------|-----|----------------------|-----|----------------------|-----|-----------------------|-----|------------------------|-----|
| Completed | 27 | 31% | 36 | 41% | 36 | 41% | 53 | 62% | 48 | 56% | 54 | 62% |
| Withdrawn | 61 | 69% | 51 | 59% | 51 | 59% | 33 | 38% | 38 | 44% | 33 | 38% |
| Lack of Efficacy | 51 | 58% | 36 | 41% | 41 | 47% | 12 | 14% | 25 | 29% | 18 | 21% |
| Other | 10 | 11% | 15 | 17% | 10 | 11% | 21 | 24% | 13 | 15% | 15 | 17% |

Observed Data



Observed Data



Central Question

What is the difference in the mean PANSS scores at week 8 between 6mg. risperidone at a specified dose level vs. placebo in the counterfactual world in which all patients were followed to that week?

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 report 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.*

ICH, EMEA and Sensitivity Analysis

- 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

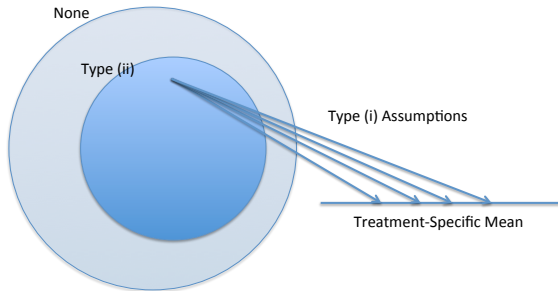
- There are an infinite number of ways of positing type (i) assumptions.
- Ultimately, however, these assumptions prescribe how missing outcomes should be "imputed."
- A reasonable way to posit these assumptions is to
 - stratify individuals with missing outcomes according to the data that we were able 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.

Global Sensitivity Analysis

Restrictions on Distribution of Observed Data



Notation

- K scheduled post-baseline assessments.
- There are $(K + 1)$ patterns representing each of the visits an individual might last be seen, i.e., $0, \dots, K$.
- 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 $Y_k^- = (Y_0, \dots, Y_k)$ and $Y_k^+ = (Y_{k+1}, \dots, Y_K)$.

- 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 $\mu^* = E[Y_K]$.

Example: $K = 2$

- 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

$$\begin{aligned}f(Y_1, Y_2 | R_1 = 0, Y_0) &= f(Y_1, Y_2 | R_1 = 1, Y_0) \\f(Y_2 | R_2 = 0, R_1 = 1, Y_1^-) &= f(Y_2 | R_2 = 1, Y_1^-)\end{aligned}$$

or

$$\begin{aligned}P[R_1 = 0 | Y_2^-] &= P[R_1 = 0 | Y_0] \\P[R_2 = 0 | R_1 = 1, Y_2^-] &= P[R_2 = 0 | R_1 = 1, Y_1^-]\end{aligned}$$

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)

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

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

- μ^* is written as a function of the distribution of the observed data.

Missing Not at Random (MNAR)

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

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; α is a sensitivity analysis parameter.
- $\alpha = 0$ is MAR.

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

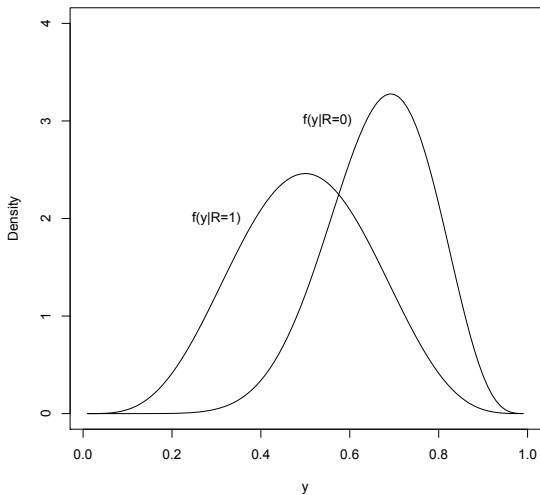
$$h_1(Y_0) = \text{logit } P[R_1 = 0 | Y_0] - \log\{E[\exp\{\alpha r(Y_1)\} | R_1 = 1, Y_0]\}$$

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

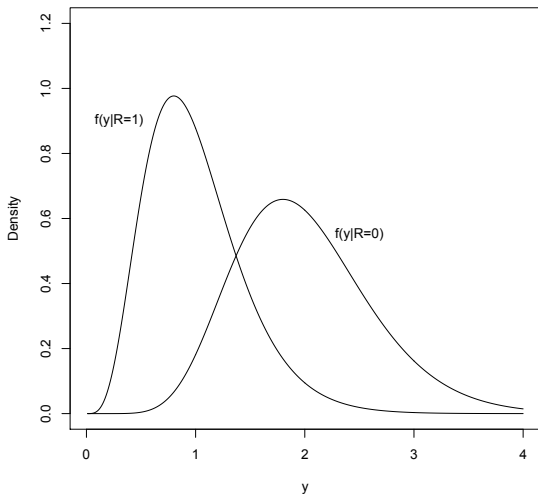
Exponential Tilting Explained

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

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



Gamma



Missing Not at Random (MNAR)

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

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

- μ^* is written as a function of the distribution of the observed data (depending on α).

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; α is a sensitivity analysis parameter.
- $\alpha = 0$ is MAR.

Missing Not at Random (MNAR) - In General

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

where

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

Missing Not at Random (MNAR) - In General

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

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

- μ^* is written as a function of the distribution of the observed data (depending on α).

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.

Type (ii) Assumptions

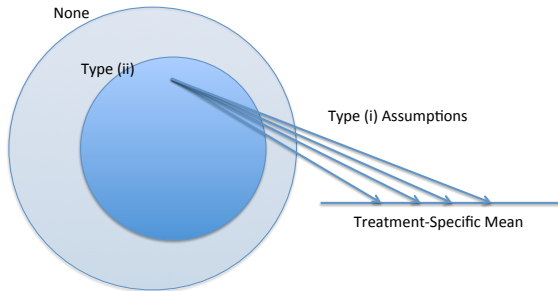
First-order Markov

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

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

Global Sensitivity Analysis

Restrictions on Distribution of Observed Data



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

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

Estimating of Smoothing Parameters

J-fold Cross-Validation

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

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

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

- 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 μ^* 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 $\psi(O; \widehat{F}, \widehat{H})$.

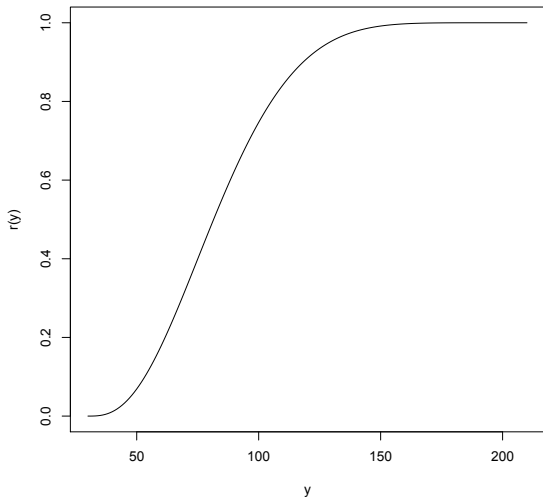
Uncertainty

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

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

- In equal-tailed studentized bootstrap, the confidence interval takes the form $[\hat{\mu} - t_{0.975}\hat{se}(\hat{\mu}), \hat{\mu} - t_{0.025}\hat{se}(\hat{\mu})]$, where t_q is the q th quantile of $\left\{ \frac{\hat{\mu}^{(b)} - \hat{\mu}}{\hat{se}(\hat{\mu}^{(b)})} : b = 1, \dots, B \right\}$
- In symmetric studentized bootstrap, the confidence interval takes the form $[\hat{\mu} - t_{0.95}^*\hat{se}(\hat{\mu}), \hat{\mu} + t_{0.95}^*\hat{se}(\hat{\mu})]$, where $t_{0.95}^*$ is selected so that 95% of the distribution of $\left\{ \frac{\hat{\mu}^{(b)} - \hat{\mu}}{\hat{se}(\hat{\mu}^{(b)})} : b = 1, \dots, B \right\}$ falls between $-t_{0.95}^*$ and $t_{0.95}^*$.
- Useful to replace influence-function based standard error estimator with jackknife standard error estimator.

RIS-INT-3: Bias Function



RIS-INT-3: Bias Function

Consider two patients who are on study through visit k and have the same history of measured factors through that visit. Suppose that the first and second patients have PANSS score at visit $k + 1$ of y_{k+1} and y_{k+1}^* , respectively ($y_{k+1} < y_{k+1}^*$).

The logarithm of the ratio of the odds of last being seen at visit k as opposed to remaining on study for the second versus the first patient is equal to $\alpha\{r(y_{k+1}^*) - r(y_{k+1})\}$.

RIS-INT-3: Bias Function

| y_{k+1}^* | y_{k+1} | Log Odds Ratio |
|-------------|-----------|----------------|
| 50 | 30 | $\alpha 0.02$ |
| 60 | 40 | $\alpha 0.07$ |
| 80 | 60 | $\alpha 0.22$ |
| 100 | 80 | $\alpha 0.30$ |
| 120 | 100 | $\alpha 0.24$ |
| 140 | 120 | $\alpha 0.12$ |
| 160 | 140 | $\alpha 0.04$ |
| 180 | 160 | $\alpha 0.01$ |
| 200 | 180 | $\alpha 0.00$ |

RIS-INT-3: Bias Function

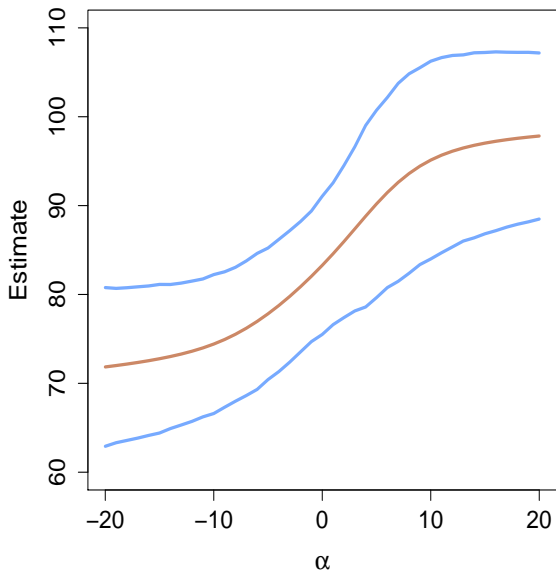
We assumed that $-20.0 \leq \alpha \leq 20.0$. Most reasonable that $\alpha \geq 0$.

When $\alpha = 4$, a patient with a PANSS score at visit $k + 1$ of 100 (120;80) vs. a patient with a PANSS score at visit $k + 1$ of 80 (100;60), has 3.3 (2.6;2.4) times the odds of last being seen at visit k vs. remaining on study.

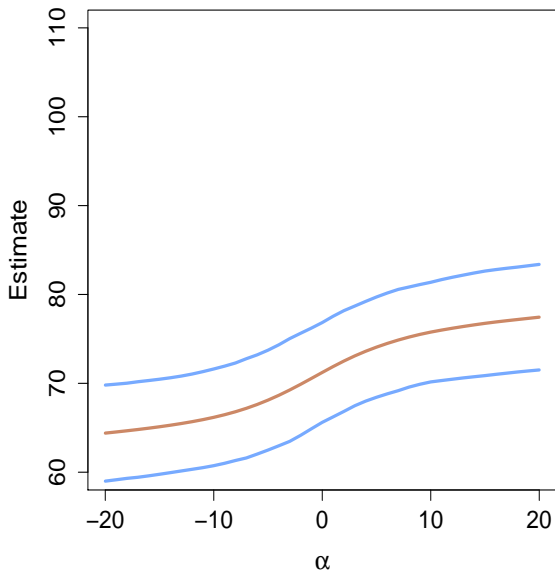
Results

| | Placebo | 6mg Risp | Difference | p-value |
|------|---------|----------|------------|---------|
| MCAR | 78.26 | 68.63 | -9.63 | 0.06 |
| LOCF | 89.42 | 75.91 | -13.51 | 0.00 |
| MAR | 83.28 | 71.23 | -12.05 | <0.05 |

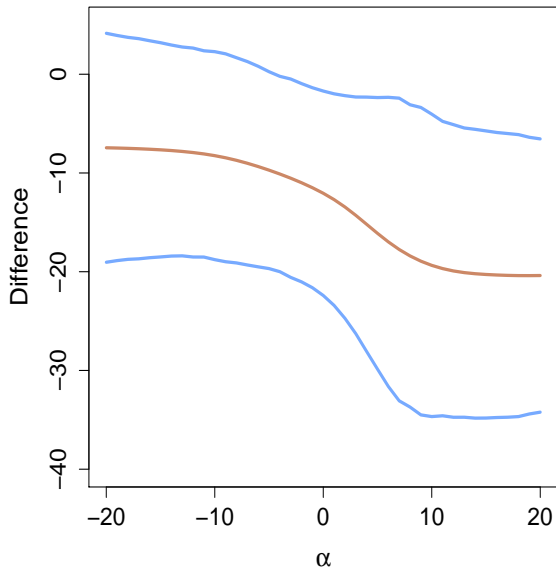
Placebo



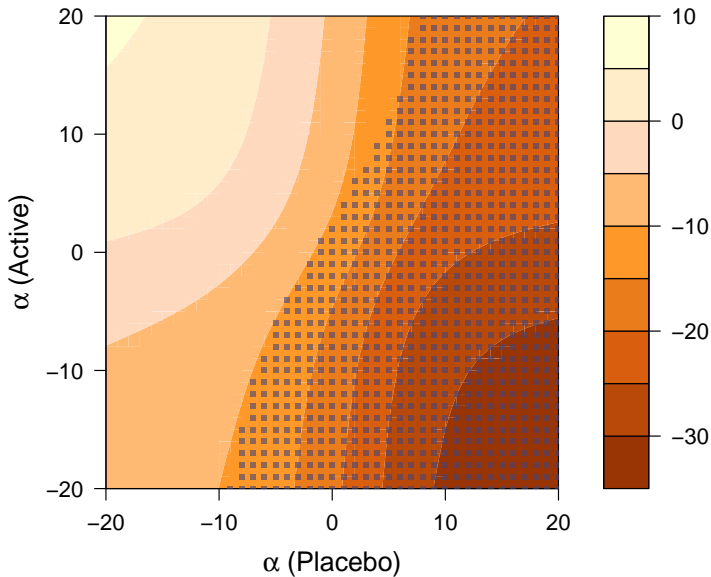
6 mg Risp



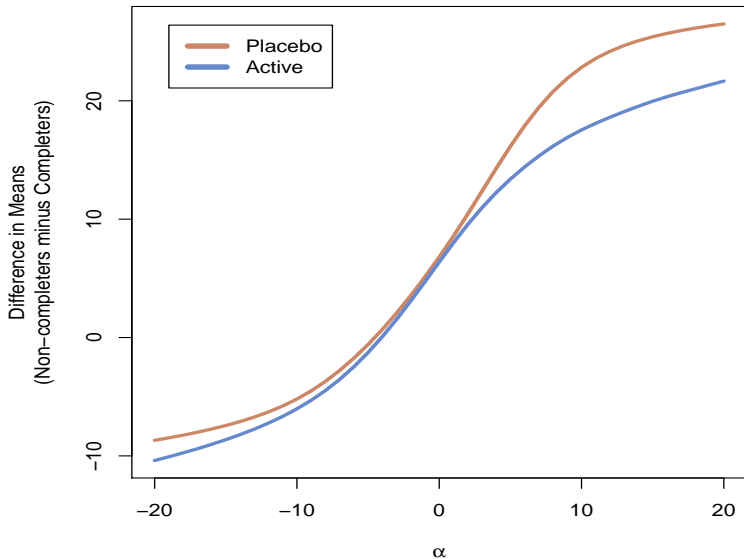
Difference



Contour



Understanding α



The evaluation of robustness needs to be based on scientific considerations.

- Is it reasonable to believe that patients with higher PANSS scores are more likely to be dropping out (i.e., $\alpha > 0$)?
- How much difference is reasonable between the mean PANSS scores for completers and non-completers? 10, 20, 50?

Next Steps

- More simulations to evaluate finite sample performance
- Faster algorithms
- Non-monotone missing data
- More case studies

Missing Data Matters

- No substitute for better trial design and procedures to minimize missing data.
- Global sensitivity analysis should be a mandatory component of trial reporting.
- Visit us at www.missingdatamatters.org or email me at dscharf@jhu.edu