Murdoch Childrens Research Institute
MONASH University
Inserm

Multiple imputation and sensitivity analysis for incomplete longitudinal data departing from the MAR assumption

Margarita Moreno-Betancur & Michel Chavance

(published in *Stat Methods Med Res*, 2013)

Biometrics by the Harbour
Hobart, November 30 2015

Longitudinal data with drop-outs

$$\mathbf{Y}_i = \left(Y_{i1}, \ldots, Y_{i(U_i-1)}, Y_{iU_i}, \ldots, Y_{iJ}\right)$$

Longitudinal data with drop-outs

$$\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{i(U_i-1)}, Y_{iU_i}, \ldots, Y_{iJ})$$

The SMI study

(sleep-maintenance insomnia)

- $n =$962 patients randomized to placebo ($X_i = 0$) or treatment ($X_i = 1$).

Longitudinal data with drop-outs

$$\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{i(U_i-1)}, Y_{iU_i}, \ldots, Y_{iJ})$$

The SMI study

(sleep-maintenance insomnia)

- n =962 patients randomized to placebo ($X_i = 0$) or treatment ($X_i = 1$).
- Scores measuring quality of sleep collected through 6 periods.

Longitudinal data with drop-outs

$$\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{i(U_i-1)}, Y_{iU_i}, \ldots, Y_{iJ})$$

The SMI study

(sleep-maintenance insomnia)

- n =962 patients randomized to placebo ($X_i = 0$) or treatment ($X_i = 1$).
- Scores measuring quality of sleep collected through 6 periods.
- Goal: Inference about the treatment effect

$$\theta = E(Y_{i5} - Y_{i0} | X_i = 1) - E(Y_{i5} - Y_{i0} | X_i = 0)$$

Longitudinal data with drop-outs

$$\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{i(U_i-1)}, Y_{iU_i}, \ldots, Y_{iJ})$$

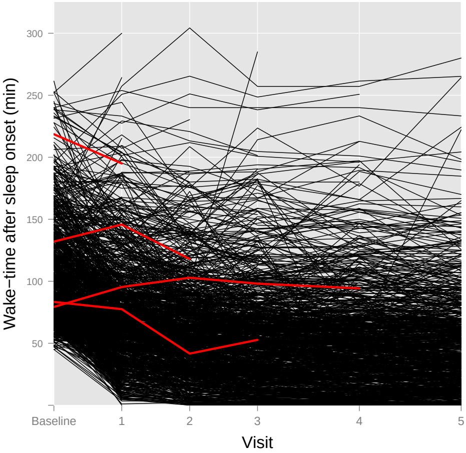
The SMI study

(sleep-maintenance insomnia)

- n =962 patients randomized to placebo ($X_i = 0$) or treatment ($X_i = 1$).
- Scores measuring quality of sleep collected through 6 periods.
- Goal: Inference about the treatment effect

$$\theta = E(Y_{i5} - Y_{i0} | X_i = 1) - E(Y_{i5} - Y_{i0} | X_i = 0)$$

- 22% dropped-out



Longitudinal data with drop-outs

$$\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{i(U_i-1)}, Y_{iU_i}, \ldots, Y_{iJ})$$

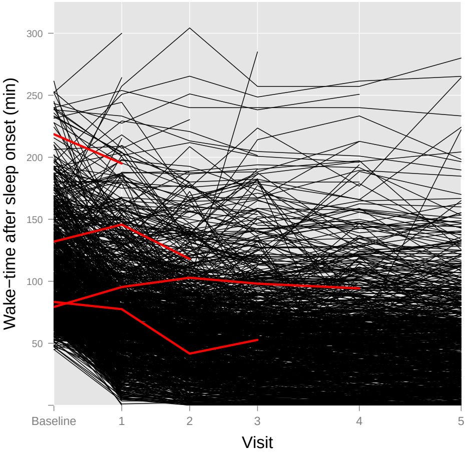
The SMI study

(sleep-maintenance insomnia)

- n =962 patients randomized to placebo ($X_i = 0$) or treatment ($X_i = 1$).
- Scores measuring quality of sleep collected through 6 periods.
- Goal: Inference about the treatment effect

$$\theta = E(Y_{i5} - Y_{i0} | X_i = 1) - E(Y_{i5} - Y_{i0} | X_i = 0)$$

- 22% dropped-out



Reasons for drop-outs: side-effects, lack of efficacy, protocol violation...

Consequences of missing outcomes

Loss of precision + Need assumptions about missingness

Consequences of missing outcomes

Loss of precision + Need assumptions about missingness

Rubin's taxonomy **(Rubin 1976)**

Consequences of missing outcomes

Loss of precision + Need assumptions about missingness

Rubin's taxonomy **(Rubin 1976)**

- **MAR** if missingness probability is conditionally independent of missing outcome given covariates and observed outcomes.
    - Missing and observed data can be assumed to share conditional distribution.
    - No further assumptions needed to draw valid inferences from available data.

Consequences of missing outcomes

Loss of precision + Need assumptions about missingness

Rubin's taxonomy (Rubin 1976)

- **MAR** if missingness probability is conditionally independent of missing outcome given covariates and observed outcomes.
  - Missing and observed data can be assumed to share conditional distribution.
  - No further assumptions needed to draw valid inferences from available data.
- **MNAR** otherwise.
  - Conditional distributions of observed and missing data differ.
  - Need to make further assumptions to address the identifiability issues.

Consequences of missing outcomes

Loss of precision + Need assumptions about missingness

Rubin's taxonomy **(Rubin 1976)**

- **MAR** if missingness probability is conditionally independent of missing outcome given covariates and observed outcomes.
  - Missing and observed data can be assumed to share conditional distribution.
  - No further assumptions needed to draw valid inferences from available data.
- **MNAR** otherwise.
  - Conditional distributions of observed and missing data differ.
  - Need to make further assumptions to address the identifiability issues.

But not possible to assess from data whether MAR or MNAR **(Molenberghs et al. 2008)**.

$\Rightarrow$ **Sensitivity Analyses**

Snapshot of incomplete longitudinal data literature

- **Covariate-dependent drop-out:** CC
- **MAR:** Direct likelihood, WGEE, MI, MI-GEE, Doubly-robust estimators...
- **MNAR:** Selection models, pattern-mixture models, shared-parameter models ('joint models')
- **Sensitivity analyses:**
  - Global and local influence diagnostics from a single model.
  - Consider a finite set of models with different structural and/or distributional assumptions.
  - Consider a family of MNAR models indexed by a parameter quantifying the distance from MAR.

    (Little 1994, Schaferstein et al. 1999, Daniels and Hogan 2000, Molenberghs et al. 2001,...)
  - *Sensitivity parameter* approach **(Daniels and Wang 2009, Hogan 2009)**

A family of PMMs for longitudinal data

A family of **Linear mixed models (LMM)** that assumes different trajectories for the observed ($R_{ij} = 0$) and missing ($R_{ij} = 1$) outcomes:

$$\boxed{Y_{ij} = \mathbf{X}_{ij}'\boldsymbol{\beta} + \mathbf{Z}_{ij}'\mathbf{b}_i + \boldsymbol{\kappa} R_{ij} + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N(0, \sigma^2), \qquad \mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G})}$$

A family of PMMs for longitudinal data

A family of **Linear mixed models (LMM)** that assumes different trajectories for the observed ($R_{ij} = 0$) and missing ($R_{ij} = 1$) outcomes:

$$Y_{ij} = \mathbf{X}'_{ij}\boldsymbol{\beta} + \mathbf{Z}'_{ij}\mathbf{b}_i + \boldsymbol{\kappa} R_{ij} + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N(0, \sigma^2), \qquad \mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G})$$

The distribution of the missing outcomes is identified up to the sensitivity parameter $\boldsymbol{\kappa}$ which quantifies:

A family of PMMs for longitudinal data

A family of **Linear mixed models (LMM)** that assumes different trajectories for the observed ($R_{ij} = 0$) and missing ($R_{ij} = 1$) outcomes:

$$Y_{ij} = \mathbf{X}'_{ij}\boldsymbol{\beta} + \mathbf{Z}'_{ij}\mathbf{b}_i + \boldsymbol{\kappa} R_{ij} + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N(0, \sigma^2), \qquad \mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G})$$

The distribution of the missing outcomes is identified up to the sensitivity parameter $\boldsymbol{\kappa}$ which quantifies:

- Degree of departure from MAR, which is equivalent to $\boldsymbol{\kappa} = 0$ ('centered' at MAR).

A family of PMMs for longitudinal data

A family of **Linear mixed models (LMM)** that assumes different trajectories for the observed ($R_{ij} = 0$) and missing ($R_{ij} = 1$) outcomes:

$$Y_{ij} = \mathbf{X}_{ij}'\boldsymbol{\beta} + \mathbf{Z}_{ij}'\mathbf{b}_i + \boldsymbol{\kappa} R_{ij} + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N(0, \sigma^2), \qquad \mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G})$$

The distribution of the missing outcomes is identified up to the sensitivity parameter $\boldsymbol{\kappa}$ which quantifies:

- Degree of departure from MAR, which is equivalent to $\boldsymbol{\kappa} = 0$ ('centered' at MAR).
- Shift in the expected trajectory of missing outcomes wrt observed outcomes.

A family of PMMs for longitudinal data

A family of **Linear mixed models (LMM)** that assumes different trajectories for the observed ($R_{ij} = 0$) and missing ($R_{ij} = 1$) outcomes:

$$\boxed{Y_{ij} = \mathbf{X}_{ij}'\beta + \mathbf{Z}_{ij}'\mathbf{b}_i + \boldsymbol{\kappa} R_{ij} + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N(0, \sigma^2), \qquad \mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G})}$$

The distribution of the missing outcomes is identified up to the sensitivity parameter $\boldsymbol{\kappa}$ which quantifies:

- Degree of departure from MAR, which is equivalent to $\boldsymbol{\kappa} = 0$ ('centered' at MAR).
- Shift in the expected trajectory of missing outcomes wrt observed outcomes.

Varying $\boldsymbol{\kappa}$ over a set of plausible values enables assessment of sensitivity.

A family of PMMs for longitudinal data

A family of **Linear mixed models (LMM)** that assumes different trajectories for the observed ($R_{ij} = 0$) and missing ($R_{ij} = 1$) outcomes:

$$Y_{ij} = \mathbf{X}_{ij}'\boldsymbol{\beta} + \mathbf{Z}_{ij}'\mathbf{b}_i + \boldsymbol{\kappa} R_{ij} + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N(0, \sigma^2), \qquad \mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G})$$

The distribution of the missing outcomes is identified up to the sensitivity parameter $\boldsymbol{\kappa}$ which quantifies:

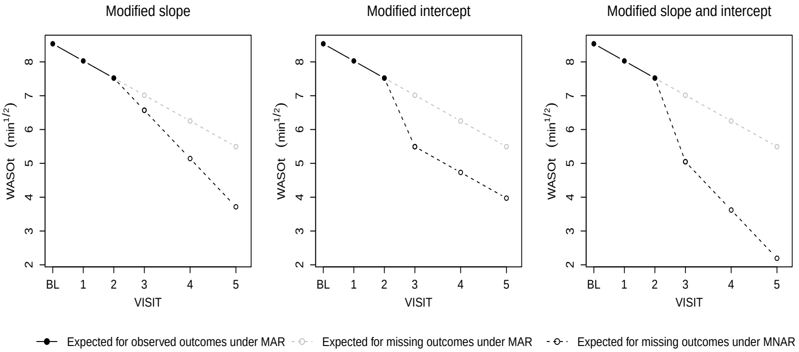
- Degree of departure from MAR, which is equivalent to $\boldsymbol{\kappa} = 0$ ('centered' at MAR).
- Shift in the expected trajectory of missing outcomes wrt observed outcomes.

Varying $\boldsymbol{\kappa}$ over a set of plausible values enables assessment of sensitivity.

Allowing $\boldsymbol{\kappa} = \boldsymbol{\kappa}(\mathbf{X}_{ij})$ we can model key characteristics of the missing outcome distribution that may affect inferences about the parameter of interest.

A picture is worth 1000 words...

SMI example: Expected trajectories for a patient in the treatment group who dropped-out after visit 2.



Practical implementation

To obtain estimates of **marginal parameters**, a practical and unified approach is MI.

Practical implementation

To obtain estimates of **marginal parameters**, a practical and unified approach is MI.

Step 1 Fit the LMM to the available data to obtain $\hat{\boldsymbol{\beta}}$, $\hat{\text{var}}(\hat{\boldsymbol{\beta}})$, $\hat{\mathbf{b}}_i$, $\hat{\text{var}}(\hat{\mathbf{b}}_i)$, $\hat{\sigma}^2$.

Practical implementation

To obtain estimates of **marginal parameters**, a practical and unified approach is MI.

Step 1 Fit the LMM to the available data to obtain $\hat{\boldsymbol{\beta}}$, $\hat{\text{var}}(\hat{\boldsymbol{\beta}})$, $\hat{\mathbf{b}}_i$, $\hat{\text{var}}(\hat{\mathbf{b}}_i)$, $\hat{\sigma}^2$.

Step 2 Impute the missing outcomes m times by drawing values from the LMM using these estimates and the chosen $\boldsymbol{\kappa}(\mathbf{X}_{ij})$. This yields m completed datasets.

Practical implementation

To obtain estimates of **marginal parameters**, a practical and unified approach is MI.

Step 1 Fit the LMM to the available data to obtain $\hat{\boldsymbol{\beta}}$, $\hat{\text{var}}(\hat{\boldsymbol{\beta}})$, $\hat{\mathbf{b}}_i$, $\hat{\text{var}}(\hat{\mathbf{b}}_i)$, $\hat{\sigma}^2$.

Step 2 Impute the missing outcomes m times by drawing values from the LMM using these estimates and the chosen $\boldsymbol{\kappa}(\mathbf{X}_{ij})$. This yields m completed datasets.

Step 3 Estimate the marginal parameter of interest $\boldsymbol{\theta}$ and its variance from each completed dataset by using a complete data method.

Practical implementation

To obtain estimates of **marginal parameters**, a practical and unified approach is MI.

Step 1 Fit the LMM to the available data to obtain $\hat{\boldsymbol{\beta}}$, $\widehat{\text{var}}(\hat{\boldsymbol{\beta}})$, $\hat{\mathbf{b}}_i$, $\widehat{\text{var}}(\hat{\mathbf{b}}_i)$, $\hat{\sigma}^2$.

Step 2 Impute the missing outcomes m times by drawing values from the LMM using these estimates and the chosen $\boldsymbol{\kappa}(\mathbf{X}_{ij})$. This yields m completed datasets.

Step 3 Estimate the marginal parameter of interest $\boldsymbol{\theta}$ and its variance from each completed dataset by using a complete data method.

Step 4 Use Rubin's formulas to combine estimates into a single inference.

Practical implementation

To obtain estimates of **marginal parameters**, a practical and unified approach is MI.

Step 1 Fit the LMM to the available data to obtain $\hat{\boldsymbol{\beta}}$, $\hat{\text{var}}(\hat{\boldsymbol{\beta}})$, $\hat{\mathbf{b}}_i$, $\hat{\text{var}}(\hat{\mathbf{b}}_i)$, $\hat{\sigma}^2$.

Step 2 Impute the missing outcomes m times by drawing values from the LMM using these estimates and the chosen $\boldsymbol{\kappa}(\mathbf{X}_{ij})$. This yields m completed datasets.

Step 3 Estimate the marginal parameter of interest $\boldsymbol{\theta}$ and its variance from each completed dataset by using a complete data method.

Step 4 Use Rubin's formulas to combine estimates into a single inference.

MI-based implementation makes it easy to do several analyses over a range of $\boldsymbol{\kappa}(\mathbf{X}_{ij})$.

Step 2: Imputation procedure for longitudinal data

Imputation procedure for $l \in \{1, \ldots, m\}$:

(a) Draw $\beta^{(l)} \sim N(\hat{\beta}, \widehat{\text{var}}\hat{\beta})$ and $\mathbf{b}_i^{(l)} \sim N(\hat{\mathbf{b}}_i, \widehat{\text{var}}\hat{\mathbf{b}}_i)$ for $i = 1, \ldots, n$.

(b) Draw $\sigma^{2(l)} \sim \hat{\sigma}^2 \times \left(\frac{d}{\chi^2_d}\right)$.
$d = n_1 - q =$ residual degrees of freedom
$n_1 =$ number of observations to fit the model
$q =$ trace of 'hat matrix' (estimate of effective # of parameters)
(Bates, 2006)

(c) Draw $\varepsilon_{ij}^{(l)} \sim N(0, \sigma^{2(l)})$.

(d) Impute each missing outcome Y_{ij} as
$\mathbf{X}'_{ij}\beta^{(l)} + \mathbf{Z}'_{ij}\mathbf{b}_i^{(l)} + \boldsymbol{\kappa}(\mathbf{X}_{ij}) + \varepsilon_{ij}^{(l)}$.

This procedure can be used for MAR analyses taking $\boldsymbol{\kappa} = 0$. Currently studying its use to impute time-dependent covariates for the Cox model.

Simulation study

Aim: To assess the approach in the realistic situation where the family of PMMs does not include the true MNAR model that generated the data.

- Design mimicked the 2-arm, 6-visit design of the SMI study, with outcomes generated from:

$$Y_{ij} = j\beta X_i + b_{0i} + jb_{1i} + \varepsilon_{ij}$$

- MNAR drop-outs were generated under a selection model.

- Target parameter θ = Expected difference in outcomes at last visit.

$$Y_{i5} = \theta_0 + \theta X_i + \epsilon_i$$

- A family of PMMs indexed by arm-specific sensitivity parameters k_0, k_1.
  - 'Best MNAR' model: $\hat{k}_0$ and $\hat{k}_1$ for which PMM$\approx$ true model.
  - Increasing departures from true model: by taking $\hat{k}_0$ and $\hat{k}_1$ as reference.

Simulation study: Some results for $\theta = 1$

- Drop-out probability lower for subjects with lower outcomes.

Analysis	k_0	k_1	% bias
MAR	0	0	-6.9
MNAR1	$\hat{k}_0/2$	$\hat{k}_1/2$	-2.6
Best MNAR	$\hat{k}_0$	$\hat{k}_1$	1.5
MNAR2	$2\hat{k}_0$	$2\hat{k}_1$	10.0
MNAR3	$\hat{k}_0/2$	$2\hat{k}_1$	88.3
MNAR4	$2\hat{k}_0$	$\hat{k}_1/2$	-80.5
CC	–	–	-35.6

- Drop-out probability lower for subjects with lower outcomes & global drop-out probability higher in treatment group.

Analysis	k_0	k_1	% bias
MAR	0	0	-41.7
MNAR1	$\hat{k}_0/2$	$\hat{k}_1/2$	-20.8
Best MNAR	$\hat{k}_0$	$\hat{k}_1$	1.9
MNAR2	$2\hat{k}_0$	$2\hat{k}_1$	44.0
MNAR3	$\hat{k}_0/2$	$2\hat{k}_1$	73.6
MNAR4	$2\hat{k}_0$	$\hat{k}_1/2$	-49.4
CC	–	–	-165.0

Simulation study: Results summary

- **When PMM $\approx$ true model:** Satisfactory CPs and type I error rates.

- **When departing from the true model:** We observed the variation in the expected value of the coefficient estimator, which depended on the missingness mechanism and other factors. The approach is thus suitable for assessing sensitivity.

- Moderate upward bias in Rubin's variance estimator ($|\text{MRB}| \leq 10\%$) resulted in conservative (yet acceptable) CPs. A consequence of misspecification and uncongeniality.

Application to the SMI study

To assess the evidence of a treatment effect: $\theta = E(Y_{i5} - Y_{i0} | X_i = 1) - E(Y_{i5} - Y_{i0} | X_i = 0)$

Application to the SMI study

To assess the evidence of a treatment effect: $\theta = E(Y_{i5} - Y_{i0} | X_i = 1) - E(Y_{i5} - Y_{i0} | X_i = 0)$

- **Primary analysis:** MAR-based MI.

$$Y_{ij} = \beta_0 + \beta_1 X_i + \beta_2 t_j + \beta_3 X_i t_j + b_{0i} + b_{1i} t_j + \varepsilon_{ij}$$

Application to the SMI study

To assess the evidence of a treatment effect: $\theta = E(Y_{i5} - Y_{i0} | X_i = 1) - E(Y_{i5} - Y_{i0} | X_i = 0)$

- **Primary analysis:** MAR-based MI.

$$Y_{ij} = \beta_0 + \beta_1 X_i + \beta_2 t_j + \beta_3 X_i t_j + b_{0i} + b_{1i} t_j + \varepsilon_{ij}$$

	$\hat{\theta}$	95% CI	p-value
WASO	−14.31	[−20.39, −8.23]	<0.001
SLREF	−0.09	[−0.17, −0.01]	0.03

WASO=Wake-time after sleep onset (minutes)
SLREF=Sleep refreshing quality (1=excellent to 4=poor)
Means of daily measures over 2/3-week periods; decrease in either is sign of improvement

Application to the SMI study

To assess the evidence of a treatment effect: $\theta = E(Y_{i5} - Y_{i0} | X_i = 1) - E(Y_{i5} - Y_{i0} | X_i = 0)$

- **Primary analysis:** MAR-based MI.

$$Y_{ij} = \beta_0 + \beta_1 X_i + \beta_2 t_j + \beta_3 X_i t_j + b_{0i} + b_{1i} t_j + \varepsilon_{ij}$$

	$\hat{\theta}$	95% CI	p-value
WASO	−14.31	[−20.39, −8.23]	<0.001
SLREF	−0.09	[−0.17, −0.01]	0.03

WASO=Wake-time after sleep onset (minutes)
SLREF=Sleep refreshing quality (1=excellent to 4=poor)
Means of daily measures over 2/3-week periods; decrease in either is sign of improvement

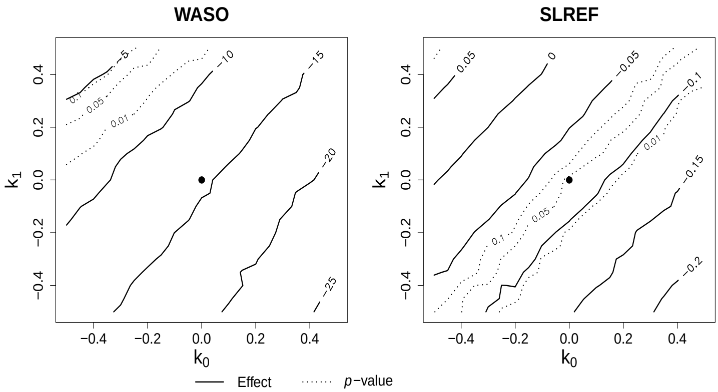
- **Secondary analyses:** Control of group-specific intercepts of the missing data distribution.

$$Y_{ij} = \beta_0 + \beta_1 X_i + \beta_2 t_j + \beta_3 X_i t_j + k_{X_i} \hat{\varsigma}_5 R_{ij} + b_{0i} + b_{1i} t_j + \varepsilon_{ij}$$

k_0, k_1 since $X_i = 0$ or 1.
$\hat{\varsigma}_5$ = sample SD of scores at visit 5.

Sensitivity analysis for θ

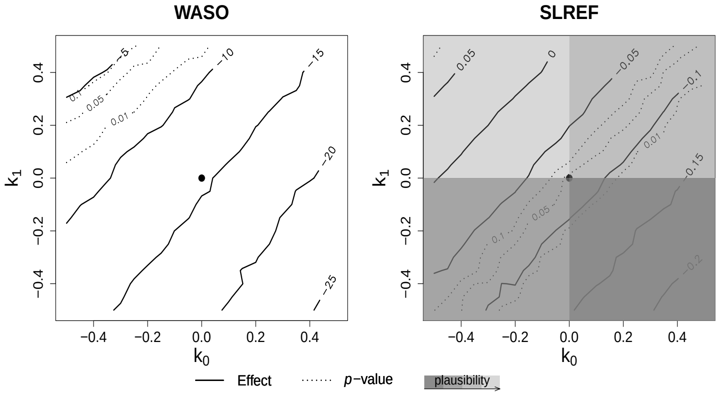
$$\theta = E(Y_{i5} - Y_{i0} | X_i = 1) - E(Y_{i5} - Y_{i0} | X_i = 0)$$



For the WASO score, there is strong evidence of a treatment effect in most scenarios. For the SLREF score, the evidence is fragile and strongly dependent on missingness assumptions.

Sensitivity analysis for θ

$$\theta = E(Y_{i5} - Y_{i0} | X_i = 1) - E(Y_{i5} - Y_{i0} | X_i = 0)$$



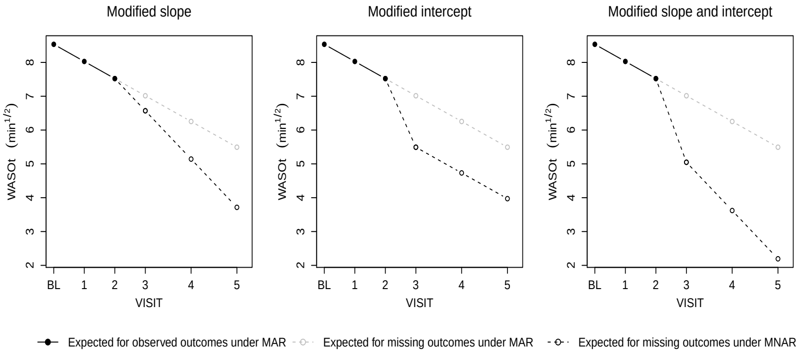
In practice, the plausibility of the scenarios studied needs to be assessed.

Sensitivity analysis for β_3=difference in time-slopes

Control group-specific intercepts and time-slopes of missing data distribution.

$$Y_{ij} = \beta_0 + \beta_1 X_i + \beta_2 t_j + \beta_3 X_i t_j + \kappa(X_i, t_j, S_i) R_{ij} + b_{0i} + b_{1i} t_j + \varepsilon_{ij}$$

$$\kappa(X_i, t_j, S_i) = k^{(1)}_{X_i} \hat{\varsigma} + k^{(2)}_{X_i} \hat{\beta}_2 (t_j - S_i)$$



Sensitivity analysis for β_3: Some results

Scenario	$k_{X_i}^{(1)}$	$k_{X_i}^{(2)}$	$\hat{\beta}_3$	95% CI	p-value
MAR	–	–	-0.031	[-0.041, -0.021]	<0.001
Modified slopes & intercepts	-0.5	0.5	-0.025	[-0.037, -0.013]	<0.001
	-1	1	-0.017	[-0.030, -0.003]	0.019
	$(-1)^{1-X_i} * 0.1$	0.5	-0.016	[-0.028, -0.004]	0.004
	$(-1)^{1-X_i} * 0.1$	1	-0.012	[-0.024, 0.000]	0.040
	-0.1	$(-1)^{X_i} * 0.25$	-0.014	[-0.026, -0.002]	0.027
	-0.25	$(-1)^{X_i} * 0.25$	-0.010	[-0.024, 0.004]	0.151
	$(-1)^{1-X_i} * 0.05$	$(-1)^{X_i} * 0.05$	-0.019	[-0.031, -0.007]	0.001
	$(-1)^{1-X_i} * 0.1$	$(-1)^{X_i} * 0.1$	-0.011	[-0.022, 0.001]	0.066

For the WASO score, there is evidence of a treatment effect under this definition too under MAR and across a large range of scenarios departing from this assumption.

Concluding remarks

- Advantages over some previous approaches **(e.g. Daniels and Hogan 2000, Ratitch et al. 2013)**:
  - Suitable for studies with large # of measurements or where planned timing and # of measurements differ across subjects.
  - Can be used with intermittent missingness
  - Sensitivity parameters with intuitive interpretations (e.g. intercepts, time-slopes) facilitating the formulation of assumptions.

Concluding remarks

- Advantages over some previous approaches **(e.g. Daniels and Hogan 2000, Ratitch et al. 2013)**:
  - Suitable for studies with large # of measurements or where planned timing and # of measurements differ across subjects.
  - Can be used with intermittent missingness
  - Sensitivity parameters with intuitive interpretations (e.g. intercepts, time-slopes) facilitating the formulation of assumptions.

- How to summarize results? **(Molenberghs et al. 2001, Vansteelandt et al. 2006)**

  Ignorance interval = Parameter regions yielded by these approaches
  e.g. for WASO score, $\hat{\theta} \in [-25, -5]$

  Uncertainty interval = Ignorance interval + Confidence interval
  = Uncertainty due to finite sampling + missing data
  e.g. for WASO score, $\hat{\theta} \in [-25 - \delta_l, -5 + \delta_u]$

Concluding remarks

- Advantages over some previous approaches **(e.g. Daniels and Hogan 2000, Ratitch et al. 2013)**:
  - Suitable for studies with large # of measurements or where planned timing and # of measurements differ across subjects.
  - Can be used with intermittent missingness
  - Sensitivity parameters with intuitive interpretations (e.g. intercepts, time-slopes) facilitating the formulation of assumptions.

- How to summarize results? **(Molenberghs et al. 2001, Vansteelandt et al. 2006)**
  Ignorance interval = Parameter regions yielded by these approaches
  e.g. for WASO score, $\hat{\theta} \in [-25, -5]$
  Uncertainty interval = Ignorance interval + Confidence interval
  = Uncertainty due to finite sampling + missing data
  e.g. for WASO score, $\hat{\theta} \in [-25 - \delta_l, -5 + \delta_u]$

- Other issues: non-continuous outcomes, more than one incomplete longitudinal variable, elicit expert opinions..

Backup

Longitudinal data and sensitivity analyses

- When modeling longitudinal data with drop-outs, we need to consider the joint distribution of $\mathbf{Y}_i = (\mathbf{Y}_i^{\mathcal{O}}, \mathbf{Y}_i^{\mathcal{M}})$ and $U_i =$ occasion of the first missing outcome:

$$f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}, u|\varphi)$$

Longitudinal data and sensitivity analyses

- When modeling longitudinal data with drop-outs, we need to consider the joint distribution of $\mathbf{Y}_i = (\mathbf{Y}_i^{\mathcal{O}}, \mathbf{Y}_i^{\mathcal{M}})$ and $U_i =$ occasion of the first missing outcome:

$$f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}, u|\varphi) = \underbrace{f(\mathbf{y}^{\mathcal{M}}|\mathbf{y}^{\mathcal{O}}, u, \varphi)}_{\text{Extrapolation}} \times \underbrace{f(\mathbf{y}^{\mathcal{O}}, u|\varphi)}_{\text{Observed}}$$

Longitudinal data and sensitivity analyses

- When modeling longitudinal data with drop-outs, we need to consider the joint distribution of $\mathbf{Y}_i = (\mathbf{Y}_i^{\mathcal{O}}, \mathbf{Y}_i^{\mathcal{M}})$ and $U_i =$ occasion of the first missing outcome:

$$f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}, u|\varphi) = \underbrace{f(\mathbf{y}^{\mathcal{M}}|\mathbf{y}^{\mathcal{O}}, u, \varphi)}_{\text{Extrapolation}} \times \underbrace{f(\mathbf{y}^{\mathcal{O}}, u|\varphi)}_{\text{Observed}}$$

- **Aim of a sensitivity analysis:** To assess the sensitivity of inferences about the target parameter $\boldsymbol{\theta} = h(\boldsymbol{\varphi})$ to assumptions about the extrapolation model, which are unverifiable.

Longitudinal data and sensitivity analyses

- When modeling longitudinal data with drop-outs, we need to consider the joint distribution of $\mathbf{Y}_i = (\mathbf{Y}_i^{\mathcal{O}}, \mathbf{Y}_i^{\mathcal{M}})$ and $U_i =$ occasion of the first missing outcome:

$$f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}, u|\varphi) = \underbrace{f(\mathbf{y}^{\mathcal{M}}|\mathbf{y}^{\mathcal{O}}, u, \varphi)}_{\text{Extrapolation}} \times \underbrace{f(\mathbf{y}^{\mathcal{O}}, u|\varphi)}_{\text{Observed}}$$

- **Aim of a sensitivity analysis:** To assess the sensitivity of inferences about the target parameter $\boldsymbol{\theta} = h(\varphi)$ to assumptions about the extrapolation model, which are unverifiable.

- **A structured and focused approach:**

  (Little 1994, Schaferstein et al. 1999, Daniels and Hogan 2000, Molenberghs et al. 2001,...)

  - **Primary analysis**: e.g. MAR model.
  - **Secondary analyses**: Consider a large family of MNAR models indexed by a parameter quantifying the distance from the primary model.

The sensitivity parameter approach

Pattern-mixture models (PMMs) assume a different response mechanism per drop-out occasion and require making explicit assumptions about the extrapolation model:

$$f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}, u|\varphi) \quad = \quad f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}|u, \varphi) \times f(u|\varphi)$$

The sensitivity parameter approach

Pattern-mixture models (PMMs) assume a different response mechanism per drop-out occasion and require making explicit assumptions about the extrapolation model:

$$\begin{aligned} f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}, u|\boldsymbol{\varphi}) &= f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}|u, \boldsymbol{\varphi}) \times f(u|\boldsymbol{\varphi}) \\ &= \underbrace{f(\mathbf{y}^{\mathcal{M}}|\mathbf{y}^{\mathcal{O}}, u, \boldsymbol{\varphi})}_{\text{Extrapolation}} \times \underbrace{f(\mathbf{y}^{\mathcal{O}}|u, \boldsymbol{\varphi}) \times f(u|\boldsymbol{\varphi})}_{\text{Observed}} \end{aligned}$$

The sensitivity parameter approach

Pattern-mixture models (PMMs) assume a different response mechanism per drop-out occasion and require making explicit assumptions about the extrapolation model:

$$\begin{aligned} f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}, u|\varphi) &= f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}|u, \varphi) \times f(u|\varphi) \\ &= \underbrace{f(\mathbf{y}^{\mathcal{M}}|\mathbf{y}^{\mathcal{O}}, u, \varphi)}_{\substack{\text{Extrapolation} \\ (\boldsymbol{\phi}, \boldsymbol{\kappa})}} \times \underbrace{f(\mathbf{y}^{\mathcal{O}}|u, \varphi) \times f(u|\varphi)}_{\substack{\text{Observed} \\ \boldsymbol{\phi}}} \end{aligned}$$

PMMs can be parametrized such that $\varphi = (\boldsymbol{\phi}, \boldsymbol{\kappa})$; $\boldsymbol{\phi}$ is identifiable but $\boldsymbol{\kappa}$ is not.

The sensitivity parameter approach

Pattern-mixture models (PMMs) assume a different response mechanism per drop-out occasion and require making explicit assumptions about the extrapolation model:

$$\begin{aligned} f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}, u|\varphi) &= f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}|u, \varphi) \times f(u|\varphi) \\ &= \underbrace{f(\mathbf{y}^{\mathcal{M}}|\mathbf{y}^{\mathcal{O}}, u, \varphi)}_{\substack{\text{Extrapolation} \\ (\phi, \boldsymbol{\kappa})}} \times \underbrace{f(\mathbf{y}^{\mathcal{O}}|u, \varphi) \times f(u|\varphi)}_{\substack{\text{Observed} \\ \phi}} \end{aligned}$$

PMMs can be parametrized such that $\varphi = (\phi, \boldsymbol{\kappa})$; ϕ is identifiable but $\boldsymbol{\kappa}$ is not.

Different values of $\boldsymbol{\kappa}$ yield same fit to the observed data
...but also imply different extrapolation models and different inferences on $\boldsymbol{\theta}$.

The sensitivity parameter approach

Pattern-mixture models (PMMs) assume a different response mechanism per drop-out occasion and require making explicit assumptions about the extrapolation model:

$$
\begin{aligned}
f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}, u|\varphi) &= f(\mathbf{y}^{\mathcal{O}}, \mathbf{y}^{\mathcal{M}}|u, \varphi) \times f(u|\varphi) \\
&= \underbrace{f(\mathbf{y}^{\mathcal{M}}|\mathbf{y}^{\mathcal{O}}, u, \varphi)}_{\substack{\text{Extrapolation} \\ (\phi, \boldsymbol{\kappa})}} \times \underbrace{f(\mathbf{y}^{\mathcal{O}}|u, \varphi) \times f(u|\varphi)}_{\substack{\text{Observed} \\ \phi}}
\end{aligned}
$$

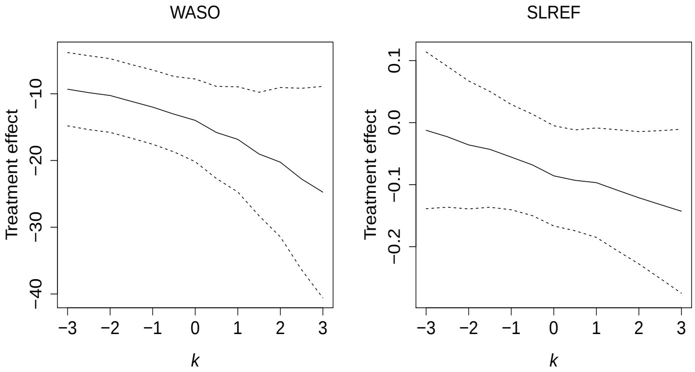
PMMs can be parametrized such that $\varphi = (\phi, \boldsymbol{\kappa})$; ϕ is identifiable but $\boldsymbol{\kappa}$ is not.

Different values of $\boldsymbol{\kappa}$ yield same fit to the observed data
...but also imply different extrapolation models and different inferences on $\boldsymbol{\theta}$.

$\boldsymbol{\kappa}$ is called a **sensitivity parameter** because it "embodies" the source of the sensitivity of inferences to different unverifiable assumptions about the extrapolation model.
(Daniels and Wang 2009, Hogan 2009)

Sensitivity analysis for θ

Scenario $k_0 = k_1 = k$



$\hat{\theta} \in [-25, -9]$ minutes $\hat{\theta} \in [-0.14, -0.01]$ points