Missing Data in Randomized Studies and the Need for Global Sensitivity Analysis

Daniel Scharfstein
Johns Hopkins University
dscharf@jhsph.edu

October 1, 2013

Missing Data Matters

- Missing outcome data are a widespread problem in randomized trials, including those used as the basis of regulatory approval of drugs and devices.

Neuronetics TMS Study

- Randomized trial of TMS System vs. sham control.
- One week no treatment phase, six week acute treatment phase, three week taper phase.
- Primary outcome: MADRS; scheduled to be measured at baseline, weeks 2, 4 and 6 of acute phase and weeks 1, 2 and 3 of taper phase.
- Primary treatment comparison: MADRS at week 4 of acute phase.
- Secondary treatment comparison: MADRS at week 6 of acute phase.
- TMS: $n = 155$; sham: $n = 146$

Neuronetics Study

Table: On-Study

	Acute			Taper		
	Wk 2	Wk 4	Wk 6	Wk 1	Wk 2	Wk 3
TMS	97%	92%	55%	41%	38%	35%
Sham	98%	92%	40%	27%	26%	24%

- After week 4 of the acute phase, treatment discontinuation was primarily due to lack of efficacy.
- Primary analysis of acute phase used LOCF.
- FDA requested alternative analyses: completers-only analysis and multiple imputation analysis

What are the Estimands?

What are the differences in the mean MADRS scores at weeks 4 and 6 of the acute phase between TMS vs. sham in the counterfactual world in which there are no missing data at these visits?

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.

Missing Data Matters

- 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.*

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.*"

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.

Neuronetics Study

- LOCF: Valid if patients outcomes don't change after dropout.
- Completers: Valid under *Missing Completely at Random*; Distribution of outcomes for patients off study at week k is the same as the distribution of outcome for patients on study at week k.
- Multiple Imputation: Valid under *Missing at Random*; For patients on-study at week $k-1$ and who share the same history of observed outcomes through week $k-1$, the distribution of outcomes after week $k-1$ is the same for those who are last seen at week $k-1$ and those who remain on-study at week k.

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.

Case Study 1: 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.

RIS-INT-3

- 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.

RSIP-INT-3

- Patients scheduled for 5 post-baseline assessements at weeks 1,2,4,6, and 8 of the double-blind phase.
- Primary efficiacy 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 ($n = 87$).

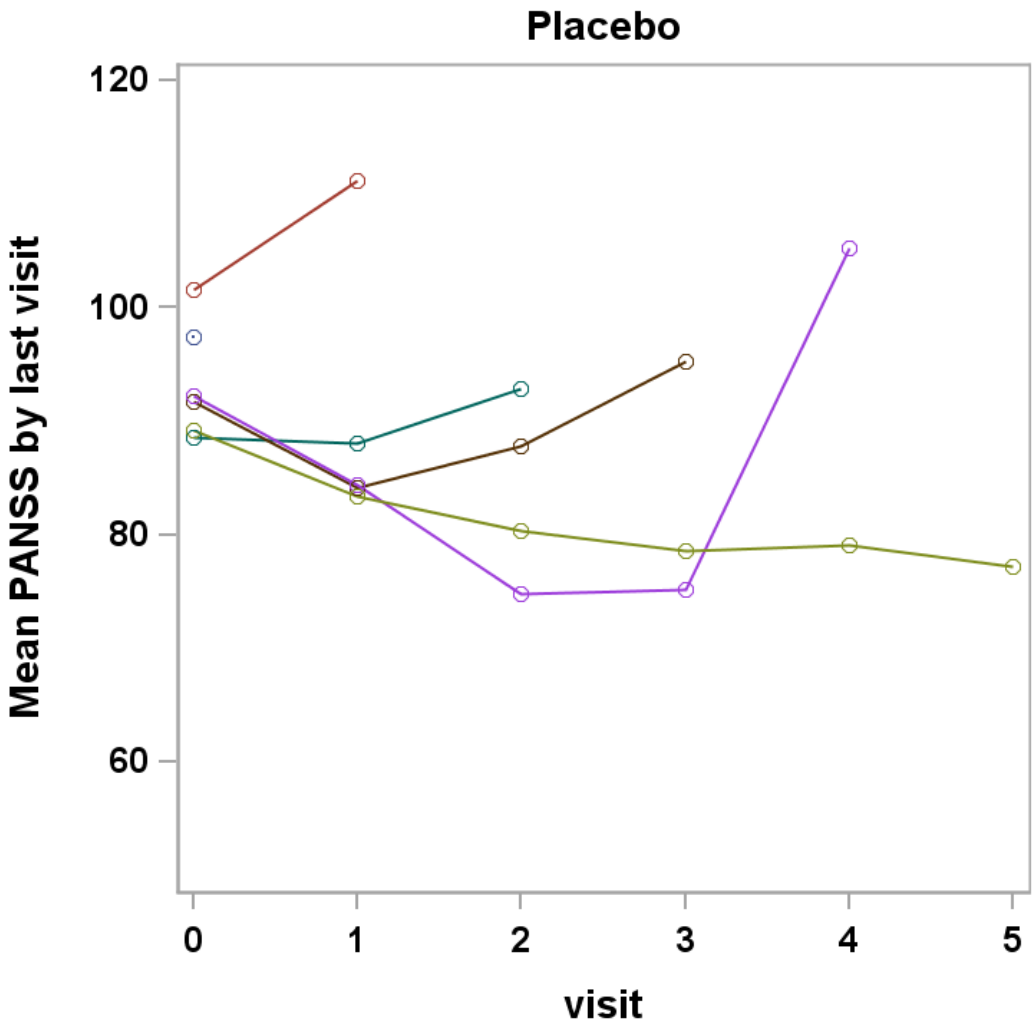
Premature Withdrawal

- 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.

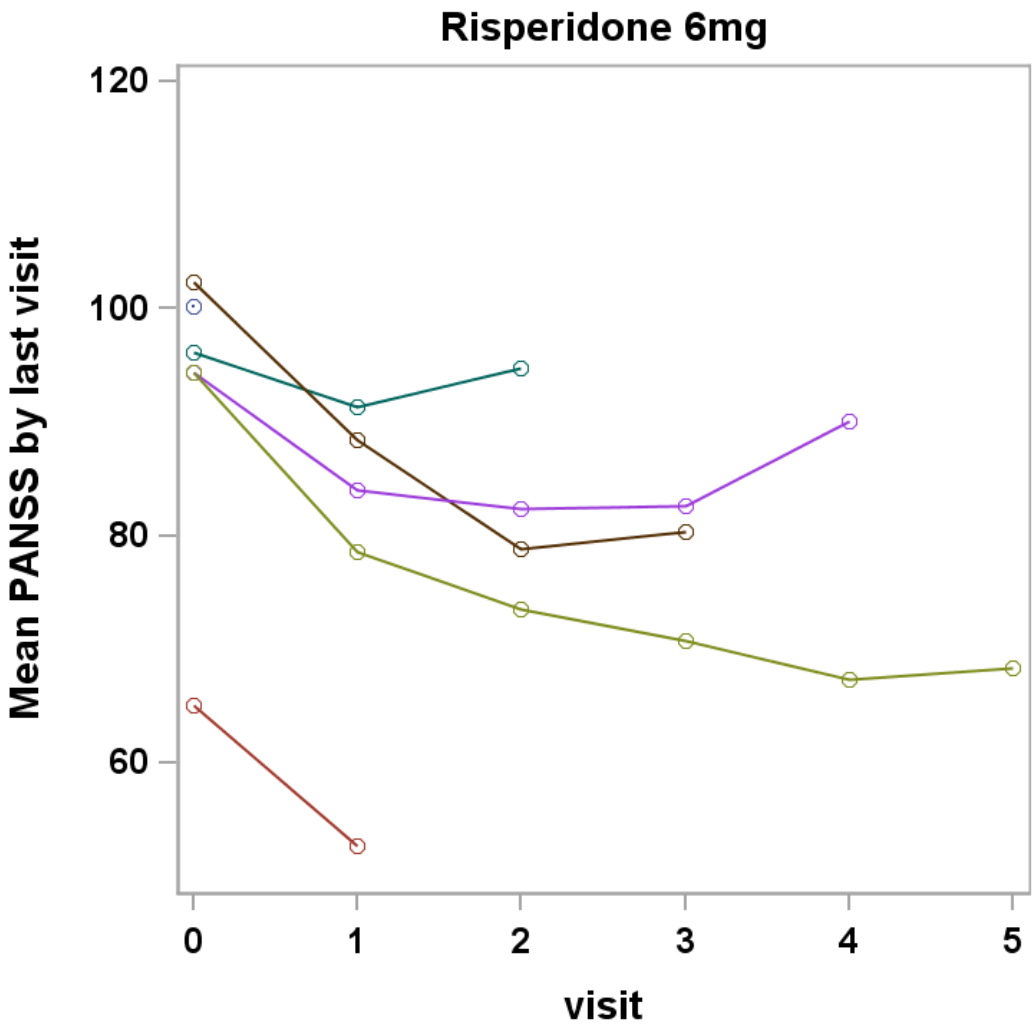
Premature Withdrawal

	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 risperidone at a specified dose level vs. placebo in the counterfactual world in which all patients were followed to that week?

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.
- 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.

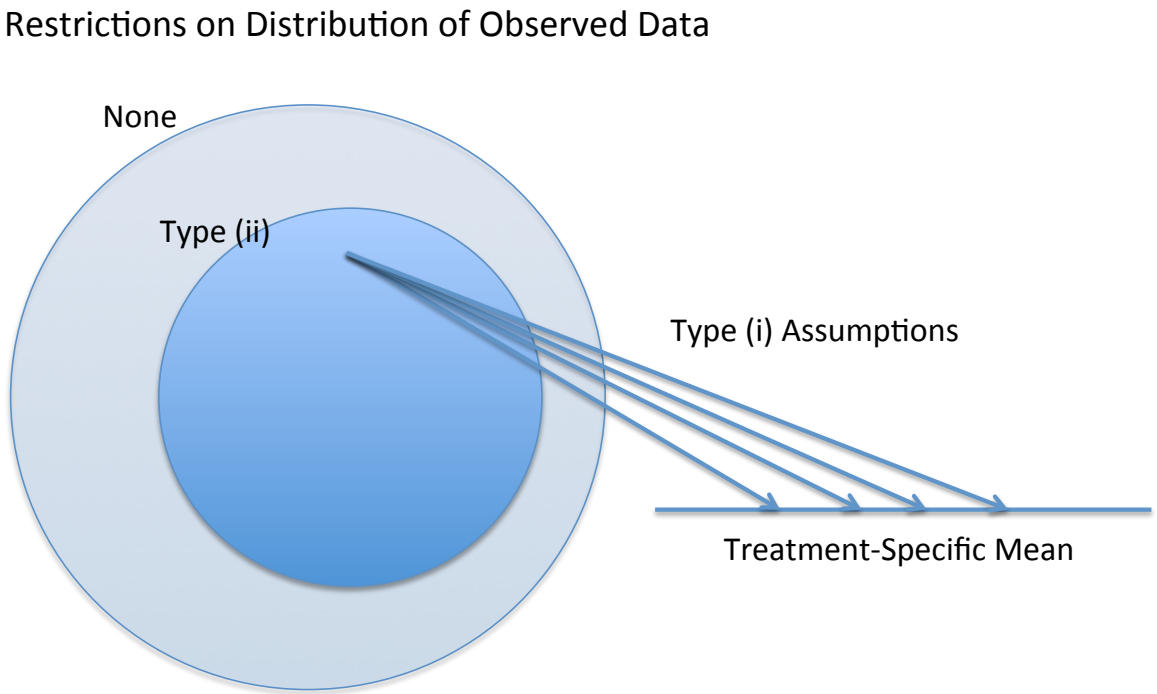
Global Sensitivity Analysis

- 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.

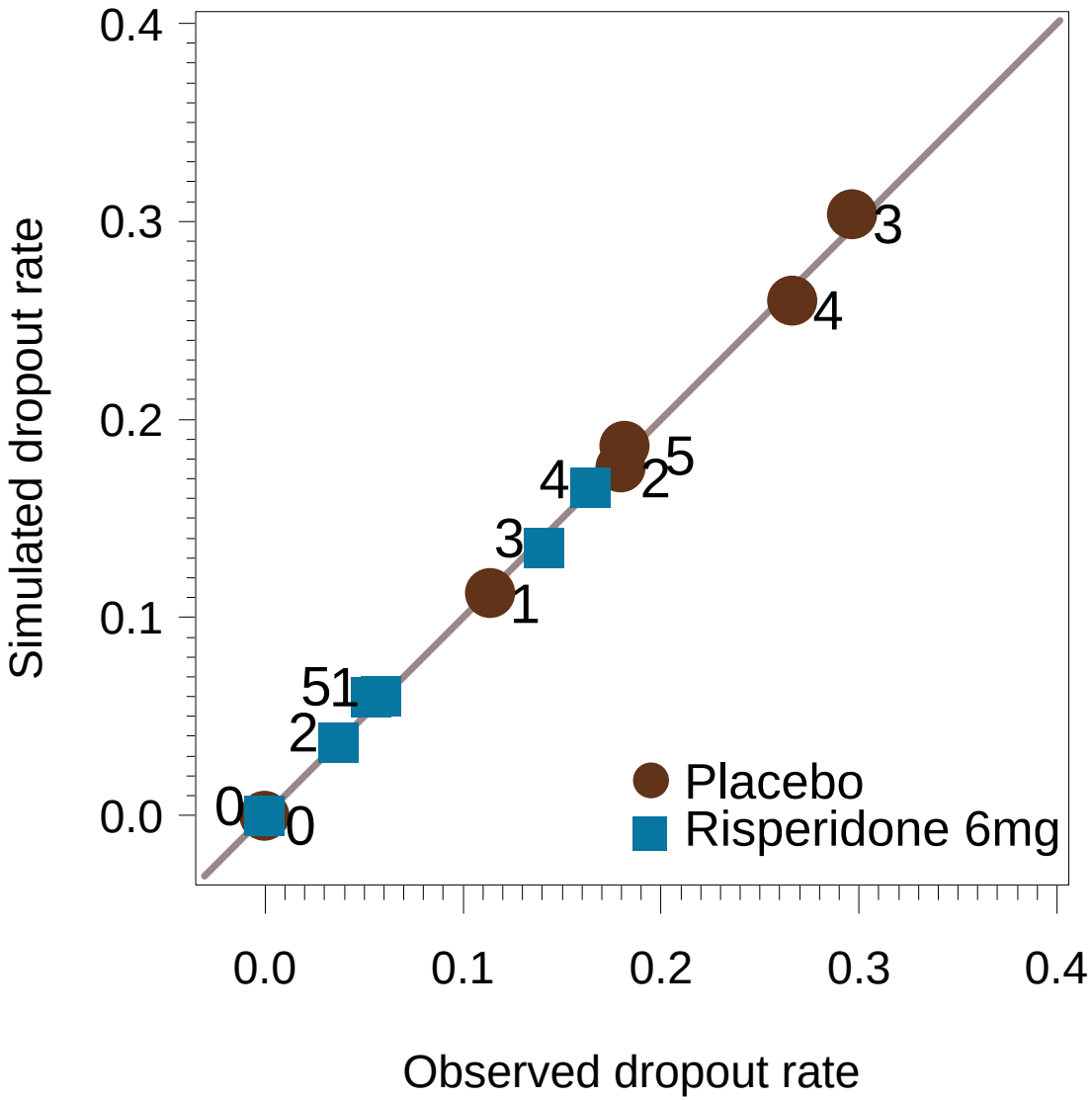
Global Sensitivity Analysis

- 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.

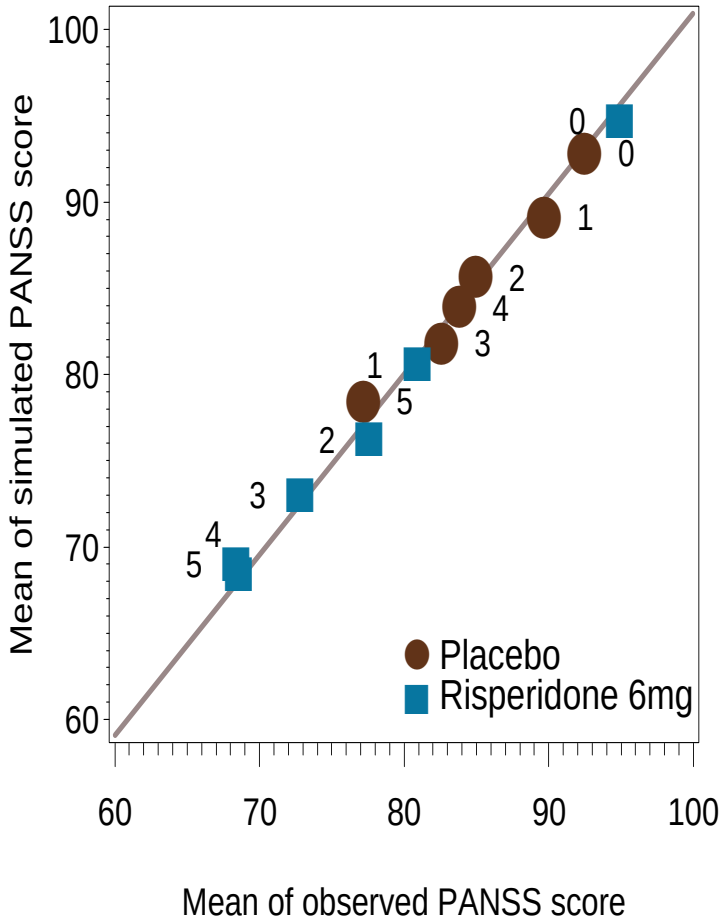
Global Sensitivity Analysis

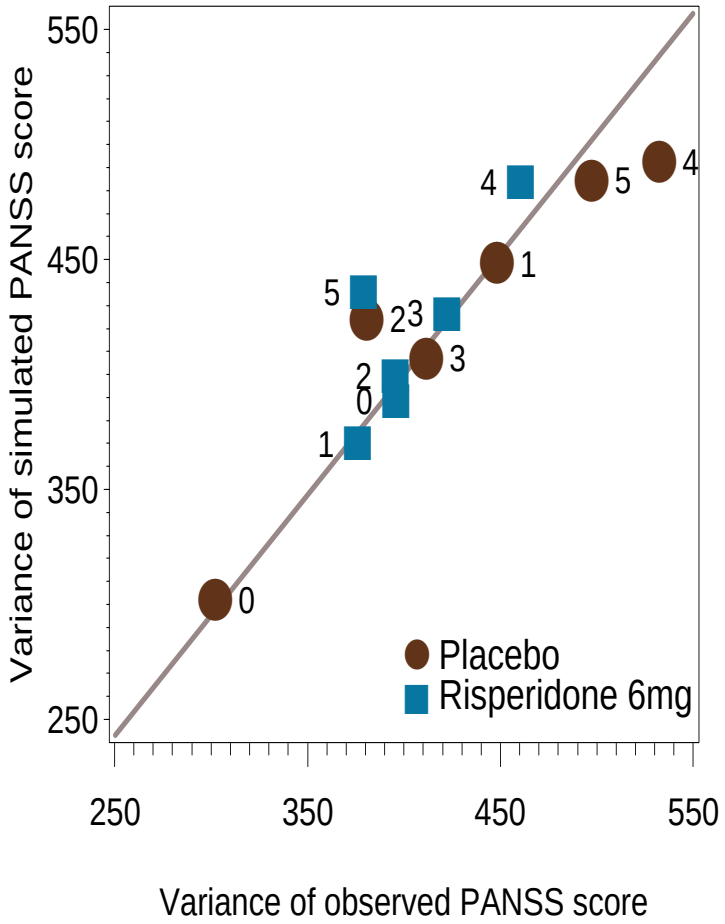


Example

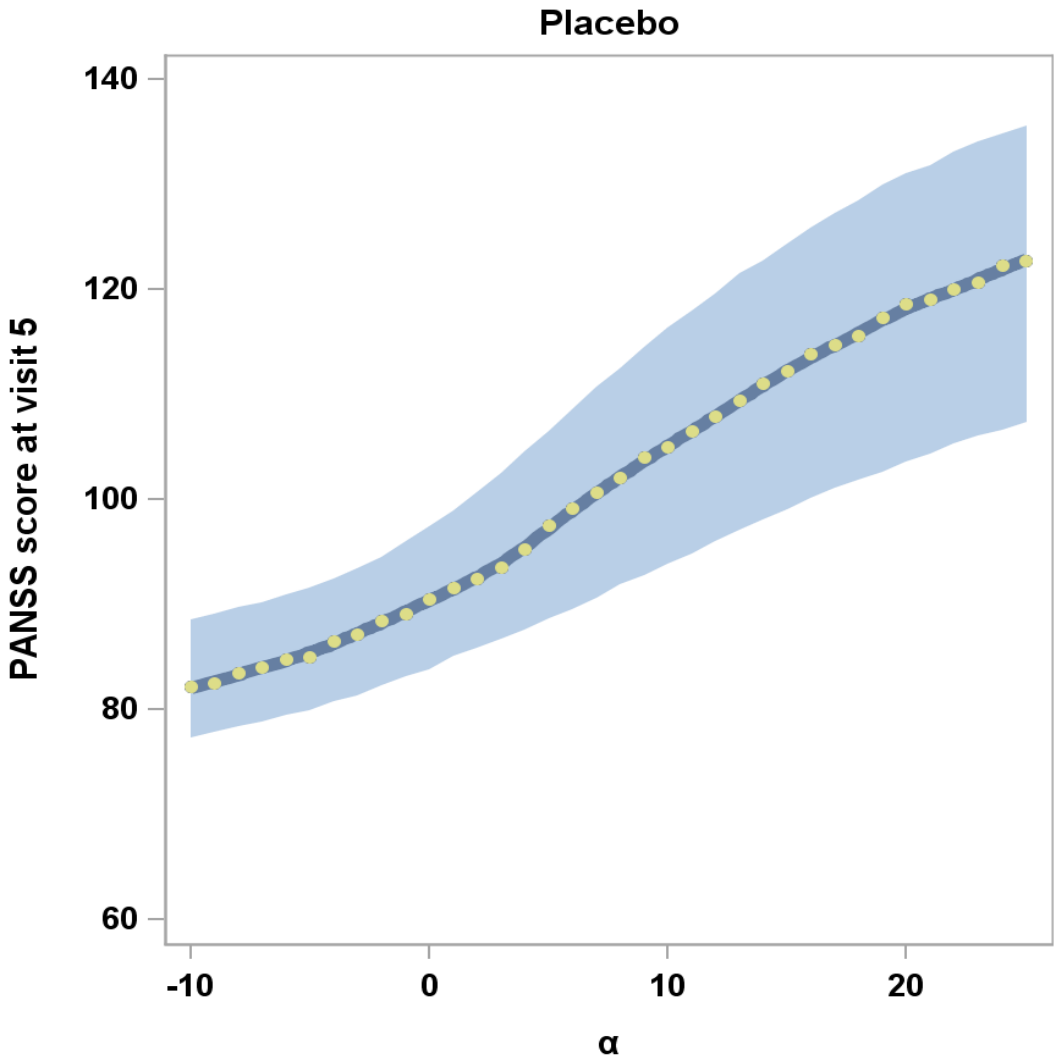

Simulated dropout rate
0.0
0.1
0.2
0.3
0.4
Observed dropout rate
0.0
0.1
0.2
0.3
0.4
Placebo
Risperidone 6mg
0
0
1
5
2
3
4
4
2
5
3

Example

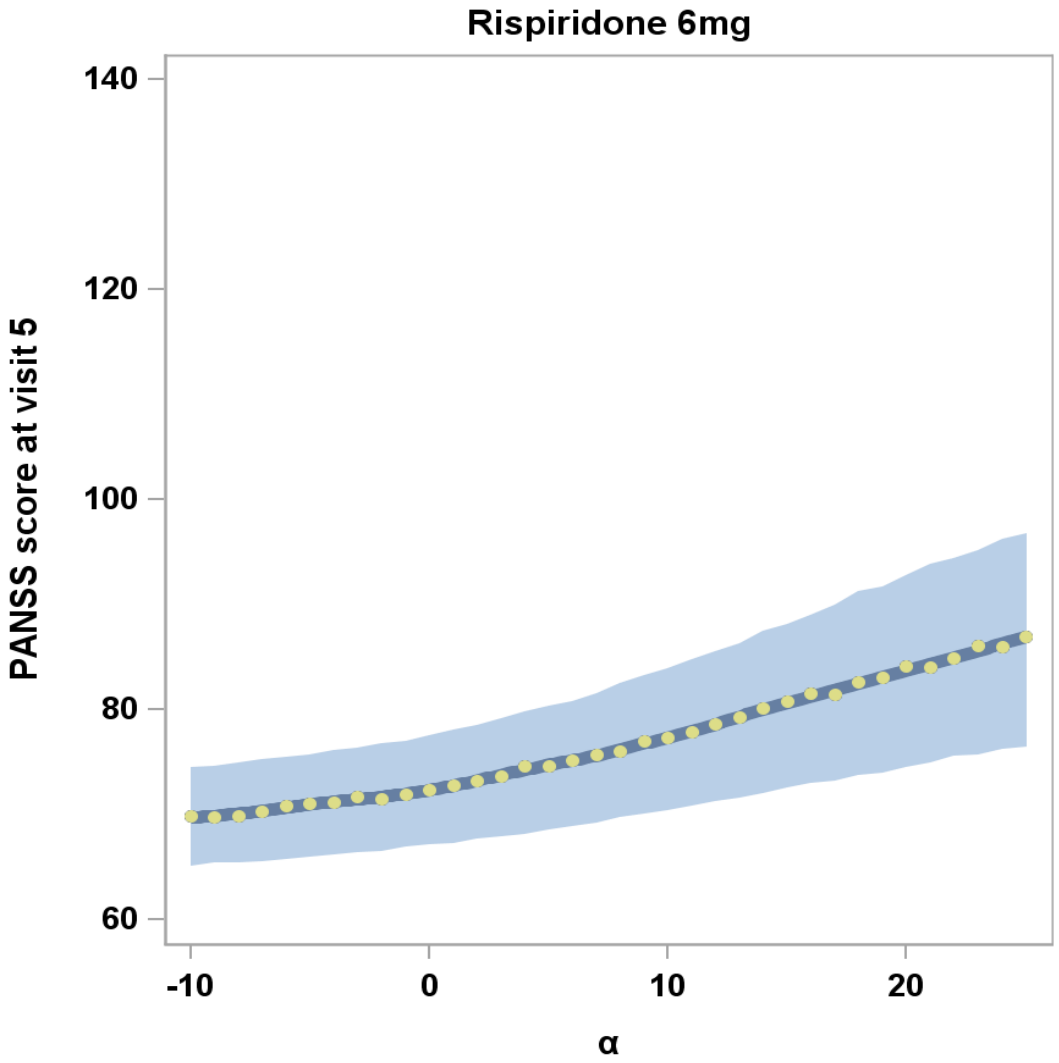

Mean of simulated PANSS score
100
90
80
70
60
60
70
80
90
100
Mean of observed PANSS score
Placebo
Risperidone 6mg


Variance of simulated PANSS score
550
450
350
250
250
350
450
550
Variance of observed PANSS score
Placebo
Risperidone 6mg

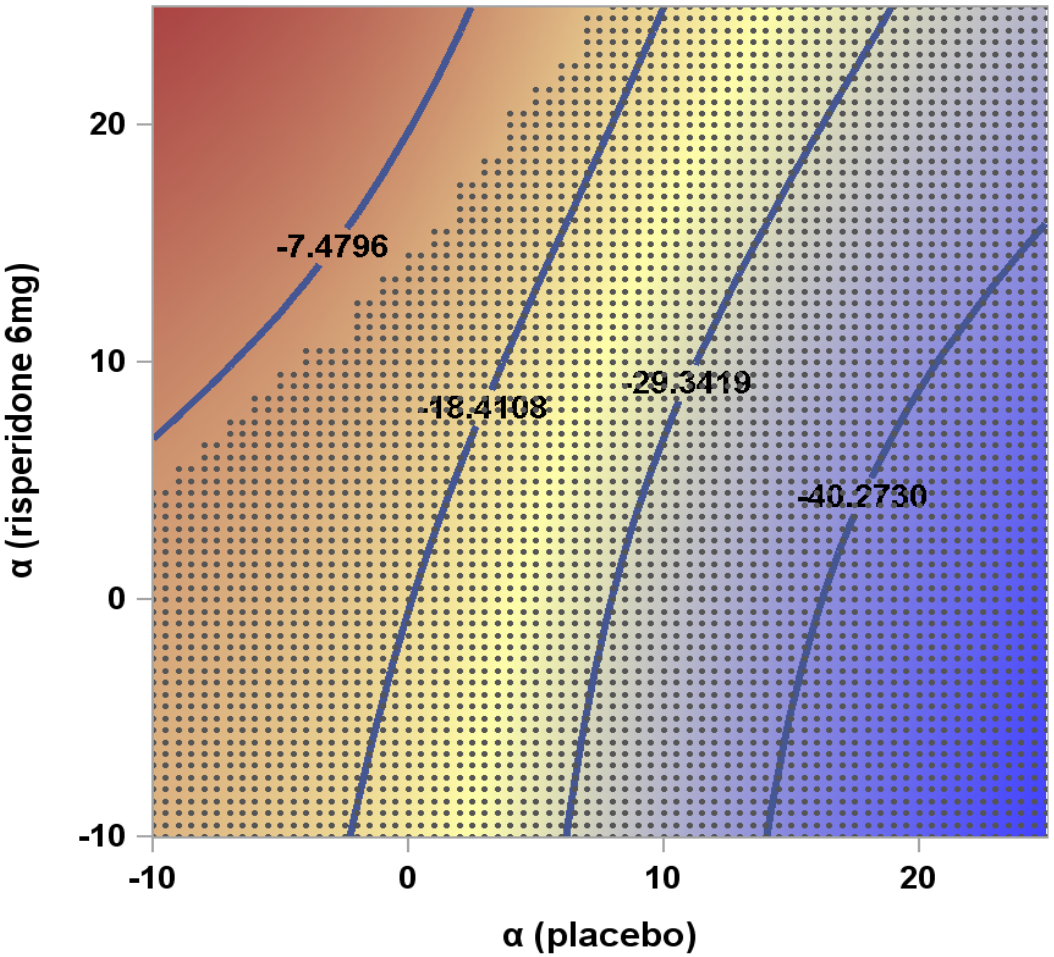
Example



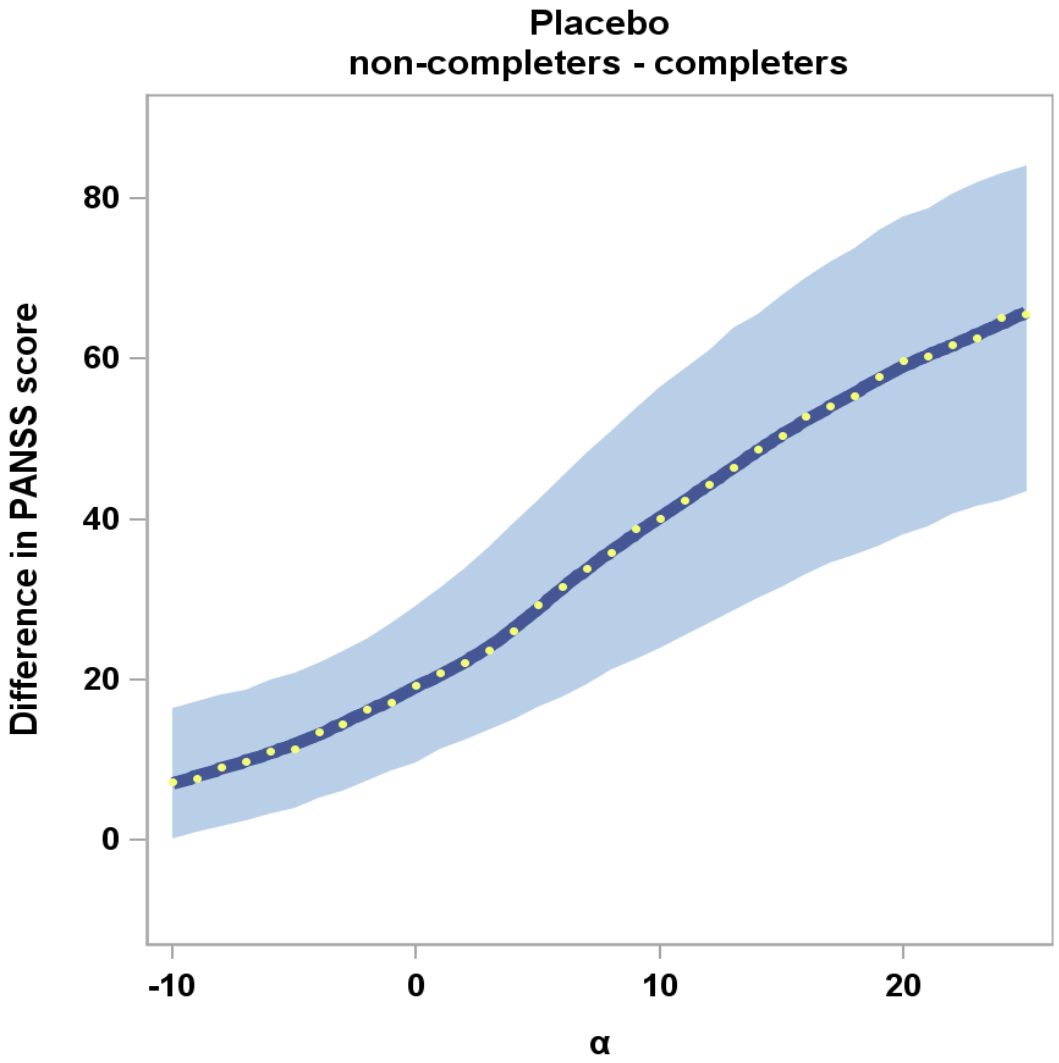
Example


Rispiridone 6mg
PANSS score at visit 5
140
120
100
80
60
-10
0
10
20
α

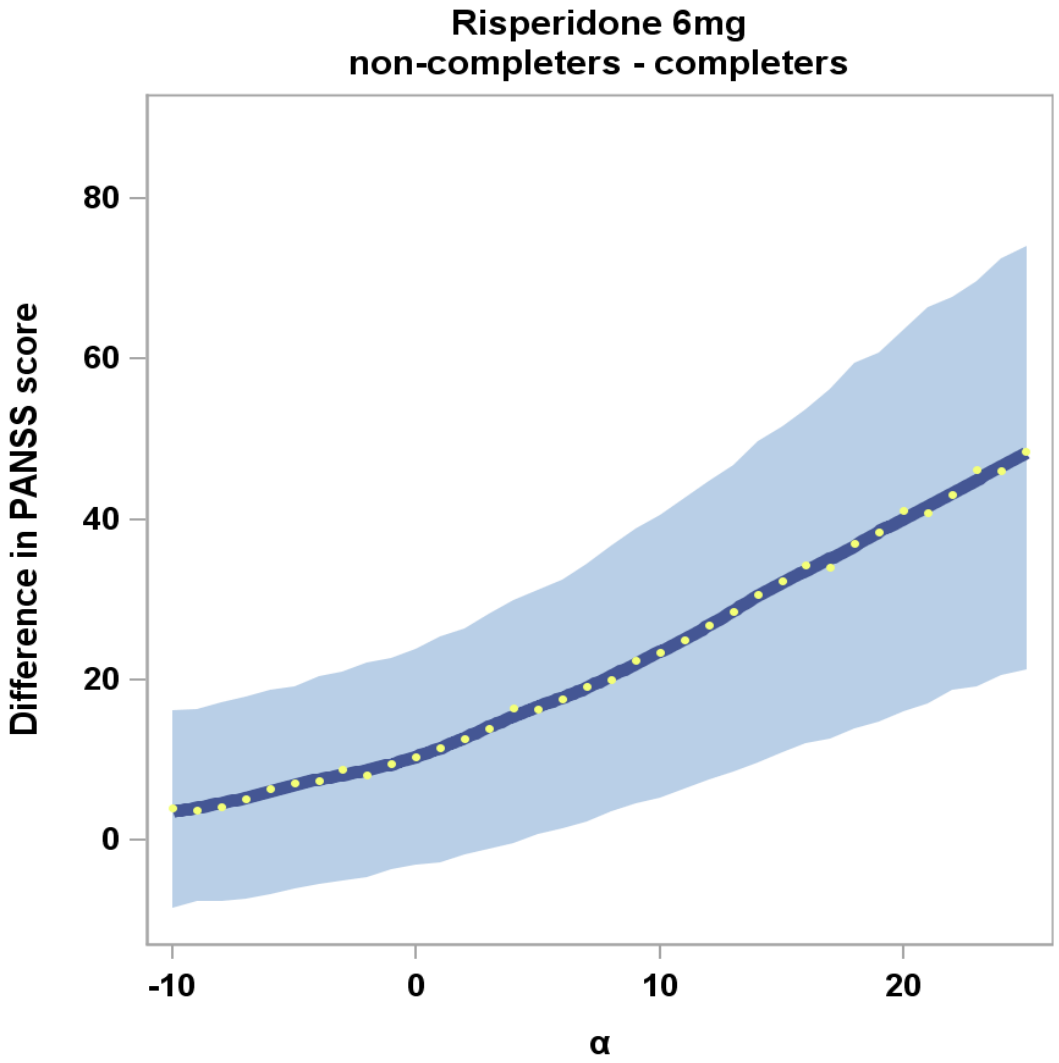
Example


-7.4796
-18.4108
-29.3419
-40.2730
20
10
0
-10
-10
0
10
20
α (risperidone 6mg)
α (placebo)

Results


Placebo
non-completers - completers
Difference in PANSS score
80
60
40
20
0
-10
0
10
20
α

Example


Risperidone 6mg
non-completers - completers
Difference in PANSS score
80
60
40
20
0
-10
0
10
20
α

Siddiqui, Hung and O'Neil

- Compared MMRM (Mixed-Effect Model Repeated Measure) to LOCF using simulation and data from 25 NDAs.
- Concluded: "MMRM analysis appears to be a superior approach in controlling Type I error rates and minimizing biases, as compared to LOCF ANCOVA analysis. In the exploratory analyses of the datasets, no clear evidence of the presence of MNAR missingness is found."
- This is NOT evidence that one should rely on MMRM. How well does MMRM fit the observed data? How does one conduct global sensitivity analysis?

Later

I will show how to

- draw inference under MAR (actually a weaker version)
- evaluate the sensitivity of inferences to deviations from MAR.
- incorporate auxiliary variables into the analysis

Discussion

- How can sensitivity analysis be integrated into the regulatory decision process?
- How can companies be encouraged to minimize missing data? Will requiring the reporting of global sensitivity analyses be useful in this regard?
- What is your perspective on intention-to-treat?

A Sensitivity Analysis Paradigm for Randomized Studies with Missing Data

Daniel Scharfstein
Johns Hopkins University
dscharf@jhsph.edu

October 1, 2013

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.

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

Sensitivity Analysis

- 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.

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.

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.

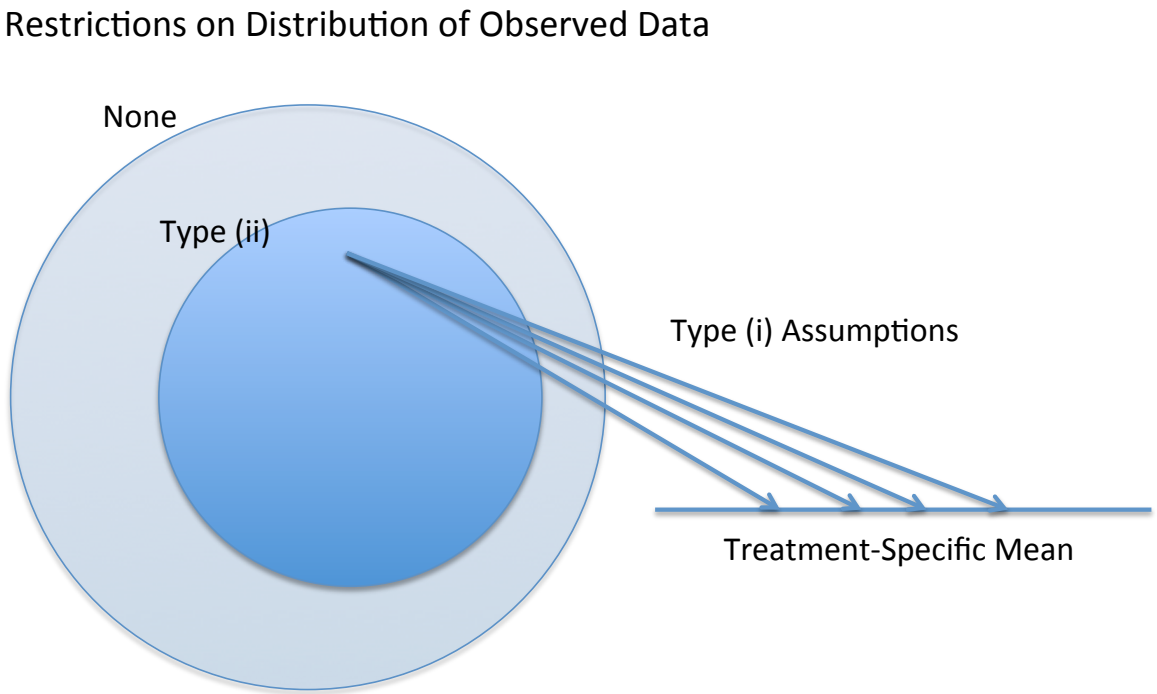
Sensitivity Analysis

- 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.

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.

Sensitivity Analysis



Example: $K = 2$

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

Missing at random (MAR)

- In this setting, MAR postulates

$$\begin{aligned} f(Y_0^+|C=0,Y_0) &= f(Y_0^+|C\geq 1,Y_0) \\ f(Y_1^+|C=1,Y_1^-) &= f(Y_1^+|C=2,Y_1^-) \end{aligned}$$

or

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

Missing at random (MAR)

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

Aside: Math Review

Suppose X and Y are random variables.

$$f(y) = \int f(y|x)dF(x)$$

$$E[Y] = E[E[Y|X]] = \int E[Y|X = x]dF(x)$$

In the special where X is an indicator variable,

$$f(y) = f(y|X = 1)P[X = 1] + f(y|X = 0)P[X = 0]$$

$$E[Y] = E[Y|X = 1]P[X = 1] + E[Y|X = 0]P[X = 0]$$

- If Y is independent of X, then $f(y|X) = f(y)$ and $E[Y|X] = E[Y]$

Aside: Math Review

Suppose there is a third variable W

$$f(y|x) = \int f(y|w,x)dF(w|x)$$

$$\begin{aligned} E[Y|X=x] &= E[E[Y|W,X=x]|X=x] \\ &= \int_w E[Y|W=w,X=x]dF(w|x) \end{aligned}$$

- If Y is independent of X given W, then $f(y|X,W) = f(y|W)$ and $E[Y|X,W] = E[Y|W]$

Missing at random (MAR)

- Under MAR, μ is identified

$$
\begin{aligned}
\mu &= \int_{y_0} E[Y_2|Y_0 = y_0]dF(y_0) \\
&= \int_{y_0} \{E[Y_2|C = 0, Y_0 = y_0]P[C = 0|Y_0 = y_0] + \\
&\qquad E[Y_2|C \geq 1, Y_0 = y_0]P[C \geq 1|Y_0 = y_0]\}\, dF(y_0) \\
&= \int_{y_0} \{E[Y_2|C \geq 1, Y_0 = y_0]P[C = 0|Y_0 = y_0] + \\
&\qquad E[Y_2|C \geq 1, Y_0 = y_0]P[C \geq 1|Y_0 = y_0]\}\, dF(y_0) \\
&= \int_{y_0} E[Y_2|C \geq 1, Y_0 = y_0]dF(y_0)
\end{aligned}
$$

Missing at random (MAR)

$$
\begin{aligned}
&= \int_{y_0}\int_{y_1} E[Y_2|C \geq 1, Y_1 = y_1, Y_0 = y_0]dF(y_1|C \geq 1, Y_0 = y_0)dF(y_0) \\
&= \int_{y_0}\int_{y_1} \{E[Y_2|C = 1, Y_1 = y_1, Y_0 = y_0]P[C = 1|C \geq 1, Y_1 = y_1, Y_0 = y_0] + \\
&\qquad E[Y_2|C = 2, Y_1 = y_1, Y_0 = y_0]P[C = 2|C \geq 1, Y_1 = y_1, Y_0 = y_0]\} \\
&\qquad dF(y_1|C \geq 1, Y_0 = y_0)dF(y_0) \\
&= \int_{y_0}\int_{y_1} \{E[Y_2|C = 2, Y_1 = y_1, Y_0 = y_0]P[C = 1|C \geq 1, Y_1 = y_1, Y_0 = y_0] + \\
&\qquad E[Y_2|C = 2, Y_1 = y_1, Y_0 = y_0]P[C = 2|C \geq 1, Y_1 = y_1, Y_0 = y_0]\} \\
&\qquad dF(y_1|C \geq 1, Y_0 = y_0)dF(y_0) \\
&= \int_{y_0}\int_{y_1} E[Y_2|C = 2, Y_1 = y_1, Y_0 = y_0]dF(y_1|C \geq 1, Y_0 = y_0)dF(y_0)
\end{aligned}
$$

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

Missing at random (MAR)

This identification formula holds under the weaker assumption:

$$\begin{aligned} f(Y_2|C=0,Y_0) &= f(Y_2|C\geq 1,Y_0) \\ f(Y_2|C=1,Y_1^-) &= f(Y_2|C=2,Y_1^-) \end{aligned}$$

or

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

More generally,

$$f(Y_K|C=k,Y_k^-) = f(Y_K|C\geq k+1,Y_k^-)$$

$$P[C=k|C\geq k,Y_K,Y_k^-] = P[C=k|C\geq k,Y_k^-]$$

Missing at random (MAR)

- Specify models for $f(Y_k|C \geq k, Y_{k-1}^-)$ (parameters η).
- Estimate η using maximum likelihood.
- Estimate μ by repeating the following simulation procedure:
  1. Simulate Y_0 from its empirical distribution. Set $k = 1$
  2. Simulate Y_k from $f(Y_k|C \geq k, Y_{k-1}^-; \widehat{\eta})$, Set $k = k + 1$.
  3. If $k > K$ then stop; otherwise repeat step 2.
- Take an average of the simulated Y_K's
- G-computation algorithm.
- Standard errors using non-parametric bootstrap.

Missing at random (MAR)

- Under MAR,

$$\mu = E\left[\frac{I(C = K)Y_K}{\prod_{k=0}^{K-1} P[C > k | C \geq k, Y_k^-]}\right]$$

- So, rather than modeling $f(Y_k | C \geq k, Y_{k-1}^-)$, one can model $P[C = k | C \geq k, Y_k^-]$
- Suppose we assume

$$\text{logit}\{P[C = k | C \geq k, Y_k^-]\} = h_k(Y_k^-; \gamma)$$

- Estimate γ by maximum likelihood
- Estimate μ by the inverse-weighted estimator

$$\tilde{\mu} = E_n\left[\frac{I(C = K)Y_K}{\prod_{k=0}^{K-1} P[C > k | C \geq k, Y_k^-; \widehat{\gamma}]}\right]$$

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:

$$f(Y_2|C=0,Y_1^-) = f(Y_2|C\geq 1,Y_1^-) \tag{1}$$

and

$$f(Y_1|C=0,Y_0) = \frac{f(Y_1|C\geq 1,Y_0)\exp\{\alpha r(Y_1)\}}{E[\exp\{\alpha r(Y_1)\}|C\geq 1,Y_0]}$$
$$f(Y_2|C=1,Y_1^-) = \frac{f(Y_2|C=2,Y_1^-)\exp\{\alpha r(Y_2)\}}{E[\exp\{\alpha r(Y_2)\}|C=2,Y_1^-]} \tag{2}$$

- $r(y)$ is a specified function of y
- α is a sensitivity analysis parameter that governs departures from the MAR assumption
- $\alpha = 0$ is MAR, $\alpha \neq 0$ is MNAR.

Exponential Tilting

- If $[Y_1|C \geq 1, Y_0] \sim N(\mu_1(Y_0), \sigma_1^2)$ and $r(Y_1) = Y_1$, then

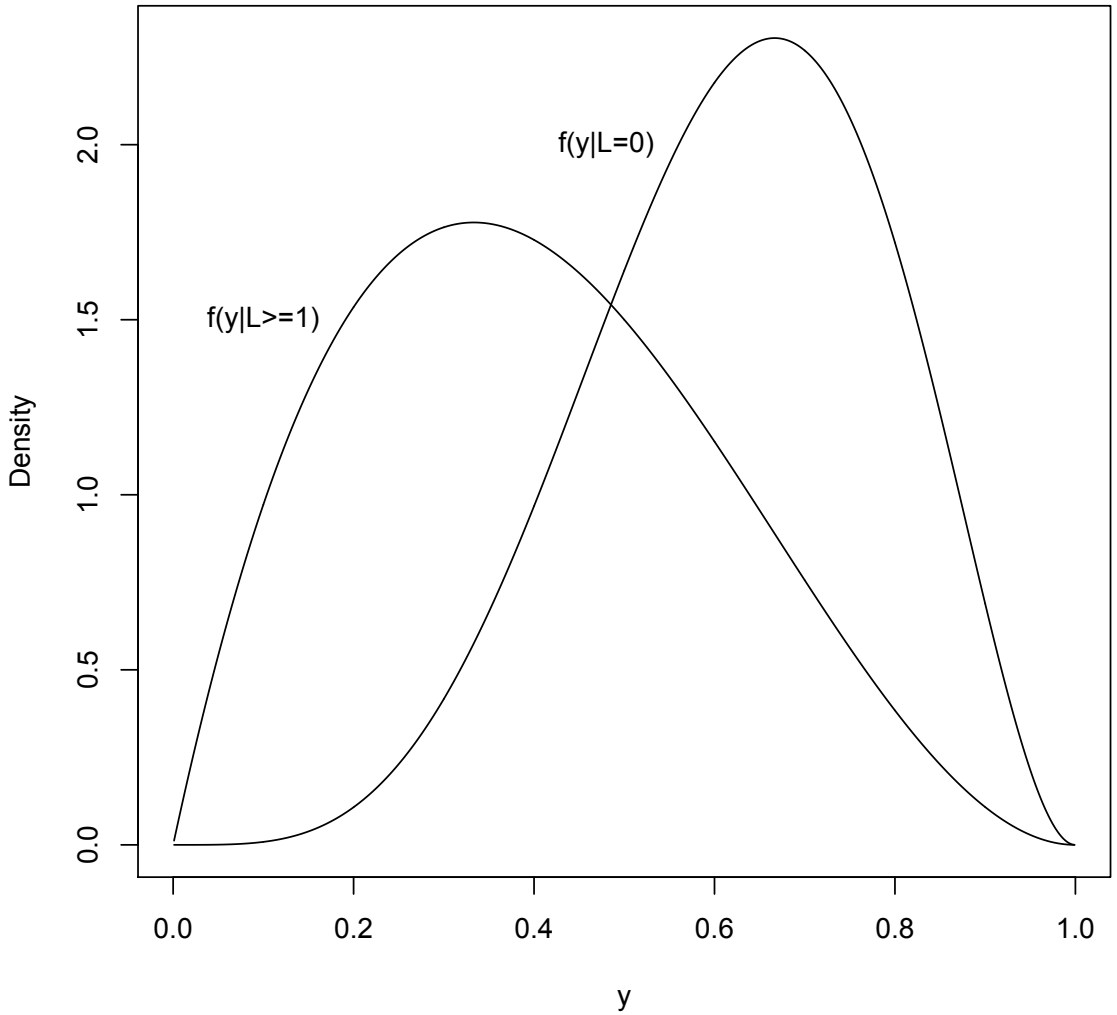
$$[Y_1|C = 0, Y_0] \sim N(\mu_1(Y_0) + \alpha\sigma_1^2, \sigma_1^2)$$

- If $[Y_1|C \geq 1, Y_0] \sim Beta(a_1(Y_0), b_1(Y_0))$ and $r(Y_1) = \log(Y_1)$, then

$$[Y_1|C = 0, Y_0] \sim Beta(a_1(Y_0) + \alpha, b_1(Y_0))$$

$\alpha > -a_1(Y_0)$

Beta



Exponential Tilting

- If $[Y_1|C \geq 1, Y_0] \sim Gamma(a_1(Y_0), b_1(Y_0))$ and $r(Y_1) = \log(Y_1)$, then

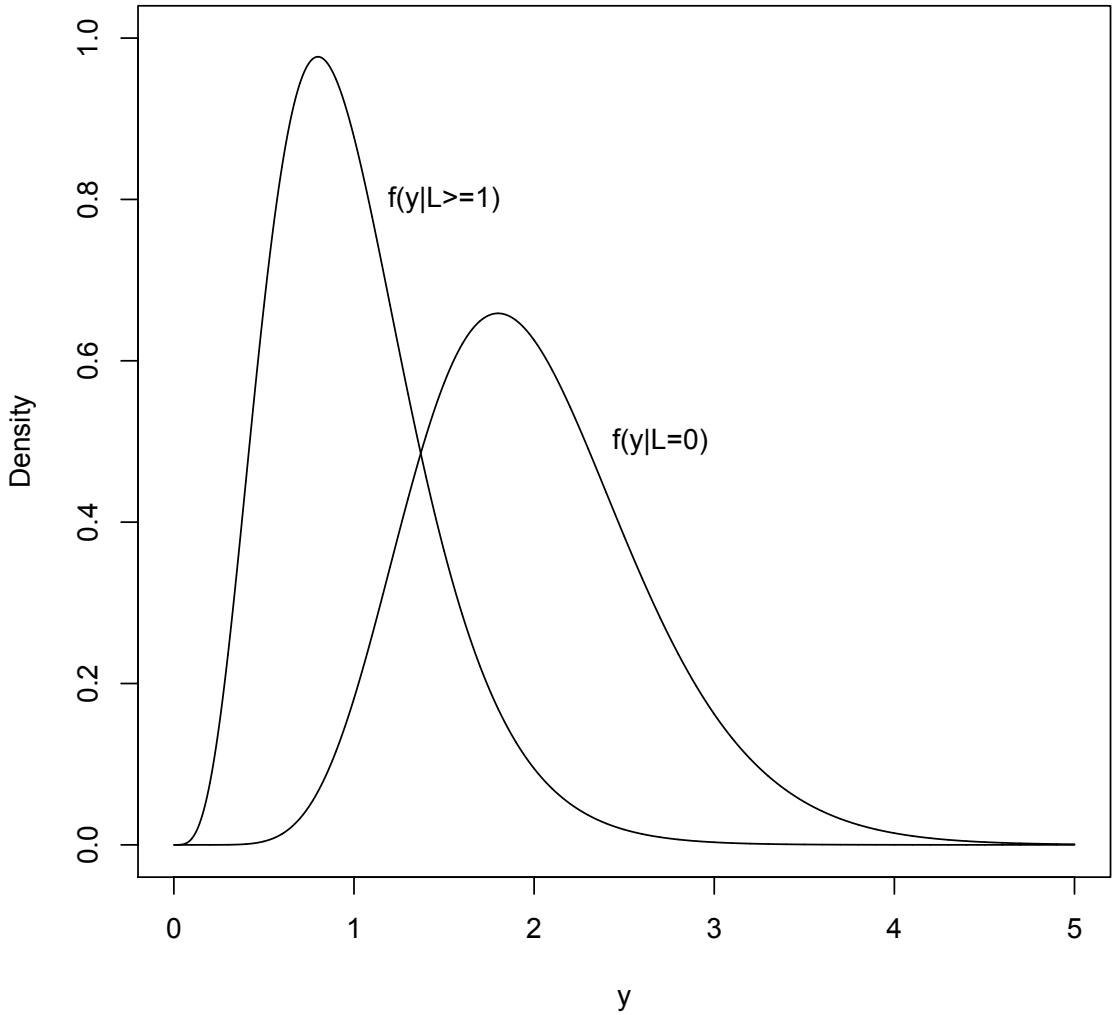
$$[Y_1|C = 0, Y_0] \sim Gamma(a_1(Y_0) + \alpha, b_1(Y_0)),$$

  $\alpha > -a_1 Y_0)$.
- If $[Y_1|C \geq 1, Y_0]] \sim Gamma(a_1(Y_0), b_1(Y_0))$ and $r(Y_1) = Y_1$, then

$$[Y_1|C = 0, Y_0] \sim Gamma(a_1(Y_0), b_1(Y_0) - \alpha),$$

  $\alpha < b_1(Y_0)$.

Gamma



Exponential Tilting

- If $[Y_1|C \geq 1, Y_0] \sim \textit{Bernoulli}(p_1(Y_0))$ and $r(Y_1) = Y_1$, then

$$[Y_1|C = 0, Y_0] \sim \textit{Bernoulli}\left(\frac{p_1(Y_0)\exp(\alpha)}{p_1(Y_0)\exp(\alpha) + 1 - p_1(Y_0)}\right)$$

Identification

- $f(Y_2|C=2, Y_1^-)$ is identified
- By (2), $f(Y_2|C=1, Y_1^-)$ is identified

$$
\begin{aligned}
&f(Y_2|C=0, Y_0) \\
&= \int_{y_1} f(Y_2|C=0, Y_1=y_1, Y_0) dF(y_1|C=0, Y_0) \\
&\overset{(1,2)}{=} \int_{y_1} f(Y_2|C\geq 1, Y_1=y_1, Y_0) \frac{dF(y_1|C\geq 1, Y_0)\exp\{\alpha r(y_1)\}}{E[exp\{\alpha r(y_1)\}|C\geq 1, Y_0]}
\end{aligned}
$$

$$
\begin{aligned}
&f(Y_2|C\geq 1, Y_1=y_1, Y_0) \\
&= f(Y_2|C=1, Y_1=y_1, Y_0)P[C=1|C\geq 1, Y_1=y_1, Y_0]+ \\
&\quad f(Y_2|C=2, Y_1=y_1, Y_0)P[C=2|C\geq 1, Y_1=y_1, Y_0]
\end{aligned}
$$

- $f(Y_2|C=0, Y_0)$ is identified

Identification

$$
\begin{aligned}
&\mu \\
&= \int_{y_0} E[Y_2|Y_0 = y_0]dF(y_0) \\
&= \int_{y_0} \{E[Y_2|C = 0, Y_0 = y_0]P[C = 0|Y_0 = y_0]+ \\
&\qquad E[Y_2|C \geq 1, Y_0 = y_0]P[C \geq 1|Y_0 = y_0]\}\, dF(y_0) \\
&= \int_{y_0} \{E[Y_2|C = 0, Y_0 = y_0]P[C = 0|Y_0 = y_0]+ \\
&\qquad \left\{\int_{y_1} E[Y_2|C \geq 1, Y_1 = y_1, Y_0 = y_0]dF(y_1|C \geq 1, Y_0 = y_0)\right\} P[C \geq 1|Y_0 = y_0]\Big\}\, dF(y_0)
\end{aligned}
$$

Estimation

- Specify models for $f(Y_k|C \geq k, Y_{k-1}^-)$ (params η).
- Specify models for $P[C = k|C \geq k, Y_k^-]$ (params γ).
- Estimate η and γ using maximum likelihood.
- Estimate μ by repeating the following simulation procedure:
  1. Simulate Y_0 from its empirical distribution.
  2. Draw from $P[C = 0|C \geq 0, Y_0; \widehat{\gamma}]$.
  3. If $C = 0$, then draw from $f(Y_2|C = 0, Y_0; \widehat{\gamma}, \widehat{\eta}; \alpha)$ and stop.
  4. If $C \neq 0$, the draw Y_1 from $f(Y_1|C \geq 1, Y_0; \widehat{\eta})$.
  5. Draw from $P[C = 1|C \geq 1, Y_1^-; \widehat{\gamma}]$
  6. If $C = 1$, then draw from $f(Y_2|C = 1, Y_1^-; \widehat{\eta}; \alpha)$ and stop
  7. If $C = 2$ then draw from $f(Y_2|C = 2, Y_1^-; \widehat{\eta})$.

Estimation

- Take an average of the simulated Y_2's
- Generalization of G-computation algorithm.
- Standard errors using non-parametric bootstrap.

Estimation

To draw from $f(Y_2|C=1, Y_0; \widehat{\eta}; \alpha)$ in Step 6, draw from

$$\frac{f(Y_2|C=2, Y_1^-; \widehat{\eta}) \exp\{\alpha r(Y_2)\}}{E[\exp\{\alpha r(Y_2)\}|C=2, Y_1^-; \widehat{\eta}]}$$

Estimation

To draw from $f(Y_2|C=0, Y_0; \widehat{\gamma}, \widehat{\eta}; \alpha)$ in Step 3, draw from

1. Draw Y_1 from

$$\frac{f(Y_1|C \geq 1, Y_0; \widehat{\eta}) \exp\{\alpha r(Y_1)\}}{E[\exp\{\alpha r(Y_1)\}|C \geq 1, Y_0; \widehat{\eta}]}$$

2. Draw from $P[C=1|C \geq 1, Y_1^-; \widehat{\gamma}]$
3. If $C=1$, then draw from $f(Y_2|C=1, Y_1^-; \widehat{\eta}; \alpha)$ (see previous slide) and stop
4. If $C=2$ then draw from $f(Y_2|C=2, Y_1^-; \widehat{\eta})$.

Recursive algorithm.

Missing not at random (MNAR)

Assumption (1) and (2) are equivalent to

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

where

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

- α is the conditional log odds ratio of last being seen at visit k between patients who differ by one unit in $r(Y_{k+1})$.
- Assuming that $r(y)$ is monotonically increasing, $\alpha > 0$ implies that patients with higher values of Y_{k+1} are more likely to withdraw than those who remain on study.

Missing not at random (MNAR)

$$\mu = E\left[\frac{I(C=K)Y_K}{\prod_{k=0}^{K-1}\{1-\text{expit}(h_k(Y_k^-)+\alpha r(Y_{k+1}))\}}\right]$$

- (Indirectly) Estimate $h_k(Y_k^-)$ by

$$\begin{aligned} &h(Y_k^-;\widehat{\gamma},\widehat{\eta};\alpha) \\ &= \text{logit}\{P[C=k|C\geq k,Y_k^-;\widehat{\gamma}]\}- \\ &\quad \log\{E[\exp\{\alpha r(Y_{k+1})\}|C\geq k,Y_k^-;\widehat{\eta}]\} \end{aligned}$$

- Estimate μ by the inverse-weighted estimator

$$\tilde{\mu} = E_n\left[\frac{I(C=K)Y_K}{\prod_{k=0}^{K-1}\{1-\text{expit}(h_k(Y_k^-;\widehat{\gamma},\widehat{\eta};\alpha)+\alpha r(Y_{k+1}))\}}\right]$$

Notes on G-Computation vs. Inverse-Weighted Estimator

- Under correct model specification, G-computation estimator is more efficient.
- Inverse-weighted estimator does not extrapolate.
- Can also directly model $h_k(Y_k^-)$. Model checking harder.

Incorporating Auxiliary Variables

- Let V_k denote auxiliary variables scheduled to be collected at assessment k
- Let $W_k = (Y_k, V_k)$
- MAR

$$f(Y_K|C = k, W_k^-) = f(Y_K|C \geq k+1, W_k^-)$$

$$P[C = k|C \geq k, Y_K, W_k^-] = P[C = k|C \geq k, W_k^-]$$

- Non-Future Dependence

$$f(Y_K|C = k, W_k^-, Y_{k+1}) = f(Y_K|C \geq k+1, W_k^-, Y_{k+1})$$

$$P[C = k|C \geq k, Y_K, Y_{k+1}, W_k^-] = P[C = k|C \geq k, Y_{k+1}, W_k^-]$$

Incorporating Auxiliary Variables

- Sensitivity Analysis Models

$$f(Y_{k+1}|C = k, W_k^-) = \frac{f(Y_{k+1}|C \geq k+1, W_k^-)\exp(\alpha r(Y_{k+1}))}{E[\exp(\alpha r(Y_{k+1})|C \geq k+1, W_k^-]}$$

$$\text{logit}\{P[C = k|C \geq k, Y_K, W_k^-, Y_{k+1}]\} = h(W_k^-) + \alpha r(Y_{k+1})$$

where

$$\begin{aligned} h(W_k^-) &= \text{logit}\{P[C = k|C \geq k, W_k^-]\} - \\ &\quad \log\{E[\exp\{\alpha r(Y_{k+1})\}|C \geq k, W_k^-]\} \end{aligned}$$

Incorporating Auxiliary Variables

- Need a model for $f(V_k|C \geq k, Y_k, W_k^-)$
- Can extend G-computation and Inverse-weighted estimation procedure

Main Idea

- Link non-identifiable to identifiable distributions using sensitivity analysis parameters
- Model the distribution of the observed data

Discussion

- Methods made seem complicated, but so are those underlying other statistical procedures such as multiple imputation and MMRM.

A Sensitivity Analysis Paradigm for Randomized Studies with Missing Data

Daniel Scharfstein
Johns Hopkins University
dscharf@jhsph.edu

October 1, 2013

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.

RIS-INT-3

- 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.

RSIP-INT-3

- Patients scheduled for 5 post-baseline assessements at weeks 1,2,4,6, and 8 of the double-blind phase.
- Primary efficiacy 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 ($n = 87$).

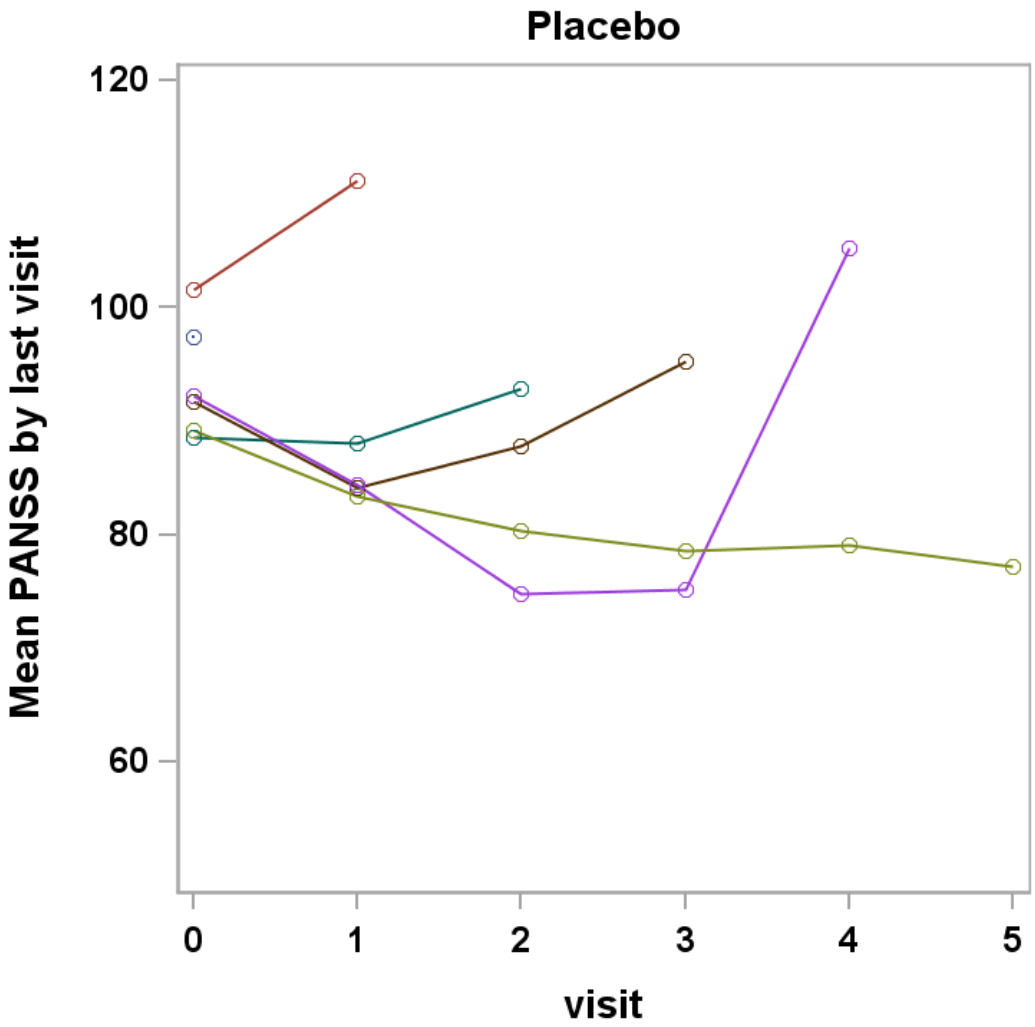
Premature Withdrawal

- 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.

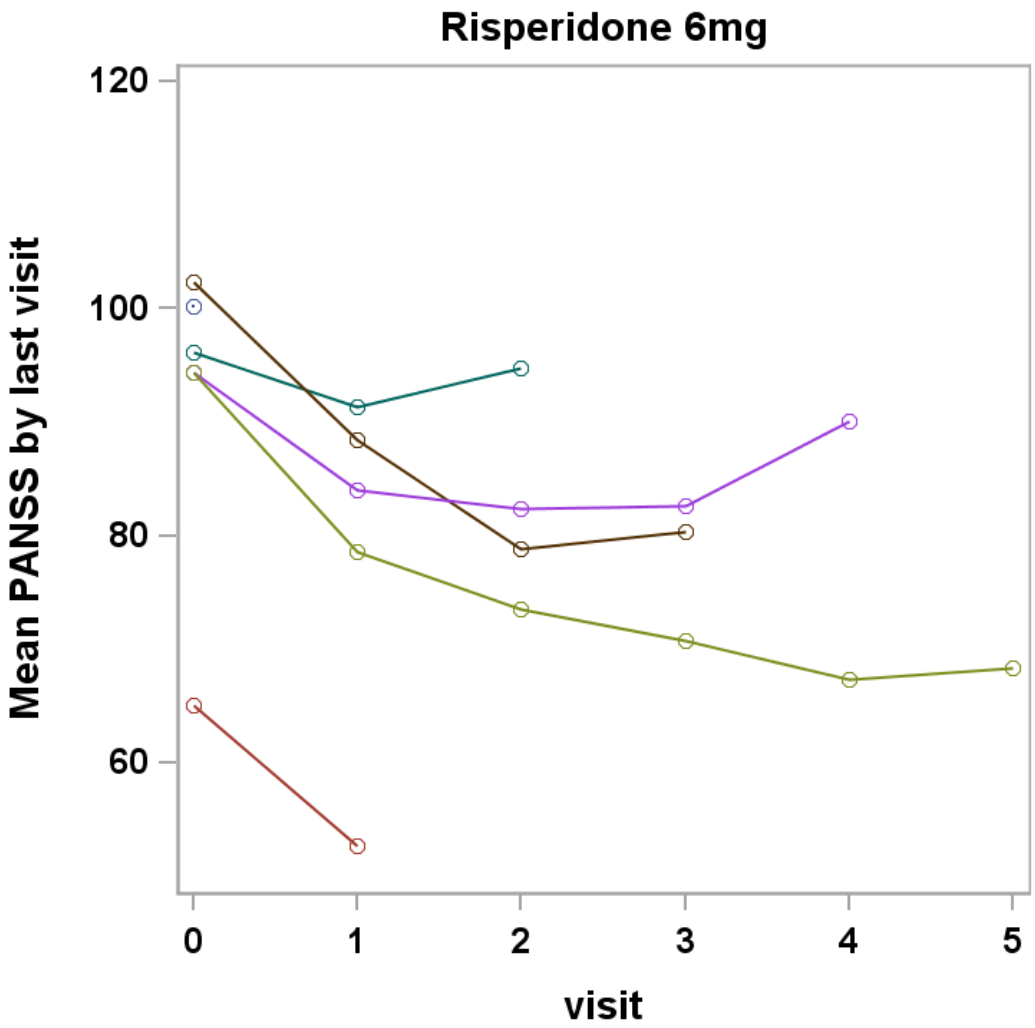
Premature Withdrawal

	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 risperidone at a specified dose level vs. placebo in the counterfactual world in which all patients were followed to that week?

Notation

- $K = 5$ scheduled post-baseline assessments.
- Y_k is PANSS score
- Higher PANSS indicates greater mental illness.

Missing At Random (Weaker Version)

$$f(Y_K|C = k, Y_k^-) = f(Y_K|C \geq k+1, Y_k^-)$$

$$P[C = k|C \geq k, Y_K, Y_k^-] = P[C = k|C \geq k, Y_k^-]$$

Missing at random (MAR)

- Specify a model for $f(Y_k|C \geq k, Y_{k-1}^-)$ (parameters η).
- Estimate η using maximum likelihood.
- Estimate μ by repeating the following simulation procedure:
  1. Simulate Y_0 from its empirical distribution. Set $k = 1$
  2. Simulate Y_k from $f(Y_k|C \geq k, Y_{k-1}^-, \widehat{\eta})$, Set $k = k + 1$.
  3. If $k > K$ then stop; otherwise repeat step 2.
- Take an average of the simulated Y_K's

Missing at random (MAR)

Model $f(Y_k|C \geq k, Y_{k-1}^{-}; \eta)$ as a truncated (between 30 and 120) normal regression model of the following form:

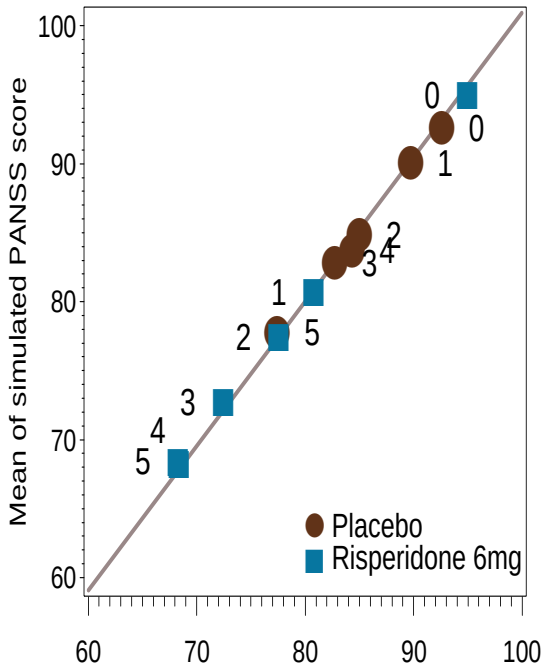
$$f(y_k|C \geq k, \overline{Y}_{k-1}) = \frac{\phi\left(\frac{y_k - \eta_{0,k} - \eta_{1,k} Y_{k-1}}{\eta_{2,k}}\right)}{\Phi\left(\frac{210 - \eta_{0,k} - \eta_{1,k} Y_{k-1}}{\eta_{2,k}}\right) - \Phi\left(\frac{30 - \eta_{0,k} - \eta_{1,k} Y_{k-1}}{\eta_{2,k}}\right)}$$

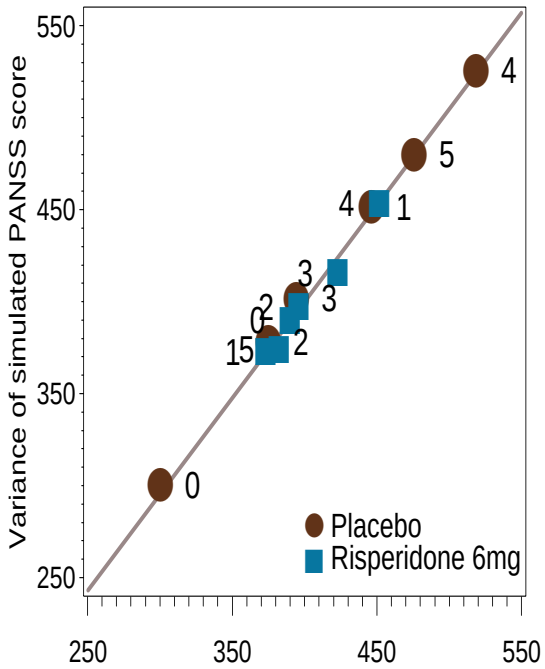
where $30 \leq y_k \leq 210$.

PANSS Model

		Placebo			Risperidone 6mg		
Outcome	**Variable**	**Estimate**	**95 % CI**		**Estimate**	**95 % CI**	
$PANSS_{t=1}$	Intercept ($\eta_{0,1}$)	11.45	-10.68	30.43	21.47	4.38	39.83
	$PANSS_{t=0}$ ($\eta_{1,1}$)	0.85	0.64	1.08	0.63	0.43	0.82
	Std. Dev. ($\eta_{2,1}$)	15.25	12.46	17.38	14.96	12.09	17.02
$PANSS_{t=2}$	Intercept ($\eta_{0,2}$)	16.80	-0.39	32.80	6.32	-4.17	17.92
	$PANSS_{t=1}$ ($\eta_{1,2}$)	0.80	0.62	1.01	0.87	0.73	1.01
	Std. Dev. ($\eta_{2,2}$)	13.24	10.10	15.65	11.64	9.60	13.45
$PANSS_{t=3}$	Intercept ($\eta_{0,3}$)	14.33	-4.46	33.56	7.68	-5.99	20.20
	$PANSS_{t=2}$ ($\eta_{1,3}$)	0.84	0.61	1.07	0.87	0.72	1.05
	Std. Dev. ($\eta_{2,3}$)	13.04	10.00	15.36	13.48	10.42	16.17
$PANSS_{t=4}$	Intercept ($\eta_{0,4}$)	23.57	2.75	53.53	-4.11	-17.99	10.10
	$PANSS_{t=3}$ ($\eta_{1,4}$)	0.77	0.44	1.00	1.00	0.79	1.20
	Std. Dev. ($\eta_{2,4}$)	17.59	8.66	26.28	12.27	9.20	14.78
$PANSS_{t=5}$	Intercept ($\eta_{0,5}$)	-2.73	-12.75	7.32	5.67	-0.88	14.07
	$PANSS_{t=4}$ ($\eta_{1,5}$)	1.01	0.89	1.16	0.93	0.81	1.02
	Std. Dev. ($\eta_{2,5}$)	7.27	3.55	9.68	6.82	4.64	8.71

PANSS Model


Mean of simulated PANSS score
100
90
80
70
60
60
70
80
90
100
Mean of observed PANSS score
Placebo
Risperidone 6mg


Variance of simulated PANSS score
550
450
350
250
250
350
450
550
Variance of observed PANSS score
Placebo
Risperidone 6mg

MAR Analysis

	Observed Mean	MAR Estimate	MAR 95% CI
Placebo	77.19	90.52	[83.82,97.43]
6mg Risperidone	68.36	72.30	[67.13,77.47]
Difference		-18.22	[-26.50,-9.22]

Non-Future Dependence

For $k = 0, \ldots, K-2$,

$$f(Y_K | C = k, Y_k^-, Y_{k+1}) = f(Y_K | C \geq k+1, Y_k^-, Y_{k+1})$$

- This assumption states that, for the cohort patients who are on study at assessment k, share the same history of outcomes through that visit and have the same outcome at assessment $k+1$, the distribution of Y_K is the same for those who are last seen at assessment k and those who are on-study at $k+1$.

Exponential Tilting

For $k = 0, \ldots, K-1$,

$$f(Y_{k+1}|C = k, Y_k^-) = \frac{f(Y_{k+1}|C \geq k+1, Y_k^-)\exp\{\alpha r(Y_{k+1})\}}{E[\exp\{\alpha r(Y_{k+1})\,|\,C \geq k+1, Y_k^-]}$$

where $r(Y_{k+1})$ is a specified function of Y_{k+1} and α is a sensitivity analysis parameter.

- This assumption relates, conditional on past history Y_k^-, the distribution of Y_{k+1} for those who drop-out between assessments k and $k+1$ to those who are on study at $k+1$.

Non-Future Dependence/Exponential Tilting

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

where

$$\begin{aligned} h_k(Y_k^{-}) \;\; = \;\; & \text{logit}\,\{P[C = k \,|\, C \geq k, Y_k^{-}]\} - \\ & \log\{E[\exp\{\alpha r(Y_{k+1})\} \,|\, C \geq k+1, Y_k^{-}]\} \end{aligned}$$

- α is the log odds ratio of drop-out between assessments k and $k+1$ per unit change in $r(Y_{k+1})$.

Non-Future Dependence/Exponential Tilting

- Specify models for $f(Y_k|C \geq k, Y_{k-1}^-)$ (params η).
- Specify models for $P[C = k|C \geq k, Y_k^-]$ (params γ)
- Estimate η and γ using maximum likelihood.
- Estimate μ by repeating the following simulation procedure:
  1. Simulate Y_0 from its empirical distribution. Set $k = 0$.
  2. Simulate R_{k+1} from $P[C > k|C \geq k, \overline{Y}_k; \widehat{\gamma}]$.
  3. If $R_{k+1} = 1$, set $k = k + 1$, simulate $f(Y_k|C \geq k, Y_{k-1}^-; \widehat{\eta})$. If $k = K$ then stop; otherwise and go to step 2.
  4. If $R_{k+1} = 0$, simulate Y_K from $f(Y_K|C = k, Y_k^-; \widehat{\eta}, \widehat{\gamma})$ and stop.

Non-Future Dependence/Exponential Tilting

To draw from $f(Y_K|C = k, Y_k^-)$ call the function $s(k, Y_k^-)$.

$s = \texttt{function}(k, Y_k^-)\{$

- If $k = K-1$, draw Y_K from

$$\frac{f(Y_K|C = K, Y_{K-1}^-; \widehat{\eta}) \exp\{\alpha r(Y_K)\}}{E[\exp\{\alpha r(Y_K)\}|C \geq K, Y_{K-1}^-; \widehat{\eta}]}$$

- If $k < K-1$,
  - Draw

$$\frac{f(Y_{k+1}|C \geq k+1, Y_k^-; \widehat{\eta}) \exp\{\alpha r(y_{k+1})\}}{E[\exp\{\alpha r(Y_{k+1})\}|C \geq k+1, Y_k^-; \widehat{\eta}]}$$

  - Draw $f(Y_K|C \geq k+1, Y_{k+1}, Y_k^-)$, i.e., call the function $g(k+1, Y_{k+1}, Y_k^-)$

$\}$

Non-Future Dependence/Exponential Tilting

To draw from $f(Y_K|C \geq k, Y_k, Y_{k-1}^-)$ call the function $g(k, Y_k, Y_{k-1}^-)$.

$g = \texttt{function}(k, Y_k, Y_{k-1}^-)\{$

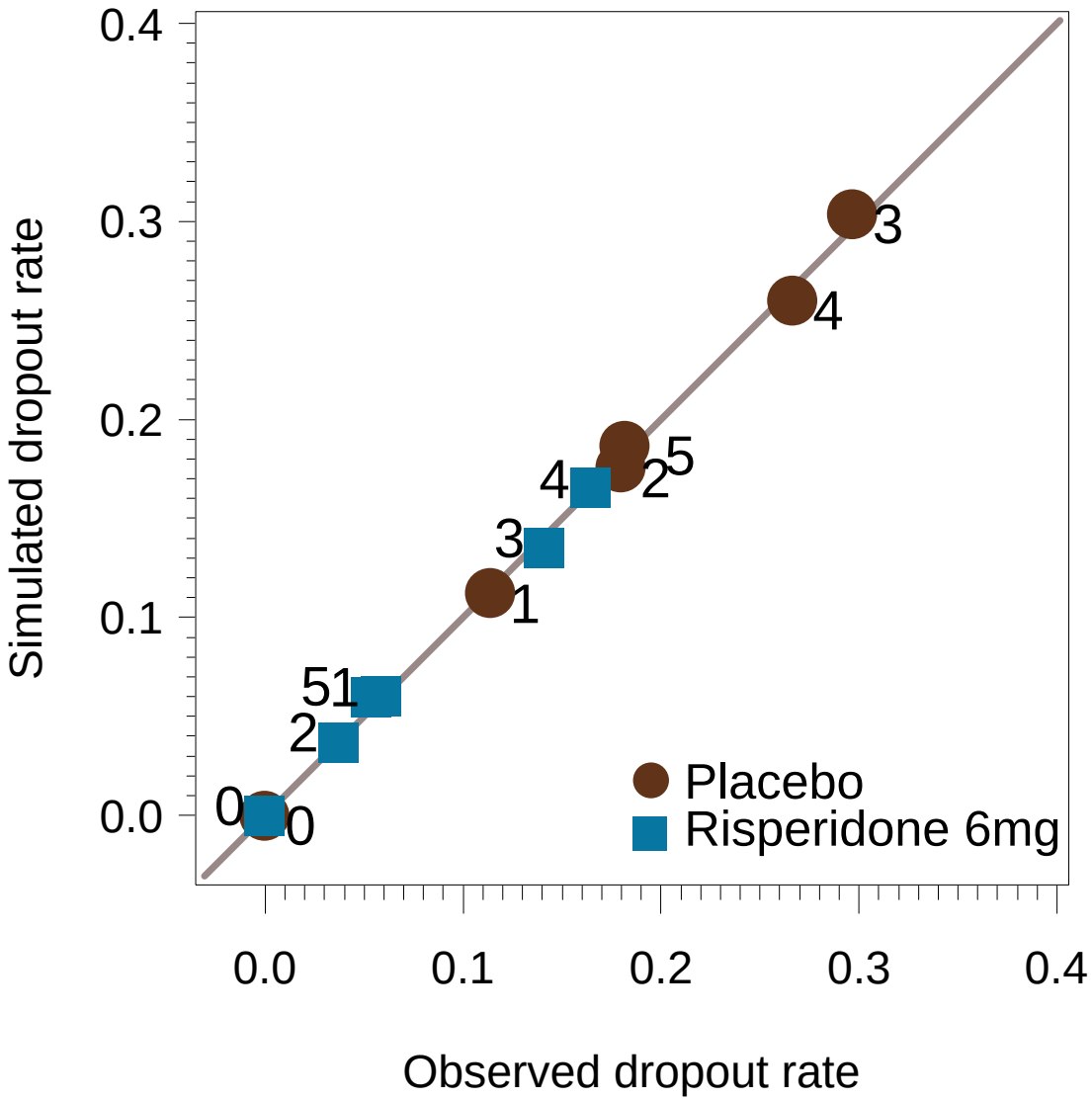
- Draw R_{k+1} from $P[C > k|C \geq k, Y_k^-; \widehat{\gamma}]$
- If $R_{k+1} = 0$, draw Y_K from $f(Y_K|C = k, Y_k^-; \widehat{\eta}, \widehat{\gamma})$, i.e., call $s(k, Y_k^-)$.
- If $R_{k+1} = 1$,
  - Draw Y_{k+1} from $f(Y_{k+1}|C \geq k+1, Y_k^-; \widehat{\eta})$
  - Draw Y_K from $f(Y_K|C \geq k+1, Y_{k+1}^-; \widehat{\eta}, \widehat{\gamma})$, i,e., call $g(k+1, Y_{k+1}^-)$.

$\}$

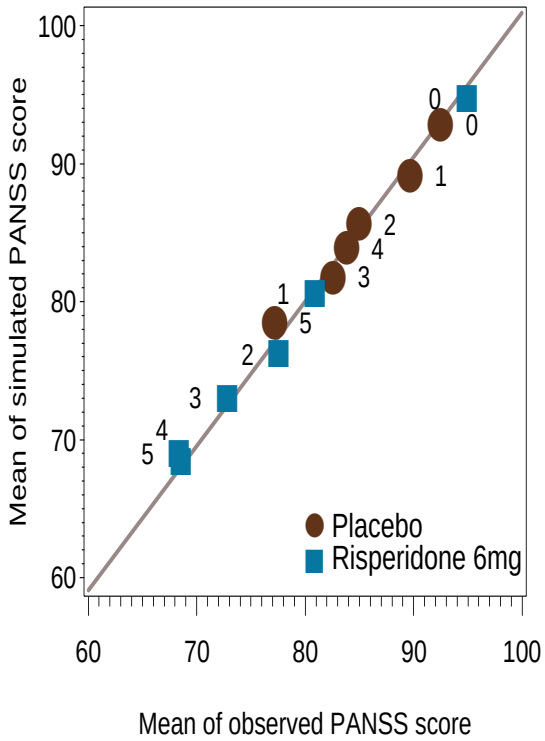
Drop-out Model

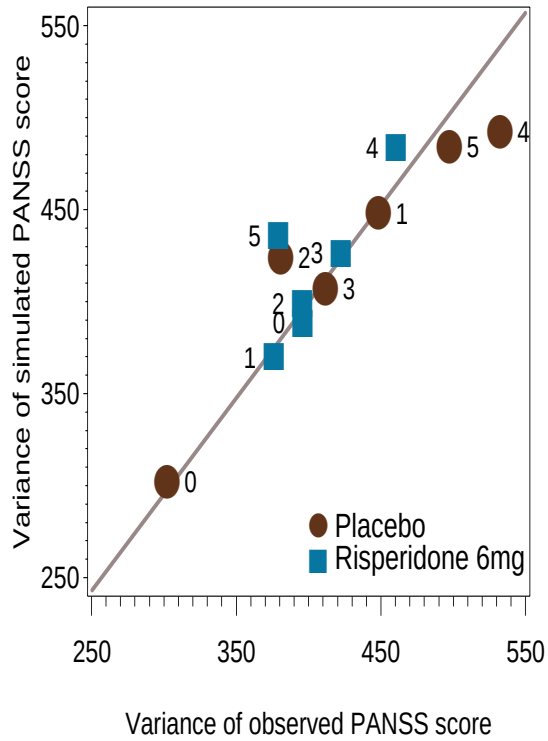
	Placebo			Risperidone 6mg		
Variable	**Estimate**	**95% CI**		**Estimate**	**95% CI**	
Visit 1 ($\gamma_{0,1}$)	-6.34	-8.94	-4.62	-5.14	-8.39	-2.93
Visit 2 ($\gamma_{0,2}$)	-5.72	-7.97	-4.22	-5.28	-17.88	-2.89
Visit 3 ($\gamma_{0,3}$)	-4.73	-6.80	-3.29	-3.73	-5.63	-2.15
Visit 4 ($\gamma_{0,4}$)	-4.82	-6.77	-3.44	-3.44	-5.42	-1.84
Visit 5 ($\gamma_{0,5}$)	-5.48	-7.62	-4.13	-4.62	-17.01	-2.98
PANSS (γ_1)	0.044	0.029	0.066	0.024	0.003	0.045

Dropout Model

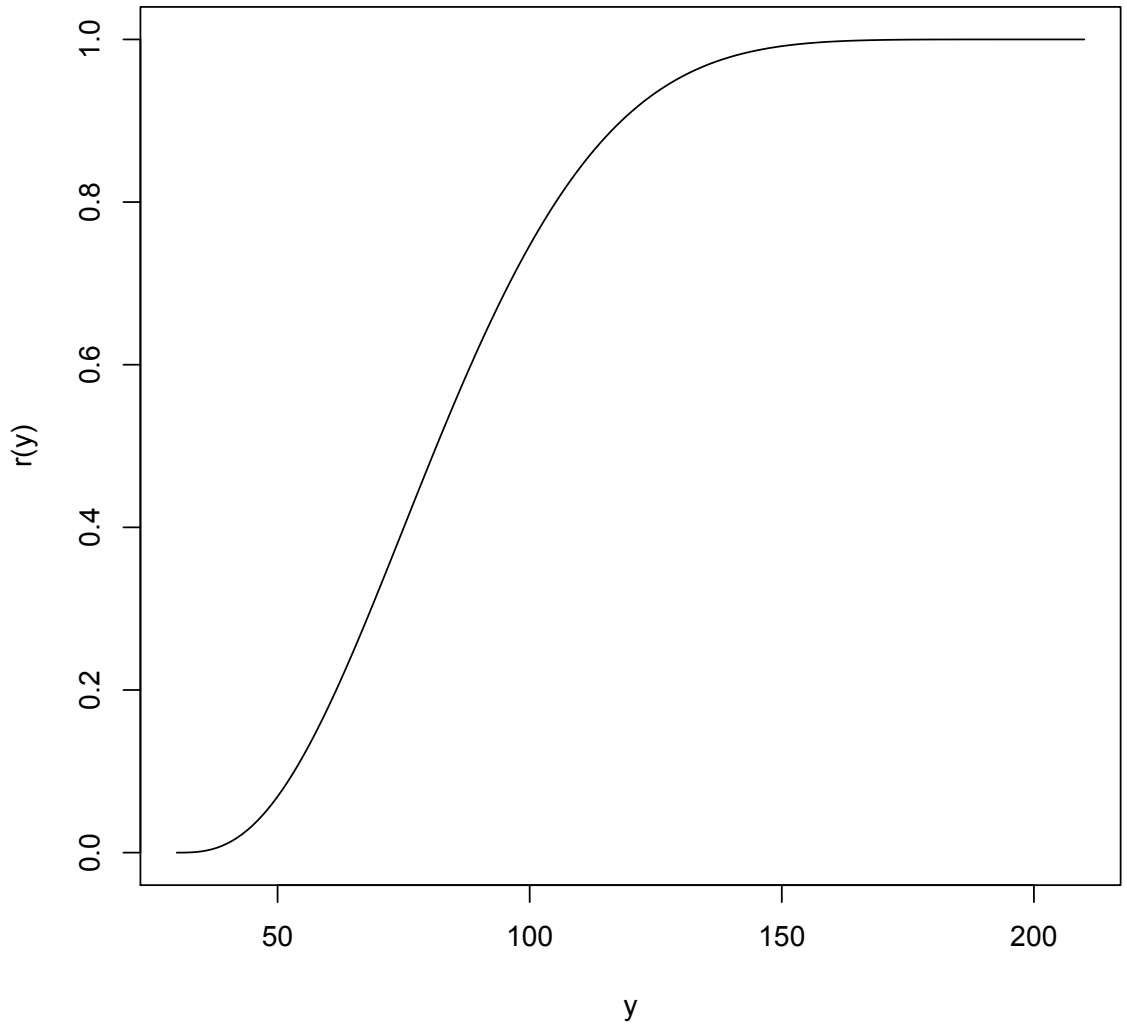


PANSS Model


Mean of simulated PANSS score
100
90
80
70
60
60
70
80
90
100
Mean of observed PANSS score
Placebo
Risperidone 6mg
0
0
1
2
4
3
1
5
2
3
4
5


Variance of simulated PANSS score
550
450
350
250
250
350
450
550
Variance of observed PANSS score
Placebo
Risperidone 6mg
4
5
4
1
5
23
3
2
0
1
0

Bias Function



Sensitivity Analysis

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 $\alpha\{r(y^*_{k+1}) - r(y_{k+1})\}$.

Sensitivity Analysis

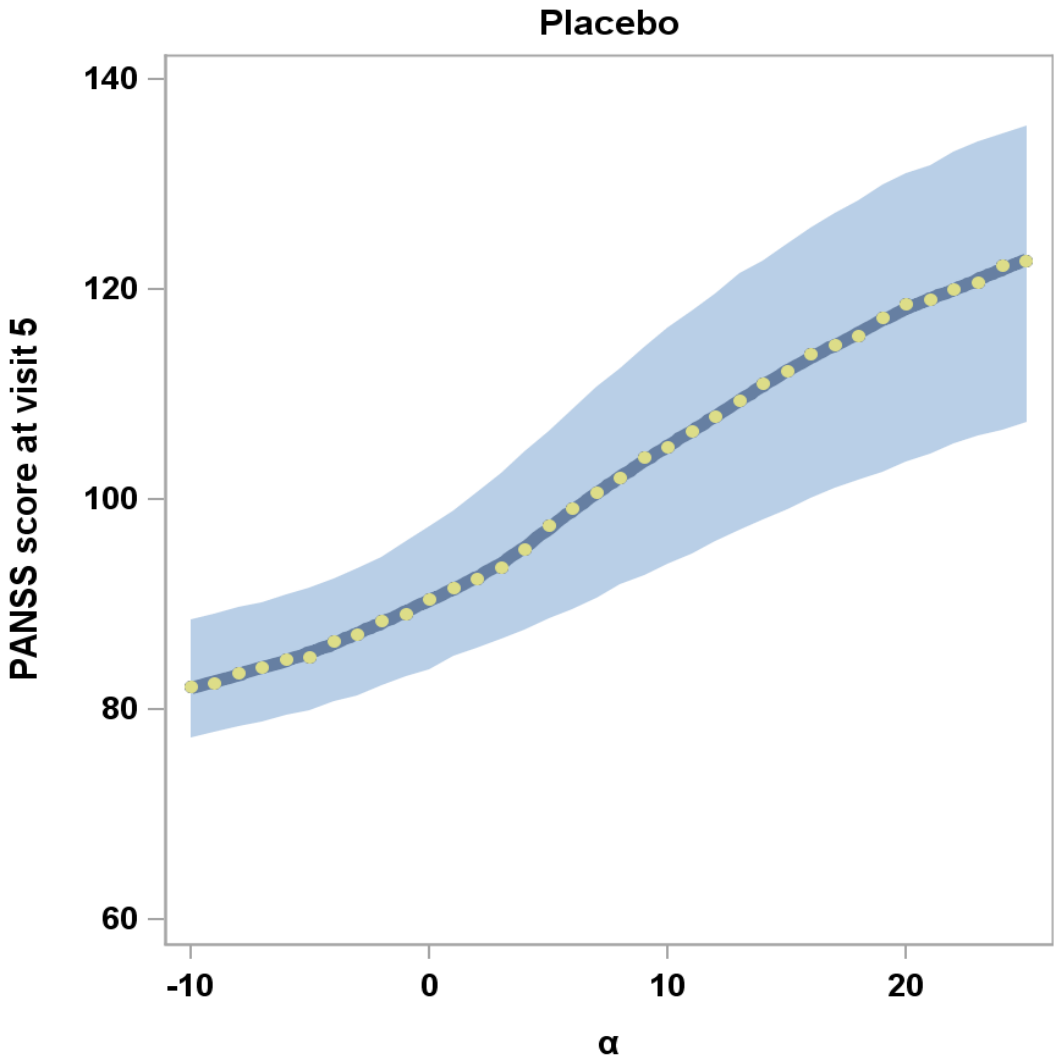
y^*_{k+1}	y_{k+1}	Log Odds Ratio
50	30	α0.02
60	40	α0.07
80	60	α0.22
100	80	α0.30
120	100	α0.24
140	120	α0.12
160	140	α0.04
180	160	α0.01
200	180	α0.00

Sensitivity Analysis

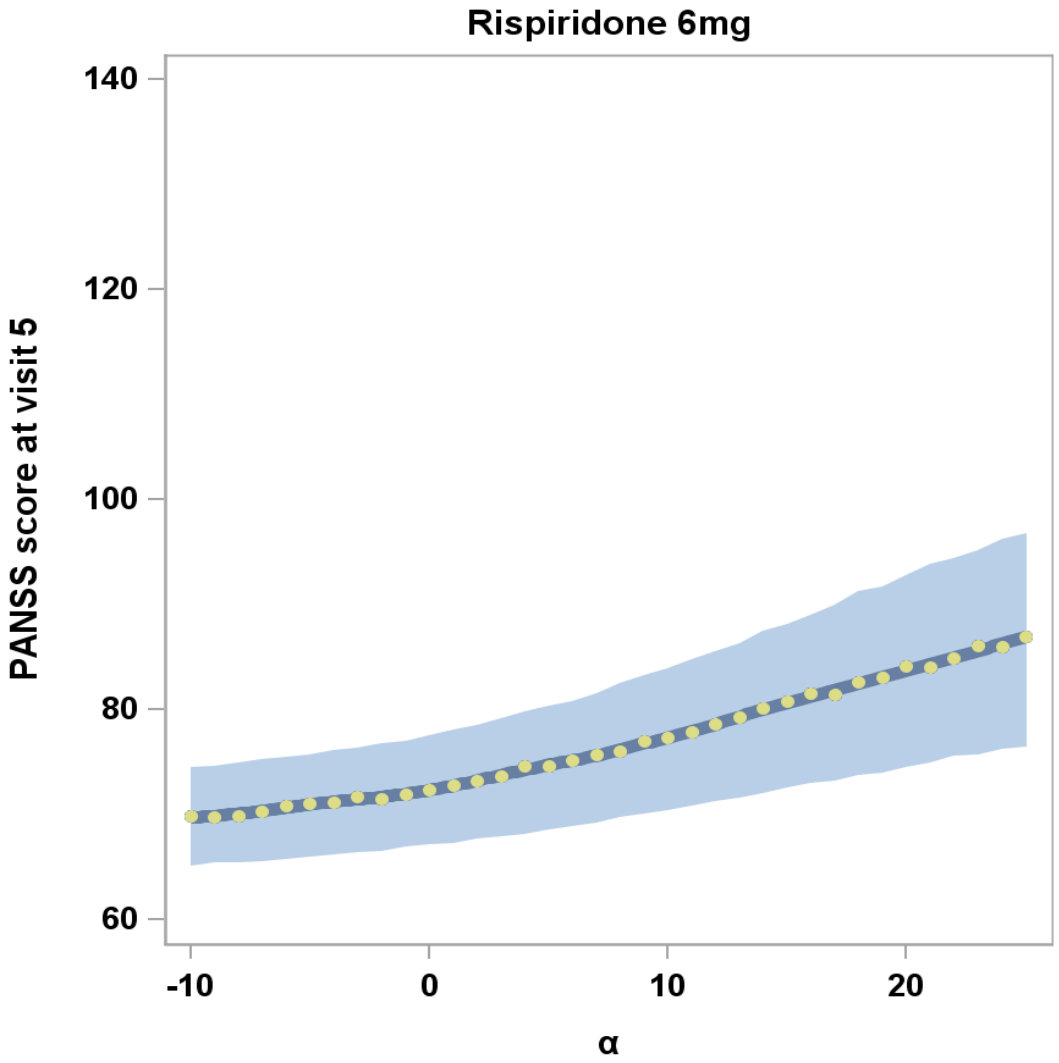
We assumed that $-10.0 \leq \alpha \leq 25.0$

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.

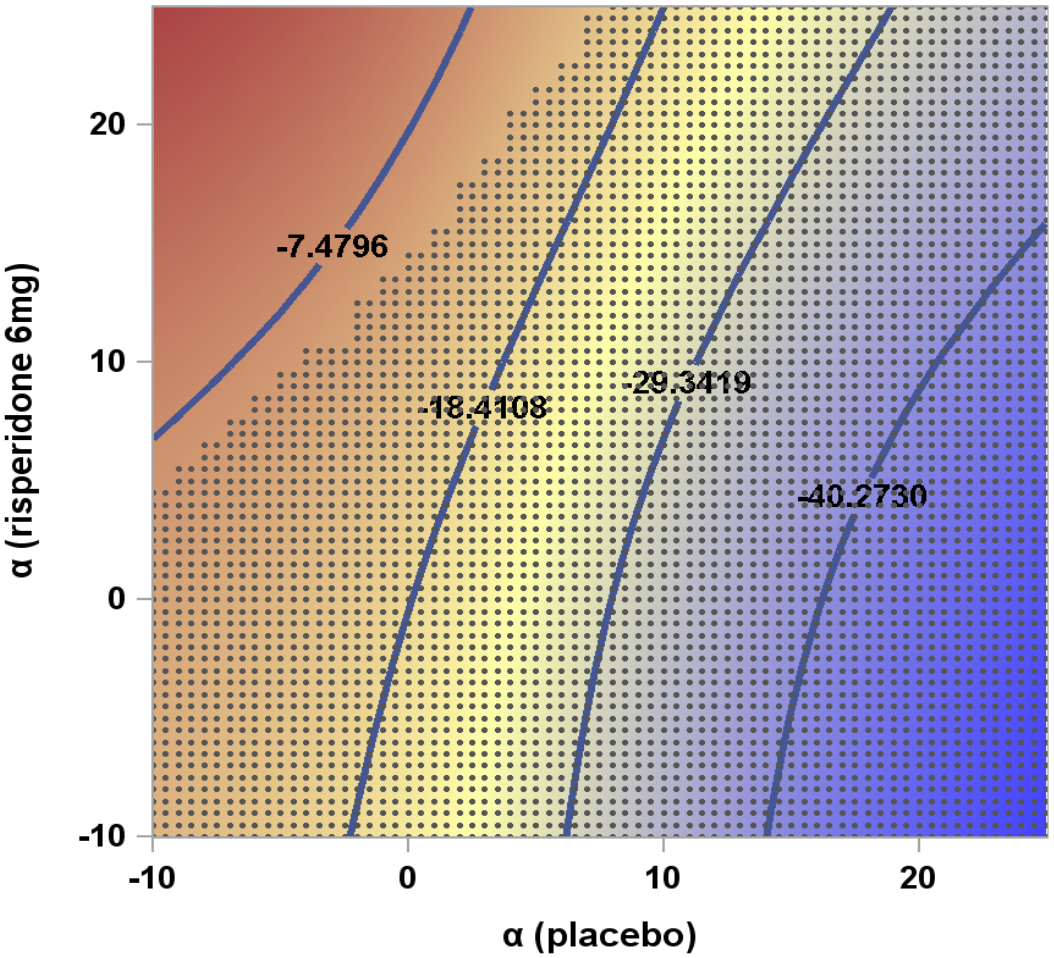
Results



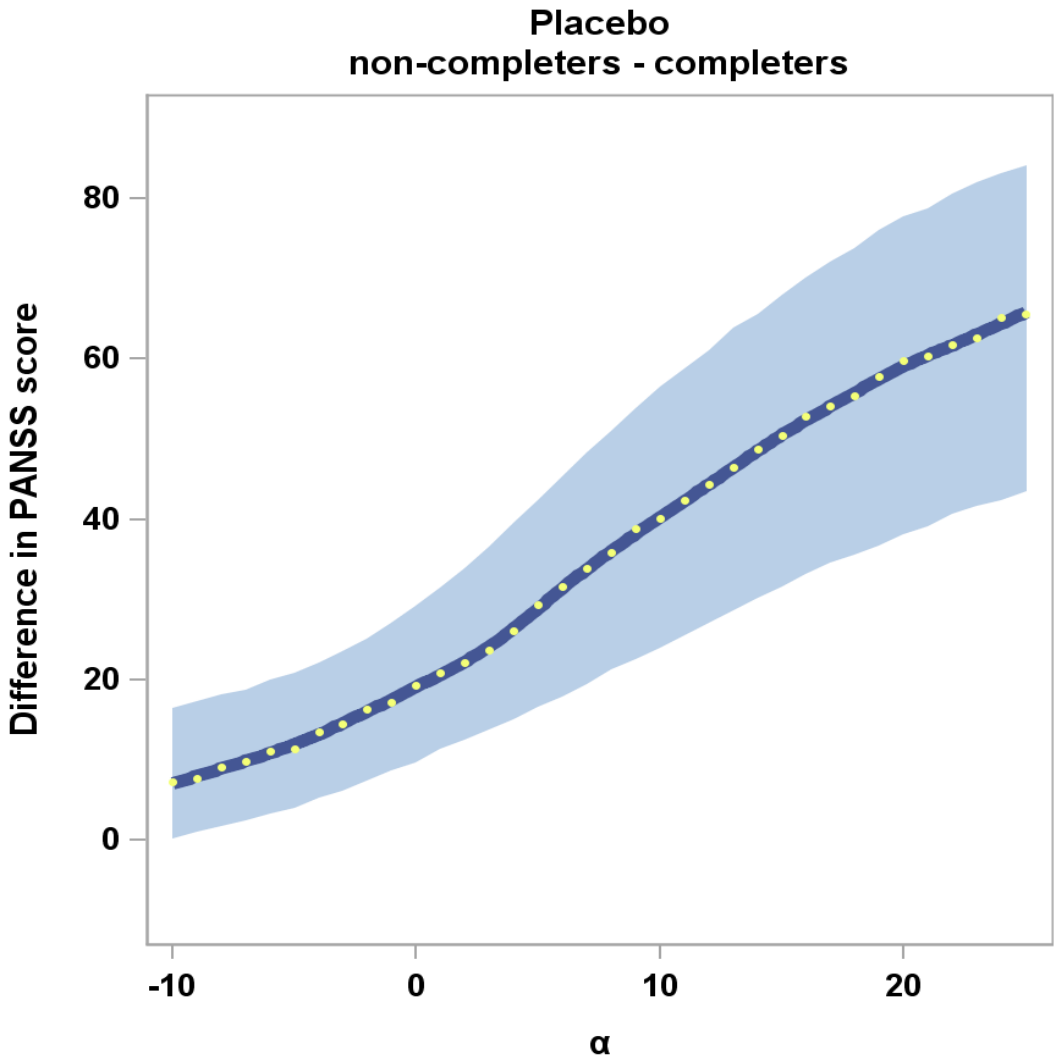
Results


Rispiridone 6mg
PANSS score at visit 5
140
120
100
80
60
-10
0
10
20
α

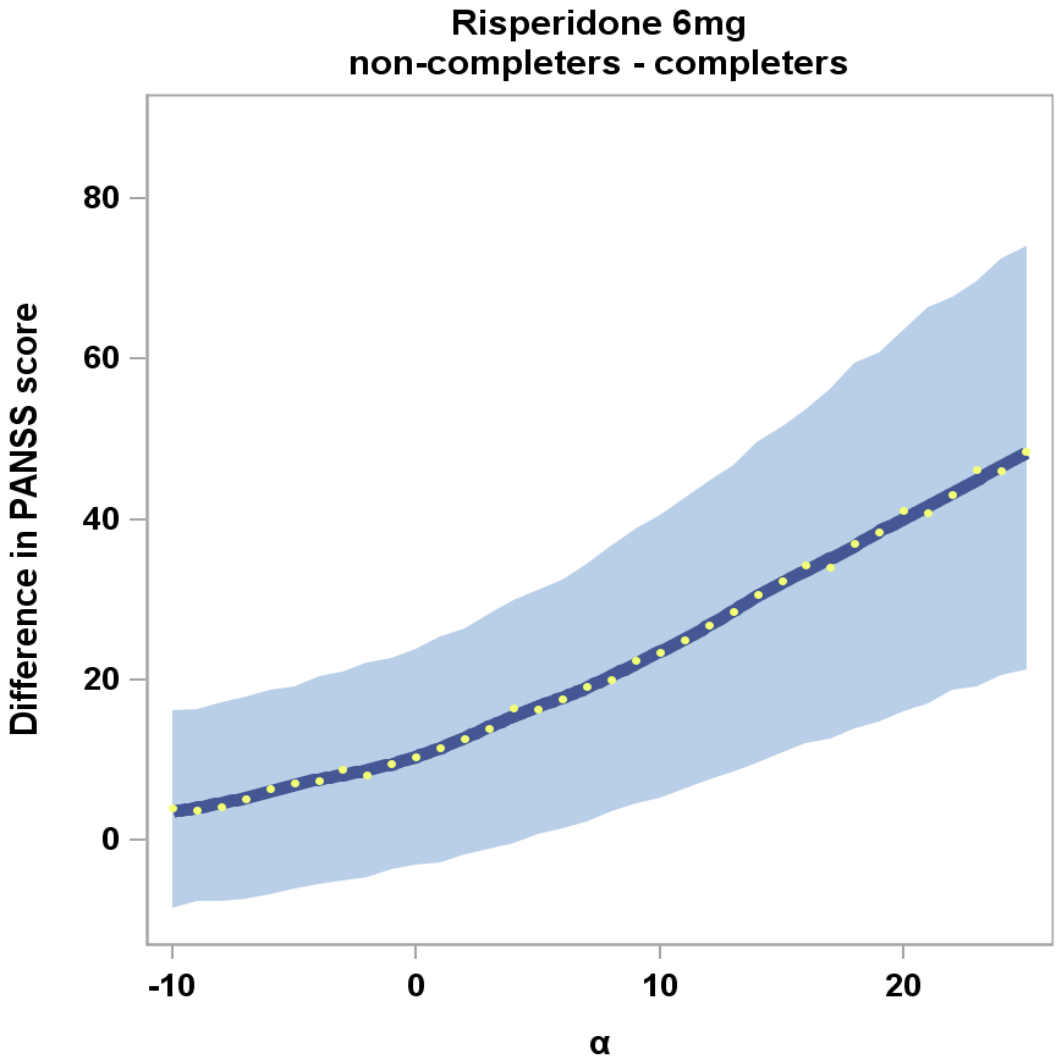
Results


-7.4796
-18.4108
-29.3419
-40.2730
20
10
0
-10
-10
0
10
20
α (risperidone 6mg)
α (placebo)

Results


Placebo
non-completers - completers
Difference in PANSS score
80
60
40
20
0
-10
0
10
20
α

Results


Risperidone 6mg
non-completers - completers
Difference in PANSS score
80
60
40
20
0
-10
0
10
20
α

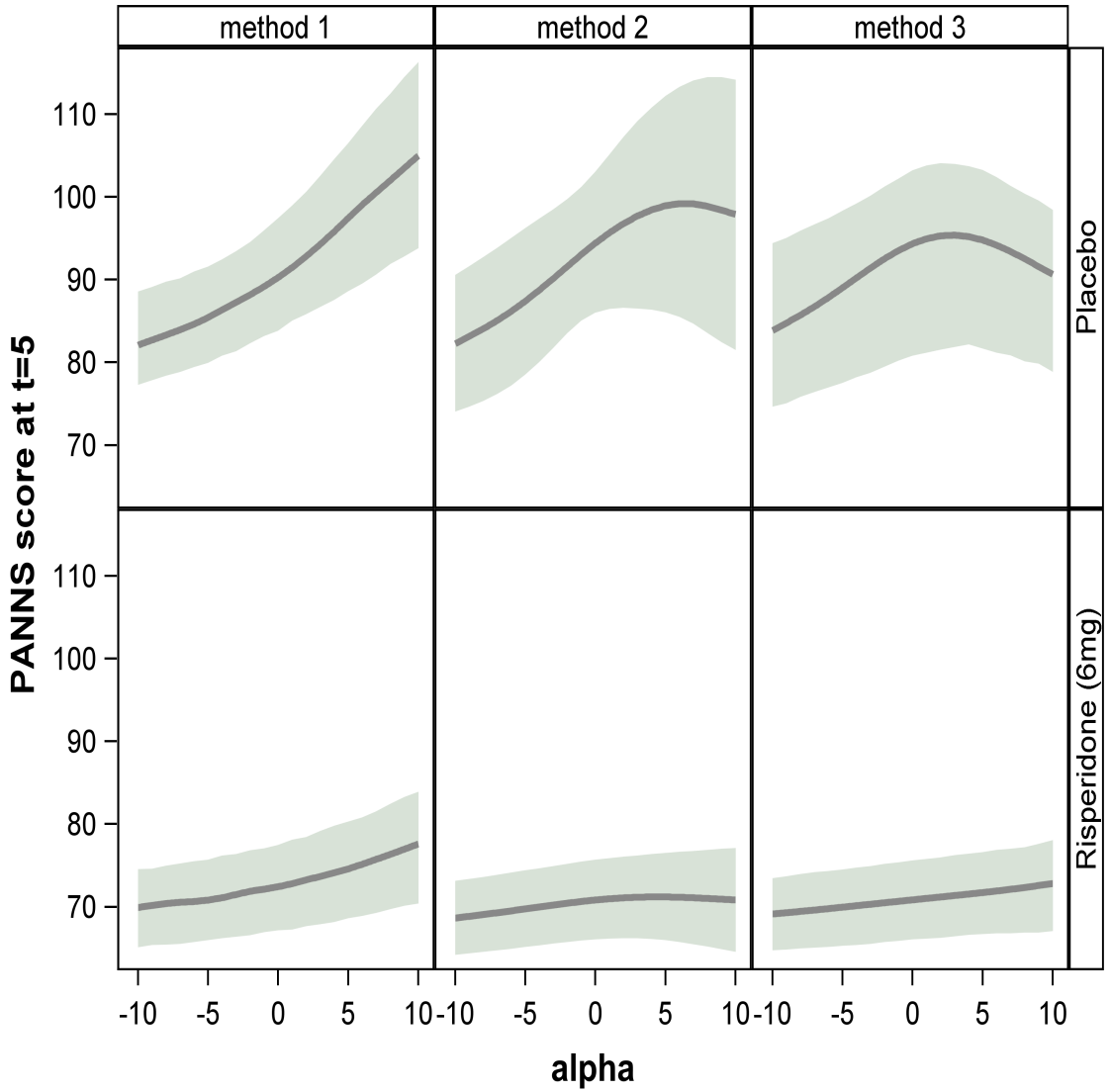
Summary

- Inference is robust to deviations from MAR.
- 6mg risperidone is superior to placebo in reducing pain.

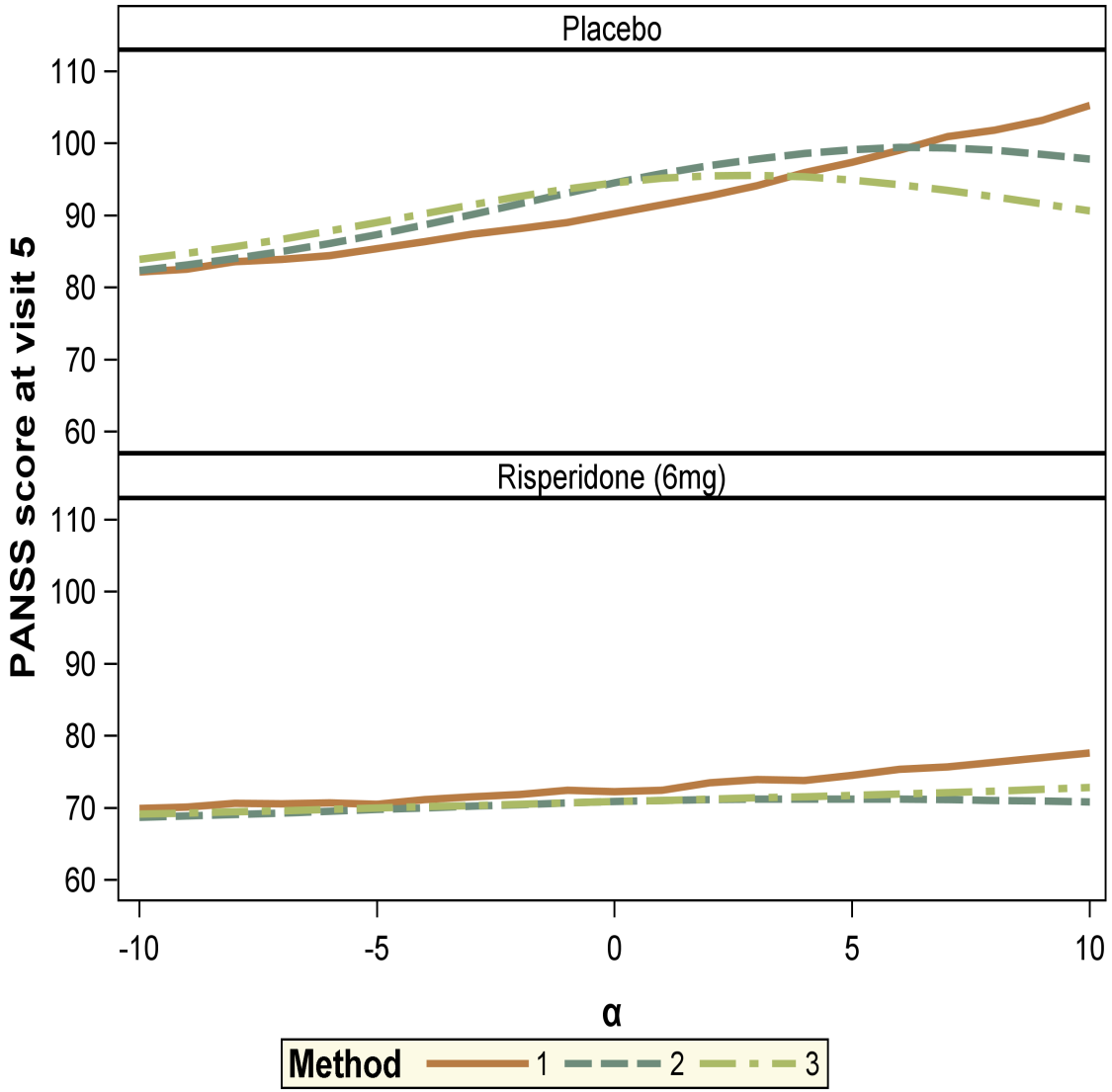
Inverse Weighting

- Indirect modeling (Method 3)
- Direct modeling (Method 2)

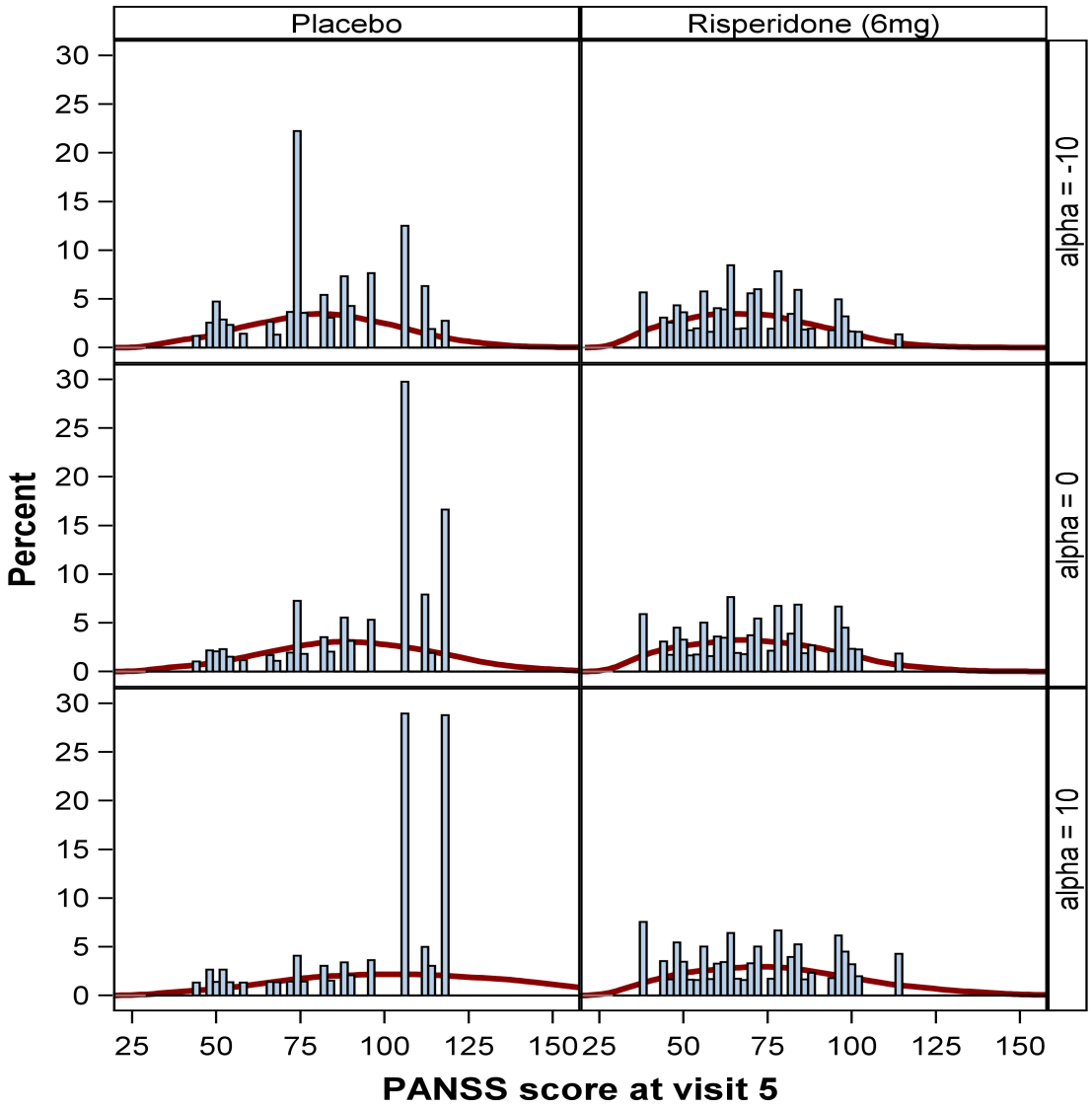
Results


method 1
method 2
method 3
Placebo
Risperidone (6mg)
PANNS score at t=5
110
100
90
80
70
-10
-5
0
5
10
alpha

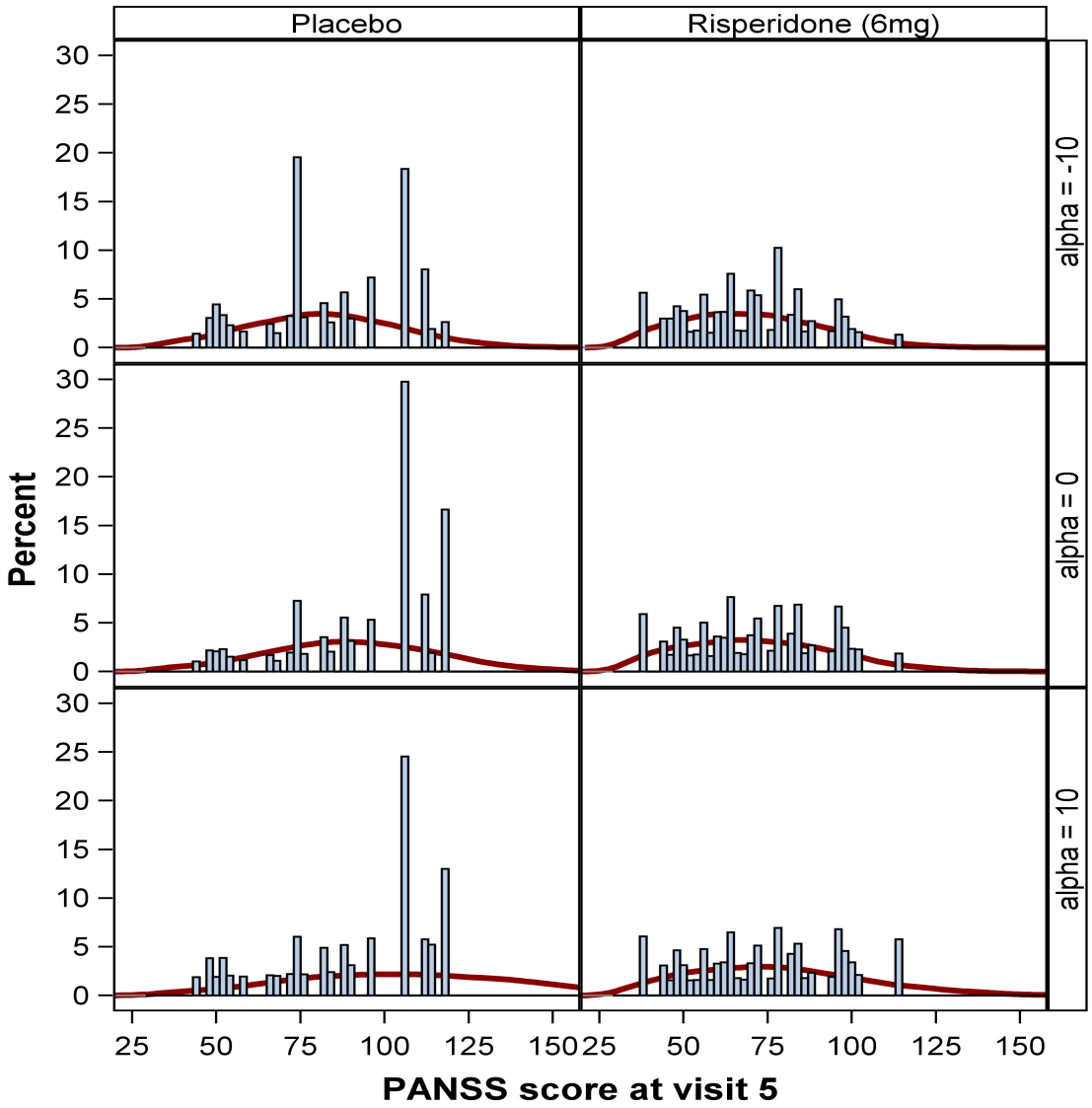
Results


Placebo
Risperidone (6mg)
PANSS score at visit 5
110
100
90
80
70
60
-10
-5
0
5
10
α
Method
1
2
3

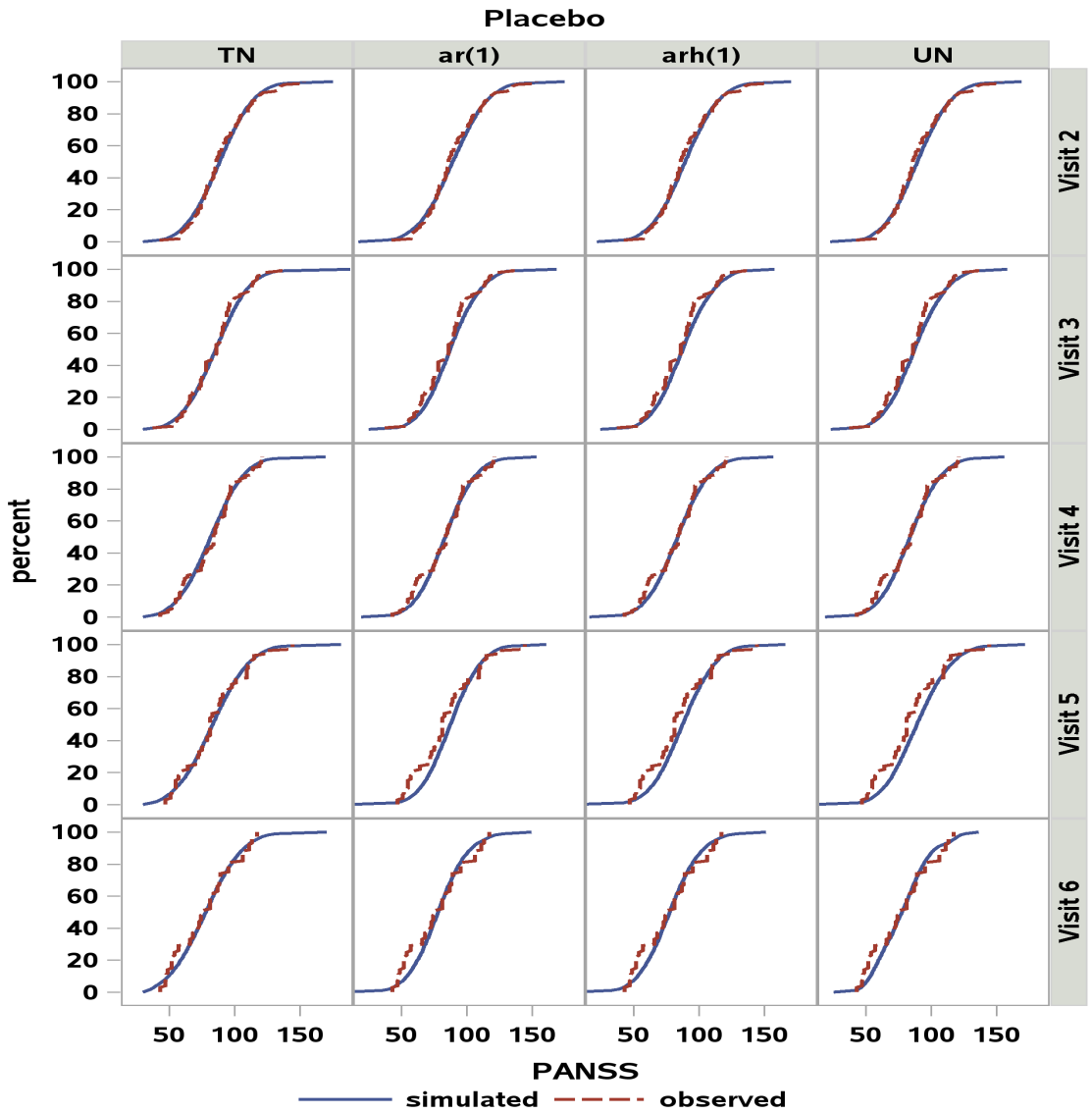
Method 1 vs. Method 2



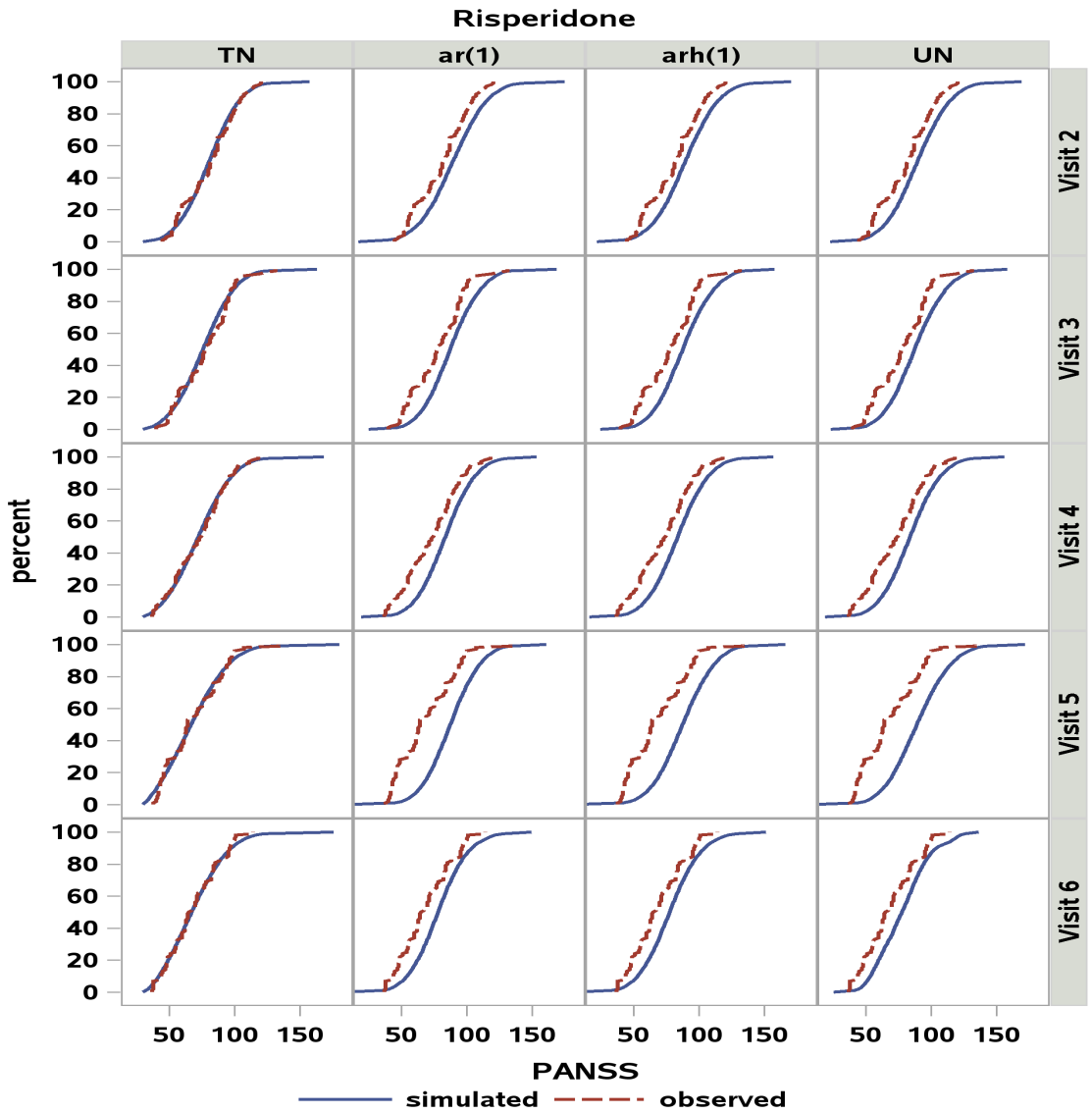
Method 1 vs. Method 3



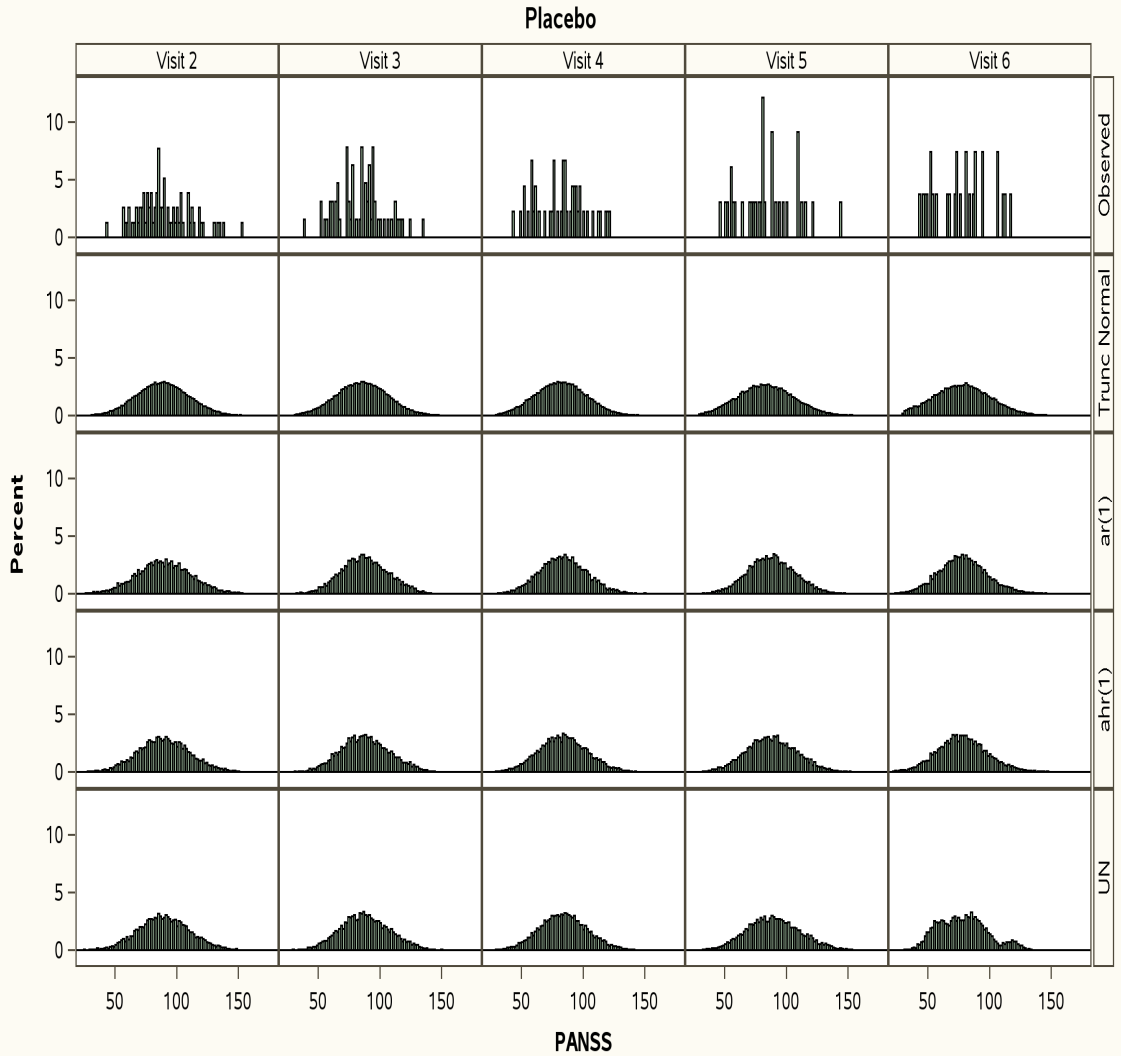
MMRM


Placebo
TN
ar(1)
arh(1)
UN
Visit 2
Visit 3
Visit 4
Visit 5
Visit 6
percent
0
20
40
60
80
100
50
100
150
PANSS
simulated
observed

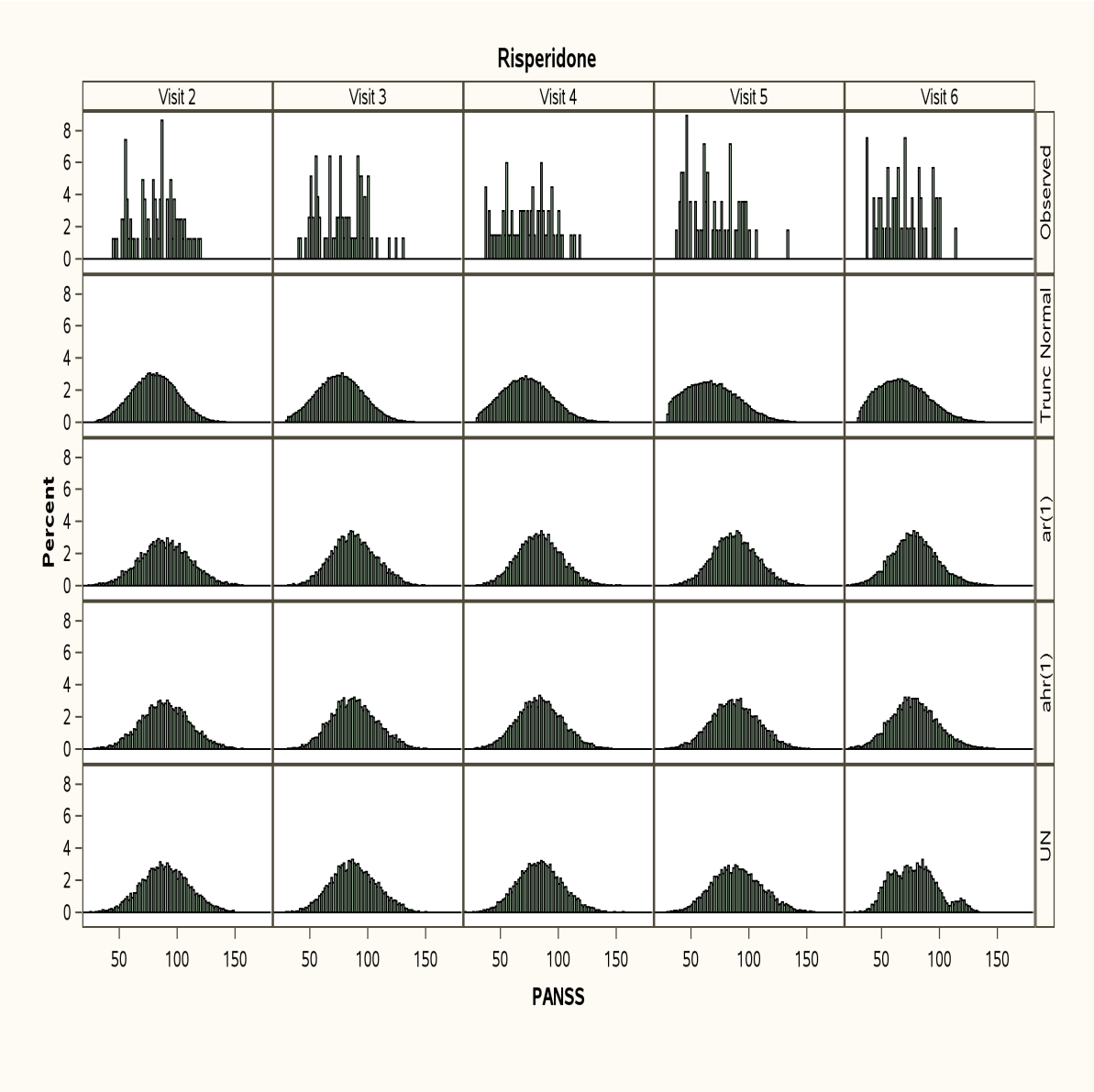
MMRM


Risperidone
TN
ar(1)
arh(1)
UN
Visit 2
Visit 3
Visit 4
Visit 5
Visit 6
percent
0
20
40
60
80
100
50
100
150
PANSS
simulated
observed

MMRM


Placebo
Visit 2
Visit 3
Visit 4
Visit 5
Visit 6
Observed
Trunc Normal
ar(1)
ahr(1)
UN
Percent
PANSS
0
5
10
50
100
150

MMRM


Risperidone
Visit 2
Visit 3
Visit 4
Visit 5
Visit 6
Observed
Trunc Normal
ar(1)
ahr(1)
UN
Percent
PANSS
0
2
4
6
8
50
100
150

MMRM, LOCF

	Difference	95% CI
AR(1)	-13.62	[-20.81,-5.42]
ARH(1)	-14.54	[-22.33,-5.49]
UN	-14.76	[-23.54,-6.07]
LOCF	-17.33	[-24.26,-10.40]
TN	-18.22	[-26.50,-9.22]

Discussion

- Can you envision regulatory applications that present sensitivity analyses in this way?
- Pre-specification. Models for observed data can be outsourced. Need they be specified in advance?
- What bells and whistles are needed? Non-monotone missing data, multiple causes of dropout etc.