

Joint specification of generalised linear  
mixed and time-to-event models:  
A robust two-stage approach

**Danilo Alvares**

MRC Biostatistics Unit, University of Cambridge

Joint work with Valeria Leiva-Yamaguchi

# Outline

Joint modelling

Estimation approaches

Simulation study

Remarks

# Outline

Joint modelling

Estimation approaches

Simulation study

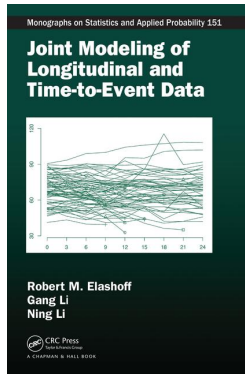
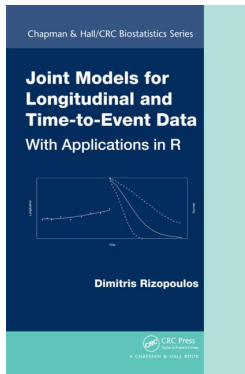
Remarks

## Joint modelling literature

Alvares, Andrinopoulou, Armero, Asar, Barrett, Brilleman, Brown, Cadarso-Suárez, Cekic, Chen, Crowther, Davidian, Diggle, Elashoff, Ibrahim, Jacqmin-Gadda, Lambert, Li, McCulloch, Molenberghs, Proust-Lima, Rizopoulos, Rondeau, Rustand, Sandler, Sinha, Sousa, Taylor, Tsiatis, van Niekerk, Wulfsohn, ...

# Joint modelling literature

Alvares, Andrinopoulou, Armero, Asar, Barrett, Brilleman, Brown, Cadarso-Suárez, Cekic, Chen, Crowther, Davidian, Diggle, Elashoff, Ibrahim, Jacqmin-Gadda, Lambert, Li, McCulloch, Molenberghs, Proust-Lima, Rizopoulos, Rondeau, Rustand, Sandler, Sinha, Sousa, Taylor, Tsiatis, van Niekerk, Wulfsohn, ...



# Standard Bayesian joint model formulation

## Longitudinal submodel

$$y_i(t | \mathbf{b}_i, \boldsymbol{\theta}_y) \sim \text{Exp-Family}(\mu_i(t | \mathbf{b}_i, \boldsymbol{\theta}_y))$$
$$(\mathbf{b}_i | \boldsymbol{\theta}_b) \sim \text{Normal}(\mathbf{0}, \boldsymbol{\Sigma})$$

## Survival submodel

$$h_i(t | \mathbf{b}_i, \boldsymbol{\theta}_y, \boldsymbol{\theta}_s) = h_0(t | \boldsymbol{\theta}_s) \exp \left\{ \mathbf{x}_i^\top \boldsymbol{\gamma} + \alpha g_i(t | \mathbf{b}_i, \boldsymbol{\theta}_y) \right\}$$

## Priors

Independent and weakly-informative

Common specifications for  $\mu_i(t | \cdot)$  and  $g_i(t | \cdot)$ :

- ▶  $\mu_i(t | \cdot)$ :  $\mathbf{x}_i^\top(t) \boldsymbol{\beta} + \mathbf{z}_i^\top(t) \mathbf{b}_i$  [Exp-Family = Normal].
- ▶  $g_i(t | \cdot)$ :  $\mu_i(t | \cdot)$ ,  $d\mu_i(t | \cdot)/dt$  and  $\int_0^t \mu_i(v | \cdot) dt$ .

# Outline

Joint modelling

Estimation approaches

Simulation study

Remarks

## Joint specification (JS) approach

### Full joint probability distribution

$$f(\mathbf{y}, \mathbf{s}, \mathbf{b}, \boldsymbol{\theta}_y, \boldsymbol{\theta}_s, \boldsymbol{\theta}_b) = f(\mathbf{y}, \mathbf{s} \mid \mathbf{b}, \boldsymbol{\theta}_y, \boldsymbol{\theta}_s) f(\mathbf{b} \mid \boldsymbol{\theta}_b) \pi(\boldsymbol{\theta}_y) \pi(\boldsymbol{\theta}_s) \pi(\boldsymbol{\theta}_b)$$

- ▶  $\mathbf{y}$ : longitudinal process with parameter vector  $\boldsymbol{\theta}_y$ .
- ▶  $\mathbf{s}$ : survival process with parameter vector  $\boldsymbol{\theta}_s$ .
- ▶  $\mathbf{b}$ : random effects with parameter vector  $\boldsymbol{\theta}_b$ .



# Joint specification (JS) approach

## Full joint probability distribution

$$f(\mathbf{y}, \mathbf{s}, \mathbf{b}, \boldsymbol{\theta}_y, \boldsymbol{\theta}_s, \boldsymbol{\theta}_b) = f(\mathbf{y}, \mathbf{s} \mid \mathbf{b}, \boldsymbol{\theta}_y, \boldsymbol{\theta}_s) f(\mathbf{b} \mid \boldsymbol{\theta}_b) \pi(\boldsymbol{\theta}_y) \pi(\boldsymbol{\theta}_s) \pi(\boldsymbol{\theta}_b)$$

- ▶  $\mathbf{y}$ : longitudinal process with parameter vector  $\boldsymbol{\theta}_y$ .
- ▶  $\mathbf{s}$ : survival process with parameter vector  $\boldsymbol{\theta}_s$ .
- ▶  $\mathbf{b}$ : random effects with parameter vector  $\boldsymbol{\theta}_b$ .

## Shared-parameter factorisation

$$f(\mathbf{y}, \mathbf{s} \mid \mathbf{b}, \boldsymbol{\theta}_y, \boldsymbol{\theta}_s) = f(\mathbf{y} \mid \mathbf{b}, \boldsymbol{\theta}_y) f(\mathbf{s} \mid \mathbf{b}, \boldsymbol{\theta}_y, \boldsymbol{\theta}_s)$$

# Standard two-stage (STS) approach

► **Stage 1:**

$$\text{MAP} \equiv (\hat{\mathbf{b}}, \hat{\boldsymbol{\theta}}_y) = \max_{\mathbf{b}, \boldsymbol{\theta}_y} \left\{ f(\mathbf{y} \mid \mathbf{b}, \boldsymbol{\theta}_y) f(\mathbf{b} \mid \boldsymbol{\theta}_b) \pi(\boldsymbol{\theta}_y) \pi(\boldsymbol{\theta}_b) \right\}$$

# Standard two-stage (STS) approach

► **Stage 1:**

$$\text{MAP} \equiv (\hat{\mathbf{b}}, \hat{\boldsymbol{\theta}}_y) = \max_{\mathbf{b}, \boldsymbol{\theta}_y} \left\{ f(\mathbf{y} \mid \mathbf{b}, \boldsymbol{\theta}_y) f(\mathbf{b} \mid \boldsymbol{\theta}_b) \pi(\boldsymbol{\theta}_y) \pi(\boldsymbol{\theta}_b) \right\}$$

► **Stage 2:**

$$\pi(\boldsymbol{\theta}_s \mid \text{data}) \propto f(\mathbf{s} \mid \hat{\mathbf{b}}, \hat{\boldsymbol{\theta}}_y, \boldsymbol{\theta}_s) \pi(\boldsymbol{\theta}_s)$$

# Novel two-stage (NTS) approach

► **Stage 1:**

$$\text{MAP} \equiv (\hat{\boldsymbol{\theta}}_y, \hat{\boldsymbol{\theta}}_b) = \max_{\boldsymbol{\theta}_y, \boldsymbol{\theta}_b} \left\{ f(\mathbf{y} \mid \mathbf{b}, \boldsymbol{\theta}_y) f(\mathbf{b} \mid \boldsymbol{\theta}_b) \pi(\boldsymbol{\theta}_y) \pi(\boldsymbol{\theta}_b) \right\}$$

# Novel two-stage (NTS) approach

► **Stage 1:**

$$\text{MAP} \equiv (\hat{\boldsymbol{\theta}}_y, \hat{\boldsymbol{\theta}}_b) = \max_{\boldsymbol{\theta}_y, \boldsymbol{\theta}_b} \left\{ f(\mathbf{y} \mid \mathbf{b}, \boldsymbol{\theta}_y) f(\mathbf{b} \mid \boldsymbol{\theta}_b) \pi(\boldsymbol{\theta}_y) \pi(\boldsymbol{\theta}_b) \right\}$$

► **Stage 2:**

$$\pi(\mathbf{b}, \boldsymbol{\theta}_s \mid \text{data}) \propto f(\mathbf{y} \mid \mathbf{b}, \hat{\boldsymbol{\theta}}_y) f(\mathbf{s} \mid \mathbf{b}, \hat{\boldsymbol{\theta}}_y, \boldsymbol{\theta}_s) f(\mathbf{b} \mid \hat{\boldsymbol{\theta}}_b) \pi(\boldsymbol{\theta}_s)$$

## Novel two-stage (NTS) approach

### Proposition

*For fixed sample size  $n$ , and as the number  $n_i$  of repeated measurements per individual in the longitudinal process  $\mathbf{y}$  increases, the maximum a posteriori (MAP) of  $\boldsymbol{\theta}_y$  and  $\boldsymbol{\theta}_s$  from the joint specification (JS) and novel two-stage (NTS) approaches will closely resemble.*

# Outline

Joint modelling

Estimation approaches

**Simulation study**

Remarks

# Joint model simulation algorithm

- 1: **Initialisation:** Set  $\theta_y$ ,  $\theta_s$ ,  $\theta_b$ ,  $n$ ,  $\Delta$ , and  $t_{\max}$ .
- 2: **Simulating survival data:**
  - Simulate  $x_i \sim \text{Bernoulli}(0.5)$  and  $\mathbf{b}_i \sim \text{Normal}(\mathbf{0}, \Sigma) \forall i$ .
  - Calculate  $T_i^* \forall i$  based on the survival submodel.
  - Simulate  $C_i \sim \text{Uniform}(0, t_{\max}) \forall i$ .
  - Set  $T_i = \min\{T_i^*, C_i\}$  and  $\delta_i = \mathbf{I}(T_i^* \leq C_i) \forall i$ .
- 3: **Simulating longitudinal data:**
  - Set  $0 = t_1, \dots, t_{n_i} \leq T_i \forall i$  such that  $t_{j+1} - t_j = \Delta$ .
  - Compute  $y_i(t_1), \dots, y_i(t_{n_i}) \forall i$  based on the longitudinal submodel.



# Exponential family specifications

## Longitudinal continuous outcome

$$y_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y) \sim \text{Normal}(\mu_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y), \sigma^2)$$

$$\mu_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y) = \beta_0 + b_{0i} + (\beta_1 + b_{1i})t + \beta_2 x_i$$

## Longitudinal count outcome

$$y_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y) \sim \text{Poisson}(\mu_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y))$$

$$\log(\mu_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y)) = \beta_0 + b_{0i} + (\beta_1 + b_{1i})t + \beta_2 x_i$$

## Longitudinal binary outcome

$$y_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y) \sim \text{Bernoulli}(\mu_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y))$$

$$\text{logit}(\mu_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y)) = \beta_0 + b_{0i} + (\beta_1 + b_{1i})t + \beta_2 x_i$$

## Survival outcome

$$h_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y, \boldsymbol{\theta}_s) = \phi t^{\phi-1} \exp\{\gamma_0 + \gamma_1 x_i + \alpha \mu_i(t \mid \mathbf{b}_i, \boldsymbol{\theta}_y)\}$$

## Setting parameters

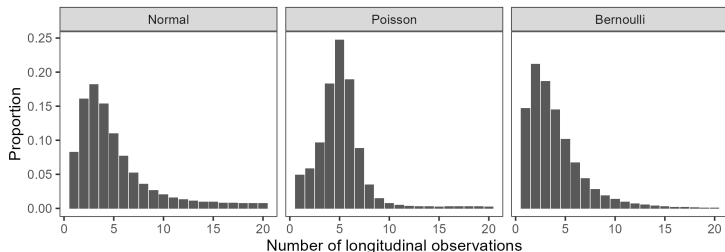
For each scenario (Normal, Poisson, and Bernoulli), we simulated 300 datasets with  $n = 500$ ,  $\Delta = 1$ , and  $t_{\max} = 20$ .

Scenario	$\beta_0$	$\beta_1$	$\beta_2$	$\sigma$	$\Sigma_{11}$	$\Sigma_{22}$	$\gamma_0$	$\gamma_1$	$\phi$	$\alpha$
Normal	-1.00	-0.10	-0.30	0.30	0.10	0.01	-1.00	-0.50	2.50	2.00
Poisson	-1.00	0.03	0.50	-	0.05	0.01	-7.00	-0.50	5.00	-2.00
Bernoulli	-3.00	2.00	-5.00	-	10.00	2.00	-3.00	-1.00	1.50	3.00

# Setting parameters

For each scenario (Normal, Poisson, and Bernoulli), we simulated 300 datasets with  $n = 500$ ,  $\Delta = 1$ , and  $t_{\max} = 20$ .

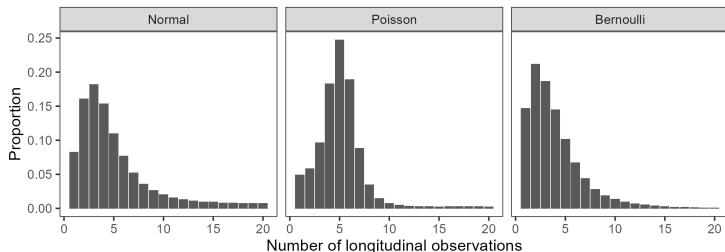
Scenario	$\beta_0$	$\beta_1$	$\beta_2$	$\sigma$	$\Sigma_{11}$	$\Sigma_{22}$	$\gamma_0$	$\gamma_1$	$\phi$	$\alpha$
Normal	-1.00	-0.10	-0.30	0.30	0.10	0.01	-1.00	-0.50	2.50	2.00
Poisson	-1.00	0.03	0.50	–	0.05	0.01	-7.00	-0.50	5.00	-2.00
Bernoulli	-3.00	2.00	-5.00	–	10.00	2.00	-3.00	-1.00	1.50	3.00



## Setting parameters

For each scenario (Normal, Poisson, and Bernoulli), we simulated 300 datasets with  $n = 500$ ,  $\Delta = 1$ , and  $t_{\max} = 20$ .

Scenario	$\beta_0$	$\beta_1$	$\beta_2$	$\sigma$	$\Sigma_{11}$	$\Sigma_{22}$	$\gamma_0$	$\gamma_1$	$\phi$	$\alpha$
Normal	-1.00	-0.10	-0.30	0.30	0.10	0.01	-1.00	-0.50	2.50	2.00
Poisson	-1.00	0.03	0.50	–	0.05	0.01	-7.00	-0.50	5.00	-2.00
Bernoulli	-3.00	2.00	-5.00	–	10.00	2.00	-3.00	-1.00	1.50	3.00



### Censoring rates:

- ▶ 26%-40% (Normal), 22%-35% (Poisson), and 16%-27% (Bernoulli).

# MCMC configuration and running time

## **Joint specification (JS):**

- ▶ 2000 iterations with warm-up of 1000.

## **Novel two-stage (NTS) and standard two-stage (STS):**

- ▶ Longitudinal submodel: 2000 iterations with warm-up of 1000.
- ▶ Survival submodel: 1000 iterations with warm-up of 500.

# MCMC configuration and running time

## **Joint specification (JS):**

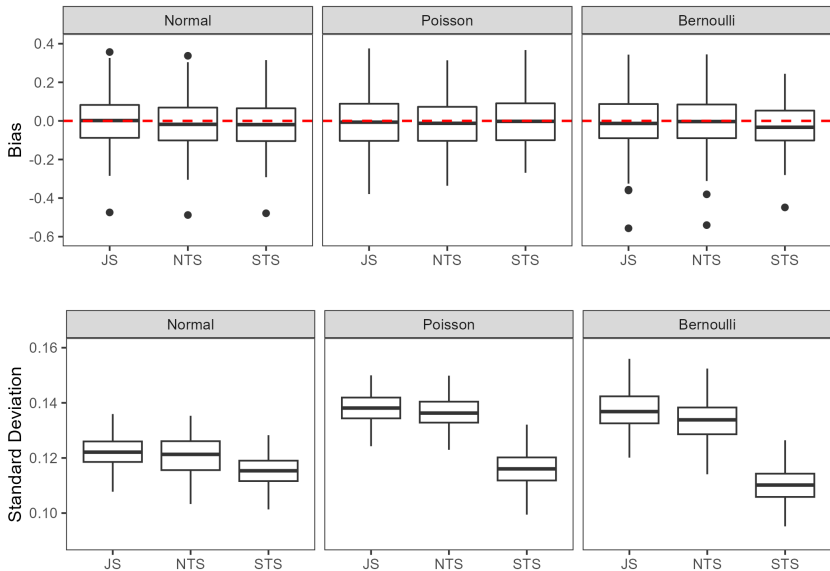
- ▶ 2000 iterations with warm-up of 1000.

## **Novel two-stage (NTS) and standard two-stage (STS):**

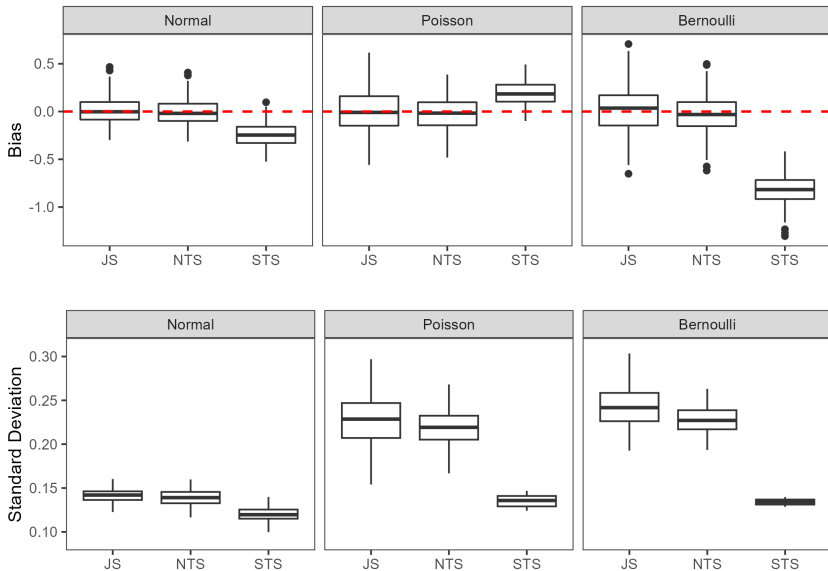
- ▶ Longitudinal submodel: 2000 iterations with warm-up of 1000.
- ▶ Survival submodel: 1000 iterations with warm-up of 500.

Scenario	JS	NTS	STS
Normal	5.65	2.70	1.65
Poisson	7.68	5.90	2.10
Bernoulli	8.45	4.25	1.38

# Posterior inference for the red parameter ( $\gamma_1$ )



# Posterior inference for the association parameter ( $\alpha$ )





# Outline

Joint modelling

Estimation approaches

Simulation study

Remarks

## Remarks

- ▶ Novel two-stage (NTS) simplifies the complex nature of joint specification (JS) by using a two-stage strategy that at the same time corrects the well-known bias of a standard two-stage (STS) approach.

## Remarks

- ▶ Novel two-stage (NTS) simplifies the complex nature of joint specification (JS) by using a two-stage strategy that at the same time corrects the well-known bias of a standard two-stage (STS) approach.
- ▶ NTS asymptotically resembles a joint inference.

## Remarks

- ▶ Novel two-stage (NTS) simplifies the complex nature of joint specification (JS) by using a two-stage strategy that at the same time corrects the well-known bias of a standard two-stage (STS) approach.
- ▶ NTS asymptotically resembles a joint inference.
- ▶ The group parameter was not significantly affected by the use of two-stage approaches.

## Remarks

- ▶ Novel two-stage (NTS) simplifies the complex nature of joint specification (JS) by using a two-stage strategy that at the same time corrects the well-known bias of a standard two-stage (STS) approach.
- ▶ NTS asymptotically resembles a joint inference.
- ▶ The group parameter was not significantly affected by the use of two-stage approaches.
- ▶ The association parameter was biased using STS. This issue was mitigated by replacing STS with NTS.

## Remarks

- ▶ Novel two-stage (NTS) simplifies the complex nature of joint specification (JS) by using a two-stage strategy that at the same time corrects the well-known bias of a standard two-stage (STS) approach.
- ▶ NTS asymptotically resembles a joint inference.
- ▶ The group parameter was not significantly affected by the use of two-stage approaches.
- ▶ The association parameter was biased using STS. This issue was mitigated by replacing STS with NTS.
- ▶ In all scenarios, NTS had a much lower computational time than the JS approach.

## Remarks

- ▶ Novel two-stage (NTS) simplifies the complex nature of joint specification (JS) by using a two-stage strategy that at the same time corrects the well-known bias of a standard two-stage (STS) approach.
- ▶ NTS asymptotically resembles a joint inference.
- ▶ The group parameter was not significantly affected by the use of two-stage approaches.
- ▶ The association parameter was biased using STS. This issue was mitigated by replacing STS with NTS.
- ▶ In all scenarios, NTS had a much lower computational time than the JS approach.
- ▶ NTS can be easily generalised to more complex longitudinal and survival submodels.

Thank you for your attention!

R codes

