Global Sensitivity Analysis for Repeated Measures Studies with Informative Drop-out: A Semi-Parametric Approach

Daniel Scharfstein
Aidan McDermott
Ivan Diaz
Johns Hopkins University
Ibrahim Turkoz
Janssen Research and Development

September 18, 2014

Andrei Yakovlev Colloquium

Dr. Jack Hall said that Dr. Yakovlev enjoyed

> *"discovering major flaws in widely used methodology and creating innovative methods to overcome them."*

Focus

- 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.

Asssumptions

- 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.
- Since type (i) assumptions are not testable, it is essential to conduct a sensitivity analysis, whereby the data analysis is repeated under different type (i) assumptions.
- There are an infinite number of ways of positing type (i) assumptions.
- Ultimately, these assumptions prescribe how missing outcomes should be "imputed."

Types of Sensitivity Analysis

- Ad-hoc
  - Try a bunch of different methods.
- Local
  - Explore sensitivity in a small neighborhood around a benchmark assumption.
- **Global**
  - Explore sensitivity in a much larger neighborhood around a benchmark assumption.

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

Benchmark Assumption: Missing at Random

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

- Type (i) Assumption

Class of Type (i) Assumptions

For $k = 0, ..., K-1$,

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

where

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

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

Class of Type (i) Assumptions

- $\alpha = 0$ is Missing at Random
- α quantifies the influence of Y_{k+1} on the decision to drop-out before visit $k+1$, among those on study at visit k with observed history Y_k^-.

Identification Formula

$$\mu(P^*) = E\left[\frac{R_K Y_K}{\prod_{k=0}^{K-1}(1+\exp(h_k(Y_k^-)+\alpha r(Y_{k+1})))^{-1}}\right]$$

where P^* is the true distribution of the observed data, characterized by

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

$$f(Y_{k+1}|R_{k+1}=1, Y_k^-) \text{ and } f(Y_0)$$

- These conditional distributions can't be estimated at fast enough rates so a plug-in estimator of μ^* will converge at $\sqrt{n}$ rates.

Type (ii) Assumptions

First-order Markov assumptions:

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

and

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

- Non-parametric smoothing with respect to the covariate Y_k using a Gaussian kernel.
- Estimate optimal smoothing parameters using a weighted squared-error loss function and 10-fold cross validation.

Estimation of μ

- Plug-in estimator, $\mu(\widehat{P})$, can suffer from non-standard asymptotics.
- To correct this problem, we use a one-step estimator:

$$\text{plug-in} + \text{average of estimated influence function}$$

Aside

- Consider a parametric submodel indexed by a finite dimensional parameter, say θ, that passes through $P \in \mathcal{P}$.
- A parametric submodel is a collection of distributions $\{P_\theta : \theta \in \Theta\} \subset \mathcal{P}$ where, WLOG, $P_{\theta=0} = P$.
- An asymptotically linear estimator of $\mu(P)$ with (mean zero) influence function, $\psi_P(O)$, will be regular at P if and only if, for all parametric submodels,

$$\left.\frac{\partial \mu(P_\theta)}{\partial \theta}\right|_{\theta=0} = E_P[\psi_P(O) S_\theta(O)] \tag{1}$$

where $S_\theta(O) = \left.\frac{\partial \log dP_\theta}{\partial \theta}\right|_{\theta=0}$.

Aside

- This implies that

$$\mu(P_\theta) - \mu(P) = E_{P_\theta}[\psi_P(O)] + O(\|\theta\|^2) \tag{2}$$

  for all parametric submodels.
- This implies that

$$\mu(Q) - \mu(P) = E_Q[\psi_P(O)] + O(\|Q - P\|^2), \tag{3}$$

  where Q is some other distribution in $\mathcal{P}$.

Aside

- With $P = \widehat{P}$, and $Q = P^*$, (3) becomes

$$\mu(\widehat{P}) - \mu^* = -E_{P^*}[\psi_{\widehat{P}}(O)] + O_{P^*}(\|\widehat{P} - P^*\|^2) \qquad (4)$$

- Adding and subtracting terms, we obtain

$$\begin{aligned} \mu(\widehat{P}) - \mu^* &= E_n[\psi_{P^*}(O)] - E_n[\psi_{\widehat{P}}(O)] + \\ &\int \{\psi_{\widehat{P}}(o) - \psi_{P^*}(o)\}\{dP_n(o) - dP^*(o)\} + \\ &O_{P^*}(\|\widehat{P} - P^*\|^2) \end{aligned}$$

Aside

- Assuming $\|\widehat{P} - P^*\|^2 = o_{P^*}(n^{-1/2})$ and additional regularity conditions,

$$\mu(\widehat{P}) - \mu^* = E_n[\psi_{P^*}(O)] - E_n[\psi_{\widehat{P}}(O)] + o_{P^*}(n^{-1/2})$$

- Consider the "one-step" estimator

$$\widehat{\mu} = \mu(\widehat{P}) + E_n[\psi_{\widehat{P}}(O)]$$

- Then

$$\sqrt{n}(\widehat{\mu} - \mu^*) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_{P^*}(O_i) + o_{P^*}(1)$$

- That is, $\widehat{\mu}$ is asymptotically linear with influence function $\psi_{P^*}(O)$.

Aside

- If no testable restrictions are placed on $\mathcal{P}$, then $\psi_{P^*}(O)$ satisfying (1) will be unique: $\psi^{np}_{P^*}(O)$.
- If testable restrictions are placed on $\mathcal{P}$, then $\psi_{P^*}(O)$ satisfying (1) will not generally be unique.
- The influence function that yields the smallest asymptotic variance, $\psi^{sp}_{P^*}(O)$, is the projection of $\psi^{np}_{P^*}(O)$ onto the tangent space of the model $\mathcal{P}$.
- The tangent space of a parametric submodel passing through $P^* \in \mathcal{P}$ is a space of random variables that can be expressed as linear combinations of the components of $S_\theta(O)$.
- The tangent space of the model $\mathcal{P}$ is the smallest, closed space that contains all the parametric submodel tangent spaces.

Uncertainty

- An influence function-based 95% confidence interval takes the form $\widehat{\mu} \pm 1.96\widehat{se}(\widehat{\mu})$, where $\widehat{se}(\widehat{\mu}) = \sqrt{E_n[\psi^{sp}_{\widehat{P}}(O)^2]/n}$.
- In studentized bootstrap, the confidence interval takes the form $[\widehat{\mu} + t_{0.025}\widehat{se}(\widehat{\mu}), \widehat{\mu} + t_{0.975}\widehat{se}(\widehat{\mu})]$, where t_q is the qth quantile of $\left\{ \frac{\widehat{\mu}^{(b)} - \widehat{\mu}}{\widehat{se}(\widehat{\mu}^{(b)})} : b = 1, \ldots, B \right\}$ and $\widehat{se}(\widehat{\mu}^{(b)})$

Uncertainty - Double Bootstrap

- For the bth 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 qth 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

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

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

Uncertainty - Fast Double Bootstrap

- The drawback of double bootstrap is that it is computationally intensive.
- To address this issue, set $S = 1$ and defined $\tilde{t}_q^{(b)} = \tilde{t}_q$ above to be qth quantile of $\left\{ \frac{\widehat{\mu}^{(b,1)} - \widehat{\mu}^{(b)}}{\widehat{se}(\widehat{\mu}^{(b,1)})} : b = 1, \ldots, B \right\}$.

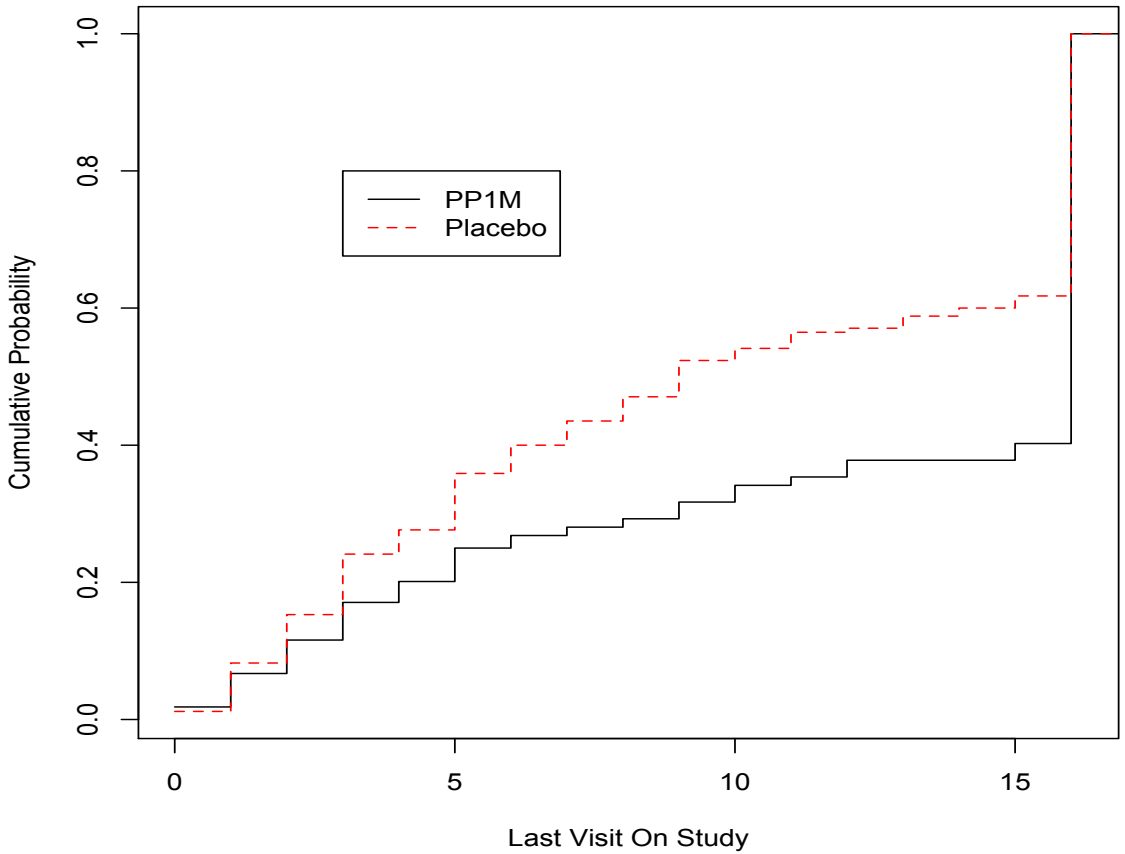
Case Study: SCA-3004

- 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



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).

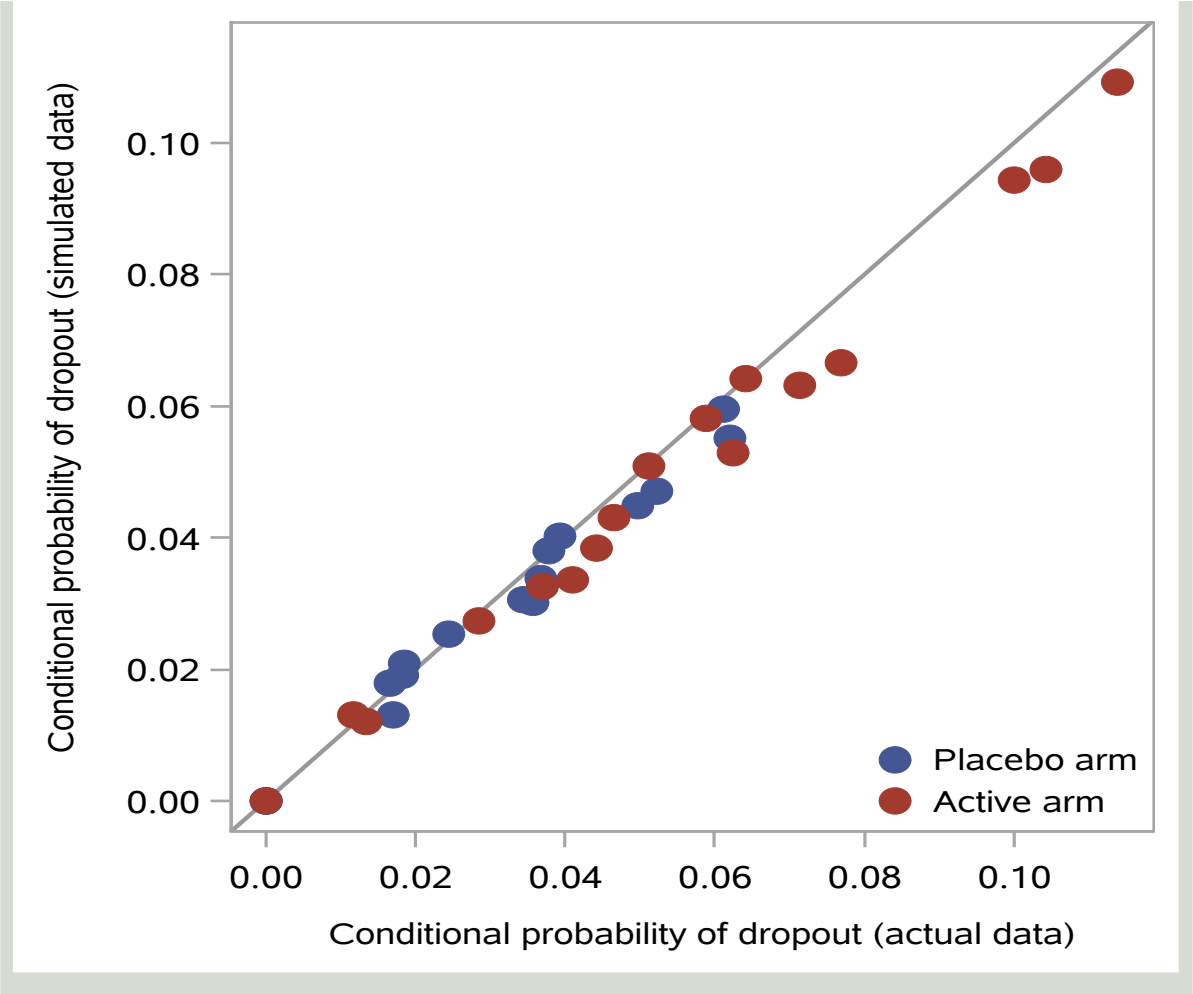
Case Study: SCA-3004

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

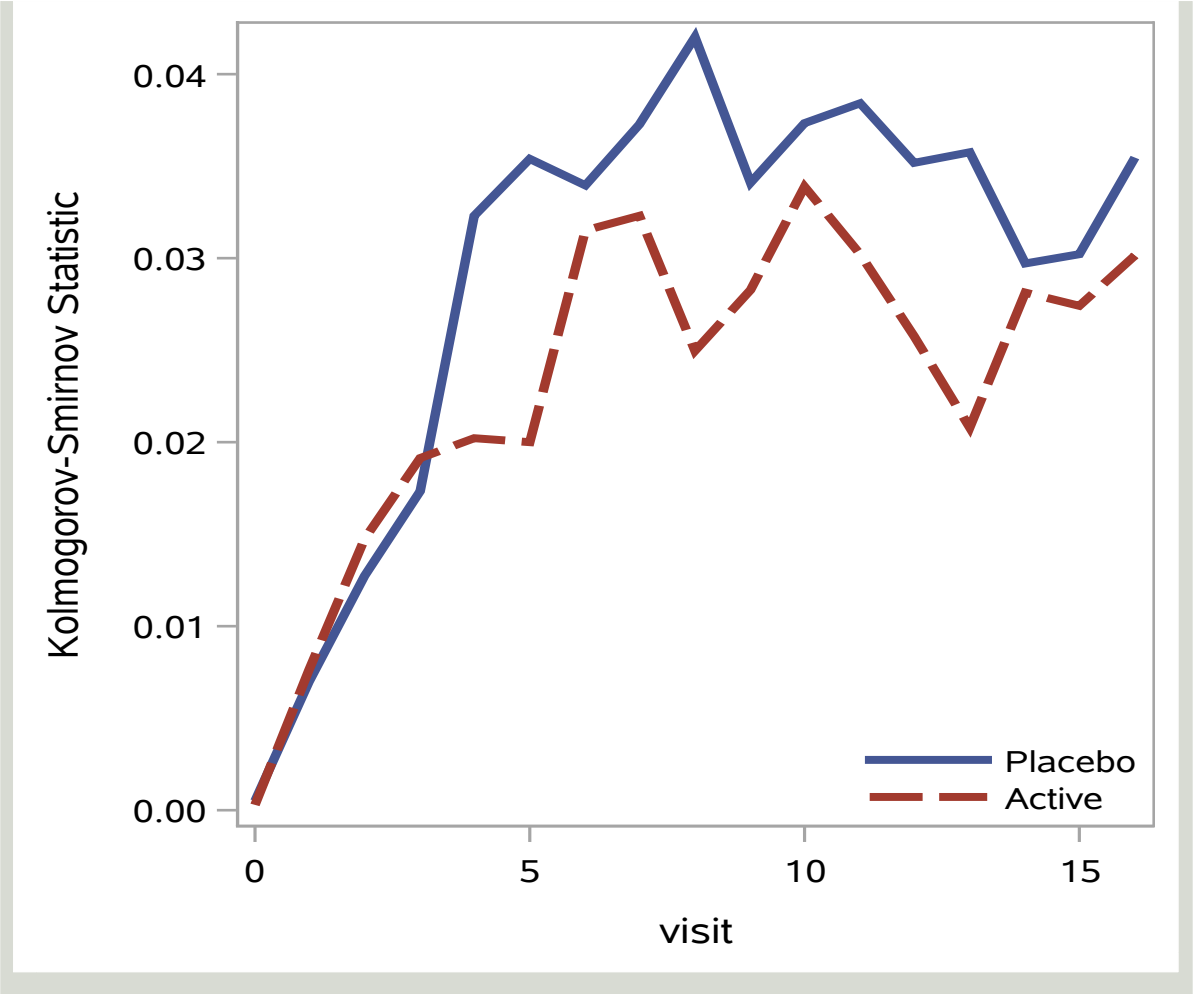
Case Study: SCA-3004

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



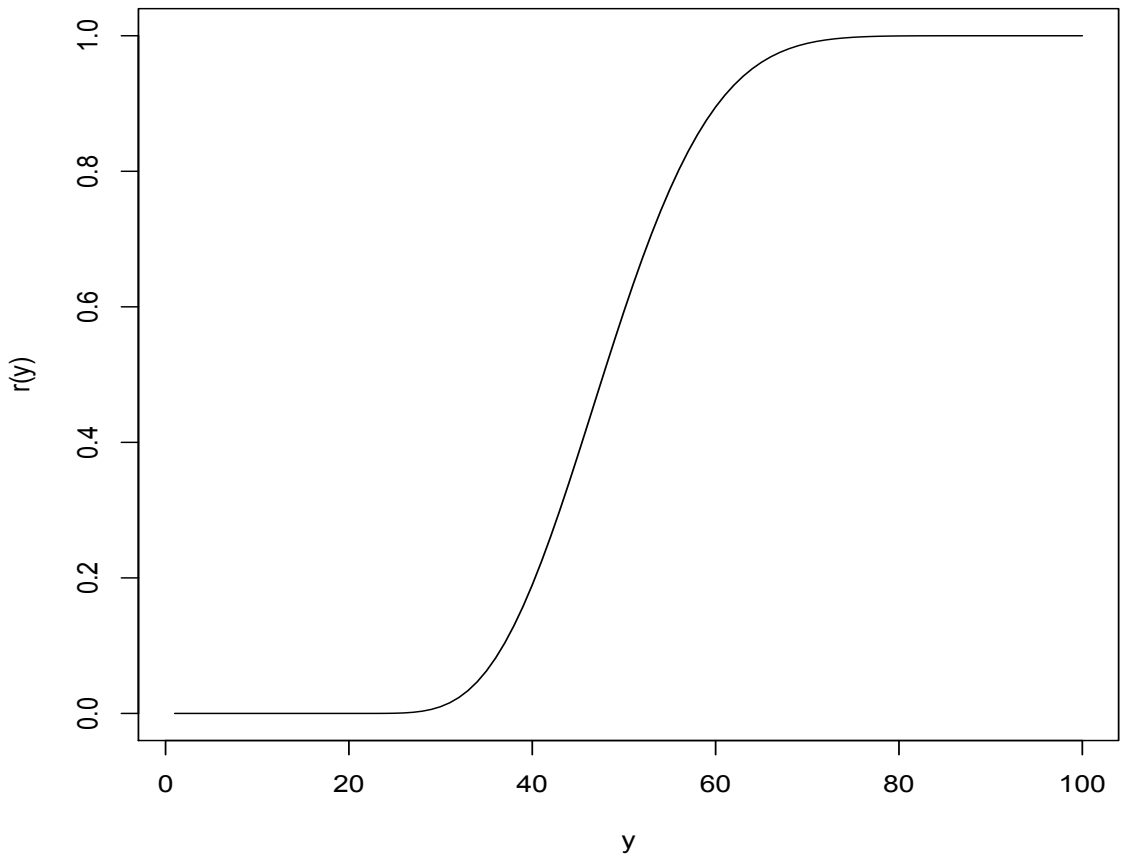
Case Study: SCA-3004



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).

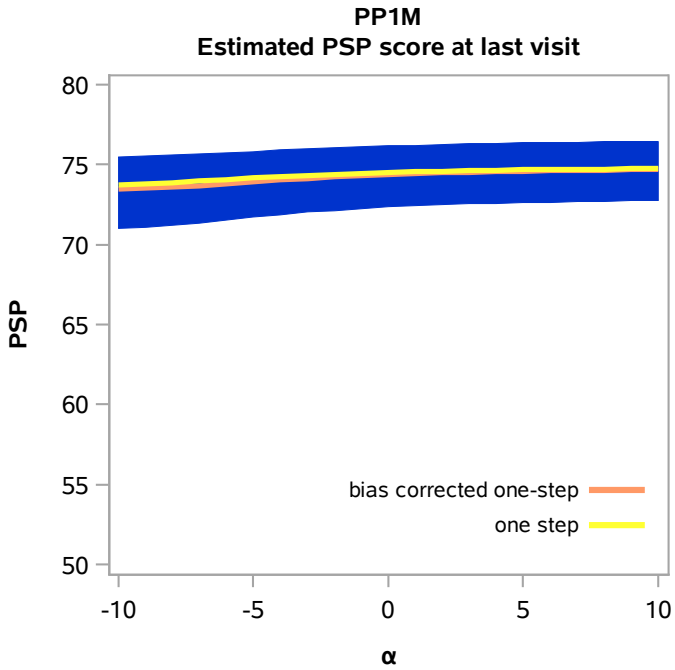
Case Study: SCA-3004

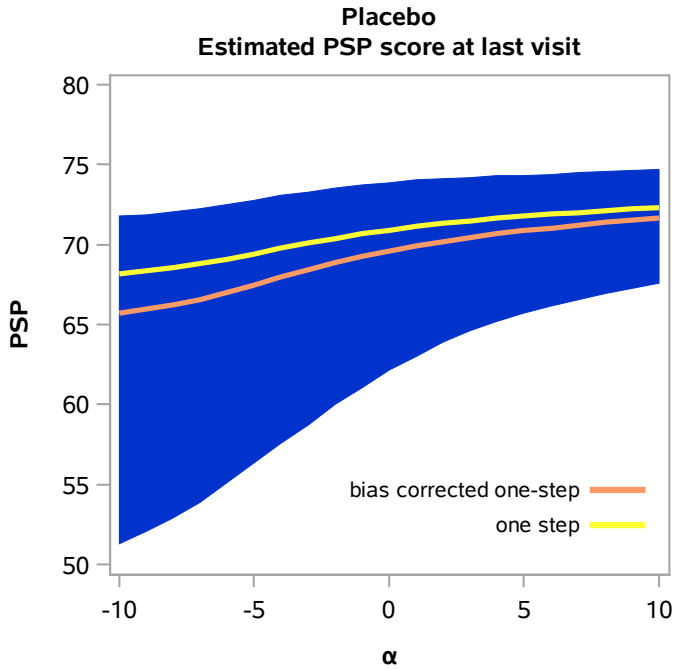


Case Study: SCA-3004

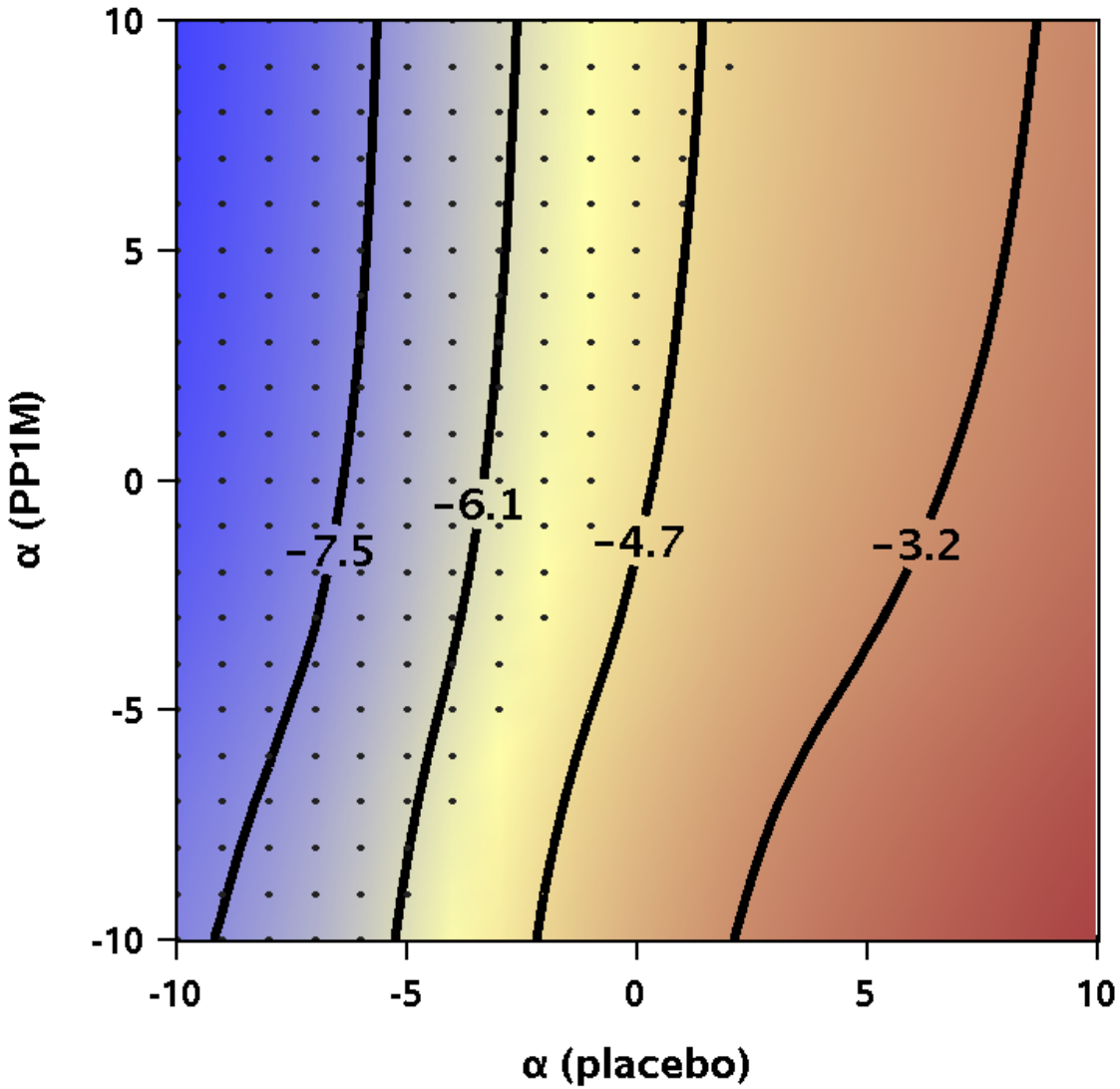
y_{k+1}	y^*_{k+1}	Log Odds Ratio
30	20	$\alpha \times 0.01$
40	30	$\alpha \times 0.18$
50	40	$\alpha \times 0.40$
60	50	$\alpha \times 0.30$
70	60	$\alpha \times 0.09$
80	700	$\alpha \times 0.01$

Case Study: SCA-3004





Case Study: SCA-3004



Simulation Study

		PP1M			PBO		
α	Estimator	μ^*	Bias	MSE	μ^*	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

Simulation Study

α	Procedure	PP1M Coverage	PBO Coverage
-10	IF	88.6%	65.8%
	SB	93.6%	90.8%
	FDB	94.3%	93.9%
-5	IF	91.3%	72.3%
	SB	94.2%	91.4%
	FDB	94.6%	93.9%
-1	IF	92.7%	81.6%
	SB	94.4%	92.2%
	FDB	94.8%	94.1%
0	IF	92.8%	83.1%
	SB	94.4%	92.6%
	FDB	94.8%	94.2%
1	IF	92.9%	84.2%
	SB	94.5%	92.8%
	FDB	94.9%	94.1%
5	IF	93.0%	87.0%
	SB	94.6%	93.5%
	FDB	94.7%	94.6%
10	IF	93.1%	88.7%
	SB	94.6%	94.1%
	FDB	94.8%	94.8%

Software - SAMON

www.missingdatamatters.org

Discussion

- Among patients on study at visit k with observed history Y_k^-, our model does not allow unmeasured predictors of R_{k+1} and Y_{k+1}.

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

- Incorporate auxiliary covariates.
- Intermittent missing data.