Handling informative dropout in longitudinal analysis of health-related quality of life:

Application on data from the oesophageal cancer clinical trial PRODIGE5/ACCORD17

Benjamin Cuer
C.Mollevi, A.Anota, E.Charton, F.Bonnetain, T.Conroy, B.Juzyna, S.Gourgou, C.Touraine

ISOQOL, Dublin

26 October 2018

benjamin.cuer@icm.unicancer.fr

Health-related quality of life (HRQoL) in cancer clinical trials

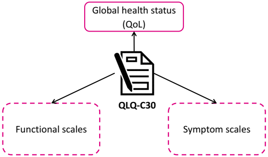
- Health-related quality of life = **PROs** used to measure the clinical benefits of new therapies

Health-related quality of life (HRQoL) in cancer clinical trials

- Health-related quality of life = **PROs** used to measure the clinical benefits of new therapies
- In clinical trials = more and more used as a **Secondary or a Co-Primary endpoint**
  $\hookrightarrow$ HRQoL assessed using questionnaires at different visit times (treatment and follow-up)

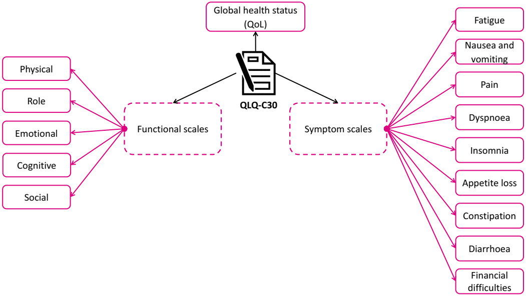
Health-related quality of life (HRQoL) in cancer clinical trials

- Health-related quality of life = **PROs** used to measure the clinical benefits of new therapies
- In clinical trials = more and more used as a **Secondary or a Co-Primary endpoint**
  ↪ HRQoL assessed using questionnaires at different visit times (treatment and follow-up)
- Europe : **EORTC QLQ-C30**



Health-related quality of life (HRQoL) in cancer clinical trials

- Health-related quality of life = **PROs** used to measure the clinical benefits of new therapies
- In clinical trials = more and more used as a **Secondary or a Co-Primary endpoint**
  ↪ HRQoL assessed using questionnaires at different visit times (treatment and follow-up)
- Europe : **EORTC QLQ-C30** = 5 functional scales + 9 symptom scales + a global health status
  ↪ **Continuous score** from 0 to 100 for each scale
  ↪ To model score **evolution over time**



Longitudinal analysis of the HRQoL score in clinical trials

To compare the score evolution between 2 treatment arms:

Longitudinal analysis of the HRQoL score in clinical trials

To compare the score evolution between 2 treatment arms:

Linear Mixed Model

For patient i and observation j, the score is as follows:

$$\begin{aligned} Y_{ij} &= \beta^T X_{ij} + b_i{}^T Z_{ij} + \varepsilon_{ij} \\ &= \beta_0 + \beta_1 \times t_{ij} + \beta_2 \times Arm_i + \beta_3 Arm_i \times t_{ij} + b_{0i} + b_{1i} \times t_{ij} + \varepsilon_{ij} \end{aligned}$$

where $\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} \sim \mathcal{N}(0, \Sigma)$ with $\Sigma \begin{pmatrix} \sigma^2_{b_0} & \sigma_{b_0 b_1} \\ \sigma_{b_0 b_1} & \sigma^2_{b_1} \end{pmatrix}$ and $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$

- Fixed effects: $\beta^T X_{ij}$ model the **average trajectory**
- Random effects: $b_i^T Z_{ij}$ represent the **individual deviation** from the average

Longitudinal analysis of the HRQoL score in clinical trials

To compare the score evolution between 2 treatment arms:

Linear Mixed Model

For patient i and observation j, the score is as follows:

$$\begin{aligned} Y_{ij} &= \beta^T X_{ij} + {b_i}^T Z_{ij} + \varepsilon_{ij} \\ &= \beta_0 + \beta_1 \times t_{ij} + \beta_2 \times Arm_i + \beta_3 Arm_i \times t_{ij} + b_{0i} + b_{1i} \times t_{ij} + \varepsilon_{ij} \end{aligned}$$

where $\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} \sim \mathcal{N}(0, \Sigma)$ with $\Sigma \begin{pmatrix} \sigma^2_{b_0} & \sigma_{b_0 b_1} \\ \sigma_{b_0 b_1} & \sigma^2_{b_1} \end{pmatrix}$ and $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$

- Fixed effects: $\beta^T X_{ij}$ model the **average trajectory**
- Random effects: $b_i^T Z_{ij}$ represent the **individual deviation** from the average

Coefficients interpretation

- β_1: **slope** in the standard arm
- β_3: **interaction effect** = difference of the score evolution in the experimental arm compared to the standard arm

Illustration on the PRODIGE5/ACCORD17 clinical trial

Illustration on the PRODIGE5/ACCORD17 clinical trial

Randomized phase II-III study comparing:
RT + FOLFOX regimen (experimental arm) *versus* ***RT + 5fu-cisplatin (control arm)***
of first line treatment in patients with inoperable oesophageal cancer

Illustration on the PRODIGE5/ACCORD17 clinical trial

Randomized phase II-III study comparing:
RT + FOLFOX regimen (experimental arm) *versus* ***RT + 5fu-cisplatin (control arm)***
of first line treatment in patients with inoperable oesophageal cancer

- Primary endpoint = **PFS** ⇒ no difference between the two treatment arms (Conroy et al., *Lancet Oncology*, 2014)
- Secondary endpoint = **HRQoL** (EORTC QLQ-C30) (Mollevi et al., *EJC*, 2017)

Illustration on the PRODIGE5/ACCORD17 clinical trial

Randomized phase II-III study comparing:
RT + FOLFOX regimen (experimental arm) *versus* ***RT + 5fu-cisplatin (control arm)***
of first line treatment in patients with inoperable oesophageal cancer

- Primary endpoint = **PFS** ⇒ no difference between the two treatment arms (Conroy et al., *Lancet Oncology*, 2014)
- Secondary endpoint = **HRQoL** (EORTC QLQ-C30) (Mollevi et al., *EJC*, 2017)

Results of the Linear Mixed Model for the Global Health Status

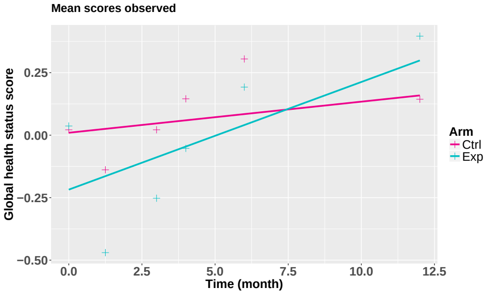
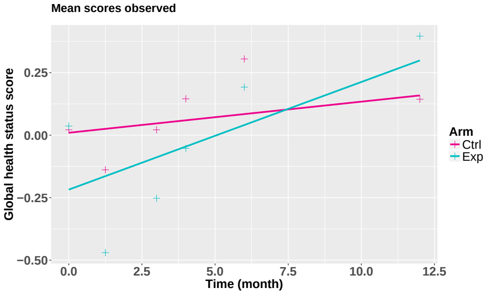


Illustration on the PRODIGE5/ACCORD17 clinical trial

Randomized phase II-III study comparing:
RT + FOLFOX regimen (experimental arm) *versus* ***RT + 5fu-cisplatin (control arm)***
of first line treatment in patients with inoperable oesophageal cancer

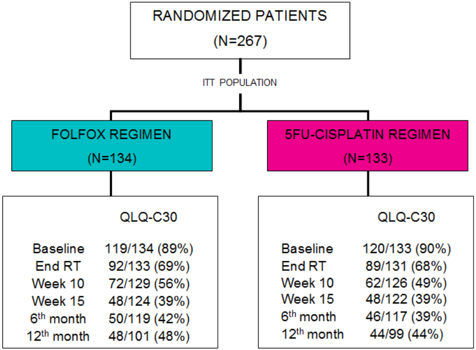
- Primary endpoint = **PFS** ⇒ no difference between the two treatment arms (Conroy et al., *Lancet Oncology*, 2014)
- Secondary endpoint = **HRQoL** (EORTC QLQ-C30) (Mollevi et al., *EJC*, 2017)

Results of the Linear Mixed Model for the Global Health Status

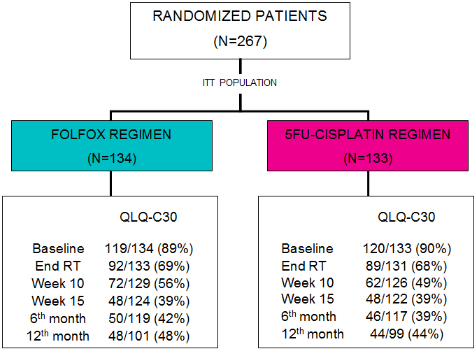


- **HRQoL improvement** over time ($\beta_1 = 0.012$, $p = 0.307$)
  ↪ **positive slope**
- **HRQoL improvement** more important in the experimental arm ($\beta_3 = 0.031$, $p = 0.076$)
  ↪ **positive arm effect**

PRODIGE5/ACCORD17 clinical trial: compliance



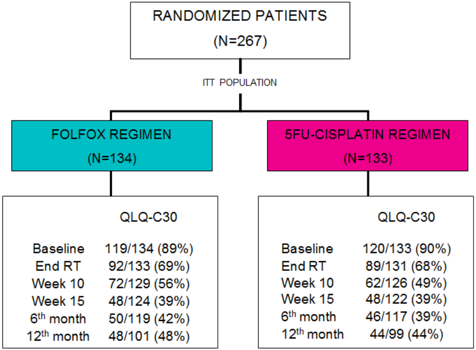
PRODIGE5/ACCORD17 clinical trial: compliance



Issue

- Estimations of the Linear Mixed Model are robust only if missing data are non informative.

PRODIGE5/ACCORD17 clinical trial: compliance



Issue

- Estimations of the Linear Mixed Model are robust only if missing data are non informative.

↪ The **objective** was to explore two types of models accounting for potentially informative missing data.

Patterns of missing data

Two patterns of missing data:

Patient	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5

Patterns of missing data

Two patterns of missing data:

- **Intermittent** missing data $\Rightarrow$ score is missing but subsequent score data can be observed

Patient	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
1	66.7	$\times$	66.7	50.0	50.0
2	33.3	41.7	$\times$	$\times$	33.3
3	$\times$	75.0	$\times$	83.3	50.0

Patterns of missing data

Two patterns of missing data:

- **Intermittent** missing data $\Rightarrow$ score is missing but subsequent score data can be observed
- **Monotone** missing data = **dropout** $\Rightarrow$ no observations are made after a certain time point
  $\hookrightarrow$ Generally corresponds to patient's dropout

Patient	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
1	66.7	×	66.7	50.0	50.0
2	33.3	41.7	×	×	33.3
3	×	75.0	×	83.3	50.0
4	83.3	100	66.7	66.7	×
5	91.7	25.0	25.0	×	×
6	66.7	66.7	×	×	×
7	50.0	×	×	×	×

Mechanisms of missing data

Missing Completely at Random (MCAR)

Missing Not at Random (MNAR)

Missing at Random (MAR)

Mechanisms of missing data

Missing Completely at Random (MCAR)

↪ Missingness is unrelated to the score

Missing at Random (MAR)

Missing Not at Random (MNAR)

Mechanisms of missing data

Missing Completely at Random (MCAR)

↪ Missingness is unrelated to the score

Missing at Random (MAR)

↪ Missingness depends on the observed score

Missing Not at Random (MNAR)

Mechanisms of missing data

Missing Completely at Random (MCAR)

↪ Missingness is unrelated to the score

Missing at Random (MAR)

↪ Missingness depends on the observed score

Missing Not at Random (MNAR)

↪ The missingness is likely to be related to the **observed and unobserved** score

Typically, if a patient does not complete a questionnaire because he/she is too tired, resulting in a missing score for the scale « fatigue », it is informative missing data.

Accounting for informative dropout in longitudinal analysis of the HRQoL score

Model jointly the score Y_i and the dropout variable D_i

Two decompositions of the joint distribution are possible according to the conditionning:

$$f_{\theta\psi}(Y_i, D_i) = \left\{ \right.$$

where θ and ψ denote the parameters vectors for the score and the missingness mechanism, respectively.

Accounting for informative dropout in longitudinal analysis of the HRQoL score

Model jointly the score Y_i and the dropout variable D_i

Two decompositions of the joint distribution are possible according to the conditionning:

$$f_{\theta\psi}(Y_i, D_i) = \begin{cases} f_\theta(Y_i \mid D_i) \times f_\psi(D_i) \\ f_\psi(D_i \mid Y_i) \times f_\theta(Y_i) \end{cases}$$

where θ and ψ denote the parameters vectors for the score and the missingness mechanism, respectively.

Accounting for informative dropout in longitudinal analysis of the HRQoL score

Model jointly the score Y_i and the dropout variable D_i

Two decompositions of the joint distribution are possible according to the conditionning:

$$f_{\theta\psi}(Y_i, D_i) = \begin{cases} f_\theta(Y_i \mid D_i) \times f_\psi(D_i) & \text{Pattern Mixture Model} \\ f_\psi(D_i \mid Y_i) \times f_\theta(Y_i) & \end{cases}$$

where θ and ψ denote the parameters vectors for the score and the missingness mechanism, respectively.

We are interested in :

1. Distribution of **the score Y_i given the dropout D_i** $\times$ marginal distribution of the dropout
   $\hookrightarrow$ Pattern Mixture Model

Accounting for informative dropout in longitudinal analysis of the HRQoL score

Model jointly the score Y_i and the dropout variable D_i

Two decompositions of the joint distribution are possible according to the conditionning:

$$f_{\theta\psi}(Y_i, D_i) = \begin{cases} f_\theta(Y_i \mid D_i) \times f_\psi(D_i) & \text{Pattern Mixture Model} \\ f_\psi(D_i \mid Y_i) \times f_\theta(Y_i) & \text{Selection Model} \end{cases}$$

where θ and ψ denote the parameters vectors for the score and the missingness mechanism, respectively.

We are interested in :

1. Distribution of **the score** Y_i **given the dropout** D_i $\times$ marginal distribution of the dropout
   $\hookrightarrow$ Pattern Mixture Model
2. Distribution of the **dropout** D_i **given the score** Y_i $\times$ marginal distribution of the score
   $\hookrightarrow$ Selection Model

Pattern Mixture Model: Definition

Definition

$$f_{\theta\psi}(Y_i, D_i) = f_\theta(Y_i \mid D_i) \times f_\psi(D_i)$$

$\Rightarrow$ **A mixture of models** for the score given the dropout weigthed by the dropout distribution

Pattern Mixture Model: Definition

Definition

$$f_{\theta\psi}(Y_i, D_i) = f_\theta(Y_i \mid D_i) \times f_\psi(D_i)$$

⇒ **A mixture of models** for the score given the dropout weigthed by the dropout distribution

To define two models :

- Multinomial model for $f(D_i)$:
- Mixed Model for $f(Y_i \mid D_i)$:

Pattern Mixture Model: Definition

Definition

$$f_{\theta\psi}(Y_i, D_i) = f_\theta(Y_i \mid D_i) \times f_\psi(D_i)$$

$\Rightarrow$ **A mixture of models** for the score given the dropout weigthed by the dropout distribution

To define two models :

- Multinomial model for $f(D_i)$: The probability of each dropout after time j is simply estimated by the **proportion** of patients which drop out after time j
- Mixed Model for $f(Y_i \mid D_i)$:

Pattern Mixture Model: Definition

Definition

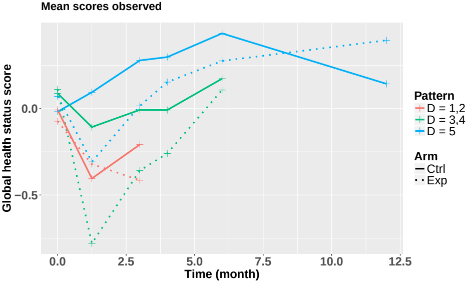
$$f_{\theta\psi}(Y_i, D_i) = f_\theta(Y_i \mid D_i) \times f_\psi(D_i)$$

$\Rightarrow$ **A mixture of models** for the score given the dropout weigthed by the dropout distribution

To define two models :

- Multinomial model for $f(D_i)$: The probability of each dropout after time j is simply estimated by the **proportion** of patients which drop out after time j
- Mixed Model for $f(Y_i \mid D_i)$: A Mixed model is assumed for the score in each pattern

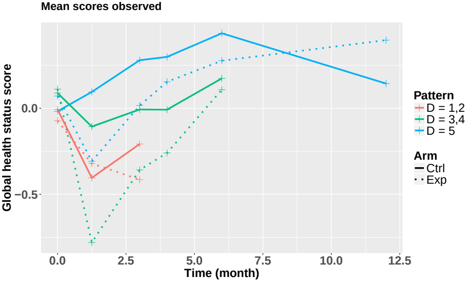
Pattern-Mixture Model: Application



First step: define 3 patterns

- Early dropout
- Medium dropout
- Late dropout

Pattern-Mixture Model: Application

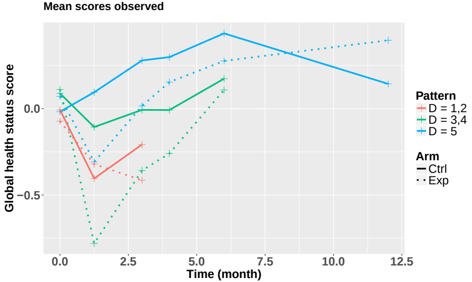


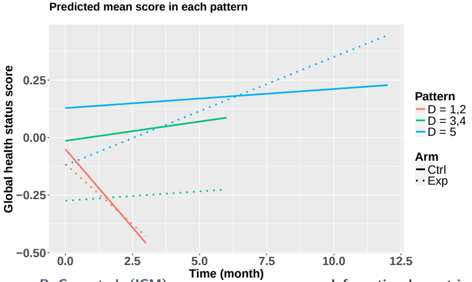
First step: define 3 patterns

Calcul of the proportion of patients in each pattern

- Early dropout $\pi^{(1,2)} = 0.34$
- Medium dropout $\pi^{(3,4)} = 0.30$
- Late dropout $\pi^{(5)} = 0.36$

Pattern-Mixture Model: Application





First step: define 3 patterns

Calcul of the proportion of patients in each pattern

- Early dropout $\pi^{(1,2)} = 0.34$
- Medium dropout $\pi^{(3,4)} = 0.30$
- Late dropout $\pi^{(5)} = 0.36$

Second step: Mixed Models

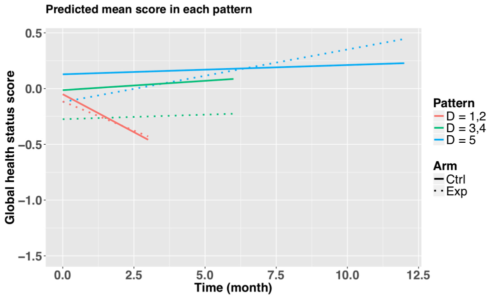
We model the score in each pattern

$$(Y_{ij} \mid D = 1,2) = X_{ij}\beta^{(1,2)} + Z_{ij} b_i + \varepsilon_{ij}$$
$$(Y_{ij} \mid D = 3,4) = X_{ij}\beta^{(3,4)} + Z_{ij} b_i + \varepsilon_{ij}$$
$$(Y_{ij} \mid D = 5) = X_{ij}\beta^{(5)} + Z_{ij} b_i + \varepsilon_{ij}$$

Pattern Mixture Model: Results

Slope and interaction effect in each pattern

	$D = 1, 2$		$D = 3, 4$		$D = 5$	
	Est. (SE)	p	Est. (SE)	p	Est. (SE)	p
β_1	-0.136 (0.108)	0.182	0.017 (0.036)	0.817	0.008 (0.014)	0.545
β_3	0.032 (0.137)	0.958	-0.009 (0.054)	0.379	0.039 (0.019)	0.047



Pattern Mixture Model: Results

Slope and interaction effect in each pattern

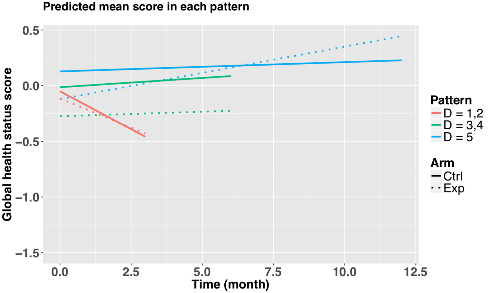
	$D = 1,2$		$D = 3,4$		$D = 5$	
	Est. (SE)	p	Est. (SE)	p	Est. (SE)	p
β_1	-0.136 (0.108)	0.182	0.017 (0.036)	0.817	0.008 (0.014)	0.545
β_3	0.032 (0.137)	0.958	-0.009 (0.054)	0.379	0.039 (0.019)	0.047

Compare 2 treatments $\Rightarrow$ Marginal parameters that is, irrespective of the pattern

Estimations in each pattern are **weighted** by the probability of dropout in each pattern:

$$\hat{\beta} = \pi^{(1,2)}\hat{\beta}^{(1,2)} + \pi^{(3,4)}\hat{\beta}^{(3,4)} + \pi^{(5)}\hat{\beta}^{(5)}$$

$\hookrightarrow$ **Extrapolation** = the score has the same trajectory after the dropout, that is, at times for which there is no available data.



Pattern Mixture Model: Results

Slope and interaction effect in each pattern

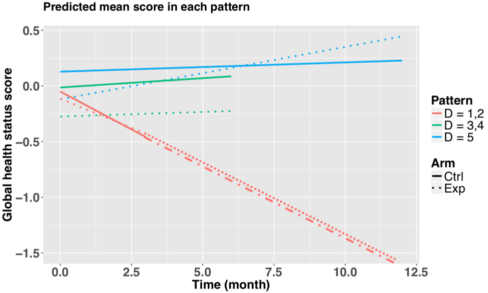
	$D=1,2$		$D=3,4$		$D=5$	
	Est. (SE)	p	Est. (SE)	p	Est. (SE)	p
β_1	-0.136 (0.108)	0.182	0.017 (0.036)	0.817	0.008 (0.014)	0.545
β_3	0.032 (0.137)	0.958	-0.009 (0.054)	0.379	0.039 (0.019)	0.047

Compare 2 treatments $\Rightarrow$ Marginal parameters that is, irrespective of the pattern

Estimations in each pattern are **weighted** by the probability of dropout in each pattern:

$$\hat{\beta} = \pi^{(1,2)}\hat{\beta}^{(1,2)} + \pi^{(3,4)}\hat{\beta}^{(3,4)} + \pi^{(5)}\hat{\beta}^{(5)}$$

$\hookrightarrow$ **Extrapolation** = the score has the same trajectory after the dropout, that is, at times for which there is no available data.



Pattern Mixture Model: Results

Slope and interaction effect in each pattern

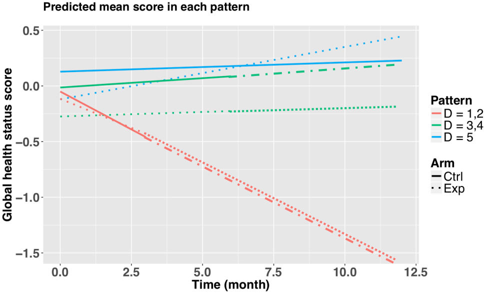
	$D=1,2$		$D=3,4$		$D=5$	
	Est. (SE)	p	Est. (SE)	p	Est. (SE)	p
β_1	-0.136 (0.108)	0.182	0.017 (0.036)	0.817	0.008 (0.014)	0.545
β_3	0.032 (0.137)	0.958	-0.009 (0.054)	0.379	0.039 (0.019)	0.047

Compare 2 treatments $\Rightarrow$ Marginal parameters that is, irrespective of the pattern

Estimations in each pattern are **weighted** by the probability of dropout in each pattern:

$$\hat{\beta} = \pi^{(1,2)}\hat{\beta}^{(1,2)} + \pi^{(3,4)}\hat{\beta}^{(3,4)} + \pi^{(5)}\hat{\beta}^{(5)}$$

$\hookrightarrow$ **Extrapolation** = the score has the same trajectory after the dropout, that is, at times for which there is no available data.



Pattern Mixture Model: Results

Slope and interaction effect in each pattern

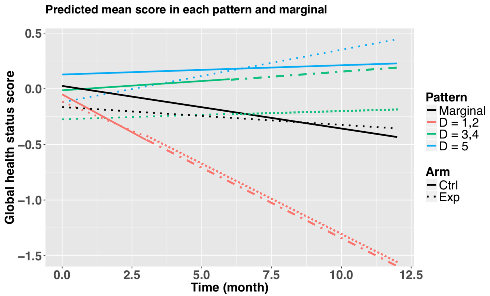
	$D = 1,2$		$D = 3,4$		$D = 5$	
	Est. (SE)	p	Est. (SE)	p	Est. (SE)	p
β_1	-0.136 (0.108)	0.182	0.017 (0.036)	0.817	0.008 (0.014)	0.545
β_3	0.032 (0.137)	0.958	-0.009 (0.054)	0.379	0.039 (0.019)	0.047

Compare 2 treatments $\Rightarrow$ Marginal parameters that is, irrespective of the pattern

Estimations in each pattern are **weighted** by the probability of dropout in each pattern:

$$\hat{\beta} = \pi^{(1,2)}\hat{\beta}^{(1,2)} + \pi^{(3,4)}\hat{\beta}^{(3,4)} + \pi^{(5)}\hat{\beta}^{(5)}$$

$\hookrightarrow$ **Extrapolation** = the score has the same trajectory after the dropout, that is, at times for which there is no available data.



Pattern Mixture Model: Results

Slope and interaction effect in each pattern

	$D = 1, 2$		$D = 3, 4$		$D = 5$		Pattern Mixture Model	
	Est. (SE)	p	Est. (SE)	p	Est. (SE)	p	Est. (SE)	p
β_1	-0.136 (0.108)	0.182	0.017 (0.036)	0.817	0.008 (0.014)	0.545	-0.038 (0.038)	0.318
β_3	0.032 (0.137)	0.958	-0.009 (0.054)	0.379	0.039 (0.019)	0.047	0.022 (0.049)	0.649

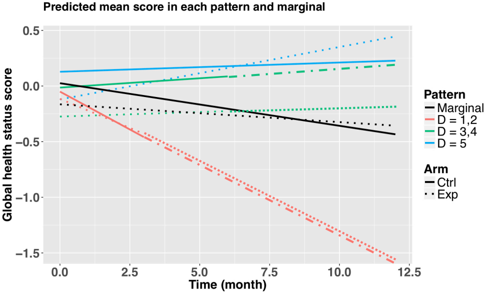
Compare 2 treatments ⇒ Marginal parameters that is, irrespective of the pattern

Estimations in each pattern are **weighted** by the probability of dropout in each pattern:

$$\hat{\beta} = \pi^{(1,2)}\hat{\beta}^{(1,2)} + \pi^{(3,4)}\hat{\beta}^{(3,4)} + \pi^{(5)}\hat{\beta}^{(5)}$$

↪ **Extrapolation** = the score has the same trajectory after the dropout, that is, at times for which there is no available data.

Pattern Mixture Model



Selection Model: Definition

Definition

$Y_i = (Y_i^{\text{mis}}, Y_i^{\text{obs}})$

$$f_{\theta\psi}(Y_i, D_i) = f_\theta(Y_i) \times f_\psi(D_i \mid Y_i)$$

$\Rightarrow$ the dropout D_i depends directly on the score Y_i observed

$\hookrightarrow$ **To test** the missingness mechanism

Selection Model: Definition

Definition

$Y_i = (Y_i^{\text{mis}}, Y_i^{\text{obs}})$

$$f_{\theta\psi}(Y_i, D_i) = f_\theta(Y_i) \times f_\psi(D_i \mid Y_i)$$
$$= \int f_\theta(Y_i) \times f_\psi(D_i \mid Y_i) \, dY_i^{\text{mis}}$$

$\Rightarrow$ the dropout D_i depends directly on the score Y_i observed and unobserved.

$\hookrightarrow$ **To test** the missingness mechanism

Selection Model: Definition

Definition

$Y_i = (Y_i^{\text{mis}}, Y_i^{\text{obs}})$

$$f_{\theta\psi}(Y_i, D_i) = f_\theta(Y_i) \times f_\psi(D_i \mid Y_i)$$
$$= \int f_\theta(Y_i) \times f_\psi(D_i \mid Y_i)\, dY_i^{\text{mis}}$$

$\Rightarrow$ the dropout D_i depends directly on the score Y_i observed and unobserved.

Model of Diggle and Kenward (Diggle & Kenward, *Applied statistics*, 1994) :

$\hookrightarrow$ **To test** the missingness mechanism

Selection Model: Definition

Definition

$Y_i = (Y_i^{\text{mis}}, Y_i^{\text{obs}})$

$$\begin{aligned} f_{\theta\psi}(Y_i, D_i) &= f_\theta(Y_i) \times f_\psi(D_i \mid Y_i) \\ &= \int f_\theta(Y_i) \times f_\psi(D_i \mid Y_i)\, dY_i^{\text{mis}} \end{aligned}$$

$\Rightarrow$ the dropout D_i depends directly on the score Y_i observed and unobserved.

Model of Diggle and Kenward (Diggle & Kenward, *Applied statistics*, 1994) :

- Mixed model for $f(Y_i)$:

$$Y_{ij} = \beta^T X_{ij} + b_i{}^T Z_{ij} + \varepsilon_{ij}$$

$\hookrightarrow$ **To test** the missingness mechanism

Selection Model: Definition

Definition

$Y_i = (Y_i^{\text{mis}}, Y_i^{\text{obs}})$

$$f_{\theta\psi}(Y_i, D_i) = f_\theta(Y_i) \times f_\psi(D_i \mid Y_i)$$
$$= \int f_\theta(Y_i) \times f_\psi(D_i \mid Y_i)\, dY_i^{\text{mis}}$$

$\Rightarrow$ the dropout D_i depends directly on the score Y_i observed and unobserved.

Model of Diggle and Kenward (Diggle & Kenward, *Applied statistics*, 1994) :

- Mixed model for $f(Y_i)$:

$$Y_{ij} = \beta^T X_{ij} + b_i{}^T Z_{ij} + \varepsilon_{ij}$$

- Logistic regression for $f(D_i \mid Y_i)$: The conditionnal probability for dropout at occasion j is:

$$\text{logit}\, \mathbb{P}(D_i = j \mid D_i \geq j, Y_{ij}, Y_{i\,j+1}) = \psi_0 + \psi_1 Y_{ij} + \psi_2 Y_{i\,j+1}$$

  where $Y_{i\,j+1}$ is the unobserved current score

$\hookrightarrow$ **To test** the missingness mechanism

Selection Model: Results

Testing the ψ parameters ($\mathcal{H}_0 : \psi = 0$)

$$\text{logit}\ \mathbb{P}(D_i = j \mid D_i \geq j, Y_{ij}, Y_{i\ j+1}) = \psi_0 + \psi_1 Y_{ij} + \psi_2 Y_{i\ j+1}$$

Selection Model: Results

Testing the ψ parameters ($\mathcal{H}_0 : \psi = 0$)

$$\text{logit}\, \mathbb{P}(D_i = j \mid D_i \geq j, Y_{ij}, Y_{i\,j+1}) = \psi_0 + \psi_1 Y_{ij} + \psi_2 Y_{i\,j+1}$$

The ψ parameters model the dropout process under a MNAR assumption

Selection Model: Results

Testing the ψ parameters ($\mathcal{H}_0 : \psi = 0$)

$$\text{logit}\ \mathbb{P}(D_i = j \mid D_i \geq j, Y_{ij}, Y_{i\ j+1}) = \psi_0 + \psi_1 Y_{ij} + \psi_2 Y_{i\ j+1}$$

The ψ parameters model the dropout process under a MNAR assumption

MAR case if $\psi_2 = 0$: the dropout depends only on the observed score

Selection Model: Results

Testing the ψ parameters ($\mathcal{H}_0 : \psi = 0$)

$$\text{logit}\,\mathbb{P}(D_i = j \mid D_i \geq j, Y_{ij}, Y_{i\,j+1}) = \psi_0 + \psi_1 Y_{ij} + \psi_2 Y_{i\,j+1}$$

The ψ parameters model the dropout process under a MNAR assumption

MAR case if $\psi_2 = 0$: the dropout depends only on the observed score

MCAR case if $\psi_2 = \psi_1 = 0$: the dropout is independent of the score

Selection Model: Results

Testing the ψ parameters ($\mathcal{H}_0 : \psi = 0$)

$$\text{logit}\, \mathbb{P}(D_i = j \mid D_i \geq j, Y_{ij}, Y_{i\,j+1}) = \psi_0 + \psi_1 Y_{ij} + \psi_2 Y_{i\,j+1}$$

The ψ parameters model the dropout process under a MNAR assumption

MAR case if $\psi_2 = 0$: the dropout depends only on the observed score

MCAR case if $\psi_2 = \psi_1 = 0$: the dropout is independent of the score

	Selection Model		
	Estimation	SE	p
ψ_0	-0.063	0.072	0.383
ψ_1	0.001	0.087	0.990
ψ_2	0.001	0.117	0.992
β_1	0.019	0.027	0.486
β_3	0.035	0.032	0.266

- The assumption of **completely random dropout** is acceptable testing the ψ parameters.

Summarising the results of the 3 models

	Selection Model			Linear Mixed Model			Pattern Mixture Model		
	Estimation	SE	p	Estimation	SE	p	Estimation	SE	p
β_1	0.019	0.027	0.486	0.012	0.012	0.307	-0.038	0.038	0.318
β_3	0.035	0.032	0.266	0.031	0.017	0.076	0.022	0.049	0.649
ψ_0	-0.063	0.072	0.383	–	–	–	–	–	–
ψ_1	0.001	0.087	0.990	–	–	–	–	–	–
ψ_2	0.001	0.117	0.992	–	–	–	–	–	–

In comparison with the Linear Mixed Model:

- Selection Model:
  $\hookrightarrow$ Estimations of the fixed effects are **similar** to the Linear Mixed Model
  $\hookrightarrow$ Missing data are non **informative** according to the test of the ψ parameters
- Pattern Mixture Model:
  $\hookrightarrow$ **Differences**: negative slope and a trend to a deterioration
  $\hookrightarrow$ Caused by the **extrapolation** step

These conclusions are valid only if the model assumptions are valid.

Conclusion

All models for non-informative data require the analyst to make **strong assumptions**:

Selection Model	Pattern Mixture Model
Model directly the quantity of interest (the marginal score)	To obtain the marginal score ⇒ extrapolation of linear trends
Test missingness mechanism (MNAR/MAR/MCAR)	Description of the HRQoL according to the dropout pattern
Untestable = **normality assumption** ↪ score observed and *unobserved* is assumed to be normally distributed	Untestable = **extrapolation** ↪ quality of life of a patient evolves in the same way after dropout

Both of theses methods assume **untestable assumptions**