Joint Modeling of (Un)bounded Longitudinal Markers, Competing Risks, and Recurrent Events in Cystic Fibrosis Data

International Society for Clinical Biostatistics

Eleni-Rosalina Andrinopoulou

PM Afonso, D Rizopoulos, AK Palipana, E Gecili, C Brokamp, JP Clancy, RD Szczesniak

July, 2024

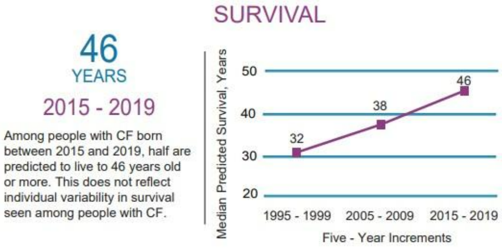
www.erandrinopoulou.com e.andrinopoulou@erasmusmc.nl @ERandrinopoulou

Introduction

Introduction: Motivation

Cystic Fibrosis

- → Genetic disorder affecting the lungs, pancreas, and other organs
- → 40,000 children and adults living with CF in the US
- → $>$ 75 percent of people with CF are diagnosed by age 2



Survival statistics for the years 2015 through 2019.

Introduction: Motivation

What to expect?

- ➔ Chronic respiratory problems $\rightarrow$ lung infections
- ➔ Poor growth $\rightarrow$ low weight
- ➔ Increased risk of death and lung transplantation

Introduction: Motivation

What to expect?

- ➔ Chronic respiratory problems $\rightarrow$ lung infections
- ➔ Poor growth $\rightarrow$ low weight
- ➔ Increased risk of death and lung transplantation

US Cystic Fibrosis Registry

- ◇ Baseline characteristics: Sex, F508del, SESlow, Enzymes
- ◇ Biomarkers: Lung function decline (ppFEV$_1$)
- ◇ Nutritional status: BMI
- ◇ Survival: Pulmonary exacerbations, death or lung transplantation

Introduction: Motivation

What to expect?

- ➔ Chronic respiratory problems $\rightarrow$ lung infections
- ➔ Poor growth $\rightarrow$ low weight
- ➔ Increased risk of death and lung transplantation

Incorporating all information could improve decisions regarding the monitoring and treatment strategies of the patients

Introduction: Research question

→ **How ppFEV$_1$ and BMI relate to the risk of recurrent pulmonary exacerbations?**

→ **How ppFEV$_1$ and BMI relate to the competing risks of death and transplantation?**

→ **Are pulmonary exacerbations related to the competing risk of death and transplantation?**

This research is supported by the National Institutes of Health / National Heart, Lung and Blood Institute (grant R01 HL141286)

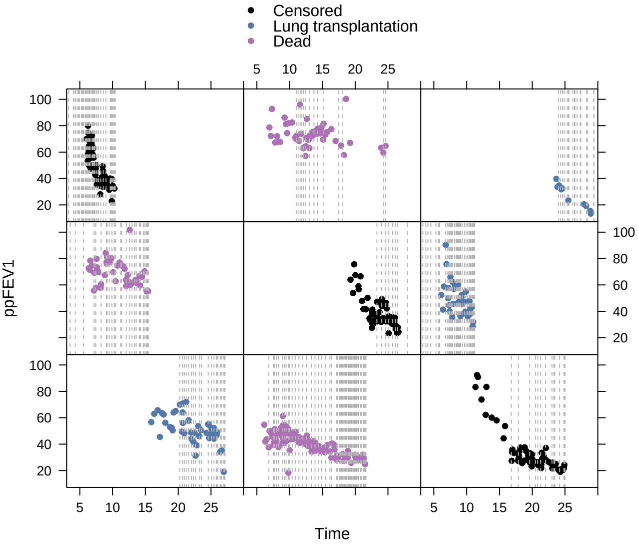
www.erandrinopoulou.com e.andrinopoulou@erasmusmc.nl @ERandrinopoulou

Introduction: Descriptive statistics

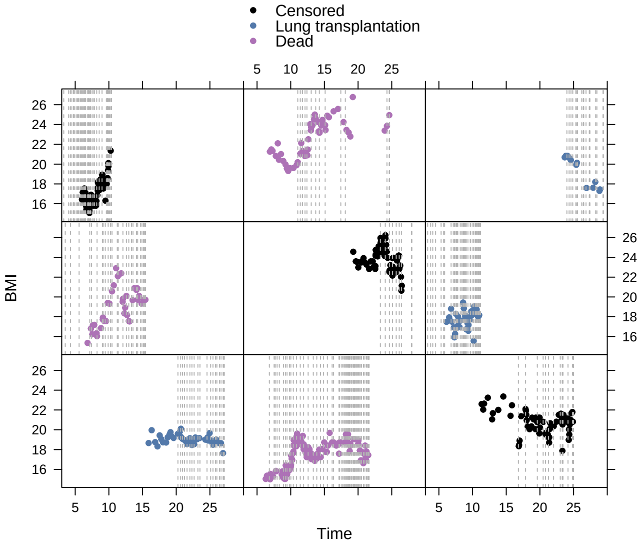
US Cystic Fibrosis Registry

- ➔ >23,000 patients
- ➔ >1,400,000 observations
- ➔ on average >10 years of follow-up
- ➔ 11% lung transplantation
- ➔ 18% died

Introduction: Descriptive statistics



Introduction: Descriptive statistics



Introduction: Challenges

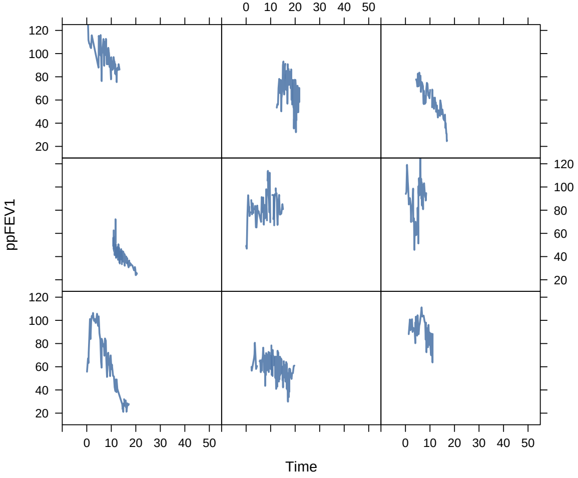
- **→ High-dimensional data**
- **→ Complex data**
  - ◇ Multiple longitudinal data
  - ◇ Recurrent events
  - ◇ Competing risks

Introduction: Challenges

Erasmus MC
University Medical Center Rotterdam

- → **High-dimensional data**
- → **Complex models**
  - ◇ Multiple longitudinal data
    - ◇ Bounded biomarkers
  - ◇ Competing risks
  - ◇ Recurrent events

Introduction: Challenges



Methods

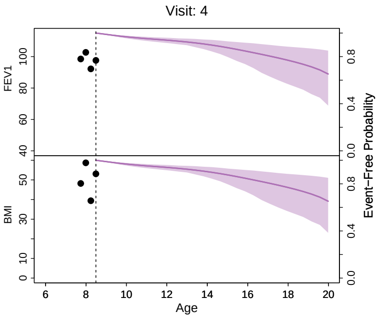
Methods: Joint Models

Common practice

➔ Separate/simplified analysis

- *FEV_1*
- *BMI*
- *Time-to-first exacerbation*

Andrinopoulou, E. R., Clancy, J. P., & Szczesniak, R. D. (2020). Multivariate joint modeling to identify markers of growth and lung function decline that predict cystic fibrosis pulmonary exacerbation onset. BMC pulmonary medicine, 20, 1-11.



Methods: Joint Models

Incorporating all information could improve decisions regarding the monitoring and treatment strategies of the patients

Methods: Joint Models

Longitudinal submodels

- ➔ ppFEV$_1$
- ➔ BMI

$$g_j[E\{Y_{ji}(t) \mid \boldsymbol{b_{ji}}\}] = \boldsymbol{x}_{ji}^{\top}(t)\beta_j + \boldsymbol{z}_{ji}^{\top}(t)\boldsymbol{b_{ji}} = \eta_{ji}(t),$$

Methods: Joint Models

Longitudinal submodels

- ➔ ppFEV$_1$
- ➔ BMI

$$g_j[E\{Y_{ji}(t) \mid \boldsymbol{b_{ji}}\}] = \boldsymbol{x}_{ji}^\top(t)\beta_j + \boldsymbol{z}_{ji}^\top(t)\boldsymbol{b_{ji}} = \eta_{ji}(t),$$

where

- ◇ $\boldsymbol{x}_{ji}^\top(t)\beta_j$ fixed effects
- ◇ $\boldsymbol{z}_{ji}^\top(t)\boldsymbol{b_{ji}}$ random effects
- ◇ $g_j[.]$ link function

Methods: Joint Models

Longitudinal submodels

- ➔ ppFEV$_1$
- ➔ BMI

$$g_j[E\{Y_{ji}(t) \mid \boldsymbol{b_{ji}}\}] = \boldsymbol{x}_{ji}^\top(t)\beta_j + \boldsymbol{z}_{ji}^\top(t)\boldsymbol{b_{ji}} = \eta_{ji}(t),$$

where

- ◇ $\boldsymbol{x}_{ji}^\top(t)\beta_j$ fixed effects
- ◇ $\boldsymbol{z}_{ji}^\top(t)\boldsymbol{b_{ji}}$ random effects
- ◇ $g_j[.]$ link function

  identity for the unbounded outcome

  logit for the bounded outcome

Methods: Joint Models

Survival submodels

➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w_i^R}^\top(t)\boldsymbol{\gamma}^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + \upsilon_i^R\right]$$

➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w_i^C}^\top(t)\boldsymbol{\gamma}_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C\{\eta_{ji}(t)\}\alpha_{kjm}^C + \upsilon_{ki}^C\right]$$

Methods: Joint Models

Survival submodel

- ➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w_i^R}^\top(t)\boldsymbol{\gamma}^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + v_i^R\right]$$

- ➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w_i^C}^\top(t)\gamma_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C\{\eta_{ji}(t)\}\alpha_{kjm}^C + v_{ki}^C\right]$$

where

- ◇ $h_0^R(t - t_{0_{li}})$ baseline hazard
- ◇ $t_{0_{li}}$ starting time of the risk interval for the lth recurrent event

Methods: Joint Models

Survival submodel

➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w_i^R}^\top(t)\gamma^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + v_i^R\right]$$

➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w_i^C}^\top(t)\gamma_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C\{\eta_{ji}(t)\}\alpha_{kjm}^C + v_{ki}^C\right]$$

where

- ◇ $\boldsymbol{w_i^R}^\top(t)$ baseline or time-varying covariates
- ◇ γ^R regression coefficients

Methods: Joint Models

Survival submodel

➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w_i^R}^\top(t)\boldsymbol{\gamma}^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + v_i^R\right]$$

➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w_i^C}^\top(t)\boldsymbol{\gamma}_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C\{\eta_{ji}(t)\}\alpha_{kjm}^C + v_{ki}^C\right]$$

where

- ◇ $H_{jm}^R\{\eta_{ji}(t)\}$ functional forms of the longitudinal outcomes
- ◇ α_{jm}^R association between longitudinal and recurrent events

Methods: Joint Models

Survival submodel

➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w_i^R}^\top(t)\gamma^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + \upsilon_i^R\right]$$

➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w_i^C}^\top(t)\gamma_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C\{\eta_{ji}(t)\}\alpha_{kjm}^C + \upsilon_{ki}^C\right]$$

where

◇ υ_i^R frailty term

Methods: Joint Models

Survival submodel

➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w_i^R}^\top(t)\gamma^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + \upsilon_i^R\right]$$

➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w_i^C}^\top(t)\boldsymbol{\gamma}_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C\{\eta_{ji}(t)\}\alpha_{kjm}^C + \upsilon_{ki}^C\right]$$

where

◇ $h_{0k}^C(t)$ cause-specific baseline hazard

Methods: Joint Models

Survival submodel

➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w_i^R}^\top(t)\gamma^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + v_i^R\right]$$

➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w_i^C}^\top(t)\gamma_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C\{\eta_{ji}(t)\}\alpha_{kjm}^C + v_{ki}^C\right]$$

where

- ◇ $\boldsymbol{w_i^C}^\top(t)$ baseline or time-varying covariates
- ◇ ${\gamma_k}^C$ regression coefficients

Methods: Joint Models

Survival submodel

➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w_i^R}^\top(t)\gamma^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + v_i^R \right]$$

➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w_i^C}^\top(t)\boldsymbol{\gamma}_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C\{\eta_{ji}(t)\}\alpha_{kjm}^C + v_{ki}^C \right]$$

where

- $H_{kjm}^C\{\eta_{ji}(t)\}$ functional forms of the longitudinal outcomes
- α_{kjm}^C association between longitudinal and the competing events

Methods: Joint Models

Survival submodel

➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w}_{\boldsymbol{i}}^{\boldsymbol{R}^\top}(t)\gamma^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + \upsilon_i^R\right]$$

➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w}_{\boldsymbol{i}}^{\boldsymbol{C}^\top}(t)\boldsymbol{\gamma}_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C\{\eta_{ji}(t)\}\alpha_{kjm}^C + \upsilon_{ki}^C\right]$$

where

◇ υ_{ki}^C frailty term

◇ $\upsilon_{ki}^C = \upsilon_i^R \alpha_k^\upsilon$

www.erandrinopoulou.com e.andrinopoulou@erasmusmc.nl @ERandrinopoulou

Methods: Joint Models

Challenges and opportunities: association

$$g_j[E\{Y_{ji}(t) \mid \boldsymbol{b_{ji}}\}] = \boldsymbol{x}_{ji}^\top(t)\beta_j + \boldsymbol{z}_{ji}(t)^\top \boldsymbol{b_{ji}} = \eta_{ji}(t)$$

When $g_j[.] \neq$ identity function

- $g^{-1}\{\eta_{ji}(t)\}$
- Beta: logit link $\rightarrow$ expit function

Methods: Joint Models

Challenges and opportunities: association

$$g_j[E\{Y_{ji}(t) \mid \boldsymbol{b_{ji}}\}] = \boldsymbol{x}_{ji}^\top(t)\beta_j + \boldsymbol{z}_{ji}(t)^\top \boldsymbol{b_{ji}} = \eta_{ji}(t)$$

When $g_j[.] \neq$ identity function

- $g^{-1}\{\eta_{ji}(t)\}$
- Beta: logit link $\rightarrow$ expit function

➔ Recurrent event times

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp\left[\boldsymbol{w_i^R}^\top(t)\gamma^R + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{jm}^R[g_j^{-1}\{\eta_{ji}(t)\}]\alpha_{jm}^R + v_i^R\right]$$

➔ Competing risks

$$h_{ki}^C(t) = h_{0k}^C(t) \exp\left[\boldsymbol{w_i^C}^\top(t)\gamma_k^C + \sum_{j=1}^{J}\sum_{m=1}^{M_j} H_{kjm}^C[g_j^{-1}\{\eta_{ji}(t)\}]\alpha_{kjm}^C + v_{ki}^C\right]$$

www.erandrinopoulou.com e.andrinopoulou@erasmusmc.nl @ERandrinopoulou

Methods: Joint Models

Challenges and opportunities: recurrent event time

$$h_i^R(t) = h_0^R(t - t_{0_{li}}) \exp \left[\boldsymbol{w}_i^{\boldsymbol{R}\top}(t)\gamma^R + \sum_{j=1}^{J} \sum_{m=1}^{M_j} H_{jm}^R\{\eta_{ji}(t)\}\alpha_{jm}^R + v_i^R \right]$$

Calendar vs gap time

- ➔ the calendar timescale uses a shared reference time for all events (e.g., study entry), $t_{0_{li}} = 0$
- ➔ the gap timescale uses the end of the previous event, assuming a renewal after each event and resetting the time to zero
- ➔ non-risk periods in which a patient is still experiencing the previous event

Application

Application: CF

Model specification

➔ **ppFEV$_1$**

- sex, birth cohort, genotype, enthicity
- percentage of green space, average annual truck, deprivation index

➔ **BMI**

- sex, birth cohort, genotype, enthicity
- deprivation index
- enzyme intake

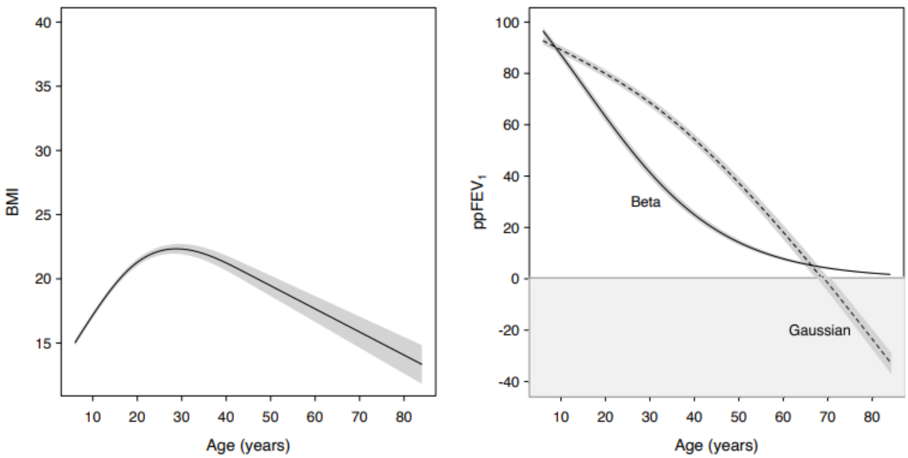
Application: CF

Model specification

- ➔ **Recurent event**
  - ◇ number of previous PEx events
  - ◇ ppFEV_1's value, standardized cumulative effect of BMI's underlying value
  - ◇ gap time scale
- ➔ **Lung transplantation/death**
  - ◇ sex, birth cohort, genotype, enthicity
  - ◇ ppFEV_1 value and rate of change, standardized cumulative effect of BMI's underlying value

Application: CF

Results: longitudinal outcomes



Application: CF

Results: association parameters

	PEx	*Transplantation*	*Death*
1-unit ppFEV$_1$ value (↑)	-3.8% (95%CI -3.9 to -3.8)	-17% (95%CI -17.5 to 16.5)	-11.6% (95%CI -11.8 to -11.3)
1-unit ppFEV$_1$ slope (↑) *(less steep)*	-	-13.7% (95%CI -16.1 to -10.9)	-9.1% (95%CI -10.8 to -7.5)
1-unit BMI area (↑)	0.04% (95%CI 0.037 to 0.042)	6% (95%CI 4.4 to 7.6)	7.1% (95%CI 5.4 to 8.7)

Simulation

Simulation: Set-up

- **Simulate**
  - Beta (bounded outcome): underlying value, transformed in original scale
  - terminal event: baseline covariate
- **Fit**
  - Beta (bounded outcome): underlying value, transformed in original scale
  - terminal event: baseline covariate

Simulation: Set-up

➔ **Simulate**

- ◇ Beta (bounded outcome): underlying value, transformed in original scale
- ◇ terminal event: baseline covariate

➔ **Fit**

- ◇ Gaussian (unbounded outcome): underlying value
- ◇ terminal event: baseline covariate

Simulation: Results

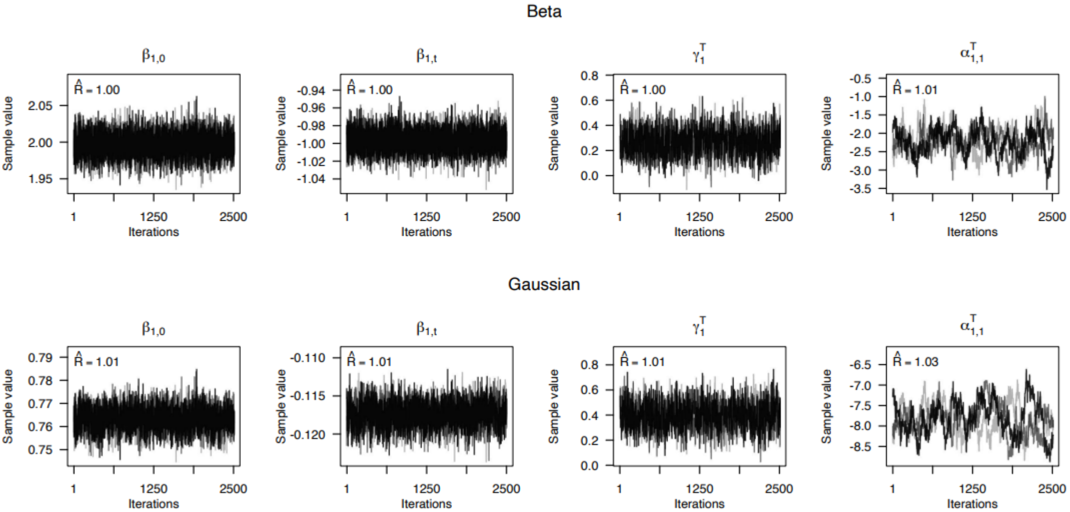
Parameters:

	True parameter (Beta)	*Correclty specified model (Beta)*	*Misspecified model (Gaussian)*
β_1	2.00	1.999	0.765
β_2	-1.00	-0.999	-0.119
γ	0.25	0.246	0.214
α	-2.00	-2.066	-7.870

www.erandrinopoulou.com e.andrinopoulou@erasmusmc.nl @ERandrinopoulou

Simulation: Results

Convergence:



www.erandrinopoulou.com e.andrinopoulou@erasmusmc.nl @ERandrinopoulou

Conclusion

Conclusion

Extended Joint Model

- ➔ multiple (un)bounded longitudinal outcomes
- ➔ recurrent events
  - ◇ gap and calendar time scales
- ➔ competing risks
- ➔ different functional forms

www.erandrinopoulou.com e.andrinopoulou@erasmusmc.nl @ERandrinopoulou

Conclusion

Extended Joint Model

- ➔ multiple (un)bounded longitudinal outcomes
- ➔ recurrent events
  - ◇ gap and calendar time scales
- ➔ competing risks
- ➔ different functional forms
- ➔ **Software:** `JMbayes2`

drizopoulos.github.io/JMbayes2/

More details:

Afonso PM, Rizopoulos D, Palipana AK, Gecili E, Brokamp C, Clancy JP, Szczesniak RD, Andrinopoulou ER. A joint model for (un) bounded longitudinal markers, competing risks, and recurrent events using patient registry data. arXiv preprint arXiv:2405.16492. 2024 May 26.

Thank you for your attention!

p.mirandaafonso@erasmusmc.nl
e.andrinopoulou@erasmusmc.nl