

# Multilevel Bayesian Joint Model in Hierarchically Structured Data

Chen (Grace) Zhou

University of Cincinnati

June 7, 2022

# Outline

1 Introduction

## 2 Multilevel Bayesian Joint Model of Longitudinal and Binary Outcomes

- ▶ Motivation
  - ▶ Model framework
  - ▶ Simulation study & Motivating data
  - ▶ Discussion

### 3 Multilevel Bayesian Joint Model of Longitudinal and Recurrent Outcomes

- ▶ Motivation
  - ▶ Model framework
  - ▶ Simulation study & Motivating data
  - ▶ Discussion

## 4 Conclusion and Future Work



## 1 Introduction

### 2 Multilevel Bayesian Joint Model of Longitudinal and Binary Outcomes

- ▶ Motivation
- ▶ Model framework
- ▶ Simulation study & Motivating data
- ▶ Discussion

### 3 Multilevel Bayesian Joint Model of Longitudinal and Recurrent Outcomes

- ▶ Motivation
- ▶ Model framework
- ▶ Simulation study & Motivating data
- ▶ Discussion

### 4 Conclusion and Future Work

# Clinical Outcomes

- **Explicit outcomes**

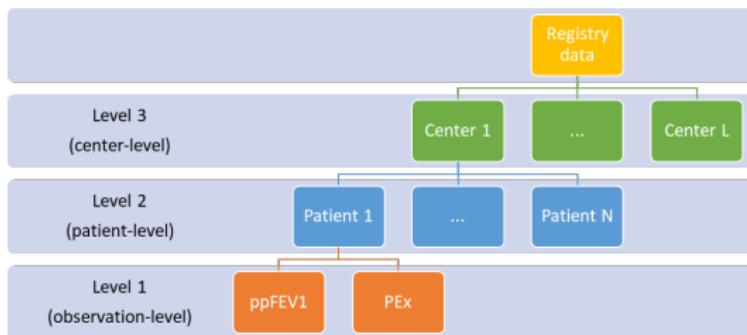
- ▶ longitudinal biomarker (e.g., blood pressure, BMI)
- ▶ time-to-event outcome (e.g., death, relapse)

- **Implicit outcomes**

- ▶ missing data (e.g., dropout)
- ▶ random visit times

## Motivating Data

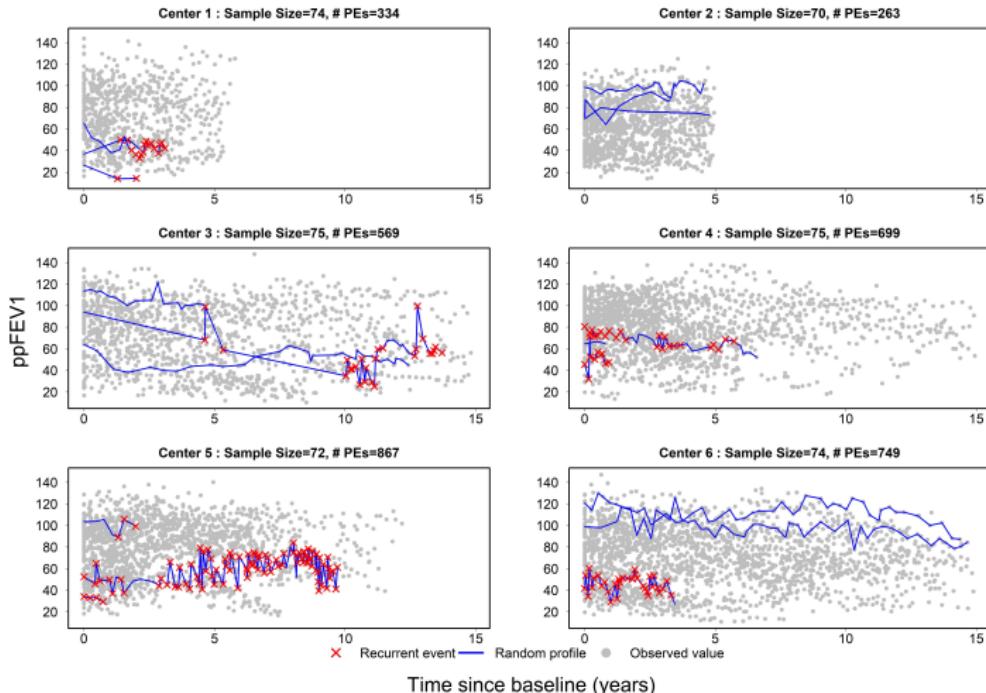
- Lung disease (cystic fibrosis) study
    - ▶ ppFEV1<sup>1</sup>: longitudinal continuous biomarker
    - ▶ PEx<sup>2</sup>: repeated binary event
  - Irregular registry data



<sup>1</sup>percent predicted forced expiratory volume in 1 second

<sup>2</sup>pulmonary exacerbation

## Motivating Data (cont'd)



# Classical Analysis

- Longitudinal outcomes

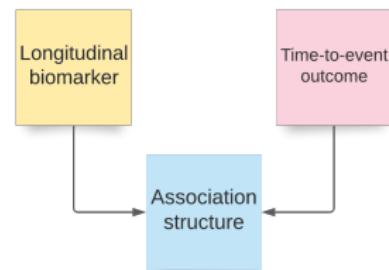
- ▶ Linear mixed effects (LME)
- ▶ Generalized linear mixed model (GLMM)
- ▶ Generalized estimating equation (GEE)
- ▶ marginal model
- ▶ ...

- Survival outcomes

- ▶ Relative risk model (e.g., Cox model)
- ▶ Accelerated failure time model
- ▶ Cure frailty model
- ▶ ...

## Joint Model (JM)

- How strong is the association between ppFEV1 and the risk of PEx?
  - Joint modeling (Rizopoulos (2011))
    - ▶ If ppFEV1 is endogenous
    - ▶ If ppFEV1's dropout is nonrandom



### Rationale of a standard joint model

## Extensions

- Categorical longitudinal outcomes
  - Multiple longitudinal outcomes
  - Multiple failure times
    - ▶ Competing risks
    - ▶ Recurrent events
    - ▶ Multi-State Process
  - Heterogenous population
    - ▶ Stratified relative risk model
    - ▶ Latent class joint model

# Estimation & Computing

- Frequentist approach
  - **Bayesian approach**

## Summary of R package implementation

R package	Method	Bayesian	Reference
joineR	Expectation Maximization algorithm	✗	Philipson et al. (2018)
joineRML	Monte Carlo Expectation Maximization algorithm	✗	Hickey et al. (2018)
JM	Maximum Likelihood Estimation	✗	Rizopoulos (2010)
lcmm	Maximum Likelihood Estimation	✗	Proust-Lima et al. (2022)
frailtypack	Maximum Penalized Likelihood Estimation	✗	Rondeau et al. (2012) Rondeau et al. (2019)
JMBayes	Monte Carlo Markov Chain (MCMC)	✓	Rizopoulos (2016)
JMBayes2	MCMC	✓	Rizopoulos et al. (2022)
rstanarm*	Hamiltonian Monte Carlo	✓	Brilleman et al. (2018)

\* function stan\_jm()

# Hamiltonian Monte Carlo (HMC) & R/Stan

- HMC (Betancourt and Girolami (2013))
  - ▶ Sophisticated but novel MCMC technique
  - ▶ Generate transitions spanning the full marginal variance
  - ▶ Eliminate random walk behavior
  - ▶ Provide efficient exploration for complex hierarchical models
- Stan (not an acronym; Stan Development Team (2019))
  - ▶ Probabilistic programming language
  - ▶ Releases computational constraints of Euclidean HMC
  - ▶ Adapts to No-U-Turn Sampler
  - ▶ R interfaces to Stan
    - **CmdStanr** (Gabry and Cešnovar (2022))
    - **RStan** (Stan Development Team (2020))

## 1 Introduction

### 2 Multilevel Bayesian Joint Model of Longitudinal and Binary Outcomes

- ▶ Motivation
- ▶ Model framework
- ▶ Simulation study & Motivating data
- ▶ Discussion

### 3 Multilevel Bayesian Joint Model of Longitudinal and Recurrent Outcomes

- ▶ Motivation
- ▶ Model framework
- ▶ Simulation study & Motivating data
- ▶ Discussion

### 4 Conclusion and Future Work



# Motivation

- Flexible link vs. common links
- Non-hierarchical joint model biases
- Prognostic utility of the optimal joint model

## Model Framework

Let  $l$ ,  $i$ ,  $j$  denote center, patient, time point with  $l = 1, \dots, L; i = 1, \dots, n_l; j = 1, \dots, n_{li}$ , respectively.

- **Longitudinal continuous submodel**

$$\text{LME}^3: Y_{lij} = \mathbf{X}_{li}(t_{lij})\boldsymbol{\alpha} + b_l + U_{li} + W_{li}(t_{lij}) + \epsilon_{lij} \quad (1)$$

where  $b_l \sim N(0, \sigma_b^2)$ ,  $U_{li} \sim N(0, \sigma_u^2)$ ,  $\epsilon_{lij} \sim N(0, \sigma^2)$ ,  $\mathbf{b} \perp \mathbf{U} \perp \epsilon$ .

The stochastic Gaussian process  $\mathbf{W}_{lj}(t)$  follows  $N(\mathbf{0}, \Sigma_W)$  with exponential covariance

$$Cov(W_{li}(t), W_{li}(s)) = \tau^2 \exp(-|t-s| \cdot \rho) \quad (2)$$

- **Longitudinal binary submodel**

$$\text{GLMM}^4 : Pr(R_{lij} = 1) = g^{-1}(\mathbf{V}_{li}(t_{lij})\boldsymbol{\beta} + \rho_1 b_l + \rho_2 U_{li}; r_l) \quad (3)$$

where  $\rho_1, \rho_2 \in (-1, 1)$ ,  $g(\cdot)$  is some known link function.

### 3 Linear Mixed Effect

#### **4 Generalized Linear Mixed Model**

## Symmetric Power Link Family (Jiang et al. (2013))

$$F_{sp}(x; r) = F_0^r\left(\frac{x}{r}\right)I_{(0,1]}(r) + \left[1 - F_0^{1/r}(-rx)\right]I_{(1,+\infty)}(r), \quad (4)$$

where covariate  $x \in (-\infty, +\infty)$ , power parameter  $r \in (0, +\infty)$ ,  $I_A(r) = 1$ , when  $r \in A$  or  $I_A(r) = 0$ , otherwise.

- symmetric power logit ( $F_{splogit}$ )

$$F_0 = F_{\text{logistic}}(x|\mu=0, s=1) = \frac{1}{1 + \exp(-x)}$$

- symmetric power exponential power ( $F_{ssep}$ )

$$F_0 = F_{\text{laplace}}(x|\mu=0, b=1) = \frac{\exp(x)}{2} I_{(-\infty, 0)}(x) + \left(1 - \frac{\exp(x)}{2}\right) I_{[0, +\infty)}(x)$$

# Conditional Independence Assumption

Random effects explain all the interdependence:

- Repeated measurements/events in each submodel are independent of each other
- Two submodels are conditionally independent

## Joint Posterior Distribution

$$\begin{aligned}
\pi(\Psi, \mathbf{b}, \mathbf{U}, \mathbf{W} | \mathcal{D}) &\propto \pi(\mathcal{D}, \mathbf{b}, \mathbf{U}, \mathbf{W} | \Psi) \pi(\Psi) \\
&\propto \pi(\mathcal{D} | \mathbf{b}, \mathbf{U}, \mathbf{W}, \Psi) \pi(\mathbf{b} | \mathbf{U}, \mathbf{W}, \Psi) \pi(\mathbf{U} | \mathbf{W}, \Psi) \pi(\mathbf{W} | \Psi) \pi(\Psi) \\
&\propto \pi(\mathbf{Y}, \mathbf{R} | \mathbf{b}, \mathbf{U}, \mathbf{W}, \Psi) \pi(\mathbf{b} | \Psi) \pi(\mathbf{U} | \Psi) \pi(\mathbf{W} | \Psi) \pi(\Psi) \\
&\propto \prod_{l=1}^L \prod_{i=1}^{n_l} I(Y_{li}, R_{li} | b_l, U_{li}, W_{li}, \Psi) \pi(b_l | \sigma_b) \pi(U_{li} | \sigma_u) \pi(W_{li} | \tau, \rho) \pi(\Psi)
\end{aligned} \tag{5}$$

where  $\mathcal{D}$  denotes observed data,  $\Psi$  represents unknown mutually independent parameters. By the assumption of conditional independence:

$$I(\mathbf{Y}_{li}, \mathbf{R}_{li} | b_l, U_{li}, \mathbf{W}_{li}, \boldsymbol{\Psi}) = I_1(\mathbf{Y}_{li} | b_l, U_{li}, \mathbf{W}_{li}, \boldsymbol{\Psi}) I_2(\mathbf{R}_{li} | b_l, U_{li}, \boldsymbol{\Psi})$$

# Prior Distributions

## Prior specifications

Priors	Simulation Study 1	Simulation Study 2	Motivating Data
$\alpha, \beta$	$N(0, 10I)$	$N(0, 100I)$	$N(0, 100I)$
$\sigma_b$	Half-Cauchy(0, 5)	Half-Cauchy(0, 5)	Half-Cauchy(0, 5)
$\sigma_u$	Truncated student-t(1, 0, 5)	Half-Cauchy(0,5)	Half-Cauchy(0,5)
$\sigma$	Half-Cauchy(0, $2.5 \cdot sd$ )	Half-Cauchy(0, $2.5 \cdot sd$ )	Half-Cauchy(0, $2.5 \cdot sd$ )
$\rho_1, \rho_2$	Uniform(-1, 1)	Uniform(-1, 1)	Uniform(-1, 1)
$r$	Exponential(1)	Gamma(2, 2)	Exponential(1)
$\tau$	-	Truncated $N(0, 5)$	Truncated $N(0, 5)$
$\rho$	-	Inv-Gamma(2, 1)	Inv-Gamma(2, 1)

Note: sd=standard deviation of longitudinal residuals; N(location,scale); Half-Cauchy(location,scale); student-t(df,location,scale); Exponential(rate); Gamma(shape,rate); Inv-Gamma(shape,scale);

## Computation

HMC	Simulation Study 1	Simulation Study 2	Motivating Data
Interface to Stan	cmdstanr	rstan	rstan
Post-warmup +	4000	4000	5000
Chains	2	2	2

<sup>+</sup> Converged samplings validated by  $\hat{R} = 1$  (Gelman and Rubin (1992))

# Dynamic Individual Prediction

- Best linear unbiased predictor (BLUP)
- Predict for a new patient from the existing center

$$E(U_{lj'} | \mathbf{Y}_{lj'}, b_l, \phi, \alpha) = \sigma_u^2 \mathbf{K}_{lj'}^T (\mathbf{V}_{lj'}(\phi))^{-1} (\mathbf{Y}_{lj'} - \mathbf{X}_{lj'} \boldsymbol{\alpha} - b_l \mathbf{K}_{lj'}) \quad (6)$$

$$E(W_{lj'}(t_{ljj}) | \mathbf{Y}_{lj'}, b_l, \phi, \alpha) = \tau^2 \mathbf{F}_{lj'}^{j^T} (\mathbf{V}_{lj'}^j(\phi))^{-1} (\mathbf{Y}_{lj'}^j - \mathbf{X}_{lj'}^j \boldsymbol{\alpha} - b_l \mathbf{K}_{lj'}^j) \quad (7)$$

where  $\mathbf{K}$  is a column vector of ones,  $\mathbf{V}(\phi) = \sigma_u^2 \mathbf{J} + \tau^2 \mathbf{R} + \sigma^2 \mathbf{I}$ ,  $\mathbf{J}$  is all ones matrix,  $\mathbf{R}$  is an exponential correlation matrix,  $\mathbf{I}$  is an identity matrix and  
 $\mathbf{F}^j = (\exp(-|t_1 - t_j|) \cdot \rho, \dots, \exp(-|t_j - t_j|) \cdot \rho)^T$

- Forecast for an existing patient

$$E(W_{lj}(t_{ljj} + u) | \mathbf{Y}_{lj}^j, b_l, U_{lj}, \phi, \alpha) = \tau^2 \mathbf{F}_{lj}^{j,u^T} (\mathbf{G}_{lj}^j(\phi))^{-1} (\mathbf{Y}_{lj}^j - \mathbf{X}_{lj}^j \boldsymbol{\alpha} - b_l \mathbf{K}_{lj}^j - U_{lj} \mathbf{K}_{lj}^j) \quad (8)$$

where  $\mathbf{G}^j(\phi) = \tau^2 \mathbf{R}^j + \sigma^2 \mathbf{I}^j$  denotes variance-covariance matrix of  $\mathbf{Y}^j$  (response history up to an observed time point  $j$ )

# Model Selection

- Watanabe-Akaike information criterion (WAIC, (Watanabe (2010)))

$$\text{WAIC} = -2 \left\{ \sum_{i=1}^n \log(p_{post}(y_i|\theta)) - \sum_{i=1}^n \text{var}(\log(p_{post}(y_i|\theta))) \right\} \quad (9)$$

- Invariant to reparameterizations contrary to DIC<sup>5</sup>(Spiegelhalter et al. (2002))
- Computed by `loo::waic()`
- The smaller the better

---

<sup>5</sup>Deviance Information Criterion



## Simulation Study 1

Compare flexible link (splogit) with common links:

$$\left\{ \begin{array}{l} g_{logit}^{-1}(x) = F_{logit}(x) = [1 + \exp(x)]^{-1} \\ g_{probit}^{-1}(x) = F_{probit}(x) = \Phi(x), \text{ where } \Phi \text{ is standard normal cdf} \\ g_{cloglog}^{-1}(x) = F_{cloglog}(x) = 1 - e^{-e^x} \\ g_{splogit}^{-1}(x, r) = F_{splogit}(x, r) = [1 + \exp(\frac{x}{r})]^{-r} I_{(0,1]}(r) + \{1 - [1 + \exp(-rx)]^{-\frac{1}{r}}\} I_{(1,+\infty)}(r) \end{array} \right.$$

# Simulation 1: Model Structures

$$\text{logit-JM} \left\{ \begin{array}{l} Y_{lij} = \alpha_0 + x_{li1}\alpha_1 + x_{li2}\alpha_2 + b_l + U_{li} + \epsilon_{lij} \\ \Pr(R_{lij} = 1) = F_{\text{logit}}(\beta_0 + \beta_1 t_{lij} + \rho_1 b_l + \rho_2 U_{li}) \end{array} \right.$$

$$\text{probit-JM} \left\{ \begin{array}{l} Y_{lij} = \alpha_0 + x_{li1}\alpha_1 + x_{li2}\alpha_1 + b_l + U_{li} + \epsilon_{lij} \\ \Pr(R_{lij} = 1) = F_{\text{probit}}(\beta_0 + \beta_1 t_{lij} + \rho_1 b_l + \rho_2 U_{li}) \end{array} \right.$$

$$\text{cloglog-JM} \left\{ \begin{array}{l} Y_{lij} = \alpha_0 + x_{li1}\alpha_1 + x_{li2}\alpha_1 + b_l + U_{li} + \epsilon_{lij} \\ \Pr(R_{lij} = 1) = F_{\text{cloglog}}(\beta_0 + \beta_1 t_{lij} + \rho_1 b_l + \rho_2 U_{li}) \end{array} \right.$$

$$\text{splogit-JM} \left\{ \begin{array}{l} Y_{lij} = \alpha_0 + x_{li1}\alpha_1 + x_{li2}\alpha_1 + b_l + U_{li} + \epsilon_{lij} \\ \Pr(R_{lij} = 1) = F_{\text{splogit}}(\beta_0 + \beta_1 t_{lij} + \rho_1 b_l + \rho_2 U_{li}; r_l) \end{array} \right.$$

## Simulation 1: WAIC

Model comparisons over simulated data sets for 50 replicates

Fitted Model	WAIC <sup>a</sup>	WAIC <sub>1</sub> <sup>b</sup>	WAIC <sub>2</sub> <sup>c</sup>
logit-JM	2239.93	1807.16	432.77
probit-JM	2244.44	1807.49	436.95
cloglog-JM	2272.57	1808.24	464.33
<b>splogit-JM</b>	<b>2203.00</b>	<b>1806.37</b>	<b>396.63</b>

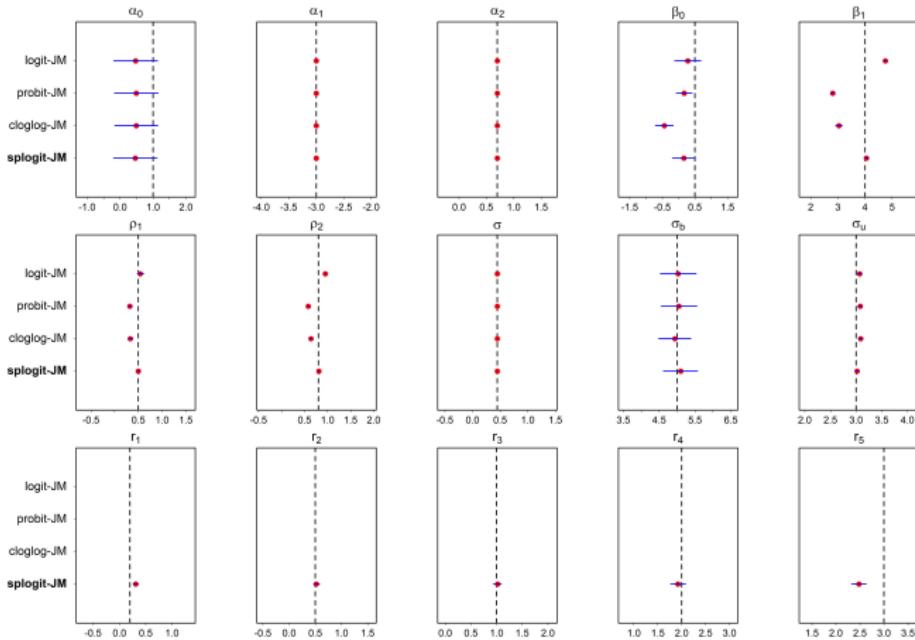
Model in boldface; true model:

<sup>a</sup> Joint model.

<sup>b</sup> Longitudinal continuous submodel:

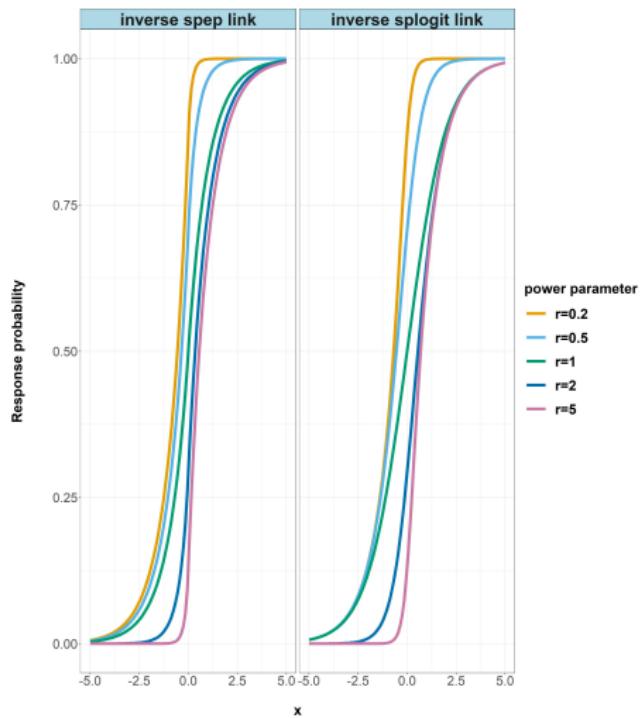
<sup>c</sup> Longitudinal binary submodel

## Simulation 1: Results



Averaged posterior mean (red dot) with **true model (boldface)**, true value (dashed line) and **95% confidence interval (blue line)** for 50 replicates via CmdStanr.

# Simulation Study 2



- **spep**
  - ▶ flexible range of skewness
  - ▶ adjustment of tail behavior
- **splotogit**
  - ▶ logit link when power parameter  $r = 1$

## Simulation Study 2: Model Structures

Compare the performance of proposed joint model with naive joint models:

$$\text{JM}_1 \left\{ \begin{array}{l} Y_{lij} = U_{li} + \epsilon_{lij} \\ \Pr(R_{lij} = 1) = F_{sep}(\beta_0 + \beta_1 t_{lij} + \rho_2 U_{li}; r) \end{array} \right.$$

$$\text{JM}_2 \left\{ \begin{array}{l} Y_{lij} = b_l + U_{li} + \epsilon_{lij} \\ \Pr(R_{lij} = 1) = F_{sep}(\beta_0 + \beta_1 t_{lij} + \rho_1 b_l + \rho_2 U_{li}; r) \end{array} \right.$$

$$\text{JM}_3 \left\{ \begin{array}{l} Y_{lij} = b_l + U_{li} + \epsilon_{lij} \\ \Pr(R_{lij} = 1) = F_{spep}(\beta_0 + \beta_1 t_{lij} + \rho_1 b_l + \rho_2 U_{li}; r_l) \end{array} \right.$$

$$\textbf{JM}_4 \left\{ \begin{array}{l} Y_{lij} = b_l + U_{li} + W_{lij} + \epsilon_{lij} \\ \Pr(R_{lij} = 1) = F_{spep}(\beta_0 + \beta_1 t_{lij} + \rho_1 b_l + \rho_2 U_{li}; r_l) \end{array} \right.$$

## Simulation 2: WAIC

## Model comparisons over simulated data sets for 50 replicates

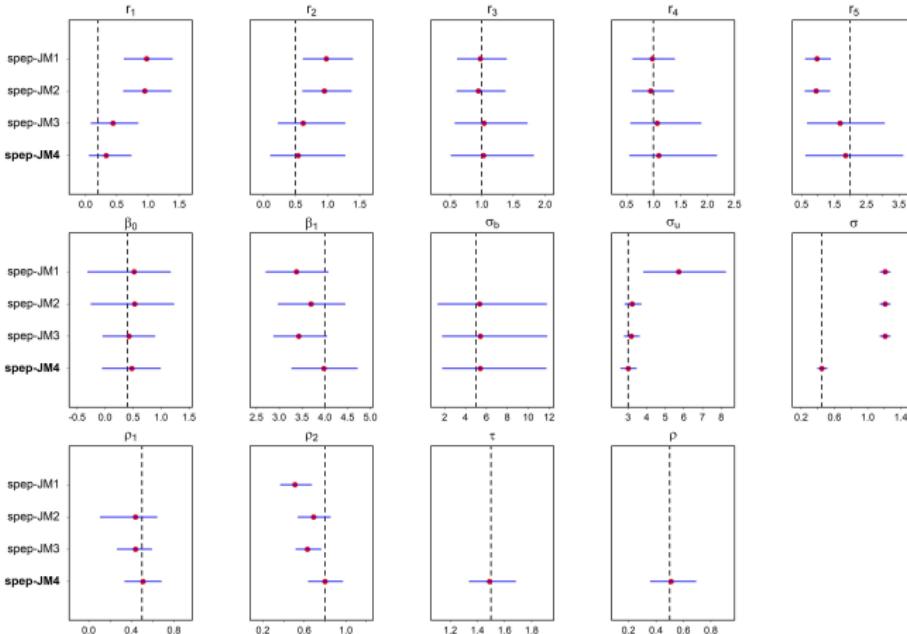
Fitted Model	WAIC <sup>a</sup>	WAIC <sub>1</sub> <sup>b</sup>	WAIC <sub>2</sub> <sup>c</sup>
Non-hierarchical (spep-JM <sub>1</sub> )	9256.99	8299.09	957.90
Common power (spep-JM <sub>2</sub> )	9179.92	8292.14	887.78
Naive LME (spep-JM <sub>3</sub> )	9157.46	8291.87	865.59
<b>Proposed (spep-JM<sub>4</sub>)</b>	<b>8352.83</b>	<b>7541.87</b>	<b>810.96</b>

Model in boldface: true model

<sup>a</sup> Joint model; <sup>b</sup> Longitudinal continuous submodel

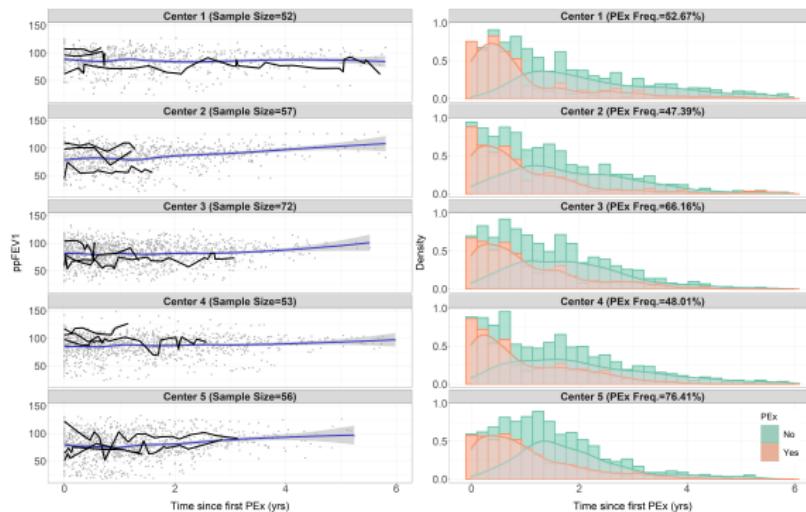
<sup>c</sup> Longitudinal binary submodel

## Simulation 2: Results



Averaged posterior means (red dot) with **true model** (boldface), true value (dashed line) and **95% credible interval** (blue line) for 50 replicates via RStan

# Motivating Data



- Censor at death/lung transplant
- Review year 2003+
- Age 6-12 years
- Num obs. 3+ spanning 6+ months
- Random 5 centers
- 381 patients

# Motivating Data: WAIC

Model comparisons for CF data with the boldface as the optimal model

Fitted Model	WAIC <sup>a</sup>	WAIC <sub>1</sub> <sup>b</sup>	WAIC <sub>2</sub> <sup>c</sup>
Non-hierarchical (srep-JM <sub>1</sub> )	49960.1	44302.0	5658.1
Common power (srep-JM <sub>2</sub> )	49867.6	44244.8	5622.8
Naive LME (srep-JM <sub>3</sub> )	49812.0	44228.0	5584.0
<b>Proposed (srep-JM<sub>4</sub>)</b>	<b>47082.3</b>	<b>42704.8</b>	<b>4377.5</b>

<sup>a</sup> Joint model;

<sup>b</sup> Longitudinal continuous submodel;

<sup>c</sup> Longitudinal binary submodel

# Motivating Data: Estimates

- **Longitudinal continuous outcome: ppFEV1**

- ▶ Positive: baseline ppFEV1, BMI percentile, Genotype F508 Heter
- ▶ Negative: Methicillin-resistant Staphylococcus aureus (MRSA), pseudomonas aeruginosa (pa) \*

- **Longitudinal binary outcome: PEx**

- ▶ Negative: time, BMI percentile, pancreatic enzymes, pa

- **Association Parameter**

- ▶ center\* strength:  $\rho_1 = -0.20(-0.81, 0.63)$
- ▶ patient strength:  $\rho_2 = -0.44(-0.68, -0.3)$

---

\* Not statistically significant

# Motivating Data: Performance

Predictive performance between training and testing cohorts

Fitted Model	Training				Testing			
	ppFEV1		PEx		ppFEV1		PEx	
	RMSE	SE	AUC	95% CI	RMSE	SE	AUC	95% CI
Non-hierarchical (ssep-JM <sub>1</sub> )	10.755	0.142	0.748	(0.734, 0.763)	10.397	0.233	0.623	(0.594, 0.651)
Common power (ssep-JM <sub>2</sub> )	10.686	0.141	0.748	(0.734, 0.763)	10.399	0.233	0.622	(0.593, 0.651)
Naive LME (ssep-JM <sub>3</sub> )	10.671	0.141	0.755	(0.741, 0.770)	10.398	0.233	0.639	(0.610, 0.667)
<b>Proposed (ssep-JM<sub>4</sub>)</b>	<b>7.768</b>	<b>0.102</b>	<b>0.882</b>	<b>(0.873, 0.892)</b>	<b>6.879</b>	<b>0.154</b>	<b>0.631</b>	<b>(0.604, 0.658)</b>

Model in boldface: optimal model; Abbreviations: RMSE=Root Mean Square Error; SE=Standard Error; AUC=Area under Curve; CI=Confidence Interval

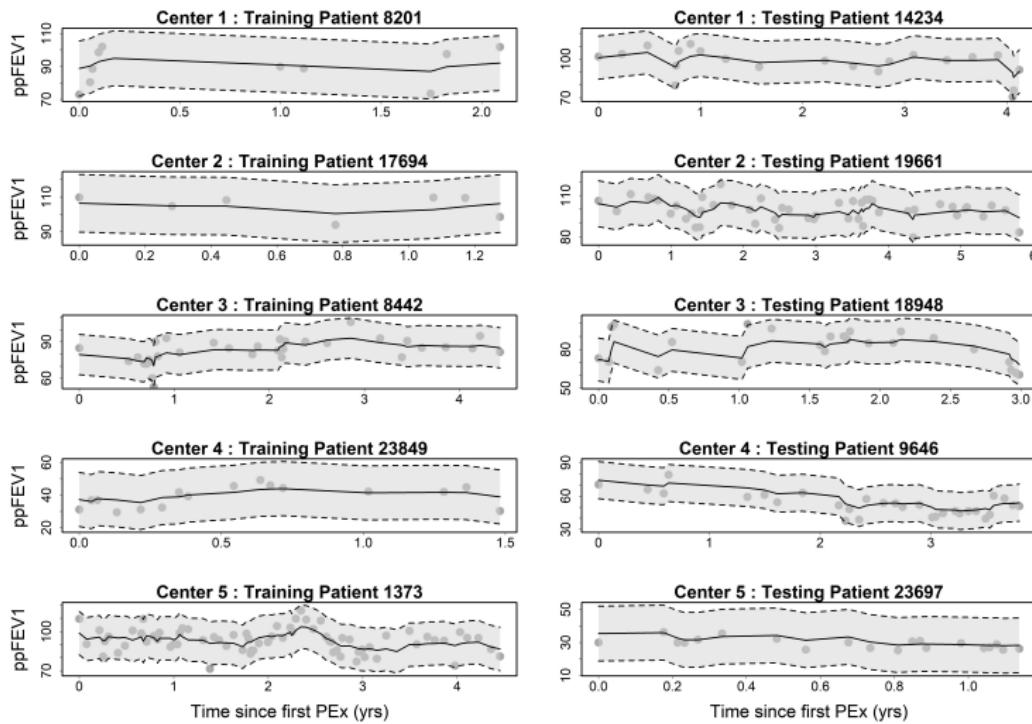
# Motivating Data: Performance (cont'd)

Forecasting performance between training and masking cohorts

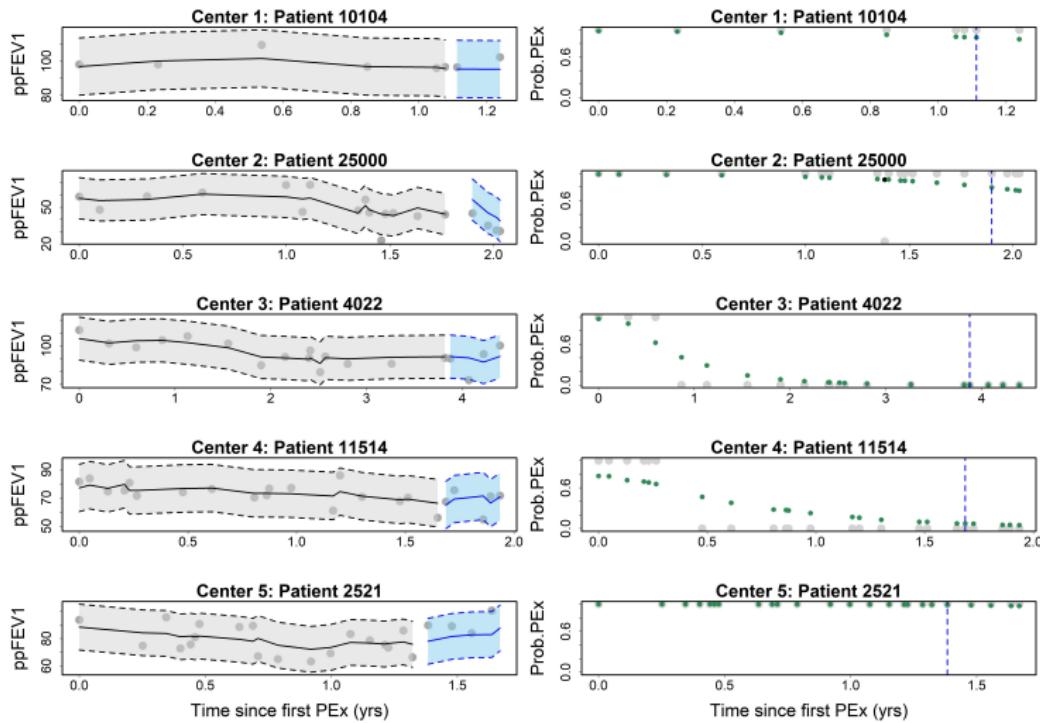
Fitted Model	Training				Masking			
	ppFEV1		PEx		ppFEV1		PEx	
	RMSE	SE	AUC	95% CI	RMSE	SE	AUC	95% CI
Non-hierarchical (ssep-JM <sub>1</sub> )	10.755	0.142	0.748	(0.734, 0.763)	10.354	0.272	0.655	(0.626, 0.683)
Common power (ssep-JM <sub>2</sub> )	10.686	0.141	0.748	(0.734, 0.763)	10.298	0.270	0.628	(0.598, 0.658)
Naive LME (ssep-JM <sub>3</sub> )	10.671	0.141	0.755	(0.741, 0.770)	10.353	0.271	0.612	(0.582, 0.642)
<b>Proposed (ssep-JM<sub>4</sub>)</b>	<b>7.768</b>	<b>0.102</b>	<b>0.882</b>	<b>(0.873, 0.892)</b>	<b>8.850</b>	<b>0.233</b>	<b>0.785</b>	<b>(0.760, 0.809)</b>

Model in boldface: optimal model; Abbreviations: RMSE=Root Mean Square Error; SE=Standard Error; AUC=Area under Curve; CI=Confidence Interval

# Motivating Data: Prediction



# Motivating Data: Forecast



# Discussion

## • Features

- ▶ Stochastic Gaussian process
- ▶ Flexible link with center-specific power parameter
- ▶ Dynamic individual prediction
- ▶ Non-hierarchical joint model biases

## • Extensions

- ▶ Predict new patients from a new center
- ▶ Flexible link family based on GEV<sup>7</sup>distribution (Wang and Dey (2010))
- ▶ Time-dependent association structure (e.g., latent Gaussian process)

---

<sup>6</sup>Generalized Extreme Value

## 1 Introduction

### 3 Multilevel Bayesian Joint Model of Longitudinal and Recurrent Outcomes

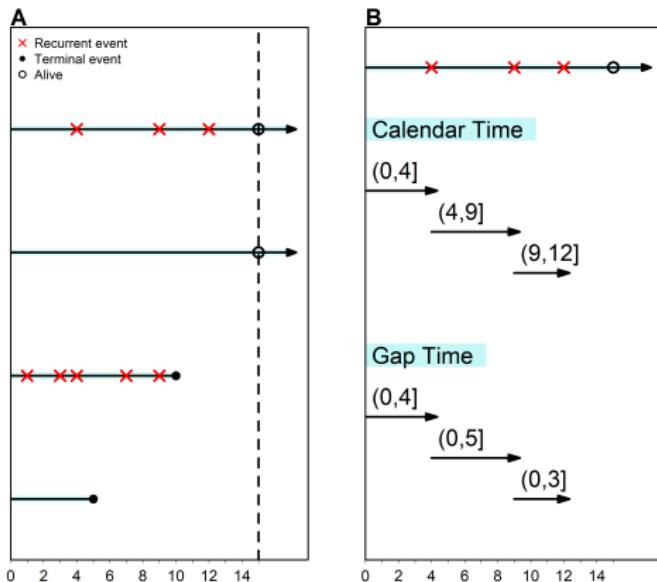
- ▶ Motivation
  - ▶ Model framework
  - ▶ Simulation study & Motivating data
  - ▶ Discussion

4 Conclusion and Future Work

# Motivation

- Time-to-recurrent PEx in two risk scales
- Center-specific time-dependent association structure
- Prognostic utility of the optimal joint model

## Risk Scales



- **calendar time**
    - ▶ effect from **entry time**
  - **gap time:**
    - ▶ effect from previous **end time**

## Model Framework

Let  $l, i, j$  be as before;  $k \in j$  denote recurrent time point.

- **Longitudinal submodel:**

$$\text{LME} \left\{ \begin{array}{l} y_{lij}(t) = m_{lij}(t) + \epsilon_{lij}(t) \\ m_{lij}(t) = \mathbf{x}_{lij}^T(t)\boldsymbol{\beta} + b_l + \mathbf{z}_{lij}^T(t)\mathbf{U}_{li} \end{array} \right. \quad (11)$$

where  $b_I \sim N(0, \sigma_b)$ ,  $\mathbf{U}_{li} \sim N(\mathbf{0}, \Sigma_U)$ ,  $\epsilon_{lij} \sim N(0, \sigma)$ ,  $\mathbf{b} \perp \mathbf{U} \perp \epsilon$ .

- Recurrent event submodel:

$$\text{ESRRFM}^8 \left\{ \begin{array}{l} h_{lik}(t) = h_{l0}(t) \exp\{\boldsymbol{\omega}_{lik}^T(t)\boldsymbol{\gamma} + f(\boldsymbol{\beta}, b_l, \mathbf{U}_{li}, \alpha_l; t) + v_{li}\} \\ h_{l0}(t) = \delta_l t^{\delta_l - 1} (\text{Weibull}) \end{array} \right. \quad (12)$$

where  $\delta_l \in (0, +\infty)$ ,  $v_{li} \sim N(0, \sigma_v)$ ,  $f(\cdot)$  denotes time-dependent association structure.

## Association Structure

linear predictor of ppFEV1:  $m_{lik}(t) = \mathbf{x}_{lik}^T(t)\beta + b_l + \mathbf{z}_{lik}^T(t)\mathbf{U}_{li}$

$$\left\{ \begin{array}{l} \text{Current value: } f(\beta, b_l, \mathbf{U}_{li}, \alpha_{v_l}; t) = \alpha_{v_l} \times m_{lik}(t) \\ \text{Current slope: } f(\beta, b_l, \mathbf{U}_{li}, \alpha_{s_l}; t) = \alpha_{s_l} \times \frac{d}{dt} m_{lik}(t) \end{array} \right. \quad (13)$$

## Bayesian Inference

- Joint posterior distribution is given by

$$p(\theta, b_l, \mathbf{U}_{li}, v_{li} | y_{li}, t_{li}, d_{li}) \propto \prod_j p(y_{lij} | b_l, \mathbf{U}_{li}, \theta) \times \prod_k p(t_{lik}, d_{lik} | b_l, \mathbf{U}_{li}, v_{li}, \theta) \\ \times p(b_l | \theta) p(\mathbf{U}_{li} | \theta) p(v_{li} | \theta) p(\theta)$$
(14)

where  $t_{ijk}$  is the observed event time and  $d_{ijk}$  is the event indicator.

- Log likelihood for event submodel
    - ▶ Calendar time

$$\log p(t_{lik}, d_{lik} | b_l, \mathbf{U}_{li}, v_{li}, \theta) = d_{lik} \cdot \log h_{li}(t_{lik}) - \int_{t_{li}(k-1)}^{t_{lik}} h_{li}(s) ds \quad (15)$$

$$\log p(t_{lik}, d_{lik} | b_l, U_{li}, v_{li}, \theta) = d_{lik} \cdot \log h_{li}(t_{lik} - t_{li(k-1)}) - \int_0^{t_{lik} - t_{li(k-1)}} h_{li}(s) ds \quad (16)$$

- ▶ Gauss-Kronrod (GK) quadrature with Q nodes (Laurie (1997))

## Computation

HMC	Simulation Study	Motivating Data
Interface to Stan	cmdstanr	cmdstanr
Post-warmup <sup>+</sup>	2000	2000
Chains	2	2

<sup>+</sup> Converged samplings validated by  $\widehat{R}$  (Gelman and Rubin (1992))

## Prior Distributions

$$\beta_0 \sim N(0, 100),$$

$$\beta_p \sim N(0, \phi_\beta), p = 1, 2, \dots, P - 1,$$

$$\sigma \sim \text{Truncated } N(0, \phi_\sigma),$$

$$\sigma_b \sim \text{Truncated } N(0, 10),$$

$\sigma_{\parallel 0}, \sigma_{\parallel 1} \sim \text{Truncated } t(1, 0, 10),$

$$\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \sim lkjCorr(2),$$

$$\gamma_0 \sim N(0, 20),$$

$$\gamma_q \sim N(0, \phi_\gamma), q = 1, 2, \dots, Q-1,$$

$$\lambda_l \sim \text{Truncated } N(0, 5), l = 1, 2, \dots, L,$$

$$\sigma_V \sim \text{Truncated } N(0, 10)$$

where  $\phi_\beta$ ,  $\phi_\gamma$  are sds of corresponding design matrix,  $\phi_\sigma$  is sd of longitudinal outcomes.  
 Note: sd=standard deviation; Normal(location,scale); student-t(df,location,scale); LKJ correlation(shape).

## Individual Prediction

- ppFEV1

$$\begin{aligned} p(\tilde{y}_{lij}(t) | \mathcal{D}) &= \int \int \int p(\tilde{y}_{lij}(t) | b_l, \mathbf{U}_{li}, \boldsymbol{\theta}) p(b_l, \mathbf{U}_{li}, \boldsymbol{\theta} | \mathcal{D}) db_l d\mathbf{U}_{li} d\boldsymbol{\theta} \\ &\approx \frac{1}{M} \sum_{m=1}^M p(\tilde{y}_{lij}(t) | b_l^{(m)}, \mathbf{U}_{li}^{(m)}, \boldsymbol{\theta}^{(m)}, \mathcal{D}) \end{aligned} \quad (17)$$

- probability of next PEx-free event

$$\begin{aligned}
S_{li}(t' | t) &= p(t_{n_{li}^*+1} \geq t' | t_{n_{li}^*+1} > t, \mathcal{D}) \\
&= \int \int \int \int p(t_{n_{li}^*+1} \geq t' | t_{n_{li}^*+1} > t, b_l, \mathbf{U}_{li}, v_{li}, \theta, \mathcal{D}) \\
&\quad \cdot p(b_l, \mathbf{U}_{li}, v_{li}, \theta | t_{n_{li}^*+1} > t, \mathcal{D}) db_l d\mathbf{U}_{li} dv_{li} d\theta \\
&\approx \frac{1}{M} \sum_{m=1}^M \exp \left[ - \int_t^{t'} h(s | b_l^{(m)}, \mathbf{U}_{li}^{(m)}, v_{li}^{(m)}, \theta^{(m)}) ds \right]
\end{aligned} \tag{18}$$

where  $\mathcal{D}$  is observed data,  $n_{ii}^*(n_{ii}^* = 0, 1, 2, \dots)$  denotes recurrent events up to the time  $t$  and  $t' > t$ .

# Predictive Performance

- Discrimination

- Time-dependent area under receiver operating characteristic curve (AUC)

- Calibration

- Time-dependent mean predictive error (MPE) on squared loss function

# Leave-one-out (loo) Cross-validation

- Pareto-smoothed importance sampling (PSIS) (Vehtari et al. (2017)) of expected log pointwise predictive density (elpd)

$$\widehat{\text{elpd}}_{\text{psis-loo}} = \sum_{l=1}^L \sum_{i=1}^{n_l} \log \left( \frac{\sum_{s=1}^S w_{li}^s p(y_{li} | \theta^s)}{\sum_{s=1}^S w_{li}^s} \right) \quad (19)$$

- LOOIC

$$\text{LOOIC} = -2 \times \widehat{\text{elpd}}_{\text{psis-loo}} \quad (20)$$

- Computed by `loo::loo()`
- The smaller the better

## Simulation Study

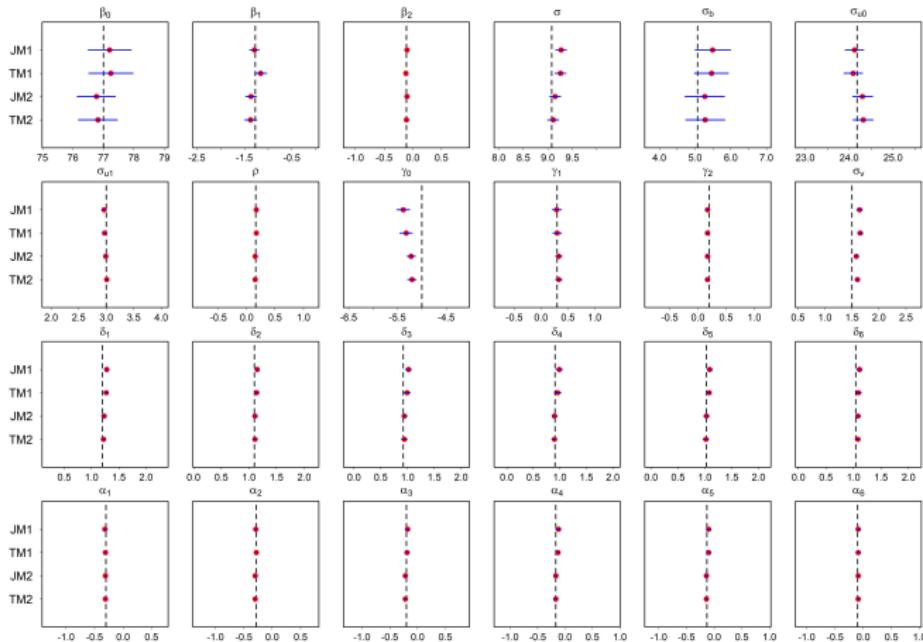
Validate Bayesian inference and compare with two-stage approach

## Simulation illustration

<b>Association + Risk scale</b>	<b>True Model</b>	<b>Fitted Model</b>
Slope + Gap	JM1	JM1 TM1
Slope + Calendar	JM2	JM2 TM2

Note: JM=Joint Model; TM=Two-stage Model

## Association Structure: Current Slope



Averaged posterior mean (red dot) with true value (dashed line) and 95% confidence interval (blue line) for 50 replicates via CmdStanr. JM1/TM1: Slope+Gap; JM2/TM2: Slope+Calendar.

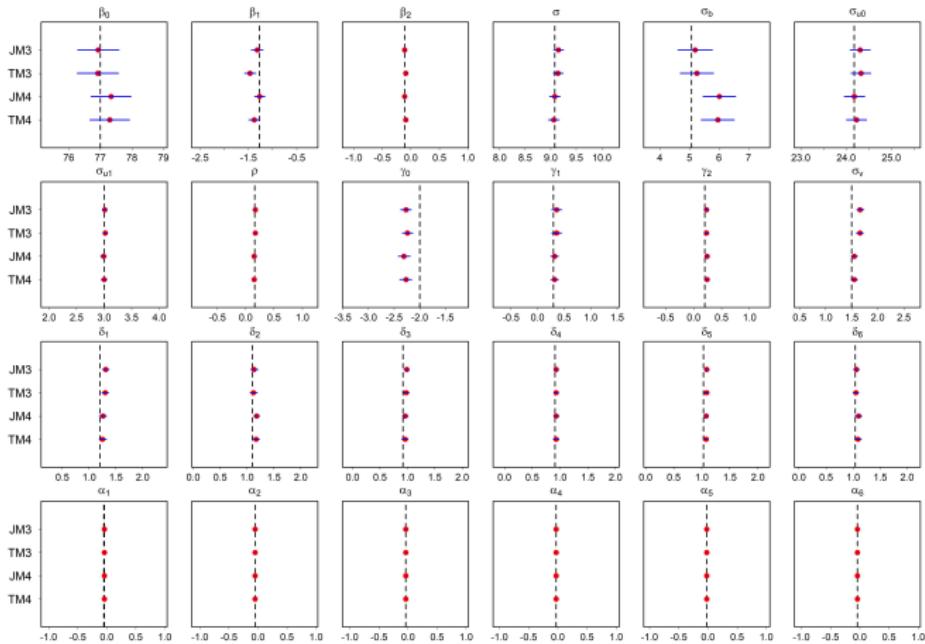
# Simulation Study (cont'd)

Simulation illustration

Association + Risk scale	True Model	Fitted Model
Value + Gap	JM3	JM3 TM3
Value + Calendar	JM4	JM4 TM4

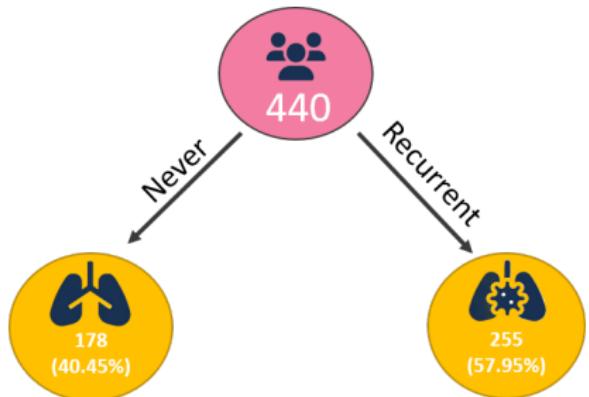
Note: JM=Joint Model; TM=Two-stage Model

## Association Structure: Current Value



Averaged posterior mean (red dot) with true value (dashed line) and 95% confidence interval (blue line) for 50 replicates via CmdStanr. JM3/TM3: Value+Gap; JM4/TM4: Value+Calendar.

## Motivating Data



- Censor at death/lung transplant/pregnancy
  - Review year 2003+
  - Age 6+ years
  - Num obs. 3+ spanning 6+ months
  - Random 6 centers by lung severity
  - Sample size 440 patients

## Motivating Data (cont'd)

Model comparisons with the boldface as the smallest LOOIC

Association + Risk scale	Model	LOOIC <sup>a</sup> (SE <sup>b</sup> )	LOOIC <sub>1</sub> <sup>c</sup> (SE <sup>b</sup> )	LOOIC <sub>2</sub> <sup>d</sup> (SE <sup>b</sup> )
Slope + Gap	Joint Model	52551.9 (383.1)	52490.3 (231.5)	61.6 (151.6)
	Two-stage Model	52578.8 (385.7)	52460.9 (232.3)	117.9 (153.4)
Slope + Calendar	Joint Model	52563.0 (389.4)	52473.1 (231.9)	89.9 (157.5)
	Two-stage Model	52575.5 (390.2)	52460.9 (232.3)	114.6 (157.9)
Value + Gap	Joint Model	52574.8 (383.4)	52491.1 (230.6)	83.7 (152.8)
	Two-stage Model	52602.4 (384.5)	52460.9 (232.3)	141.5 (152.2)
<b>Value + Calendar</b>	<b>Joint Model</b>	<b>52537.3 (386.2)</b>	52486.4 (231.1)	50.9 (155.1)
	Two-stage Model	52551.3 (388.3)	52460.9 (232.3)	90.4 (156)

<sup>a</sup> Joint model; <sup>b</sup> Standard error approximated as byproduct of *loo* package; <sup>c</sup> Longitudinal submodel;

<sup>d</sup> Event submodel

## Motivating Data (cont'd)

- **Longitudinal outcome: ppFEV1**
    - ▶ Positive: baseline ppFEV1; num PEx within prior year, birth cohort, male
    - ▶ Negative: time, time<sup>2</sup>, pa
  - **Time-to-recurrent outcome: PEx**
    - ▶ Positive: num previous PEx
    - ▶ Negative: baseline insurance, male
  - **Association structure**
    - ▶ one percentage predicted **increase** in 'true and unobserved' ppFEV1 would **decrease** PEx risk for

■ Center 1: 3.92%

#### ■ Center 4: 3.92%

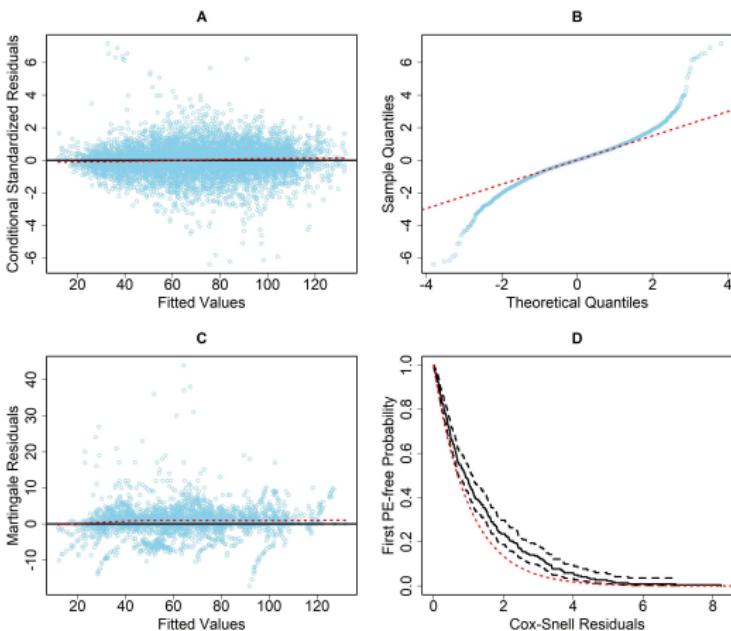
#### ■ Center 2: 4.88%

#### ■ Center 5: 2.96%

■ Center 3: 3.92%

■ Center 6: 3.92%

## Motivating Data (cont'd)



(A) subject-specific standardized residuals versus fitted values; (B) normal Q-Q plot; (C) subject-specific martingale residuals versus fitted values; (D) Cox-Snell residuals. Red dashed lines: loess curve (A & C); normal curve (B); exponential curve (D)

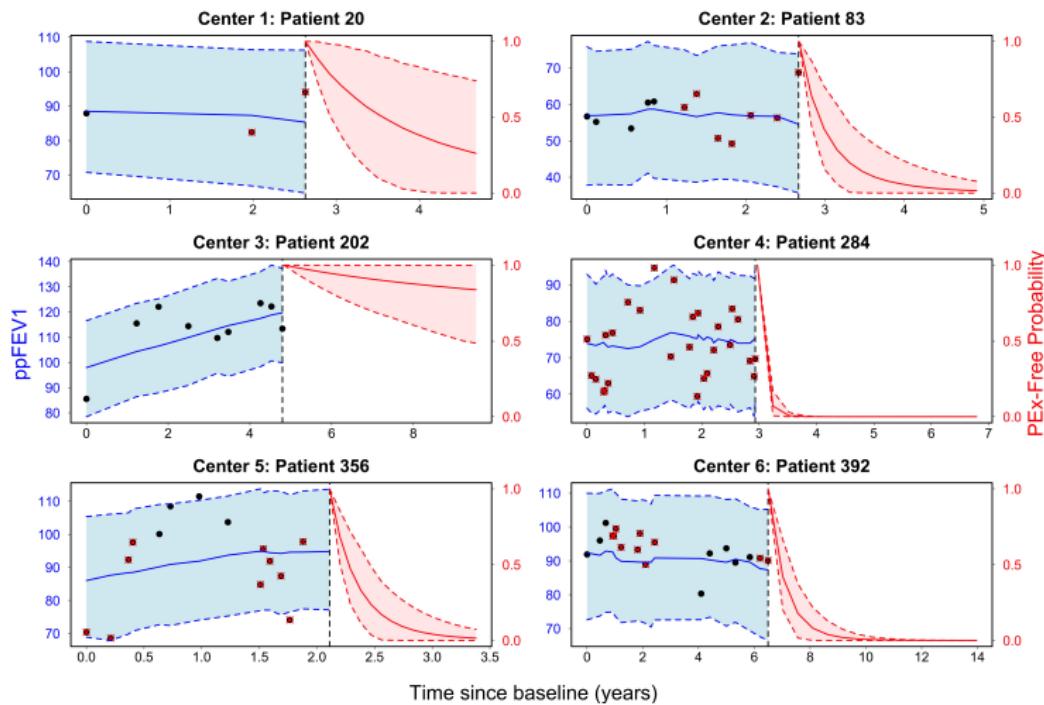
## Motivating Data (cont'd)

## Predictive performance of proposed joint models

$t_{II}$	$\Delta t$	$t'$	Num at risk	Joint Model	AUC	MPE
$[0, 2.73)$	1.41	2.73	42	Slope + Gap	0.61	0.28
				Slope + Calendar	0.64	0.27
				Value + Gap	0.66	0.27
				Value + Calendar	0.68	0.26
$[0, 5.10)$	1.56	5.10	65	Slope + Gap	0.89	0.14
				Slope + Calendar	0.88	0.14
				Value + Gap	0.90	0.14
				Value + Calendar	0.88	0.14
$[0, 7.84)$	2.76	7.84	31	Slope + Gap	0.92	0.12
				Slope + Calendar	0.91	0.13
				Value + Gap	0.92	0.12
				Value + Calendar	0.92	0.13

Note:  $t$ =individual prediction start time in year;  $\Delta t$ : averaged prediction window in year;  $t'$ : future time in year; Num at risk: Number of patients at risk at  $t'$ ; AUC=area under curve; MPE=mean predictive error based on squared loss function

## Motivating Data (cont'd)



# Discussion

## • Features

- ▶ Time-to-recurrent events in two risk scales
- ▶ Center-specific time-dependent association structure
- ▶ Applicable Stan programs
- ▶ Joint model diagnostics

## • Extensions

- ▶ Presence of terminal event
- ▶ Dynamic individual prediction
- ▶ Alternative baseline hazard candidates

## 1 Introduction

### 3 Multilevel Bayesian Joint Model of Longitudinal and Recurrent Outcomes

## 4 Conclusion and Future Work

# Conclusion

- Two novel multilevel Bayesian joint models for PEx risk in hierarchically structured data
  - ▶ PEx as binary outcome: Gaussian process + Flexible link
  - ▶ PEx as recurrent outcome: Current value + Calendar time
- Rationale for (dynamic) individual prediction
- Center-specific association strength
- Applicable Stan programs & R codes  
([https://github.com/GraceChenZhou/2022\\_DISSSERTATION.git](https://github.com/GraceChenZhou/2022_DISSSERTATION.git))

# Future Work

- RShiny app
- Left-truncation problem
- parallel-HMC algorithm

## Acknowledgements

- Advisor: Dr. Seongho Song
  - Co-advisor: Dr. Rhonda D. Szczesniak
  - Committee Members: Dr. Won Chang, Dr. Joon Hang Kim and Dr. Xia Wang



## Appendix

## LKJ correlation distribution

- $\text{lkjCorr}(\Sigma | \eta) \propto \det(\Sigma)^{\eta - 1}$
  - $\Sigma$  is a positive-definite, symmetric matrix with unit diagonal correlation matrix (i.e., a correlation matrix) with shape parameter  $\eta \in \mathbb{R}^+$  (Lewandowski et al. (2009))
  - Practically, Stan provides an implicit parameterization of the LKJ correlation matrix density in terms of its Cholesky factor
    - ▶ Let  $L_u$  denote a Cholesky factor of the correlation matrix  $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$
    - ▶  $L_u \sim \text{lkj\_corr\_cholesky}(2)$  implies  $\Sigma = L_u \cdot L_u^T \sim \text{lkjCorr}(2)$

## Conditional Survival Probability

$$\begin{aligned}
S_{li}(t' | t) &= p(t_{n_{li}+1} \geq t' | t_{n_{li}+1} > t, \mathcal{D}) \\
&= \int \int \int \int p(t_{n_{li}+1} \geq t' | t_{n_{li}+1} > t, b_l, \mathbf{U}_{li}, v_{li}, \theta, \mathcal{D}) p(b_l, \mathbf{U}_{li}, v_{li}, \theta | t_{n_{li}+1} > t, \mathcal{D}) db_l d\mathbf{U}_{li} dv_{li} d\theta \\
&= \int \int \int \int \frac{p(t_{n_{li}+1} \geq t' | b_l, \mathbf{U}_{li}, v_{li}, \theta)}{p(t_{n_{li}+1} > t | b_l, \mathbf{U}_{li}, v_{li}, \theta)} \cdot p(b_l, \mathbf{U}_{li}, v_{li}, \theta | t_{n_{li}+1} > t, \mathcal{D}) db_l d\mathbf{U}_{li} dv_{li} d\theta \\
&= \int \int \int \int \frac{S(t' | b_l, \mathbf{U}_{li}, v_{li}, \theta)}{S(t | b_l, \mathbf{U}_{li}, v_{li}, \theta)} \cdot p(b_l, \mathbf{U}_{li}, v_{li}, \theta | t_{n_{li}+1} > t, \mathcal{D}) db_l d\mathbf{U}_{li} dv_{li} d\theta \\
&= \int \int \int \int \frac{\exp \left[ - \int_0^{t'} h(s | b_l, \mathbf{U}_{li}, v_{li}, \theta) ds \right]}{\exp \left[ - \int_0^t h(s | b_l, \mathbf{U}_{li}, v_{li}, \theta) ds \right]} \cdot p(b_l, \mathbf{U}_{li}, v_{li}, \theta | t_{n_{li}+1} > t, \mathcal{D}) db_l d\mathbf{U}_{li} dv_{li} d\theta \\
&= \int \int \int \int \exp \left[ - \int_t^{t'} h(s | b_l, \mathbf{U}_{li}, v_{li}, \theta) ds \right] \cdot p(b_l, \mathbf{U}_{li}, v_{li}, \theta | t_{n_{li}+1} > t, \mathcal{D}) db_l d\mathbf{U}_{li} dv_{li} d\theta \\
&\approx \frac{1}{M} \sum_{m=1}^M \exp \left[ - \int_t^{t'} h(s | b_l^{(m)}, \mathbf{U}_{li}^{(m)}, v_{li}^{(m)}, \theta^{(m)}) ds \right]
\end{aligned} \tag{21}$$

## Time-dependent AUC

## Time-dependent AUC

- 1 Define individual-specific start time  $t_i = t_{start,i}$  and a common future stop time  $t'$ . To conform the prediction data, we only include individuals who are still at risk of the event at  $t$ . For longitudinal data, we adopt observations observed until  $t_i$ .
  - 2 Calculate event-free ('survival') probability at  $t'$  and observed  $tstop_i$  for each individual based Equation (18) to obtain  $S_i(t'|t_i)$  and  $S_i(tstop_i|t_i)$
  - 3 Sort individuals by their observed  $tstop$  in an increasing order and group each two by combinations without replacement.
  - 4 AUC is calculated by accounting for weights caused by censoring conditions. For each combination  $c = 1, \dots, C$ , assume that  $tstop_i < tstop_j$ :
    - If only individual  $i$  is censored at  $t'$ , which means  $tstop_i \leq t'$  &  $status_i = 0$ , then weight  $w = 1 - S_i(tstop_i|t_i)$
    - If only individual  $j$  is censored at  $t'$ , which means  $tstop_j \leq t'$  &  $status_j = 0$ , then weight  $w = S_j(tstop_j|t_j)$
    - If both individuals are censored at  $t'$ , then weight  $w = (1 - S_i(tstop_i|t_i)) \times S_j(tstop_j|t_j)$
    - If it does not belong to above cases,  $w = 1$
    - Let  $S_i = S_i(t'|t_i)$ ,  $S_j = S_j(t'|t_j)$ , compute  $A_c = I_{S_i < S_j} \cdot w$ ;  $D_c = I_{S_i > S_j} \cdot w$ ;  $T_c = I_{S_i = S_j} \cdot w$ , where  $I_x$  denotes a indicator function with 1 when  $x$  is true and 0, otherwise.
  - 5 Repeat Step 4 until  $C$  times
  - 6  $AUC = \sum_{c=1}^C \left( \frac{A_c + 0.5 \cdot T_c}{A_c + D_c + T_c} \right)$

## Time-dependent MPE

### Time-dependent mean predictive error (MPE)

- 1 For each individual  $i$  ( $i = 1, \dots, N$ ):
    - ▶ If individual  $i$  died or censored after  $t'$ ,  $\text{Error}_i = (1 - S_i)^2$
    - ▶ If individual  $i$  died before  $t'$ ,  $\text{Error}_i = (0 - S_i)^2$
    - ▶ If individual  $i$  censored before  $t'$ ,  $\text{Error}_i = S_i(t_{\text{stop}}|t_i) \times (1 - S_i)^2 + (1 - S_i(t_{\text{stop}}|t_i)) \times (0 - S_i)^2$
  - 2 Mean Predictive Error (MPE) =  $\frac{\sum_{i=1}^N \text{Error}_i}{N}$

## Project I: System & Time

	<b>Mac</b>	<b>PC</b>	<b>BMI</b>
Platform	x86_64-apple-darwin17.0 (64-bit)	x86_64-w64-mingw32/x64 (64-bit)	x86_64-pc-linux-gnu
Running under	macOS Big Sur 10.16	Windows 10 x64 (build 19043)	x86_64, linux-gnu
R version	4.0.5 (2021-03-31)	4.0.2 (2020-06-22)	3.6.1 (2019-07-05)
CmdStan	v2.28.2	v2.29.1	-
cmdstanr	v0.4.0	v0.5.0	-
rstan	-	-	2.19.2

Note: i). Simulation A: Mac & PC; ii) Simulation B & Real data: Biomedical Informatics (BMI) at CCHMC

Joint Model	Flexible link	Simulation A (mins/rep)	Simulation B (hrs/rep)	Motivating Data (hrs/rep)
JM1	splogit	-	0.20	-
	spep	-	0.16	1.42
JM2	splogit	-	0.25	-
	spep	-	0.22	2.86
JM3	splogit	0.31	0.19	-
	spep	-	0.18	7.62
JM4	splogit	-	0.61	-
	spep	-	0.70	13.7

mins=minutes; rep=replicate; hrs=hours

## Project II: System & Time

	<b>Simulated data</b>	<b>Real data</b>
Platform	x86_64-apple-darwin17.0 (64-bit)	x86_64-w64-mingw32/x64 (64-bit)
Running under	macOS Big Sur 10.16	Windows 10 x64 (build 19043)
R version	4.0.5 (2021-03-31)	4.0.2 (2020-06-22)
CmdStan	v2.28.2	v2.29.1
cmdstanr	v0.4.0	v0.5.0

Association + Time scale	Model	Simulated data (hrs/rep)	Real data (hrs)
Slope + Gap	Joint Model	0.18	9.16
	Two-stage Model	0.05	2.86
Slope + Calendar	Joint Model	0.26	3.82
	Two-stage Model	0.08	2.59
Value + Gap	Joint Model	0.18	8.17
	Two-stage Model	0.04	5.31
Value + Calendar	Joint Model	0.15	7.10
	Two-stage Model	0.04	4.90

rep=replicate; hrs=hours

# References I

- Betancourt, M. J. and Girolami, M. (2013). Hamiltonian monte carlo for hierarchical models.
- Brilleman, S., Crowther, M., Moreno-Betancur, M., Buros Novik, J., and R Wolfe, R. (2018). Joint longitudinal and time-to-event models via Stan. StanCon 2018. 10-12 Jan 2018. Pacific Grove, CA, USA.
- Gabry, J. and Cešnovar, R. (2022). *cmdstanr: R Interface to 'CmdStan'*. <https://mc-stan.org/cmdstanr/>, <https://discourse.mc-stan.org>.
- Gelman, A. and Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–511.
- Hickey, G. L., Philipson, P., and Jorgensen, A. (2018). joinerml: a joint model and software package for time-to-event and multivariate longitudinal outcomes. *BMC Medical Research Methodology*, 18(1).
- Jiang, X., Dey, D. K., Prunier, R., Wilson, A. M., and Holsinger, K. E. (2013). A new class of flexible link functions with application to species co-occurrence in cape floristic region. *Ann. Appl. Stat.*, 7(4):2180–2204.
- Laurie, D. (1997). Calculation of gauss-kronrod quadrature rules. *Math Comput*, 66(219):1133–45.
- Lewandowski, D., Dorota, K., and Harry, J. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100:1989–2001.
- Philipson, P., Sousa, I., Diggle, P. J., Williamson, P., Kolamunnage-Dona, R., Henderson, R., and Hickey, G. L. (2018). *joineR: Joint Modelling of Repeated Measurements and Time-to-Event Data*. R package version 1.2.6.
- Proust-Lima, C., Philipps, V., Diakite, A., and Liquet, B. (2022). *lcmm: Extended Mixed Models Using Latent Classes and Latent Processes*. R package version: 1.9.5.
- Rizopoulos, D. (2010). Jm: An r package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software*, 35(9):1–33.
- Rizopoulos, D. (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics*, 67(3):819–829.

## References II

- Rizopoulos, D. (2016). The r package jmbayes for fitting joint models for longitudinal and time-to-event data using mcmc. *Journal of Statistical Software*, 72(7):1–46.

Rizopoulos, D., Papageorgiou, G., and Miranda Afonso, P. (2022). *JMbayes2: Extended Joint Models for Longitudinal and Time-to-Event Data*. <https://drizopoulos.github.io/JMbayes2/>, <https://github.com/drizopoulos/JMbayes2>.

Rondeau, V., Gonzalez, J. R., Mazroui, Y., Mauguen, A., Diakite, A., Laurent, A., Lopez, M., Król, A., and Sofeu, C. L. (2019). *frailtypack: General Frailty Models: Shared, Joint and Nested Frailty Models with Prediction; Evaluation of Failure-Time Surrogate Endpoints*. R package version 3.0.3.

Rondeau, V., Mazroui, Y., and Gonzalez, J. R. (2012). frailtypack: An R package for the analysis of correlated survival data with frailty models using penalized likelihood estimation or parametrical estimation. *Journal of Statistical Software*, 47(4):1–28.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. Roy. Stat. Soc. B*, 64(4):583–639.

Stan Development Team (2011-2019). *Stan Modeling Language Users Guide and Reference Manual*. Version 2.29.

Stan Development Team (2020). *Rstan: the R interface to Stan*. R Foundation for Statistical Computing, Vienna, Austria. R package version 2.21.2.

Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Stat Comput*, 27:1413–1432.

Wang, X. and Dey, D. (2010). Generalized extreme value regression for binary response data: an application to b2b electronic payments system adoption. *The Annals of Applied Statistics*, 4:2000–2023.

Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11:3571–3594.