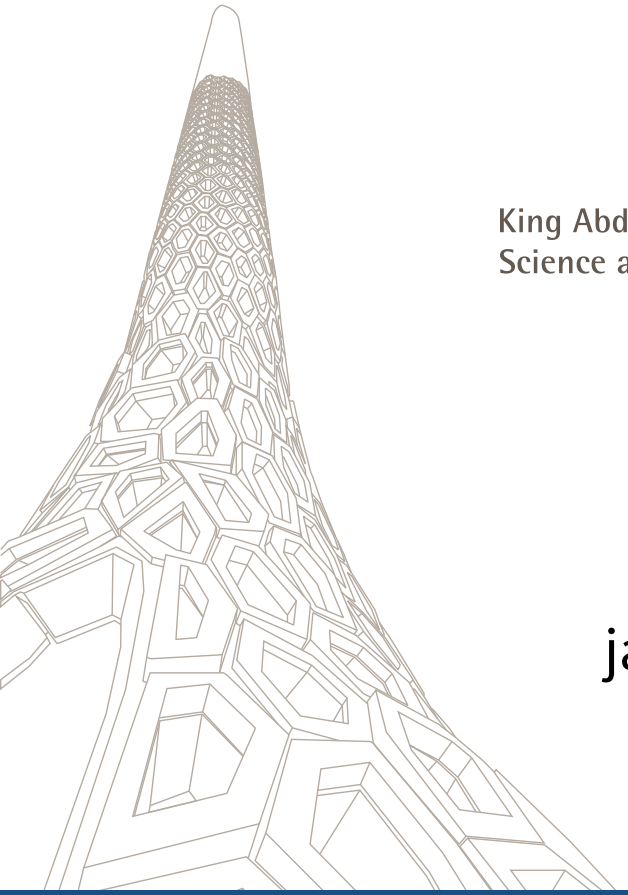
Bayesian Survival and Joint modeling with INLA - fast and accurate



جامعة الملك عبدالله
للعلوم والتقنية

King Abdullah University of
Science and Technology

Janet van Niekerk
janet.vanNiekerk@kaust.edu.sa
July 2025


SCAN ME

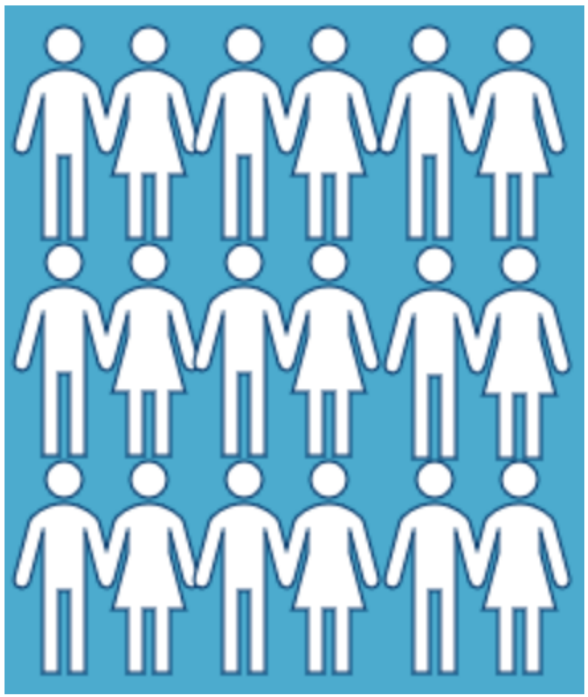
1. Survival analysis
   - Proportional Hazards regression model
   - Complex survival models
   - Computational approaches for inference
2. Time-dependent covariates in survival models
3. Longitudinal modeling
   - Computational approaches for inference
4. Joint models
   - Association structure
   - Computational approaches for inference
5. INLA
   - INLAjoint - user interface for common joint models
6. Examples
   - Posteriors and dynamic prediction
7. Joint models with INLA but not INLAjoint
8. Summary
9. Current work
10. Extra examples of other datasets

Survival analysis/Reliability/Time-to-event modeling/Lifetime modeling

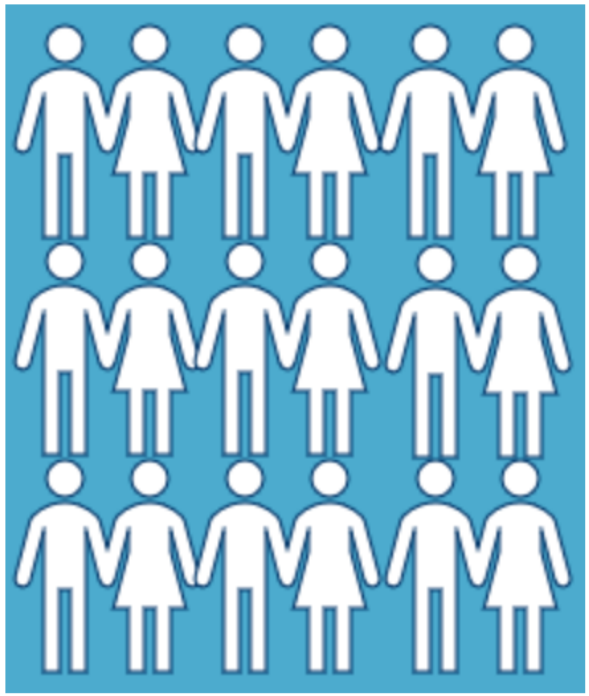
Example: treatment evaluation in a clinical trial

The evaluation of new therapeutics requires a **criterion** in order to **compare them to a placebo or standard of care** (i.e., best treatment available).

New treatment



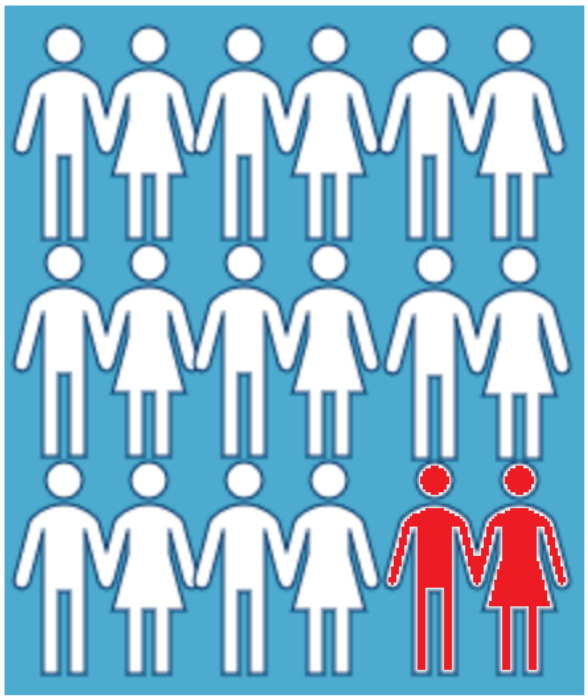
Standard treatment



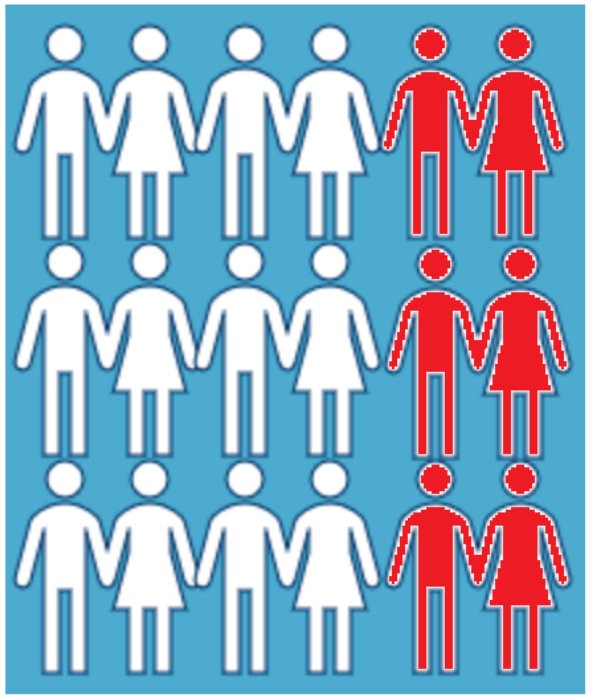
Example: treatment evaluation in a clinical trial

A **criterion** is required in order to **compare** them to a **placebo or standard of care** .

New treatment



Standard treatment

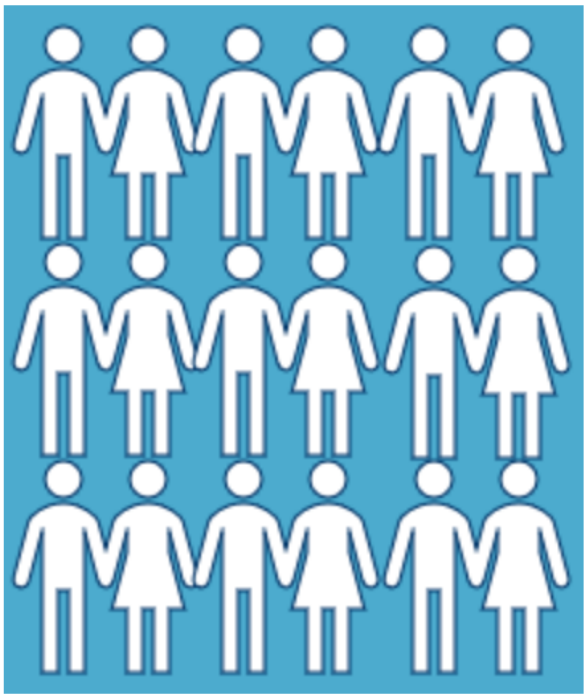


Count how many people died in each group?

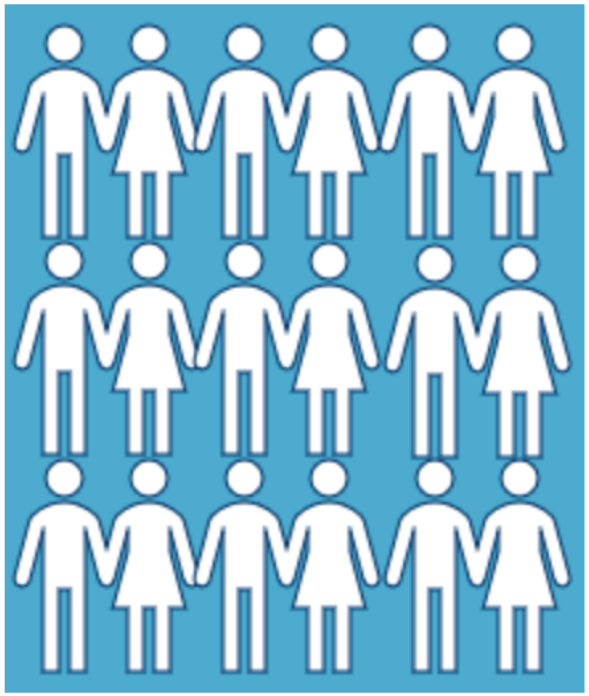
Duration of follow-up

How lo

New treatment



Standard treatment



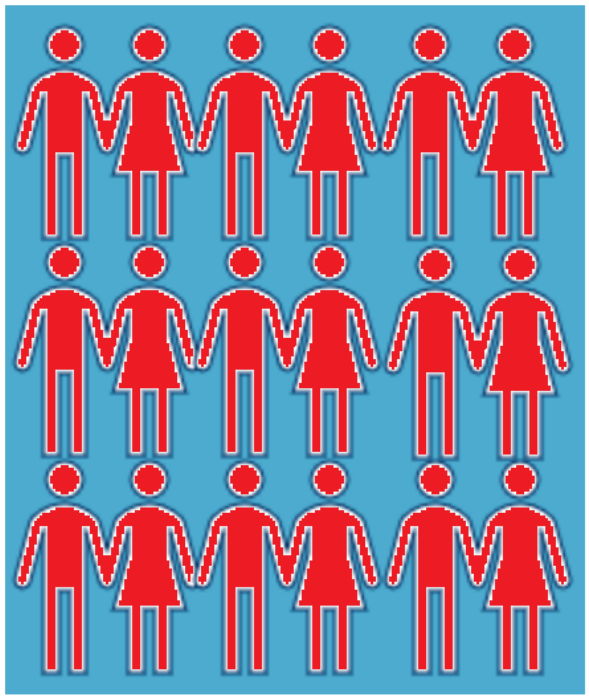
ns?

Follow-up too short = no deaths = no comparison possible

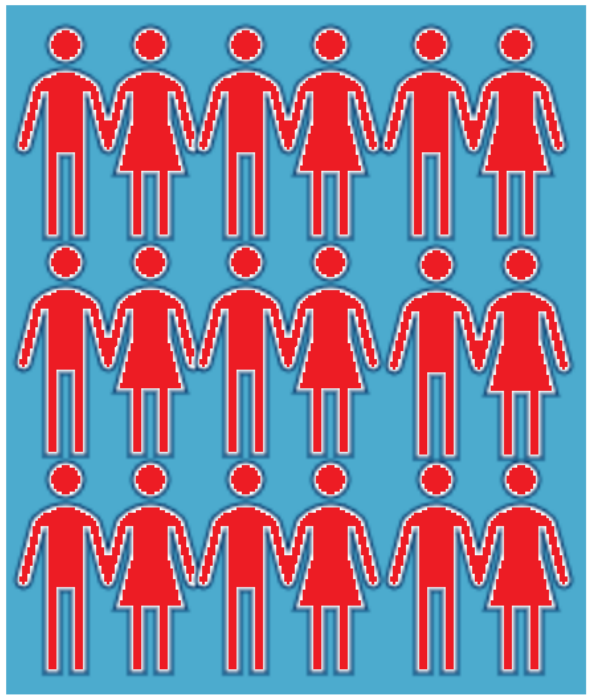
Duration of follow-up

How l

New treatment



Standard treatment

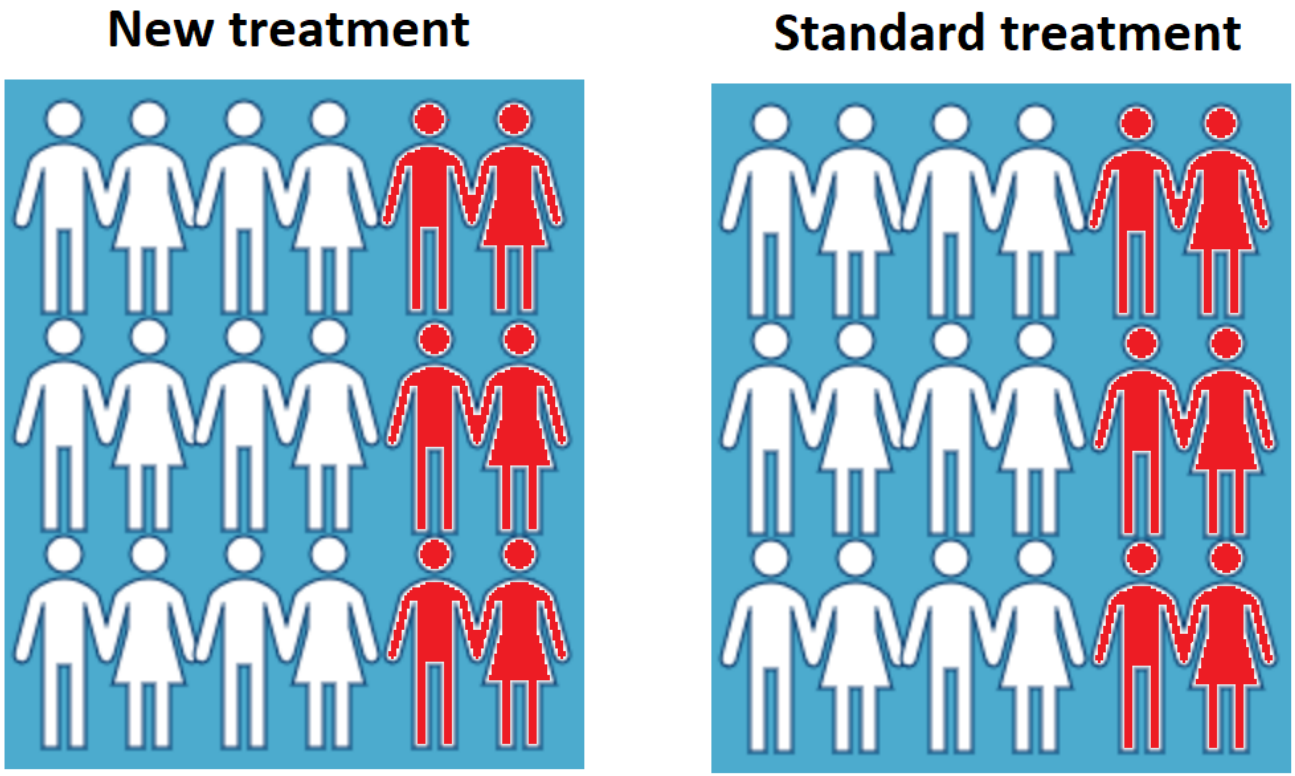


ıs?

Follow-up too long = all patients die = no comparison possible

Long follow-up also means **longer delay before the new treatment is available**!

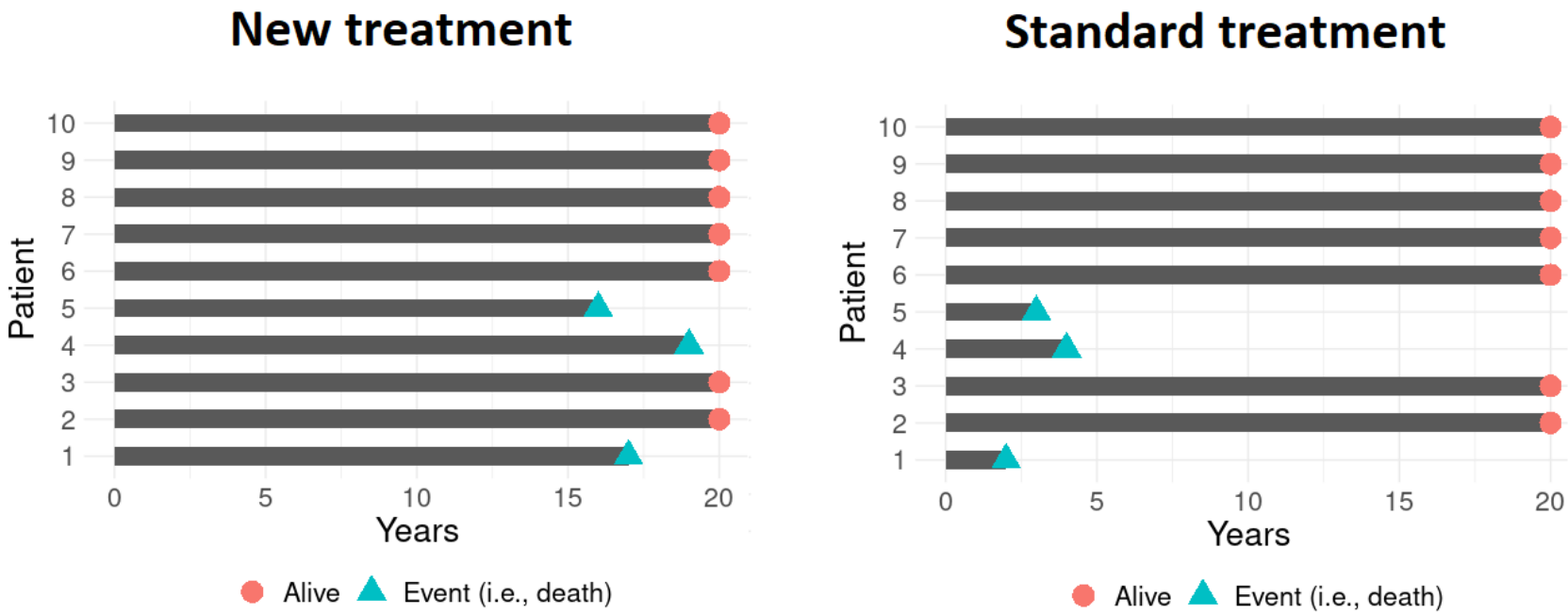
Time to event



What about the **time elapsed** between the beginning of the trial and the event?

Time to event

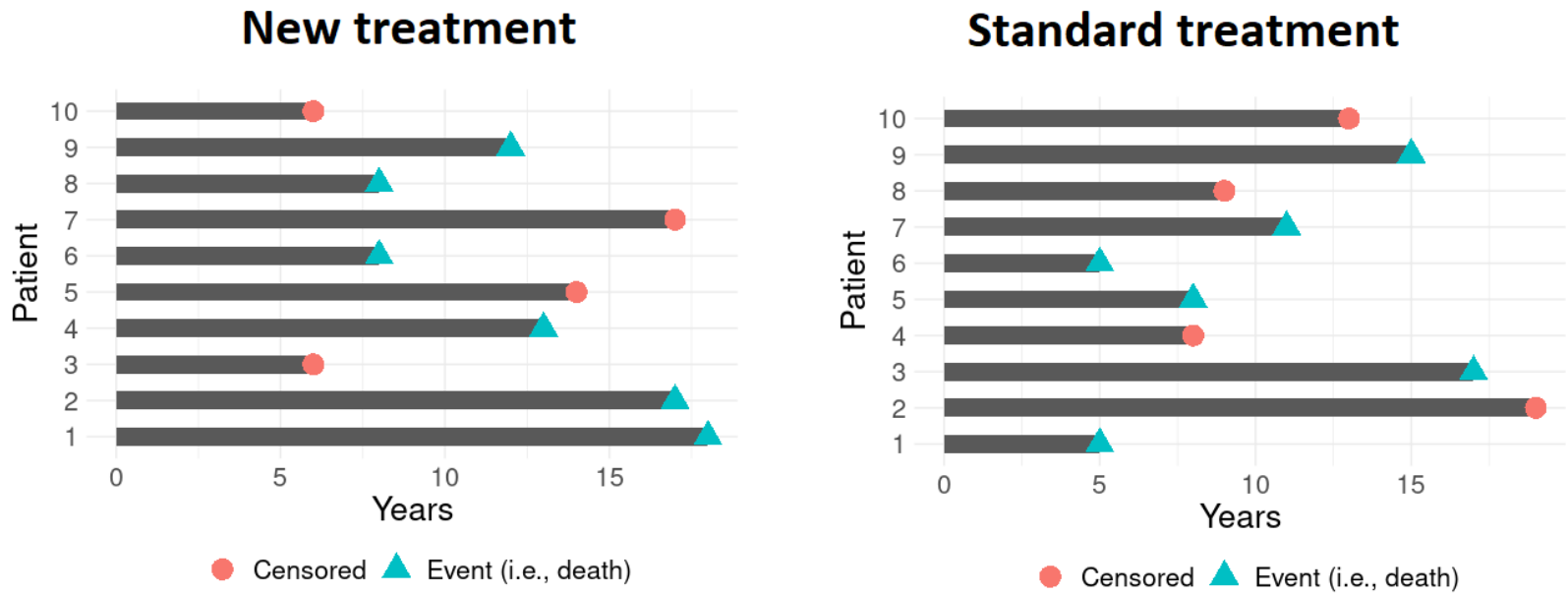
What about the **time elapsed** between the beginning of the trial and the event?



Censoring

Censoring

The event (i.e., death) is not observed for some individuals...



Censoring occurs when a patient **withdraws** from the study, is **lost to follow-up**, or is **alive without event occurrence at last follow-up**.

=> We don't know what happened to these people after censoring.

Random censoring / informative censoring

- **Random (or non-informative)** censoring: Patients are censored due to reasons unrelated to the study.

Examples: The event has not occurred at the end of the follow-up; Patient moves to another country and cannot continue its participation to the trial.

- **Informative censoring**: Patients are censored due to reasons related to the study.

Examples: Patient decides to stop taking the treatment because of side effects; Patient is too sick and clinicians decide to switch to another treatment *(assuming independent censoring in this case would lead to underestimation of the hazard)*; Patient feels cured by an effective treatment and may no longer feel the need to follow-up *(assuming independent censoring in this case would lead to overestimation of the hazard)*.

In survival analysis, we usually need to assume censoring is random!

Random censoring / informative censoring

- **Random (or non-informative)** censoring: Patients are censored due to reasons unrelated to the study.

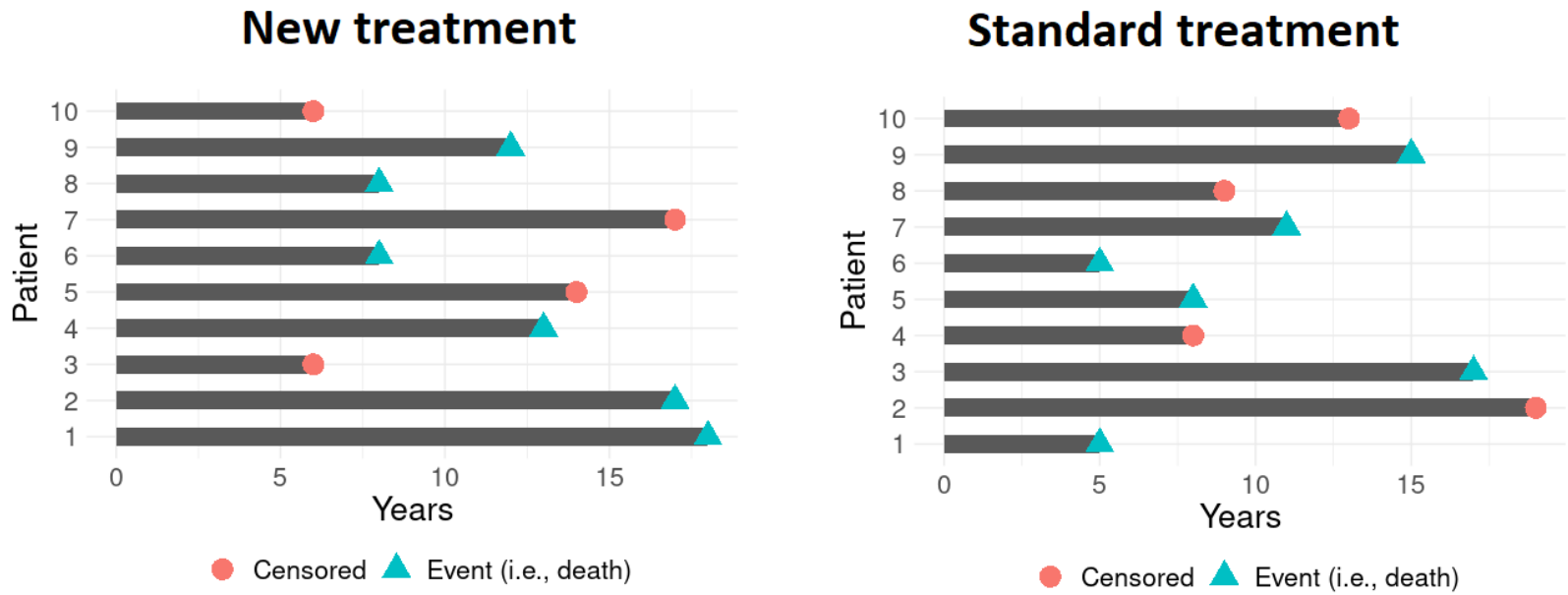
Examples: The event has not occurred at the end of the follow-up; Patient moves to another country and cannot continue its participation to the trial.

- **Informative censoring**: Patients are censored due to reasons related to the study.

Examples: Patient decides to stop taking the treatment because of side effects; Patient is too sick and clinicians decide to switch to another treatment *(assuming independent censoring in this case would lead to underestimation of the hazard)*; Patient feels cured by an effective treatment and may no longer feel the need to follow-up *(assuming independent censoring in this case would lead to overestimation of the hazard)*.
In survival analysis, we usually need to assume censoring is random!

Survival data - Survival analysis

Survival analysis is used to analyze data in which **the time until the event is of interest**. The response is often referred to as a failure time, survival time, or event time.



=> Not easy to compare!

Moreover clinical trials usually include hundreds of patients.

Example - PBC data 1/2

Primary Biliary Cholangitis (PBC)

- **312 PBC patients** were followed at the Mayo Clinic between 1974 and 1988 and received either a **placebo or D-penicillamine**.
- This data is **publicly available** in the R package 'JM'.
- During the follow-up, **140 patients died**.

Example - PBC data 2/2

PBC dataset (available in R package JM):

id	years	death	drug	sex
1	1.095170	1	D-penicil	female
2	14.152338	0	D-penicil	female
3	2.770781	1	D-penicil	male
4	5.270507	1	D-penicil	female
5	4.120578	0	placebo	female
6	6.853028	1	placebo	female
⋮	⋮	⋮	⋮	⋮

Question of interest: Effect of drug and sex on the risk of death?

Quantities of interest in survival analysis

Let ***T*** denote the positive continuous response variable that represents the **elapsed time between the beginning of the follow-up and the event of interest**. There are many ways to represent and describe the distribution of T.

- **Hazard function:** Instantaneous **risk of event.**

$$\lambda(t) = \lim_{h \to 0^+} \frac{P(t \leq T < t + h | T > t)}{h}$$

- **Survival function:** Probability of being alive up to time t (i.e., dying after t).

$$S(t) = P(T > t) = \exp(-\int_0^t \lambda(u)du)$$

Survival function properties

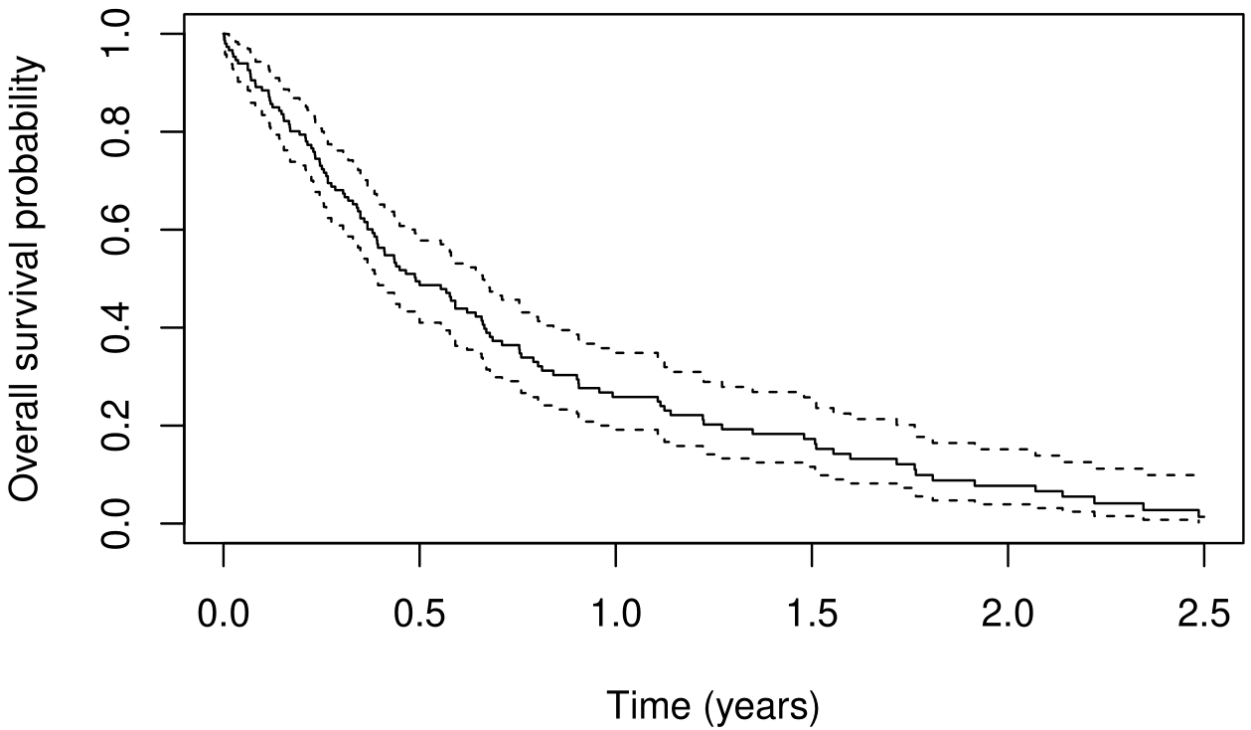
- **Survival function:** Probability of being alive up to time t (i.e., dying after t).

$$S(t) = P(T > t) = \exp(-\int_0^t \lambda(u)du)$$

$S(0) = 1$ (probability of being alive at origin time = 1)
$\lim_{n \to +\infty} S(t) = 0$ (probability of eternal life = 0)
The survival function is **only decreasing** (no resurrection)

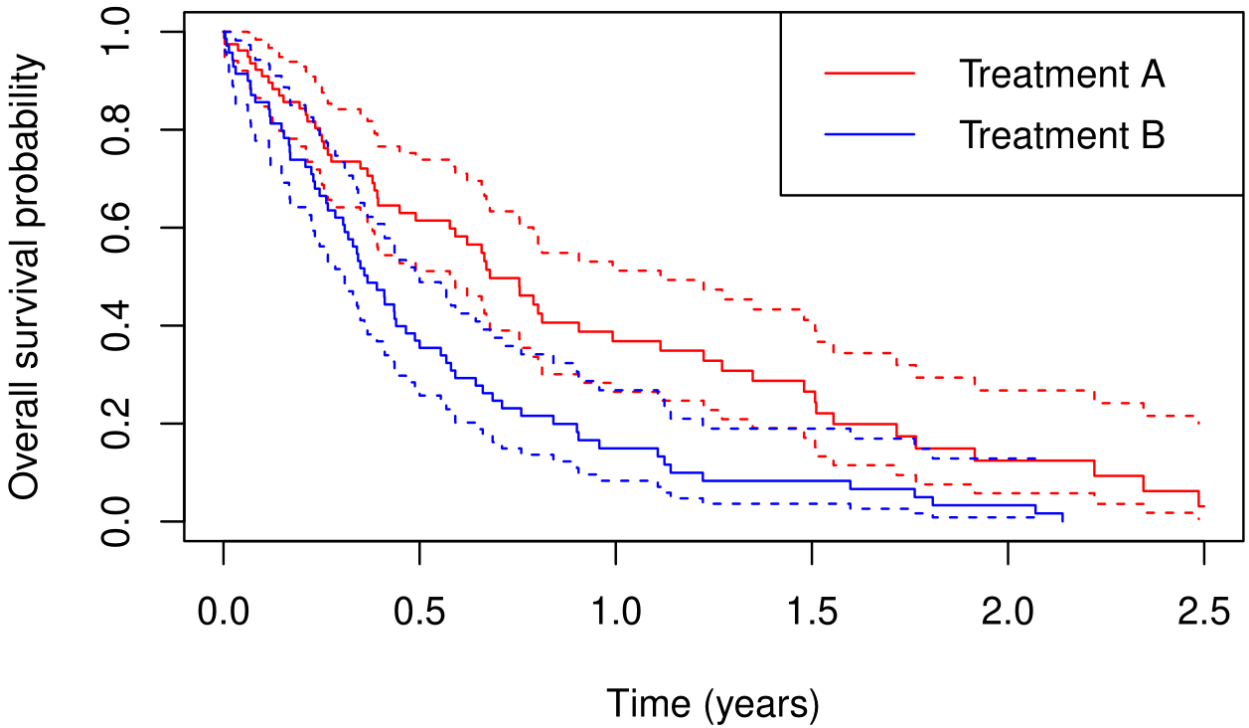
Survival curves (Kaplan-Meier)

The **survival curve** is one of several ways to describe and **display survival data**. It has the advantage of taking into account censoring over time.



The **horizontal axis** (x-axis) represents **time** and the **vertical axis** (y-axis) shows the **probability of surviving** (or the proportion of people surviving).

Survival curves - Example

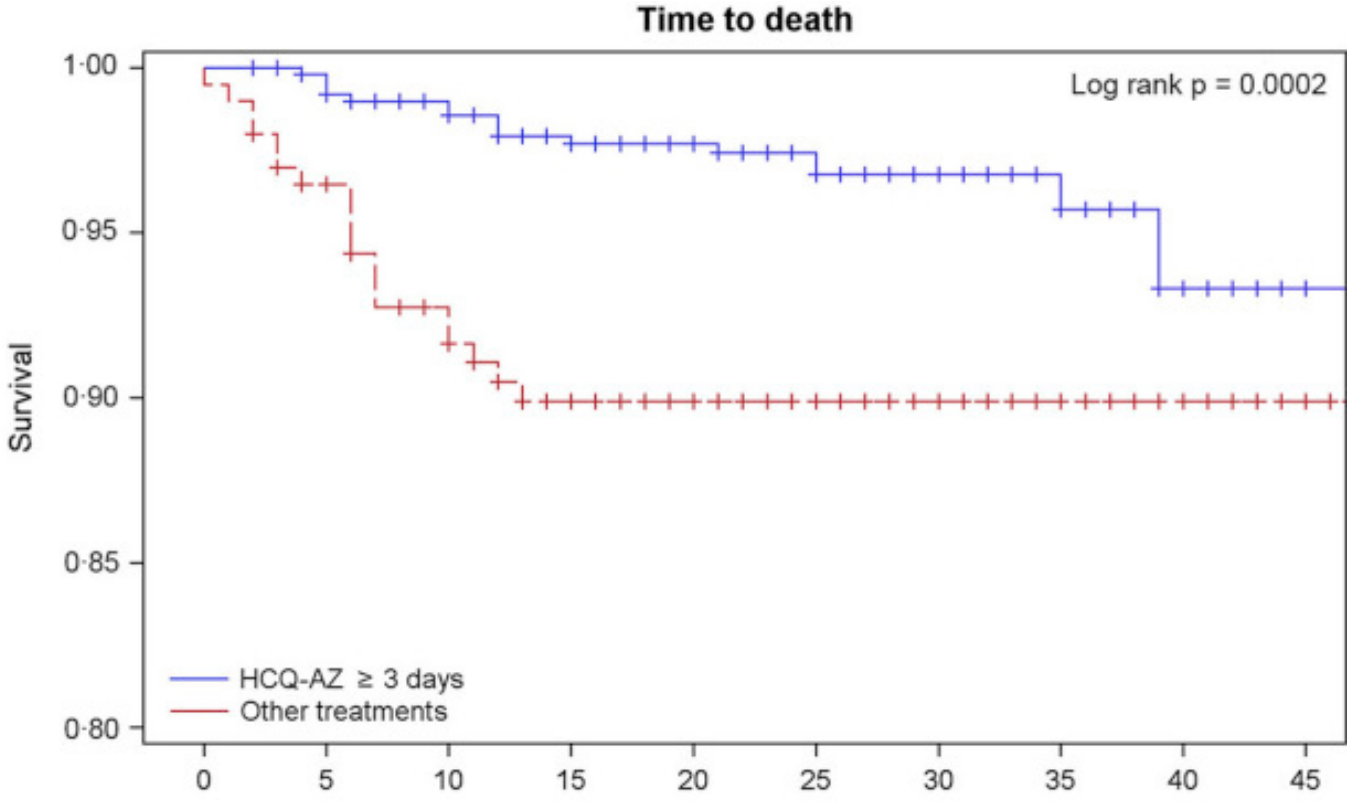


There appears to be a **survival advantage** for treatment A compared to treatment B. However, a **formal statistical test** is required to evaluate whether this difference is **statistically significant**.

The **log-rank test** is the most widely used method for **comparing two or more survival curves**.

Survival analysis: what could go wrong?

Example: “proof" of HCQ+AZ efficacy to treat COVID 19.



Number at risk	Days since inclusion									
HCQ-AZ ≥ 3 days	503	494	470	431	373	296	193	91	27	7
Other treatments	199	187	166	143	115	85	55	39	24	9

Survival analysis: what could go wrong?

- **Immortal time bias**: HCQ + AZ $\geq$ 3 days ; anyone dying in the 3 first days is assumed to be in the "other treatment group" => artificially overestimating the effect of the HCQ–AZI.

- **Imbalance between groups**: Randomization bias
  Age 18-44: 52.8% in HCQ+AZ vs. 36.4% in control
  Age >74: 3.6% in HCQ+AZ vs. 16% in control
  Patients with chronic heart diseases: 4% in HCQ+AZ vs. 15.2% in control
  etc.

  Moreover, patients with contraindications to HCQ or AZI were included in the control group (they should have been excluded)

Confusion bias & Simpson paradox 1/4

Example: choice of a **treatment** for cancer.

	Chemotherapy	*Surgery*
Cured	761	658
Not cured	239	342
Rate	76%	66%

Reasonable choice: **Chemotherapy**.
What happens if we account for the **tumors size**?

Confusion bias & Simpson paradox 2/4

Big tumors (>2cm):

	Chemotherapy	*Surgery*
Cured	90	564
Not cured	92	331
Rate	49%	63%

Small tumors (<=2cm):

	Chemotherapy	*Surgery*
Cured	671	94
Not cured	147	11
Rate	82%	90%

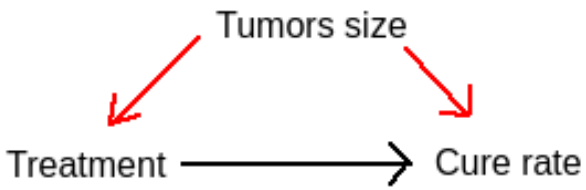
Surgery is better regardless of tumor size?!

Confusion bias & Simpson paradox 3/4

The fact that Chemotherapy looks better on average is an “illusion”.

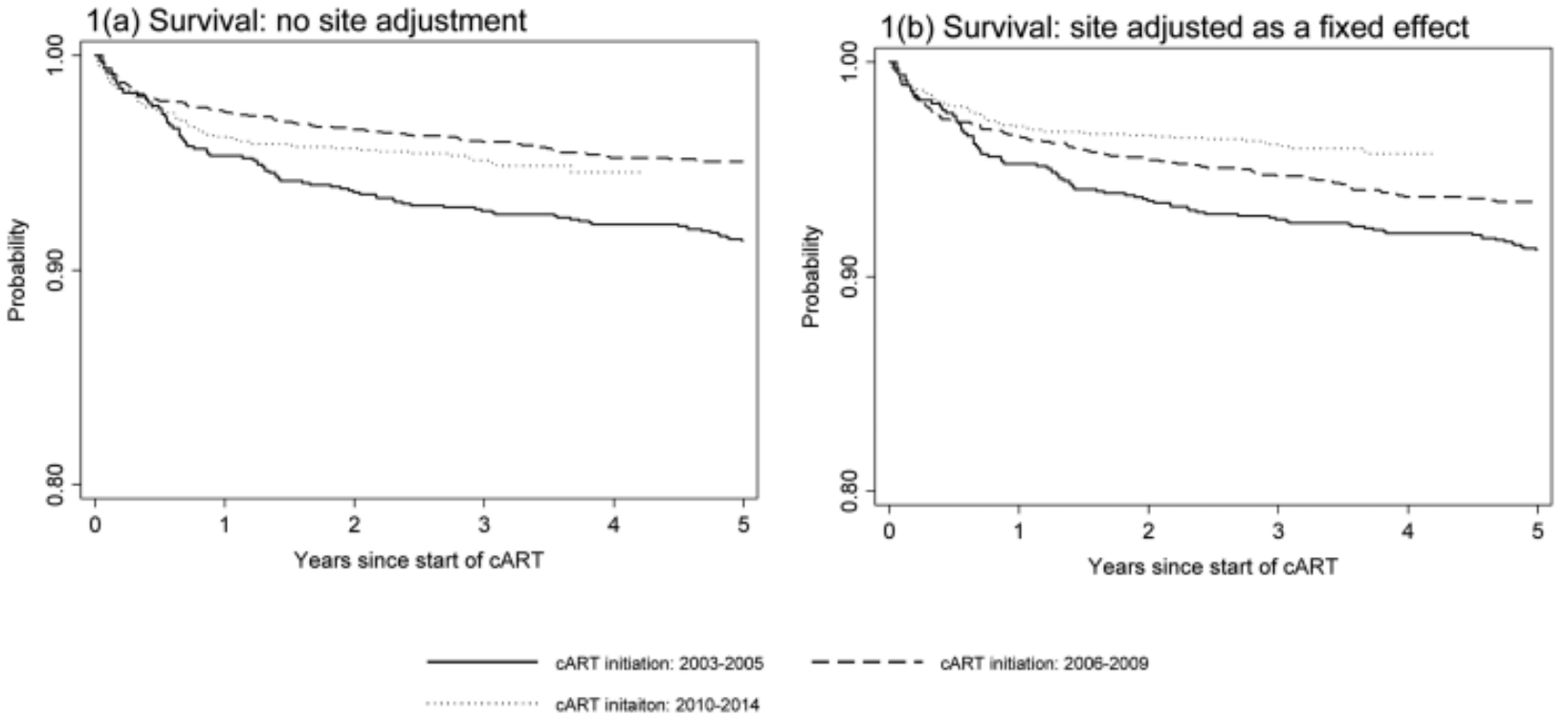
- **Big tumors** have a **lower cure rate** compared to small tumors, regardless of treatment
- **Surgery** is preferred for **big tumors**

Therefore, surgery has to deal with the “tough” cases more often.
=> Tumors size is a **confounder** when studying the association between treatment and cure rate.



Confusion bias & Simpson paradox 4/4

Example with survival curves: combination antiretroviral therapy in HIV observational cohorts



Survival regression models

The survival curves and the log-rank tests are **limited to categorical variables**. Moreover, the heterogeneity of the population within groups is ignored.
Regression models that describe **survival** as function of **explanatory variables** have been introduced.
Multiple independent prognosis factors can be analyzed simultaneously and **treatment differences** can be assessed while adjusting for **heterogeneity** and **imbalances in baseline characteristics**.

Survival regression models

The **shape of the distribution of survival times** justifies the requirement for **specific models**. (**positive** values, **skewed** shapes of distribution, **censoring**).
=> Statistical methods that rely on **normality** are not directly applicable and may produce **invalid results** with survival data.
Proportional hazards model: Most commonly used statistical model to **study the relationship** between the **survival time** of patients and **predictor variables**.

Survival regression models

The **shape of the distribution of survival times** justifies the requirement for **specific models**. (**positive** values, **skewed** shapes of distribution, **censoring**).
=> Statistical methods that rely on **normality** are not directly applicable and may produce **invalid results** with survival data.
Proportional hazards model: Most commonly used statistical model to **study the relationship** between the **survival time** of patients and **predictor variables**.

Proportional Hazards regression model

Proportional hazards regression model

The **PH regression model** can include both **categorical** and **quantitative** variables and study their **effect on the risk of event** (i.e., death).

$$\lambda_i(t) = \lambda_0(t) \exp\left(X_i(t)\,\gamma\right)$$

Baseline hazard $\lambda_0(t)$ = time-dependent intercept.
The rest of the equation is a **multiple linear regression** of the **logarithm** of the hazard on the variables X_i.
Interpretation (example):
Patients in the new treatment group at any time point during the study period were 10% [8% - 12%] less likely to die than patients in the control group.

Proportional hazards regression model

The **PH regression model** can include both **categorical** and **quantitative** variables and study their **effect on the risk of event** (i.e., death).

$$\lambda_i(t) = \lambda_0(t) \exp\left(X_i(t)\,\gamma\right)$$

Baseline hazard $\lambda_0(t)$ = time-dependent intercept.
The rest of the equation is a **multiple linear regression** of the **logarithm** of the hazard on the variables X_i.
Interpretation (example):
Patients in the new treatment group at any time point during the study period were 10% [8% - 12%] less likely to die than patients in the control group.

Proportional hazards regression model

The PH regression model can also answer **more complex questions**:

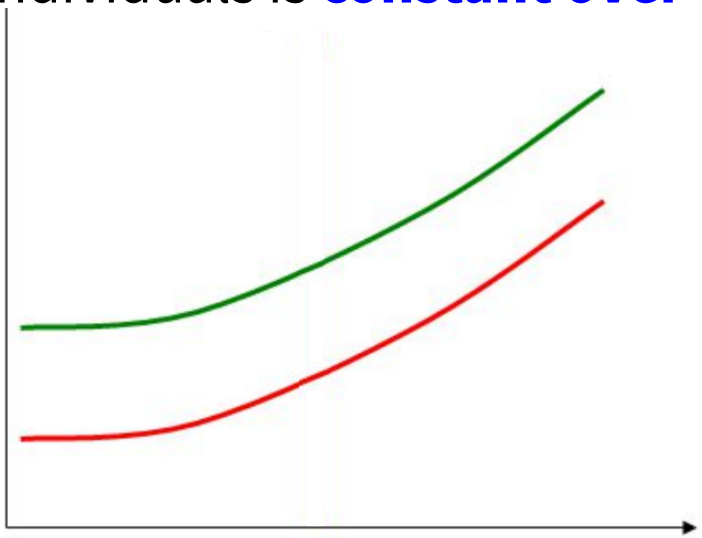
- Effect of weight on survival? (**continuous variable** which we do not want to categorize).
- Effect of treatment if we **control for other variables**? (e.g., weight, age, history of other diseases, etc.)
- Identify **specific subgroups** that benefit from the treatment.

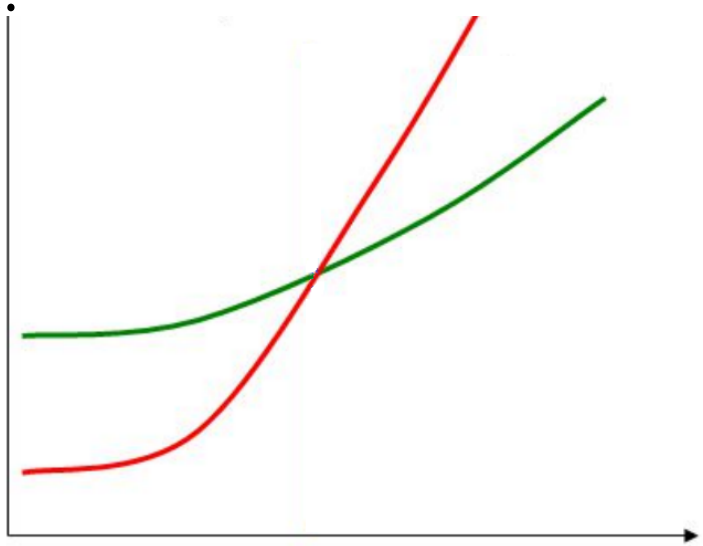
Assumptions:

- Non-informative missing data (censoring)
- Proportional hazards

Proportional hazards assumption

Proportional hazards assumption: The ratio of the hazards for any two individuals is **constant over time**.





We often use PH in health research because a new treatment tends to change the **overall mortality rate**, rather than change the pattern of mortality over time. However, the PH assumption is not always verified...

Survival regression models - Likelihood 1/2

Likelihood = **probability of the observed data** given a statistical model.
Observed data in survival:

1. Time for individuals for which **we observe the event**
2. Time for individuals that are **censored** (either during the follow-up or because the follow-up is over)

Therefore,

- Likelihood for 1: Probability of being **alive until time T** * probability of the **event at time T**
- Likelihood for 2: Probability of being **alive until time T**

Survival regression models - Likelihood 2/2

- **Hazard function:** Instantaneous **risk of event.**

$$\lambda(t) = \lim_{h \to 0^+} \frac{P(t \leq T < t + h | T > t)}{h}$$

- **Survival function:** Probability of being alive up to time t (i.e., dying after t).

$$S(t) = P(T > t) = \exp(-\int_0^t \lambda(u)du)$$

- **Likelihood for observed events**: Survival function $S(t)$ * Hazard function $\lambda(t)$
- **Likelihood for censored**: Survival function $S(t)$

Survival function for all individuals and hazard function only for those who observe the event:
Likelihood = $\prod_{i=1}^{N} S_i(t)\lambda_i(t)^{[Eventindicator]}$

Parametric baseline risk (Exponential)

$$\lambda_i(t) = \lambda_0(t) \exp\left(\beta_{drug} * drug_i + \beta_{sex} * sex_i\right)$$

Where $\lambda_0(t) = \lambda_0$ (constant risk)

Interpretation:

The effect of covariates can have an intuitive interpretation with **hazard ratios**. For example the effect of drug for the reference sex category:

$$HR = \frac{\lambda_i(t|drug = 1, sex = 0)}{\lambda_i(t|drug = 0, sex = 0)} = \exp\left(\beta_{drug}\right)$$

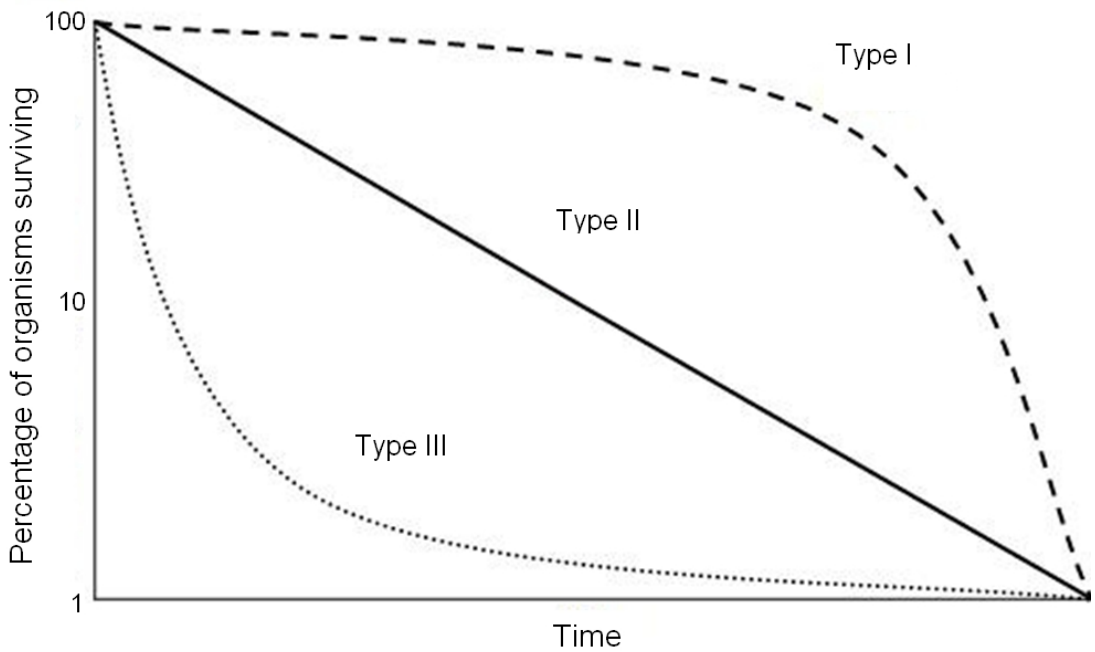
Covariates acts multiplicatively on the hazard ratio, not on the survival time.

- If **HR = 1**, then the hazard is the same regardless of treatment.
- If **HR**$>$**1**, then the risk of death increases with treatment (decrease if HR$<$1).
- **HR = 1.2**: "The risk of event is increased by 20% for patients who received the drug, compared to those who did not, given sex=0.

Parametric baseline risk (Weibull)

Weibull distribution allows the hazard to increase (or decrease) **proportionally with time**.

- **Shape $>$ 1:** Hazard increases the longer you survive (Type I)
- **Shape = 1:** Special case, exponential (Type II)
- **Shape $<$ 1:** Hazard decreases the longer you survive (Type III)



Semi-parametric approach using partial likelihood

If we are only interested in **hazard ratios**, we do not need to estimate the baseline risk, because:

$$\frac{\lambda_i(t)|x=1}{\lambda_i(t)|x=0} = \frac{\lambda_0(t)\exp(\beta_x * 1)}{\lambda_0(t)\exp(\beta_x * 0)} = \exp(\beta_x)$$

No survival curves in this case! (partial likelihood)
It is referred to as "semi-parametric" because the baseline hazard is left **completely unspecified** but the rest of the model is parametric in terms of its relationship to the unspecified baseline.

Smooth spline baseline hazard

In practice, **survival distributions are complex** and **not well fit by any common parametric distribution**, that's what makes the Cox proportional hazards model popular.
However, we may **need to estimate the baseline hazard function**, either for technical reasons (because the "partial likelihood" technique for the Cox PH model does not apply for more complex models) or for practical reasons (because we want to look at survival curves).
We can add more flexibility to the **baseline hazard** to let it **fluctuate with time** using functions like splines **BUT** we need to avoid overfitting as we aim for the population baseline risk, not the sample!
=> smooth functions to avoid **parametric assumptions**

Replicate a Cox model using Poisson regression

It is possible to write the **Cox PH** model as a **Poisson regression**.
Although the parameter estimates and standard errors are identical, the models are not technically identical. In the Poisson regression model we assume the **hazards are constant within event times**, an assumption that is not made with the Cox model.
However, it is possible to have more flexibility and replicate the Cox model by **splitting the follow-up in many small intervals**, and hence Poisson modeling of datasets with many records, each representing a small piece of the follow-up time for a person.

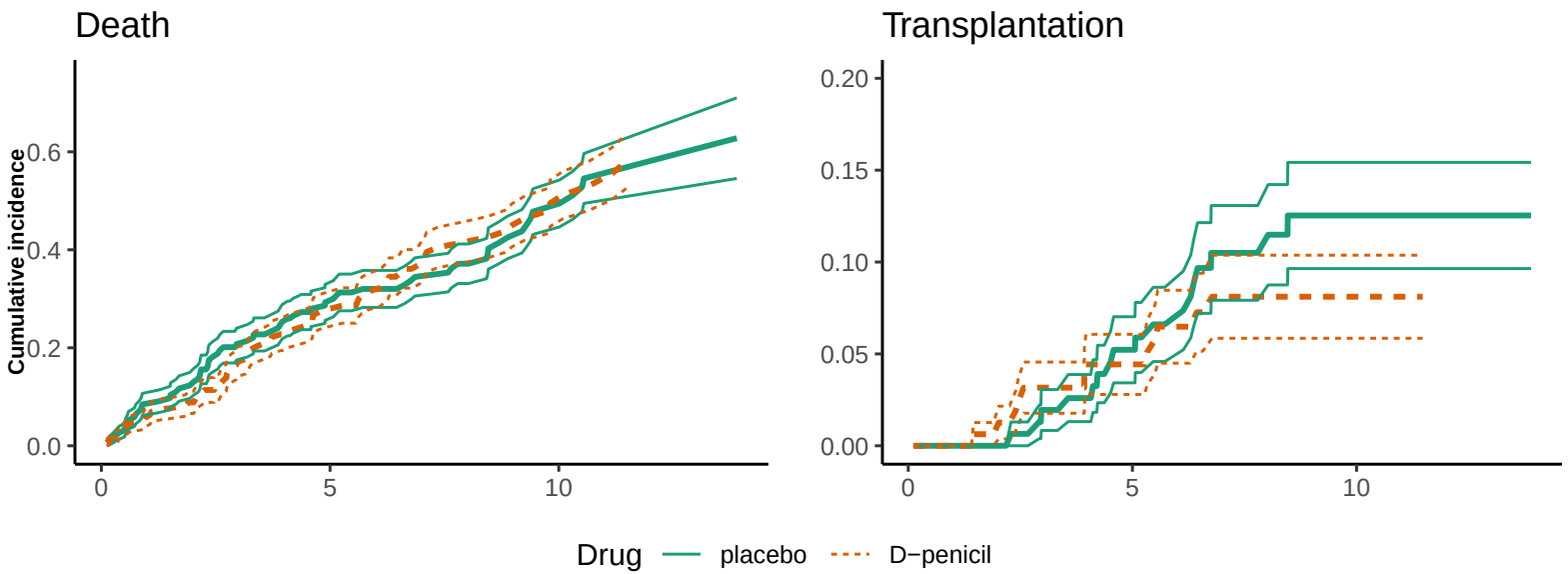
Semi-parametric proportional hazards

$$\lambda_i(t) = \lambda_0(t) \exp\left(\beta_{drug} * drug_i + \beta_{sex} * sex_i\right)$$

Where $\lambda_0(t)$ is fitted with **Bayesian smooth splines** (random walk order 1 or 2).
When the true **shape** of the baseline hazard is **unknown**, it is better to use semi-parametric approach.

Complex survival models

Competing risks



Competing risks! (i.e., an individual can only experience one event type: **death** or **liver transplantation**)

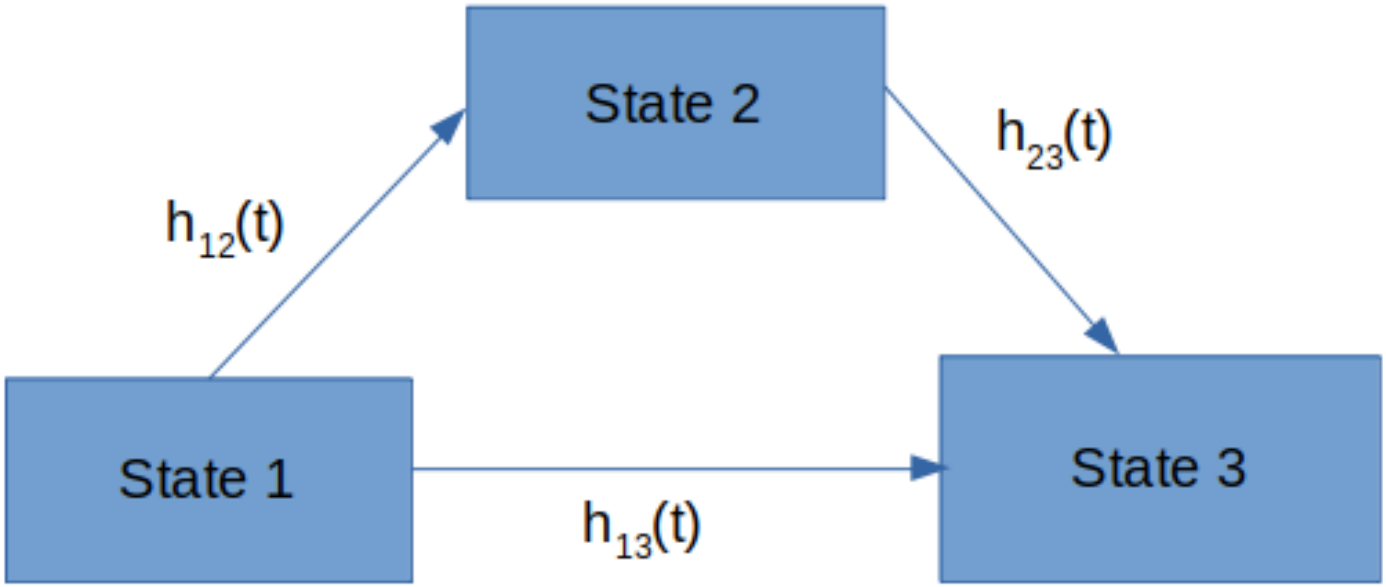
PBC data - Competing risks

id	years	death	transplantation	drug	sex
1	1.095170	1	0	D-penicil	female
2	14.152338	0	0	D-penicil	female
3	2.770781	1	0	D-penicil	male
4	5.270507	1	0	D-penicil	female
5	4.120578	0	1	placebo	female
6	6.853028	1	0	placebo	female
⋮	⋮	⋮	⋮	⋮	

Cause-specific hazard functions are modeled from a PH specification:

$$\begin{cases} \lambda_{i1}^{death}(t) & = \lambda_{01}(t)\exp\left(X_{i1}(t)^\top \gamma_1\right) \\ \lambda_{i2}^{transpl.}(t) & = \lambda_{02}(t)\exp\left(X_{i2}(t)^\top \gamma_2\right) \end{cases}$$

Multi-state model



$$\begin{cases} h_{i,12}(t) = h_{0,12}(t)\exp(\gamma_{12}X_i) & \text{(S1)} \\ h_{i,13}(t) = h_{0,13}(t)\exp(\gamma_{13}X_i) & \text{(S2)} \\ h_{i,23}(t) = h_{0,23}(t)\exp(\gamma_{23}X_i) & \text{(S3)} \end{cases}$$

Mixture cure model

bmt dataset (**smcure** R package)

The **bmt** dataset contains 91 patients with refractory acute lymphoblastic leukemia who participated in a bone marrow transplant study. Each patient has the following variables:

- **Time**: time to death (in days).
- **Status**: censoring indicator (1: if patient is uncensored; 0: otherwise).
- **TRT**: treatment group indicator (1: autologous; 0: allogeneic).

Mixture cure model

Model specification

Let Z be a binary random variable defined as 0 for susceptible and 1 for cured or immune individuals. The incidence submodel is given by:

$$Z \sim \text{Bernoulli}(\eta), \tag{1}$$

$$\text{logit}(\eta) = \beta_{C1} + \beta_{C2}\texttt{TRT}, \tag{2}$$

The latency submodel is expressed through a PH specification:

$$\lambda(t) = \lambda_0(t)\exp\left\{\beta_U \texttt{TRT}\right\}, \quad t > 0, \tag{3}$$

Shared frailty model

readmission dataset (**frailtypack** R package)
The **frailty model** can fit recurrent events with a gaussian distribution for the frailty term (i.e., lognormal frailty model).

$$\lambda_{i1}(t) = \lambda_{01}(t) \exp\left(\gamma_1 Chemo_i + \omega_i\right)$$

$$\omega_i \sim \mathcal{N}(\mu, \sigma^2)$$

Joint recurrent and terminal event

readmission dataset (**frailtypack** R package)
The **joint frailty model** deals with recurrent events with a frailty model and shares the frailty term in a proportional hazards model for a terminal event.
The parameter φ_1 is **scaling the frailty term in the terminal event submodel**.

$$\begin{cases} \lambda_{i1}(t) = \lambda_{01}(t)\exp\left(\gamma_{11}Chemo_i + \omega_i\right) \text{ (S1)} \\ \lambda_{i2}(t) = \lambda_{02}(t)\exp\left(\gamma_{12}Chemo_i + \varphi_1\omega_i\right) \text{(S2)} \end{cases}$$

$$\omega_i \sim log\mathcal{N}(\mu, \sigma^2)$$

Software packages in R

Maximum likelihood - *survreg, frailtypack*
Bayesian inference - MCMC-based *bayesSurv*, *INLA*

Time-dependent covariates in survival models

Time-dependent covariates

It is possible to include **time-dependent covariates** in a survival model, they can be classified into two categories:

- **Exogeneous** (or external) covariates remain measurable and their distribution is unchanged after the occurrence of the event.
- **Endogeneous** (or internal) covariates' distribution is affected by the event.

The proportional hazards model can handle exogeneous time-dependent covariates but the likelihood requires **knowing the value of these covariates for all subjects at risk for each event time**.

Time-dependent covariates

When covariates measurements does not coincide with event times in the sample, **models are required to impute values** at the times of events.
Likelihood = $\prod_{i=1}^{N} S_i(t)\lambda_i(t)^{[Eventindicator]}$
where $S(t) = \exp(-\int_0^t \lambda(u)du)$ and $\lambda_i(t) = \lambda_0(t)\exp(X_i(t)\gamma)$
However, most biomarkers of interest in clinical research are endogeneous variables, their **values are affected** by a **change in the risk** of occurrence of the event.
For example in a cancer clinical trial, if a **treatment reduces** both the **risk of death** and the **size of tumors**, adjusting a survival model on the tumors size may severely **bias the effect of treatment** on the risk of death.

Longitudinal models - Mixed effects

```
# Longitudinal data
id year serBilir drug
1 0.0000000 14.5 D-penicil
1 0.5256817 21.3 D-penicil
2 0.0000000 1.1 D-penicil
2 0.4983025 0.8 D-penicil
2 0.9993429 1.0 D-penicil
2 2.1027270 1.9 D-penicil
2 4.9008871 2.6 D-penicil
2 5.8892783 3.6 D-penicil
2 6.8858833 4.2 D-penicil
2 7.8907020 3.6 D-penicil
```

Longitudinal models - Mixed effects

Example: mixed-effects regression model for a longitudinal biomarker.

$$Y_{ij} = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11})year_{ij} + \beta_{drug} * drug_i + \varepsilon_{ij}$$

- $Y1$: longitudinal outcome (continuous - lognormal)
- β_0: fixed intercept
- β_1: fixed slope
- β_2: fixed effect of drug
- b_{i0}: random individual intercept
- b_{i1}: random individual slope
- ε_{ij}: residual error

Two longitudinal markers

$$\begin{cases} Y_{i1j} & = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11}) year_{ij} + \varepsilon_{ij} \\ \mathrm{E}[\log(Y_{i2j})] & = \beta_{20} + b_{i20} + (\beta_{21} + b_{i21}) year_{ij} \end{cases}$$

- $Y1$: longitudinal outcome (continuous - Gaussian)
- $Y2$: second longitudinal outcome (count - Poisson)
- β_0: fixed intercept
- β_1: fixed slope
- b_{i0}: random individual intercept
- b_{i1}: random individual slope
- ε_{ij}: residual error

The two submodels are linked by **correlated random effects**.

Software packages in R

Maximum likelihood - *lme4,geepack*
Bayesian inference - MCMC-based *BayesRGMM*, *INLA*

Joint models

Regression models

Regression: Effect of covariates on an outcome (continuous, counts, binary, ...).
Example: Tumors size in cancer clinical trials (outcome)
=> Effect of treatment on tumors size?
Survival analysis: time-to-event.
Example: time to death in cancer clinical trials.
Survival regression model: Evaluate the effect of covariates on the risk of event.
Example: Effect of treatment on the risk of death?

Joint modeling - underlying idea

Can we have a **statistical model** that evaluates **simultaneously** the effect of treatment on the tumors size and the risk of death?
Useful because when death occurs => no more measurements of tumors size!
Maybe we never observe big tumors because a patient with big tumors die (**informative censoring**).
The risk of death for a patient with big tumor is higher compared to a patient with small tumors (**recorded heterogeneity of the population**).

Joint modeling - underlying idea

Joint modeling consists in **simultaneously** modeling **multiple outcomes** while taking into account their **association**. The outcomes are either **longitudinal** or **time to an event** of interest (e.g., death in health research). Joint models are popular in health research because it is common to observe **longitudinal markers censored by a terminal event** with interest in the analysis of the longitudinal marker trajectory and the risk of the terminal event as well as their relationship.

Joint modeling - Examples

A few examples of recent applications of joint models:

- Cancer tumor dynamics and the risk of death
- Analysis of CD4 lymphocytes counts and AIDS survival
- Prostate-specific antigen dynamics and the risk of cancer recurrence
- Dynamics of aortic gradient and aortic regurgitations and their relationship with the competing risks of death or reoperation in the field of cardiac surgery
- Cognitive markers's relationship with the time to onset of Alzheimer's disease
- Jointly modeling forest fires ignition, number of fires and the proportion of burned area in mainland Portugal aggregated by years and regions

Joint modeling Longitudinal and Survival data

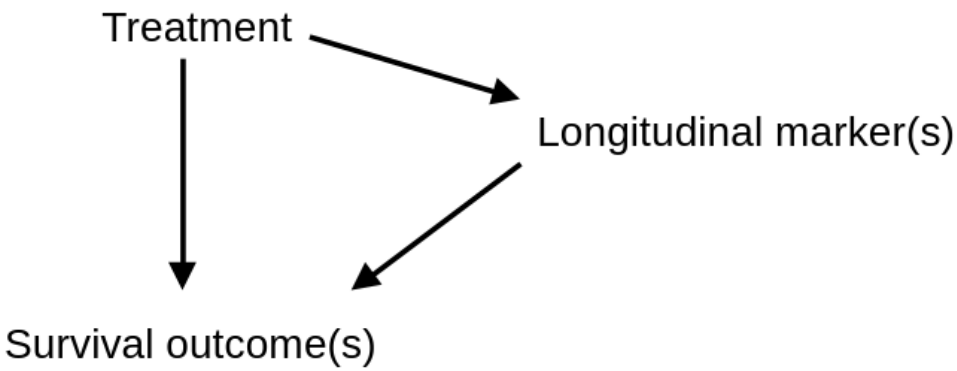
- Can **efficiently utilize all available information** to limit costs and optimize outcomes.
- Dynamic risk predictions for personalized care.
- Have interesting properties for **mediation analysis** (decomposition of a treatment effect into **direct and indirect effects** (could be a promising tool in the shift toward causality in clinical trials evaluation, see ICH E9 addendum recommendations).
- Evaluation of **surrogate markers**.

Joint modeling Longitudinal and Survival data

- Can **efficiently utilize all available information** to limit costs and optimize outcomes.
- Dynamic risk predictions for personalized care.
- Have interesting properties for **mediation analysis** (decomposition of a treatment effect into **direct and indirect effects** (could be a promising tool in the shift toward causality in clinical trials evaluation, see ICH E9 addendum recommendations).
- Evaluation of **surrogate markers**.

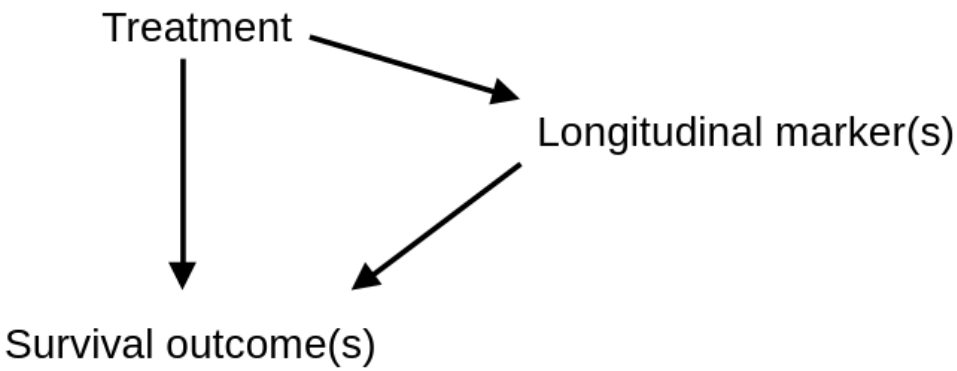
Joint modeling Longitudinal and Survival data

- Can **efficiently utilize all available information** to limit costs and optimize outcomes.
- Dynamic risk predictions for personalized care.
- Have interesting properties for **mediation analysis** (decomposition of a treatment effect into **direct and indirect effects** (could be a promising tool in the shift toward causality in clinical trials evaluation, see ICH E9 addendum recommendations).
- Evaluation of **surrogate markers**.



Joint modeling Longitudinal and Survival data

- Can **efficiently utilize all available information** to limit costs and optimize outcomes.
- Dynamic risk predictions for personalized care.
- Have interesting properties for **mediation analysis** (decomposition of a treatment effect into **direct and indirect effects** (could be a promising tool in the shift toward causality in clinical trials evaluation, see ICH E9 addendum recommendations).
- Evaluation of **surrogate markers**.



Parameterizations

$$\begin{cases} Y_{ij} = \eta_{ij}(year_{ij}) + \varepsilon_{ij} = \beta_{10} + b_{i0} + (\beta_1 + b_{i1})year_{ij} + \varepsilon_{ij} \\ \lambda_i(t) = \lambda_0(t)\exp\left(b_{i0}\varphi_1 + b_{i1}\varphi_2\right) \end{cases}$$

- Y: longitudinal outcome (continuous - Gaussian)
- $\lambda_i(t)$: Terminal event risk
- $\lambda_0(t)$: Baseline risk
- β_0: fixed intercept
- b_{i0}: random individual intercept
- b_{i1}: random individual slope
- β_1: fixed slope
- ε_{ij}: residual error
- φ: association parameters

Association structure

Joint Model - Individual deviation association

$$\begin{cases} Y_{ij} = \eta_{ij}(year_{ij}) + \varepsilon_{ij} = \beta_0 + b_{i0} + (\beta_1 + b_{i1})year_{ij} + \varepsilon_{ij} \\ \lambda_i(t) = \lambda_0(t)\exp\left((b_{i0} + b_{i1}t)\varphi\right) \end{cases}$$

- Y: longitudinal outcome (continuous - Gaussian)
- $\lambda_i(t)$: Terminal event risk
- $\lambda_0(t)$: Baseline risk
- β_0: fixed intercept
- b_{i0}: random individual intercept
- b_{i1}: random individual slope
- β_1: fixed slope
- ε_{ij}: residual error
- φ: association parameter

Joint Model - Current value association

$$\begin{cases} Y_{ij} = \eta_{ij}(year_{ij}) + \varepsilon_{ij} = \beta_0 + b_{i0} + (\beta_1 + b_{i1})year_{ij} + \varepsilon_{ij} \\ \lambda_i(t) = \lambda_0(t)\exp\left(\eta_{ij}(t)\varphi\right) \end{cases}$$

- Y: longitudinal outcome (continuous - Gaussian)
- $\lambda_i(t)$: Terminal event risk
- $\lambda_0(t)$: Baseline risk
- β_0: fixed intercept
- b_{i0}: random individual intercept
- b_{i1}: random individual slope
- β_1: fixed slope
- ε_{ij}: residual error
- φ: association parameter

Joint Model - Current slope association

$$\begin{cases} Y_{ij} = \eta_{ij}(year_{ij}) + \varepsilon_{ij} = \beta_0 + b_{i0} + (\beta_1 + b_{i1})year_{ij} + \varepsilon_{ij} \\ \lambda_i(t) = \lambda_0(t)\exp\left(\eta_{ij}(t)'\varphi\right) \end{cases}$$

- Y: longitudinal outcome (continuous - Gaussian)
- $\lambda_i(t)$: Terminal event risk
- $\lambda_0(t)$: Baseline risk
- β_0: fixed intercept
- b_{i0}: random individual intercept
- b_{i1}: random individual slope
- β_1: fixed slope
- ε_{ij}: residual error
- φ: association parameter

Software packages in R

Maximum likelihood - *JM, joineRML, frailtypack,*
Bayesian inference - MCMC-based *JMbayes2*

Lack of efficient algorithms!

For frequentist inference the multiple random effects included in joint models needs to be integrated out in the likelihood => **multidimensional integral** that requires numerical approximation.

Iterative algorithms (e.g., Newton-like, Monte-Carlo) have **slow convergence properties**, joint modeling has been **limited** so far by the available inference techniques and associated statistical software.

Hickey, G. L., Philipson, P., Jorgensen, A., Kolamunnage-Dona, R. *Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues.* BMC medical research methodology, 16(1), 1-15. (2016)

Lack of efficient algorithms!

For frequentist inference the multiple random effects included in joint models needs to be integrated out in the likelihood => **multidimensional integral** that requires numerical approximation.
Iterative algorithms (e.g., Newton-like, Monte-Carlo) have **slow convergence properties**, joint modeling has been **limited** so far by the available inference techniques and associated statistical software.

Hickey, G. L., Philipson, P., Jorgensen, A., Kolamunnage-Dona, R. *Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues.* BMC medical research methodology, 16(1), 1-15. (2016)

INLA

Integrated Nested Laplace Approximations (INLA)

INLA is a **deterministic alternative** to the Markov chain Monte Carlo sampling methods for Bayesian inference of latent Gaussian models.

$\rightarrow$ Approximate Bayesian inference **without sampling**.

A latent Gaussian model is where the data $\boldsymbol{Y}_i$ depends on the latent field χ only through the linear predictor η_i and the latent field has a **Gaussian distribution** (with sparse precision).

$$\pi(\chi, \boldsymbol{\theta}) \propto \pi(\boldsymbol{y}|\chi, \boldsymbol{\theta})\pi(\chi|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$$

Likelihood (any distribution), latent field (Gaussian), prior (any distribution).

Integrated Nested Laplace Approximations (INLA)

INLA is a **deterministic alternative** to the Markov chain Monte Carlo sampling methods for Bayesian inference of latent Gaussian models.
$\rightarrow$ Approximate Bayesian inference **without sampling**.

A latent Gaussian model is where the data $\boldsymbol{Y}_i$ depends on the latent field χ only through the linear predictor η_i and the latent field has a **Gaussian distribution** (with sparse precision).

$$\pi(\chi, \boldsymbol{\theta}) \propto \pi(\boldsymbol{y}|\chi, \boldsymbol{\theta})\pi(\chi|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$$

Likelihood (any distribution), latent field (Gaussian), prior (any distribution).

Integrated Nested Laplace Approximations (INLA)

INLA is a **deterministic alternative** to the Markov chain Monte Carlo sampling methods for Bayesian inference of latent Gaussian models.
$\rightarrow$ Approximate Bayesian inference **without sampling**.
A latent Gaussian model is where the data $\boldsymbol{Y}_i$ depends on the latent field $\boldsymbol{\chi}$ only through the linear predictor η_i and the latent field has a **Gaussian distribution** (with sparse precision).

$$\pi(\boldsymbol{\chi}, \boldsymbol{\theta}) \propto \pi(\boldsymbol{y}|\boldsymbol{\chi}, \boldsymbol{\theta})\pi(\boldsymbol{\chi}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$$

Likelihood (any distribution), latent field (Gaussian), prior (any distribution).

Integrated Nested Laplace Approximations (INLA)

INLA is a **deterministic alternative** to the Markov chain Monte Carlo sampling methods for Bayesian inference of latent Gaussian models.
$\rightarrow$ Approximate Bayesian inference **without sampling**.
A latent Gaussian model is where the data $\boldsymbol{Y}_i$ depends on the latent field $\boldsymbol{\chi}$ only through the linear predictor η_i and the latent field has a **Gaussian distribution** (with sparse precision).

$$\pi(\boldsymbol{\chi}, \boldsymbol{\theta}) \propto \pi(\boldsymbol{y}|\boldsymbol{\chi}, \boldsymbol{\theta})\pi(\boldsymbol{\chi}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$$

Likelihood (any distribution), latent field (Gaussian), prior (any distribution).

Integrated Nested Laplace Approximations (INLA)

Many models fit with the LGM framework and thus can be fitted with INLA:

- **Survival models** (AFT, Cox models, competing risks, multi-state, cure models, frailty, censoring and truncation, parametric and nonparametric hazards)
- **Longitudinal models** (GLMM, zero-inflated models, proportional odds, ...)
- **Spatial and spatio-temporal models** (CAR, ICAR, MCAR, SPDE)
- **Joint models that include multiple longitudinal and survival outcomes**
- and many more...

With a few exceptions (e.g., PK/PD models using ODE or SDE without analytical solution).

Integrated Nested Laplace Approximations (INLA)

Many models fit with the LGM framework and thus can be fitted with INLA:

- **Survival models** (AFT, Cox models, competing risks, multi-state, cure models, frailty, censoring and truncation, parametric and nonparametric hazards)
- **Longitudinal models** (GLMM, zero-inflated models, proportional odds, ...)
- **Spatial and spatio-temporal models** (CAR, ICAR, MCAR, SPDE)
- **Joint models that include multiple longitudinal and survival outcomes**
- and many more...

With a few exceptions (e.g., PK/PD models using ODE or SDE without analytical solution).

Integrated Nested Laplace Approximations (INLA)

Many models fit with the LGM framework and thus can be fitted with INLA:

- **Survival models** (AFT, Cox models, competing risks, multi-state, cure models, frailty, censoring and truncation, parametric and nonparametric hazards)
- **Longitudinal models** (GLMM, zero-inflated models, proportional odds, ...)
- **Spatial and spatio-temporal models** (CAR, ICAR, MCAR, SPDE)
- **Joint models that include multiple longitudinal and survival outcomes**
- and many more...

With a few exceptions (e.g., PK/PD models using ODE or SDE without analytical solution).

Integrated Nested Laplace Approximations (INLA)

Many models fit with the LGM framework and thus can be fitted with INLA:

- **Survival models** (AFT, Cox models, competing risks, multi-state, cure models, frailty, censoring and truncation, parametric and nonparametric hazards)
- **Longitudinal models** (GLMM, zero-inflated models, proportional odds, ...)
- **Spatial and spatio-temporal models** (CAR, ICAR, MCAR, SPDE)
- **Joint models that include multiple longitudinal and survival outcomes**
- and many more...

With a few exceptions (e.g., PK/PD models using ODE or SDE without analytical solution).

Integrated Nested Laplace Approximations (INLA)

Many models fit with the LGM framework and thus can be fitted with INLA:

- **Survival models** (AFT, Cox models, competing risks, multi-state, cure models, frailty, censoring and truncation, parametric and nonparametric hazards)
- **Longitudinal models** (GLMM, zero-inflated models, proportional odds, ...)
- **Spatial and spatio-temporal models** (CAR, ICAR, MCAR, SPDE)
- **Joint models that include multiple longitudinal and survival outcomes**
- and many more...

With a few exceptions (e.g., PK/PD models using ODE or SDE without analytical solution).

Integrated Nested Laplace Approximations (INLA)

Many models fit with the LGM framework and thus can be fitted with INLA:

- **Survival models** (AFT, Cox models, competing risks, multi-state, cure models, frailty, censoring and truncation, parametric and nonparametric hazards)
- **Longitudinal models** (GLMM, zero-inflated models, proportional odds, ...)
- **Spatial and spatio-temporal models** (CAR, ICAR, MCAR, SPDE)
- **Joint models that include multiple longitudinal and survival outcomes**
- and many more...

With a few exceptions (e.g., PK/PD models using ODE or SDE without analytical solution).

Integrated Nested Laplace Approximations (INLA)

Many models fit with the LGM framework and thus can be fitted with INLA:

- **Survival models** (AFT, Cox models, competing risks, multi-state, cure models, frailty, censoring and truncation, parametric and nonparametric hazards)
- **Longitudinal models** (GLMM, zero-inflated models, proportional odds, ...)
- **Spatial and spatio-temporal models** (CAR, ICAR, MCAR, SPDE)
- **Joint models that include multiple longitudinal and survival outcomes**
- and many more...

With a few exceptions (e.g., PK/PD models using ODE or SDE without analytical solution).

Integrated Nested Laplace Approximations (INLA)

This **computationally efficient** and **accurate** algorithm is implemented in the freely available open-source R package *INLA*.

R-INLA project (`www.r-inla.org`)

Rue, Håvard, Sara Martino, and Nicolas Chopin. *Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations.* Journal of the royal statistical society: Series b (statistical methodology) 71.2 (2009): 319-392.
Van Niekerk, Janet, Elias Krainski, Denis Rustand, and Håvard Rue. "A new avenue for Bayesian inference with INLA." Computational Statistics & Data Analysis 181 (2023): 1-14

Comparison with MLE, STAN and JMbayes2

Comparison with Levenberg-Marquardt algorithm (Newton-like algorithm):
Rustand D, van Niekerk J, Rue H, Tournigand C, Rondeau V, Briollais L. *Bayesian estimation of two-part joint models for a longitudinal semicontinuous biomarker and a terminal event with INLA: Interests for cancer clinical trial evaluation.* Biometrical Journal. 2023 Apr;65(4):2100322.

Comparison with MCMC and MCEM:
Rustand, D., van Niekerk, J., Krainski, E. T., Rue, H., Proust-Lima, C. *Fast and flexible inference for joint models of multivariate longitudinal and survival data using integrated nested Laplace approximations*, Biostatistics, 2023 Aug; kxad019, https://doi.org/10.1093/biostatistics/kxad019.

R package: INLAjoint

We recently developed a **flexible** and **user-friendly** interface to fit models for survival and/or longitudinal data with INLA. The **INLAjoint** package allows to build various models with a lot of flexibility, using its main function ***joint()***:

- Univariate and multivariate **survival** outcomes (i.e., frailty models, competing risks, multi-state) fitted with **proportional hazards models**
- Univariate and multivariate **longitudinal** outcomes fitted with **mixed effects models**
- Multivariate longitudinal and survival outcomes

R: Example with **INLAjoint**

Rustand, D., Teixeira Krainski, E., Rue, H., van Niekerk, J. INLAjoint: Multivariate joint modeling for longitudinal and time-to-event outcomes with INLA.
https://github.com/DenisRustand/INLAjoint

Examples in R

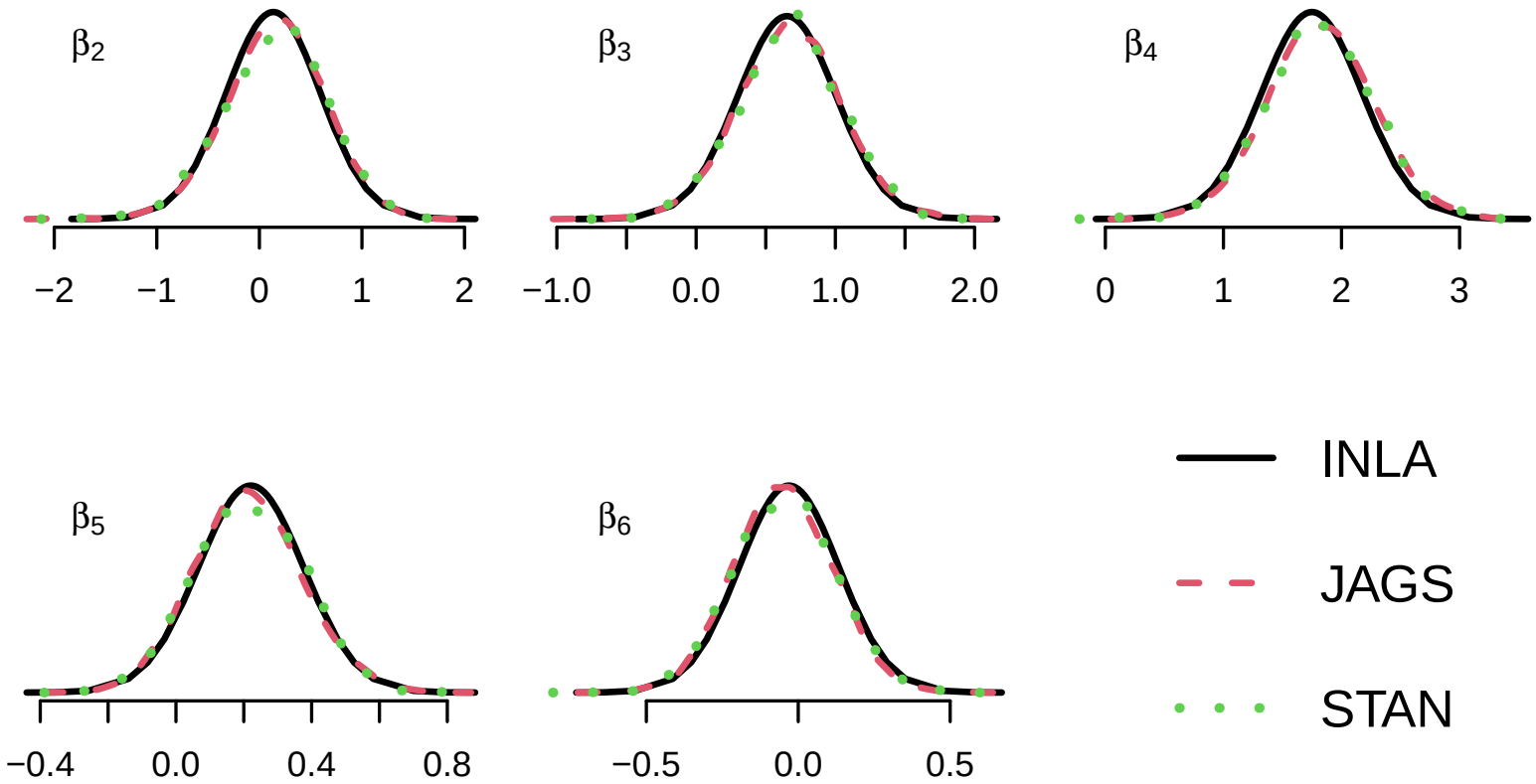
PBC data analysis

- Parametric and semi-parametric survival analysis
- Competing risks model for transplant and death
- Joint model using serum bilirubin
- Joint competing risks model using serum bilirubin
- Bivariate joint model (+competing risks)
- Multivariate joint model (+competing risks)

Survival analysis

$$\lambda_i(t) = \lambda_0(t) \exp\left(\beta_{drug} * drug_i + \beta_{sex} * sex_i\right)$$

Semi-parametric proportional hazards

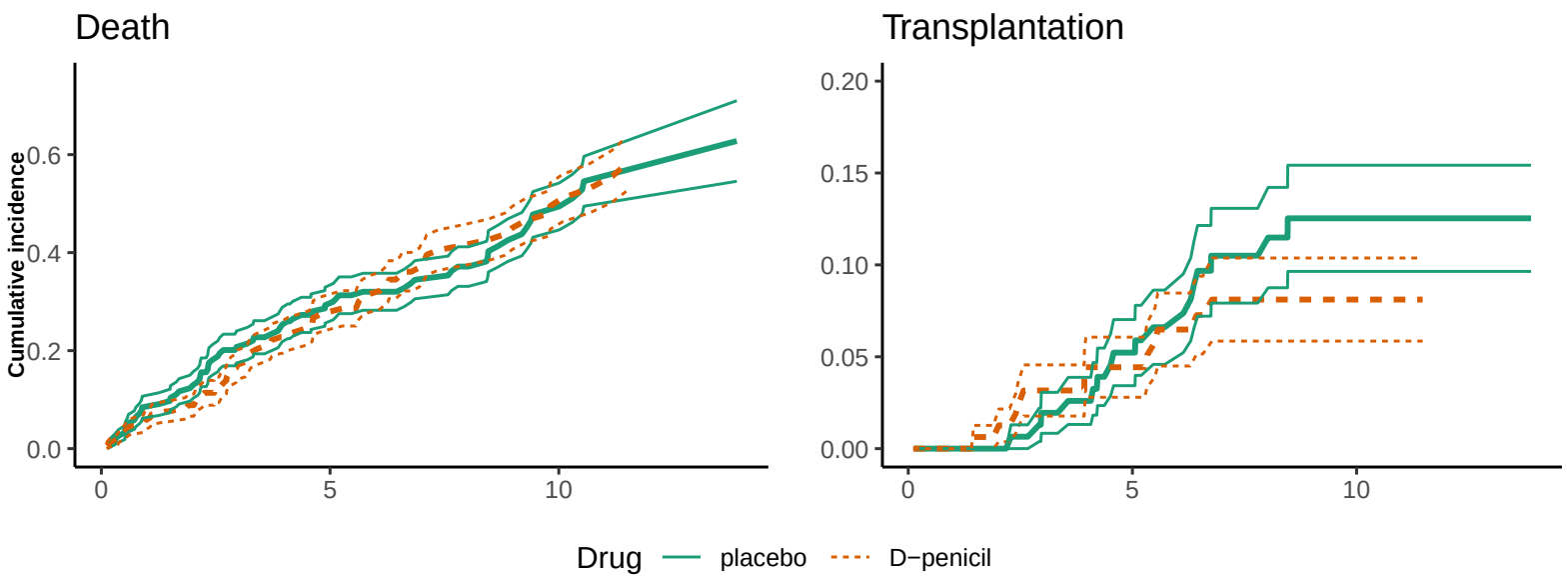


Computation time: INLA = 0.6sec. ; JAGS = 24sec. ; STAN = 67sec.

JAGS: 4 chains, 1000 burn-in and 4000 iterations

STAN: 4 chains, 1000 burn-in and 1000 iterations

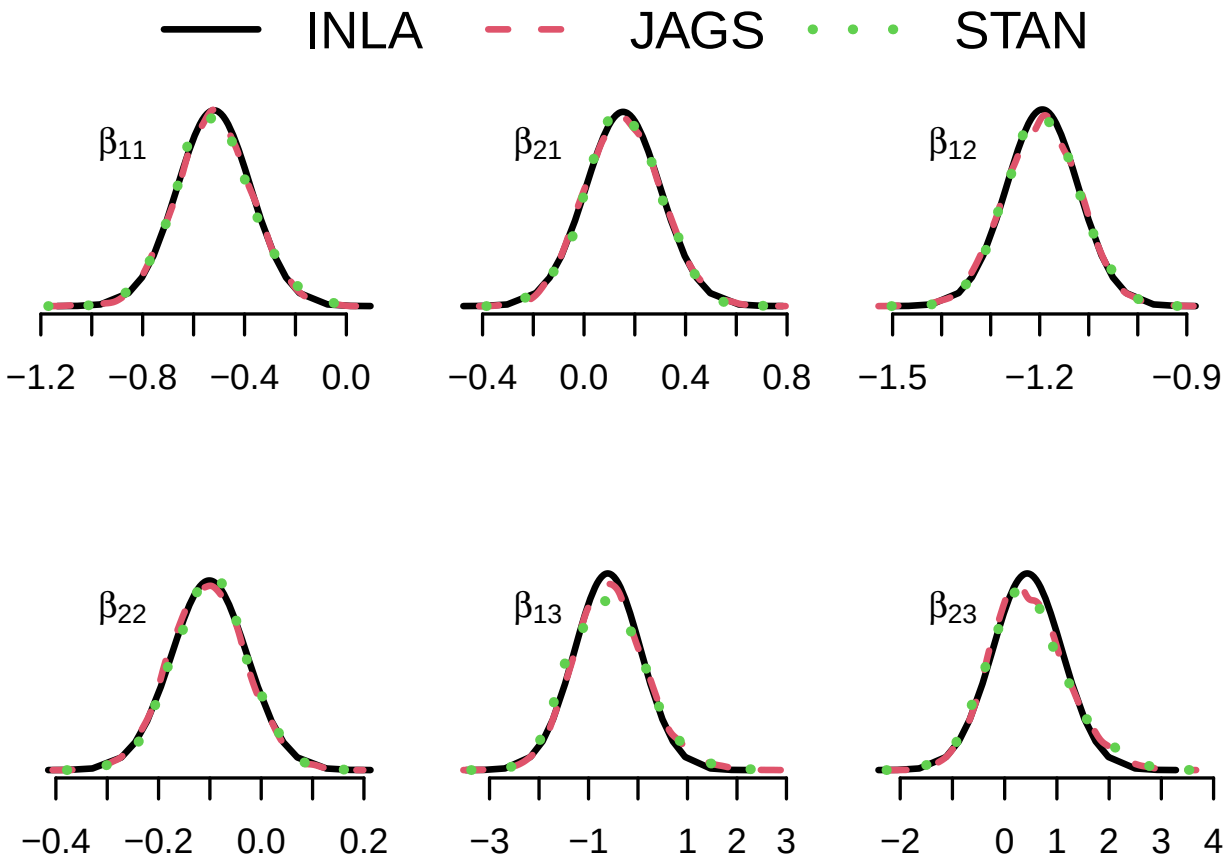
Competing risks



Cause-specific hazard functions are modelled from a PH specification:

$$\begin{cases} \lambda_{i1}^{death}(t) & = \lambda_{01}(t)\exp\left(X_{i1}(t)^\top \gamma_1\right) \\ \lambda_{i2}^{transpl.}(t) & = \lambda_{02}(t)\exp\left(X_{i2}(t)^\top \gamma_2\right) \end{cases}$$

Competing risks



Computation time: INLA = 1.3sec. ; JAGS = 1272sec. ; STAN = 350sec.

JAGS: 4 chains, 1000 burn-in and 4000 iterations

STAN: 4 chains, 1000 burn-in and 1000 iterations

Longitudinal models - Mixed effects

Example: mixed-effects regression model for a serum bilirubin

$$Y_{ij} = \beta_{10} + b_{i10} + (\beta_{11} + b_{i11})year_{ij} + \beta_{drug} * drug_i + \varepsilon_{ij}$$

- $Y1$: longitudinal outcome (continuous - lognormal)
- β_0: fixed intercept
- β_1: fixed slope
- β_2: fixed effect of drug
- b_{i0}: random individual intercept
- b_{i1}: random individual slope
- ε_{ij}: residual error

Joint modeling 1/2

Association: **Shared random effects**

$$\begin{cases} Y_{ij} = X_{ij}^\top \beta + Z_{ij}^\top \boldsymbol{b}_i + \epsilon_{ij} & (\text{Biomarker}) \\ \lambda_i(t) = \lambda_0(t) \exp\left(X_i(t)\,\gamma + \varphi^\top \boldsymbol{b}_i\right) & (\text{Time-to-event}) \end{cases}$$

We only share the variability not explained by fixed effects (therefore adjusting the survival model on population heterogeneity not related to covariates of interest).

Advantages: Hazard ratio interpretation ; account for unobserved confounders.

Joint modeling 2/2

Association: **Current value**

$$\begin{cases} Y_{ij} = \eta_{ij} = X_{ij}^\top \beta + Z_{ij}^\top \boldsymbol{b}_i + \epsilon_{ij} & (\text{Biomarker}) \\ \lambda_i(t) = \lambda_0(t) \exp\left(X_i(t)\,\gamma + \varphi^\top \eta_i(t)\right) & (\text{Time-to-event}) \end{cases}$$

The risk of event depends on the individual value of the longitudinal marker.

Similar to shared random effects but shares also fixed effects.

Joint model for longitudinal and competing risks

$$
\begin{cases}
\begin{aligned}
\log(serBilir_{ij}) &= \eta_i(t_{ij}) + \varepsilon_{ij} && \text{(L1)}\\
&= \beta_0 + b_{i0} + (\beta_1 + b_{i1})year_{ij} + \beta_2 drug_i + \beta_3 sex_i \\
&\quad + \beta_4 year_{ij} drug_i + \beta_5 year_{ij} sex_i + \varepsilon_{ij}
\end{aligned}\\
\lambda_{i1}^{death}(t) = \lambda_{01}(t)\exp\left(\gamma_{11} sex_i + \gamma_{12} drug_i + \varphi_{11}(b_{i0} + b_{i1}t)\right) && \text{(S1)}\\
\begin{aligned}
\lambda_{i2}^{transpl.}(t) &= \lambda_{02}(t)\exp\left(\gamma_{21} edema\_no_i + \gamma_{22} edema\_de_i\right.\\
&\quad + \gamma_{23} sex_i + \gamma_{24} edema\_no_i sex_i\\
&\quad \left. + \gamma_{25} edema\_de_i sex_i + \varphi_{21} b_{i0} + \varphi_{22} b_{i1}\right)
\end{aligned} && \text{(S2)}
\end{cases}
$$

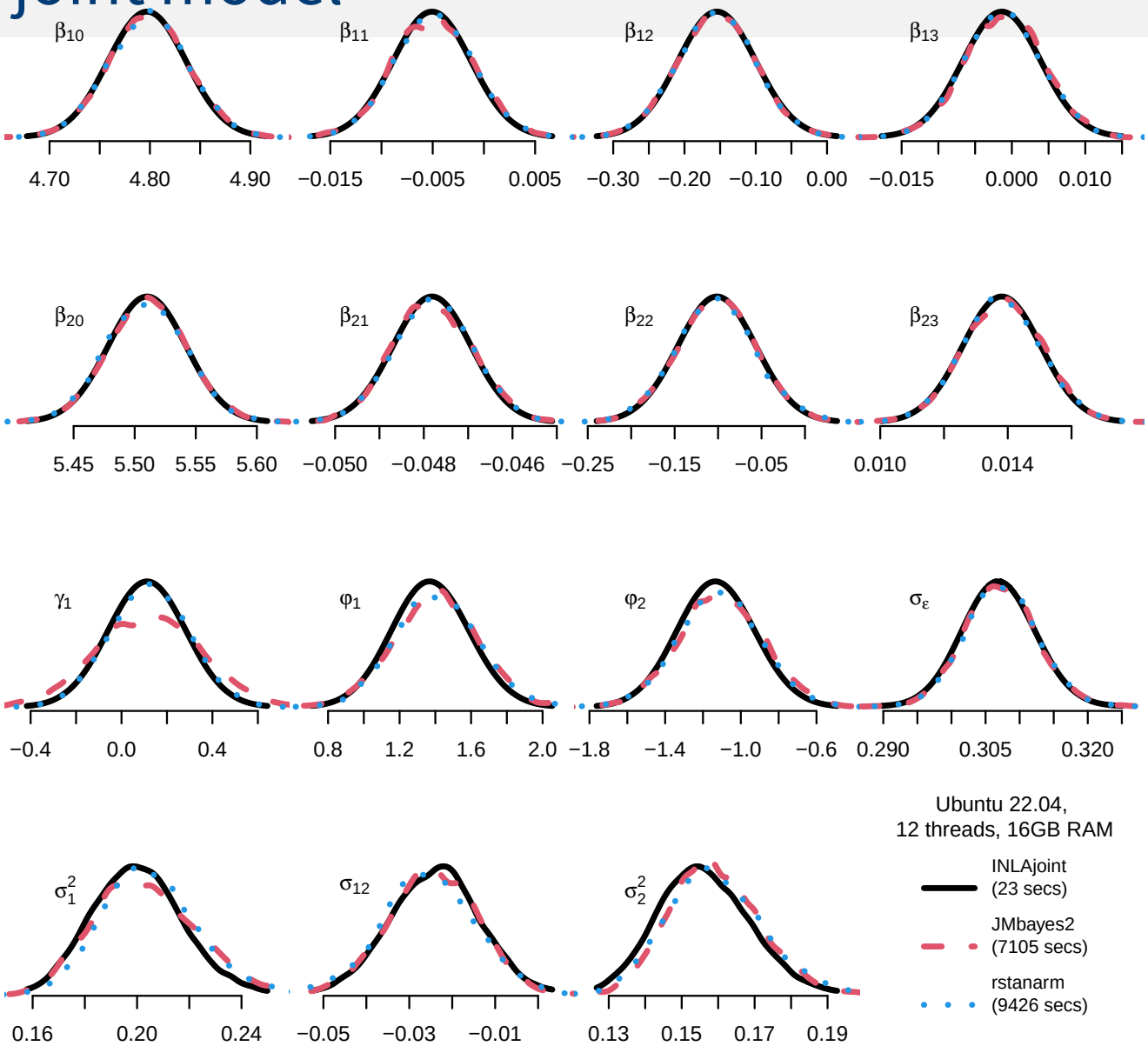
R: Joint modeling longitudinal and competing risks

Bivariate joint model

$$
\begin{cases}
\log(Y_{i1}(t)) &= \eta_{i1}(t) + \varepsilon_{i1}(t) \qquad \textbf{(Serum bilirubin - lognormal)} \\
&= \beta_{10} + b_{i1} + \beta_{11} t + \beta_{12} drug_i + \beta_{13} tdrug_i + \varepsilon_{i1}(t) \\
\log(E[Y_{i2}(t)]) &= \eta_{i2}(t) \qquad \textbf{(platelet - Poisson)} \\
&= \beta_{20} + b_{i2} + \beta_{21} t + \beta_{22} drug_i + \beta_{23} tdrug_i \\
\lambda_i(t) &= \lambda_0(t) \exp\left(\gamma_1 drug_i + \eta_{i1}(t)\varphi_1 + \eta_{i2}(t)\varphi_2\right) \qquad \textbf{(event risk)}
\end{cases}
$$

Objective: comparison with software specialized in joint modeling.

Bivariate joint model



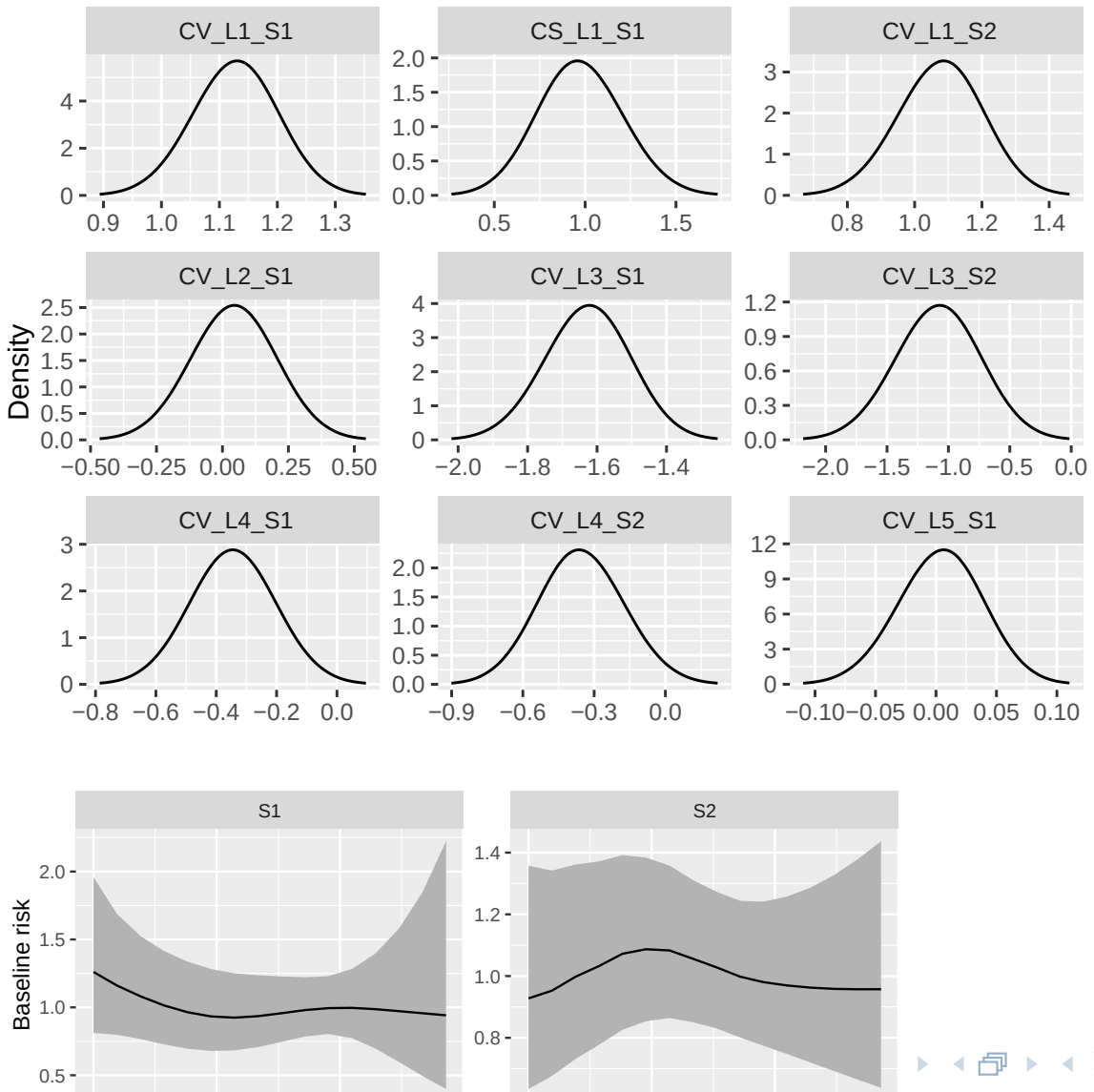
Multivariate joint model: Application of 10.1093/biostatistics/kxad019

$$
\begin{cases}
\log(Y_{i1}(t)) &= \eta_{i1}(t) + \varepsilon_{i1}(t) \qquad \textit{(bilirubin - lognormal - L1)} \\
&= (\beta_{10} + b_{i10}) + \beta_{11}X_i + (\beta_{12} + b_{i11})\mathrm{NS}_1(t) + (\beta_{13} + b_{i12})\mathrm{NS}_2(t) + \\
&\quad (\beta_{14} + b_{i13})\mathrm{NS}_3(t) + \beta_{15}X_i\mathrm{NS}_1(t) + \beta_{16}X_i\mathrm{NS}_2(t) + \beta_{17}X_i\mathrm{NS}_3(t) + \varepsilon_{i1}(t) \\
\log(Y_{i2}(t)) &= \eta_{i2}(t) + \varepsilon_{i2}(t) \qquad \textit{(aspartate aminotransferase - lognormal - L2)} \\
&= (\beta_{20} + b_{i20}) + \beta_{21}X_i + (\beta_{22} + b_{i21})\mathrm{NS}_1(t) + (\beta_{23} + b_{i22})\mathrm{NS}_2(t) + \\
&\quad (\beta_{24} + b_{i23})\mathrm{NS}_3(t) + \beta_{25}X_i\mathrm{NS}_1(t) + \beta_{26}X_i\mathrm{NS}_2(t) + \beta_{27}X_i\mathrm{NS}_3(t) + \varepsilon_{i2}(t) \\
Y_{i3}(t) &= \eta_{i3}(t) + \varepsilon_{i3}(t) \qquad \textit{(albumin - normal - L3)} \\
&= (\beta_{30} + b_{i30}) + \beta_{31}X_i + (\beta_{32} + b_{i31})t + \beta_{33}X_it + \varepsilon_{i3}(t) \\
\log(E[Y_{i4}(t)]) &= \eta_{i4}(t) \qquad \textit{(platelet - Poisson - L4)} \\
&= (\beta_{40} + b_{i40}) + \beta_{41}X_i + (\beta_{42} + b_{i41})\mathrm{NS}_1(t) + (\beta_{43} + b_{i42})\mathrm{NS}_2(t) + \\
&\quad (\beta_{44} + b_{i43})\mathrm{NS}_3(t) + \beta_{45}X_i\mathrm{NS}_1(t) + \beta_{46}X_i\mathrm{NS}_2(t) + \beta_{47}X_i\mathrm{NS}_3(t) \\
\mathrm{logit}(E[Y_{i5}(t)]) &= \eta_{i5}(t) \qquad \textit{(spiders - binomial - L5)} \\
&= (\beta_{50} + b_{i50}) + \beta_{51}X_i + (\beta_{52} + b_{i51})t + \beta_{53}X_it \\
\lambda_{i1}(t) &= \lambda_{01}(t)\exp\left(\gamma_1 X_i + \eta_{i1}(t)\varphi_1 + \eta'_{i1}(t)\varphi_3 + \eta_{i2}(t)\varphi_4 + \right. \qquad \textit{(death risk - S1)} \\
&\qquad \left. \eta_{i3}(t)\varphi_5 + \eta_{i4}(t)\varphi_7 + \eta_{i5}(t)\varphi_9\right) \\
\lambda_{i2}(t) &= \lambda_{02}(t)\exp\left(\gamma_2 X_i + \eta_{i1}(t)\varphi_2 + \eta_{i3}(t)\varphi_6 + \eta_{i4}(t)\varphi_8\right) \qquad \textit{(transplantation risk - S2)}
\end{cases}
$$

η: linear predictors, β: fixed effects, b_i: random effects, NS: natural cubic splines, X_i: drug,

φ: longitudinal-survival association parameters, ε_i: residual error.

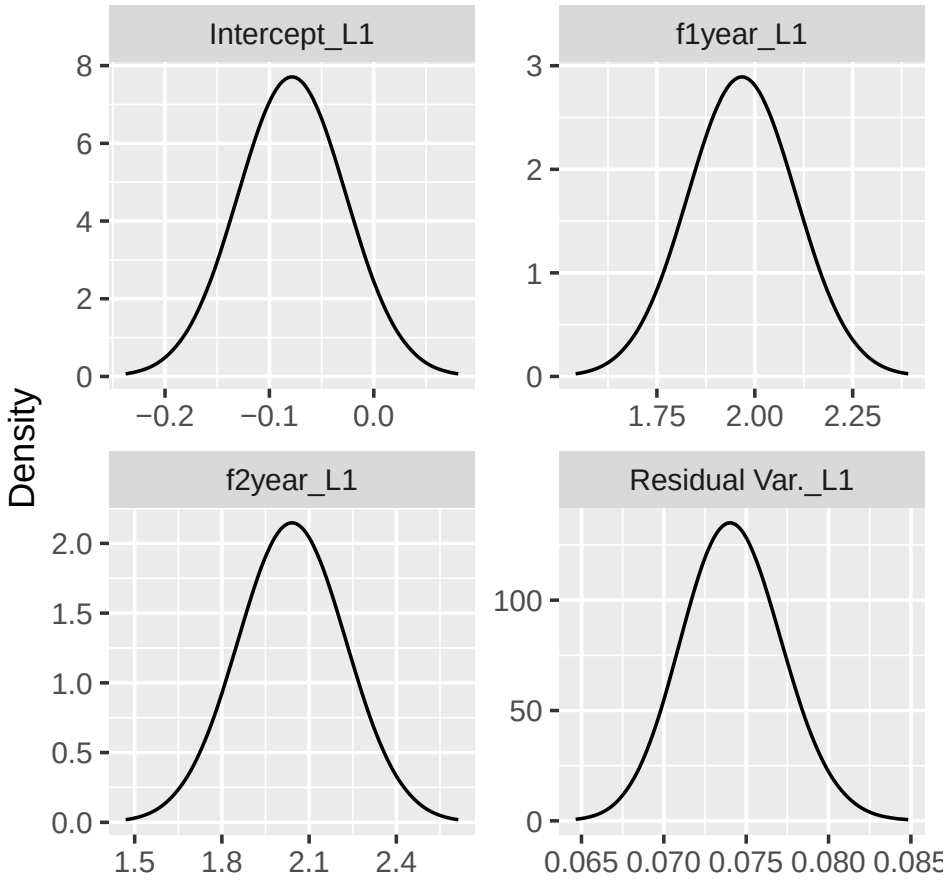
Multivariate joint model: Application of 10.1093/biostatistics/kxad019



Posteriors and dynamic prediction

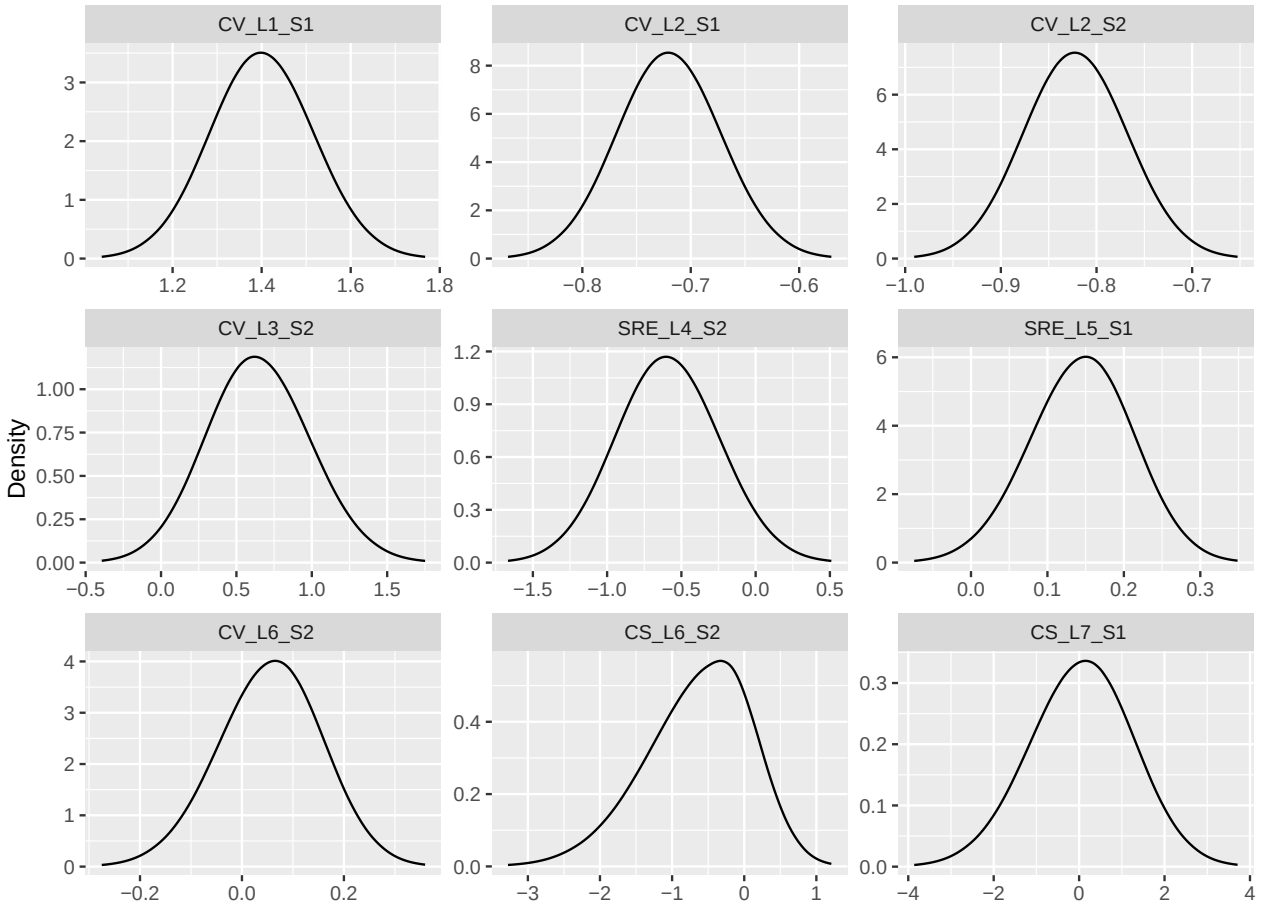
Multivariate joint model

Posteriors marginals of all parameters can be plotted with the *plot()* function.
plot(MD)$Outcomes$L1



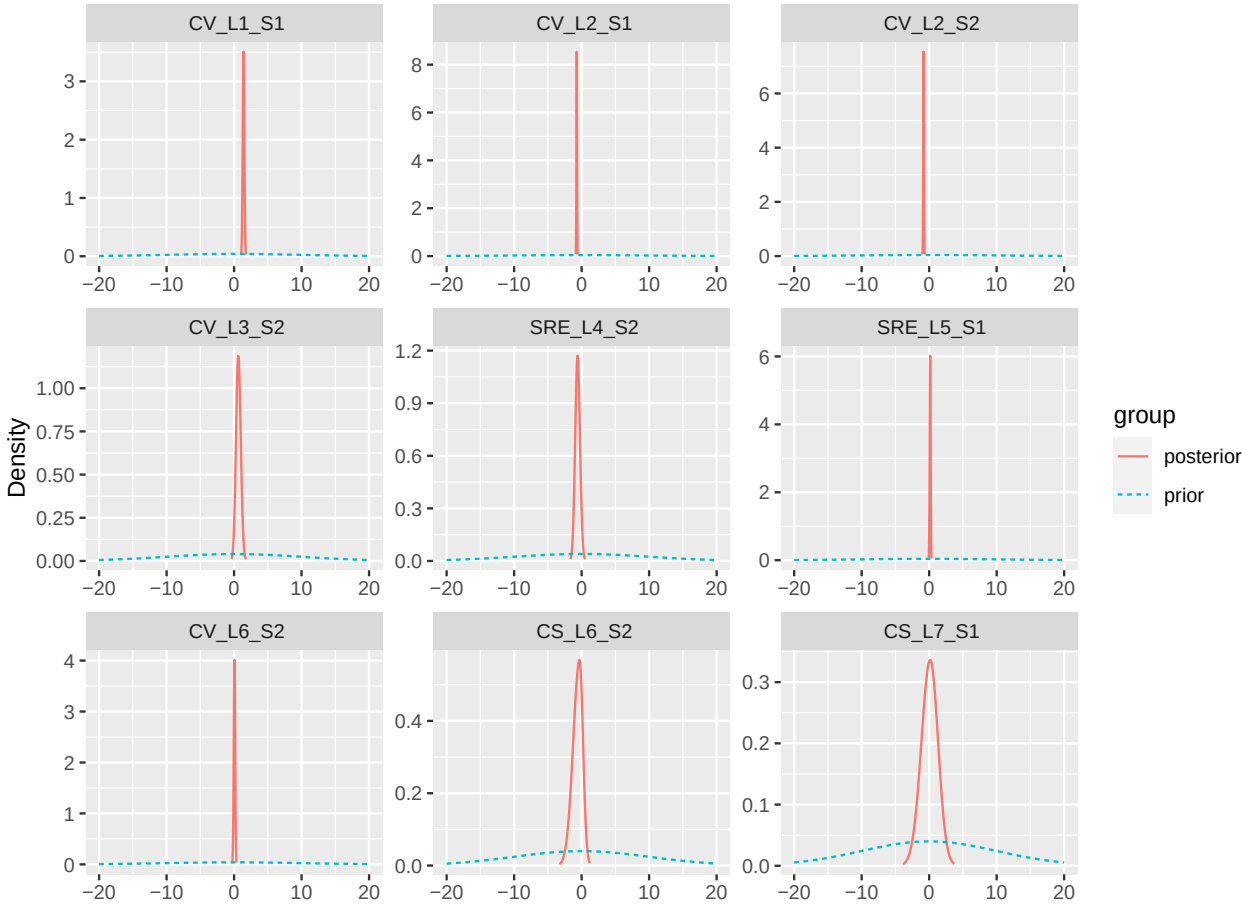
Multivariate joint model

plot(MD)$Associations



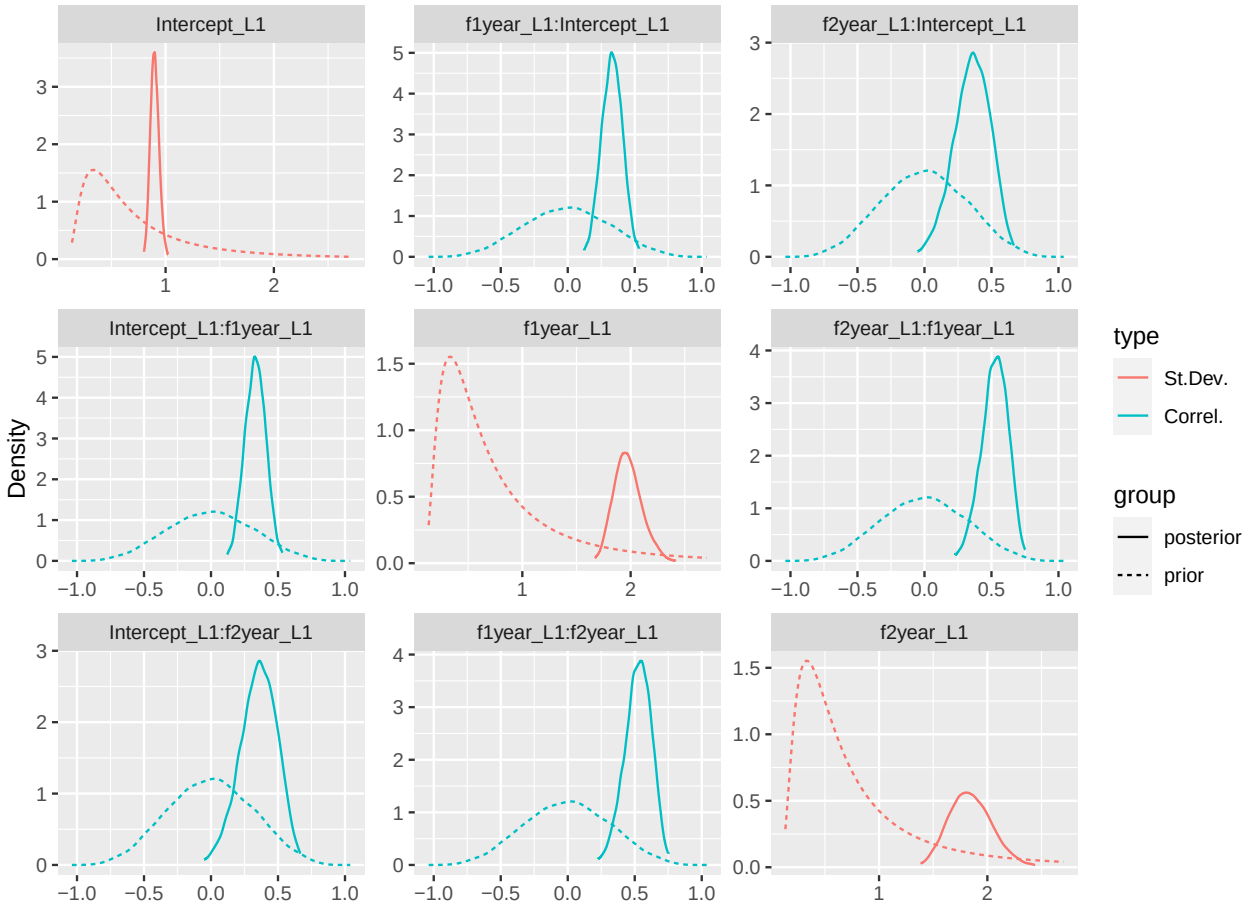
Multivariate joint model

It is possible to add prior distributions on the plot (e.g., to evaluate how data informs about parameters values and for prior sensitivity analysis).
plot(MD, priors=TRUE)$Associations



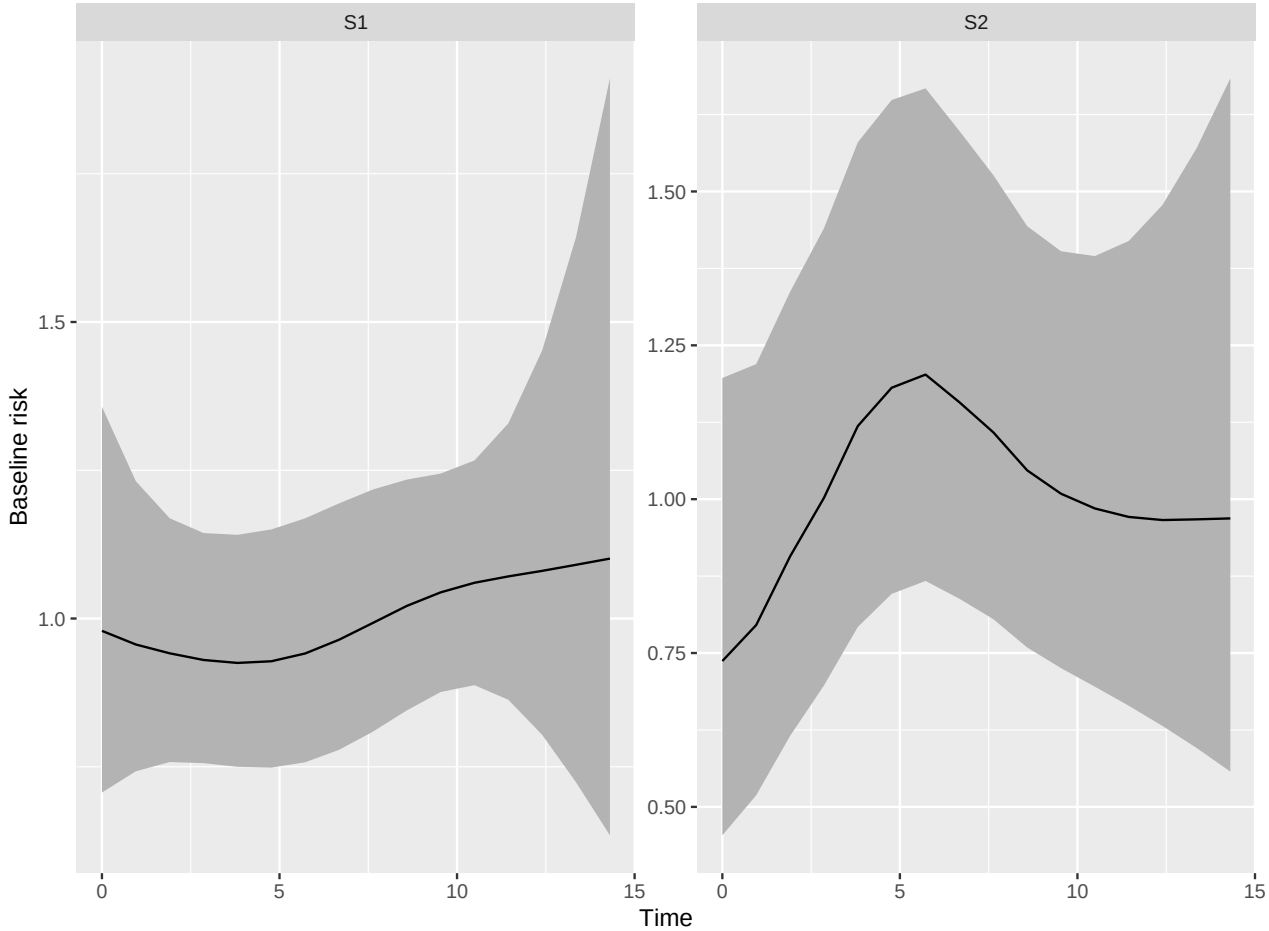
Multivariate joint model

plot(MD, sdcor=TRUE, priors=TRUE)$Covariances$L1



Multivariate joint model

plot(MD)$Baseline



Multivariate joint model

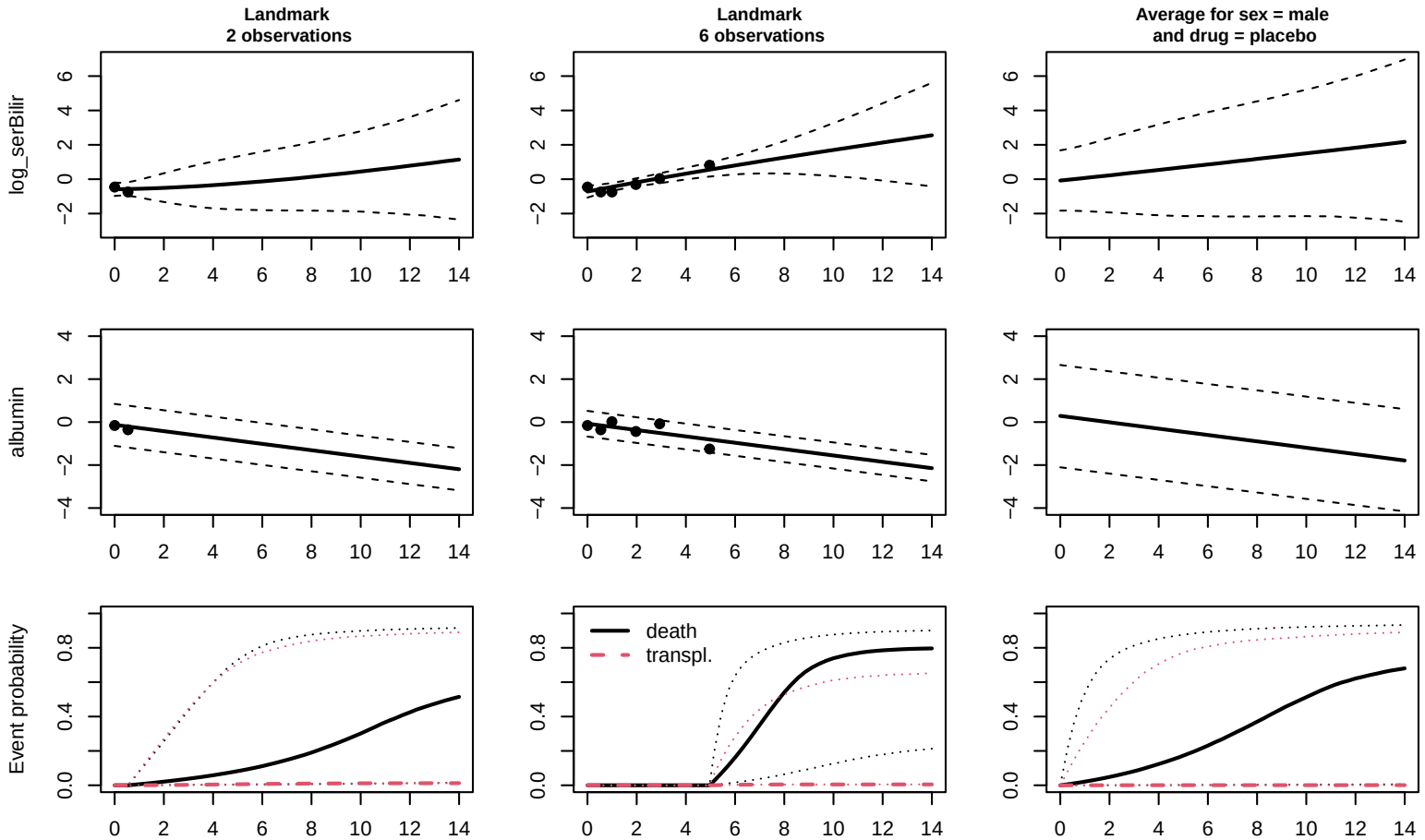
Finally, we can do imputation, prediction and inference with the *predict()* function. First set up new dataset:

id	drug	sex	year	ascites	spiders	log_serBilir	albumin	log_SGOT	platelets	log_prothrombin
1	0	1	0.0	0	0	-0.46	-0.16	-0.19	142	0.6
1	0	1	0.5	0	0	-0.74	-0.36	0.08	120	0.6
2	0	1	0.0	0	0	-0.46	-0.16	-0.19	142	0.6
2	0	1	0.5	0	0	-0.74	-0.36	0.08	120	0.6
2	0	1	1.0	0	0	-0.74	0.02	-0.36	121	-0.1
2	0	1	2.0	1	0	-0.31	-0.44	-0.51	81	0.6
2	0	1	2.9	0	0	0.03	-0.08	-0.42	64	1.5
2	0	1	5.0	0	0	0.81	-1.25	-0.22	59	2.3
3	0	0	0.0	NA	NA	NA	NA	NA	NA	NA

We ask for prediction for an individual at 2 different landmark times and average prediction for the reference individual.

Multivariate joint model

predict(m5, NewDat, horizon=14, inv.link=TRUE, CIF=TRUE)



Joint models with INLA but not INLAjoint
Adherence study (Quantile joint model)

Introduction

From the AARDEX Group - medication adherence and patient persistence is crucial for successful treatment regimes and regularized evaluation[1]. Adherence is a proportion and persistence is defined based on a set of criteria unique to each drug.

Research problem - outliers in the adherence bias the model, and can cause false positives for intervention.

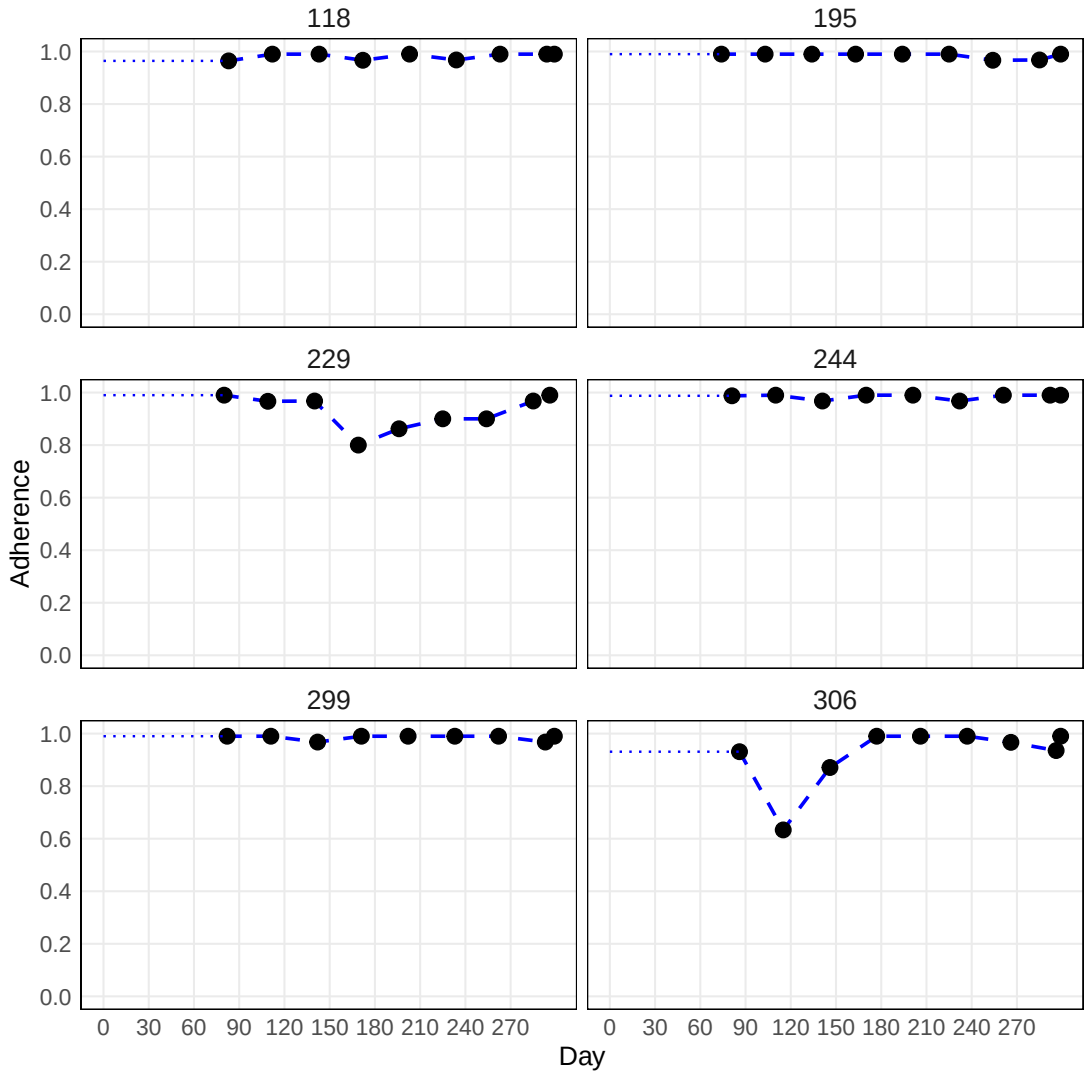
[1] Burger, D.A., Van der Merwe, S., Van Niekerk, J., Lesaffre, E. and Pironet, A. Joint quantile regression of longitudinal continuous proportions and time-to-event data: application in medication adherence and persistence, *Statistical Methods in Medical Research*, Accepted

Introduction

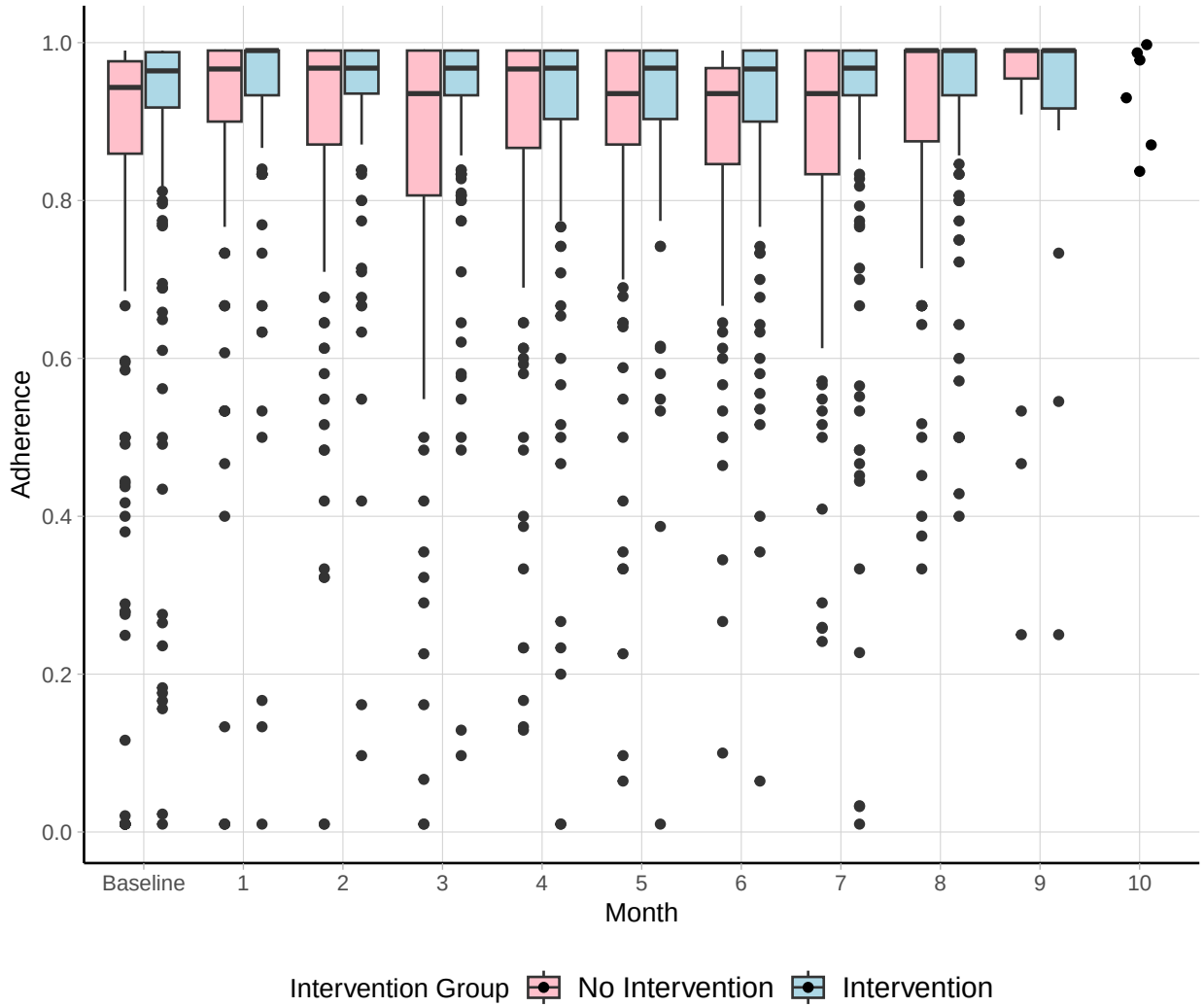
From the AARDEX Group - medication adherence and patient persistence is crucial for successful treatment regimes and regularized evaluation[1].
Adherence is a proportion and persistence is defined based on a set of criteria unique to each drug.
Research problem - outliers in the adherence bias the model, and can cause false positives for intervention.

[1] Burger, D.A., Van der Merwe, S., Van Niekerk, J., Lesaffre, E. and Pironet, A. Joint quantile regression of longitudinal continuous proportions and time-to-event data: application in medication adherence and persistence, *Statistical Methods in Medical Research*, Accepted

Dataset


118
195
229
244
299
306
Adherence
1.0
0.8
0.6
0.4
0.2
0.0
0 30 60 90 120 150 180 210 240 270
Day

Dataset


1.0
0.8
0.6
0.4
0.2
0.0
Adherence
Baseline
1
2
3
4
5
6
7
8
9
10
Month
Intervention Group
No Intervention
Intervention

Model

Longitudinal model:

$$
\begin{aligned}
f\left(y_{ij};\kappa_{q,i}\left(t_{ij}\right),\psi\right) &= \alpha_{ij_1}\alpha_2 y_{ij}^{\alpha_{ij_1}-1}\left(1-y_{ij}^{\alpha_{ij_1}}\right)^{\alpha_2-1}, \\
\kappa_{q,i}\left(t_{ij}\right) &= g^{-1}\left(\eta_{q,i}\left(t_{ij}\right)\right) \\
\eta_{q,i}\left(t_{ij}\right) &= \left(\beta_0+b_{0_i}\right)+\left(\beta_{\text{time}}+b_{\text{time}_i}+\boldsymbol{z}_i'\beta_{\text{tx}}\right)t_{ij}+\boldsymbol{x}_i'\beta_{\text{cov}}
\end{aligned}
$$

with $\alpha_{ij_1}=\frac{\log\left(1-(1-q)^{\frac{1}{\alpha_2}}\right)}{\log\left(\kappa_{q,i}\left(t_{ij}\right)\right)}$ and $\alpha_2=\frac{\log(1-q)}{\log\left(1-e^{-\psi}\right)}$.

Survival model:

$$
h_i\left(t\right)=h_0\left(t\right)\exp\left(\phi\eta_i\left(t\right)+\boldsymbol{w}_i'\gamma\right).
$$

Model

Longitudinal model:

$$
\begin{aligned}
f\left(y_{ij}; \kappa_{q,i}\left(t_{ij}\right), \psi\right) &= \alpha_{ij_1}\alpha_2 y_{ij}^{\alpha_{ij_1}-1}\left(1-y_{ij}^{\alpha_{ij_1}}\right)^{\alpha_2-1}, \\
\kappa_{q,i}\left(t_{ij}\right) &= g^{-1}\left(\eta_{q,i}\left(t_{ij}\right)\right) \\
\eta_{q,i}\left(t_{ij}\right) &= \left(\beta_0 + b_{0_i}\right) + \left(\beta_{\text{time}} + b_{\text{time}_i} + \boldsymbol{z}_i'\beta_{\text{tx}}\right) t_{ij} + \boldsymbol{x}_i'\beta_{\text{cov}}
\end{aligned}
$$

with $\alpha_{ij_1} = \frac{\log\left(1-(1-q)^{\frac{1}{\alpha_2}}\right)}{\log\left(\kappa_{q,i}\left(t_{ij}\right)\right)}$ and $\alpha_2 = \frac{\log(1-q)}{\log\left(1-e^{-\psi}\right)}$.

Survival model:

$$
h_i\left(t\right) = h_0\left(t\right)\exp\left(\phi\eta_i\left(t\right) + \boldsymbol{w}_i'\gamma\right).
$$

Model

Longitudinal model:

$$
\begin{aligned}
f\left(y_{ij};\kappa_{q,i}\left(t_{ij}\right),\psi\right) &= \alpha_{ij_1}\alpha_2 y_{ij}^{\alpha_{ij_1}-1}\left(1-y_{ij}^{\alpha_{ij_1}}\right)^{\alpha_2-1},\\
\kappa_{q,i}\left(t_{ij}\right) &= g^{-1}\left(\eta_{q,i}\left(t_{ij}\right)\right)\\
\eta_{q,i}\left(t_{ij}\right) &= \left(\beta_0+b_{0_i}\right)+\left(\beta_{\text{time}}+b_{\text{time}_i}+\boldsymbol{z}_i'\boldsymbol{\beta}_{\text{tx}}\right)t_{ij}+\boldsymbol{x}_i'\boldsymbol{\beta}_{\text{cov}}
\end{aligned}
$$

with $\alpha_{ij_1}=\frac{\log\left(1-(1-q)^{\frac{1}{\alpha_2}}\right)}{\log\left(\kappa_{q,i}\left(t_{ij}\right)\right)}$ and $\alpha_2=\frac{\log(1-q)}{\log\left(1-e^{-\psi}\right)}$.

Survival model:

$$h_i(t)=h_0(t)\exp\left(\phi\eta_i(t)+\boldsymbol{w}_i'\boldsymbol{\gamma}\right).$$

Model

Longitudinal model:

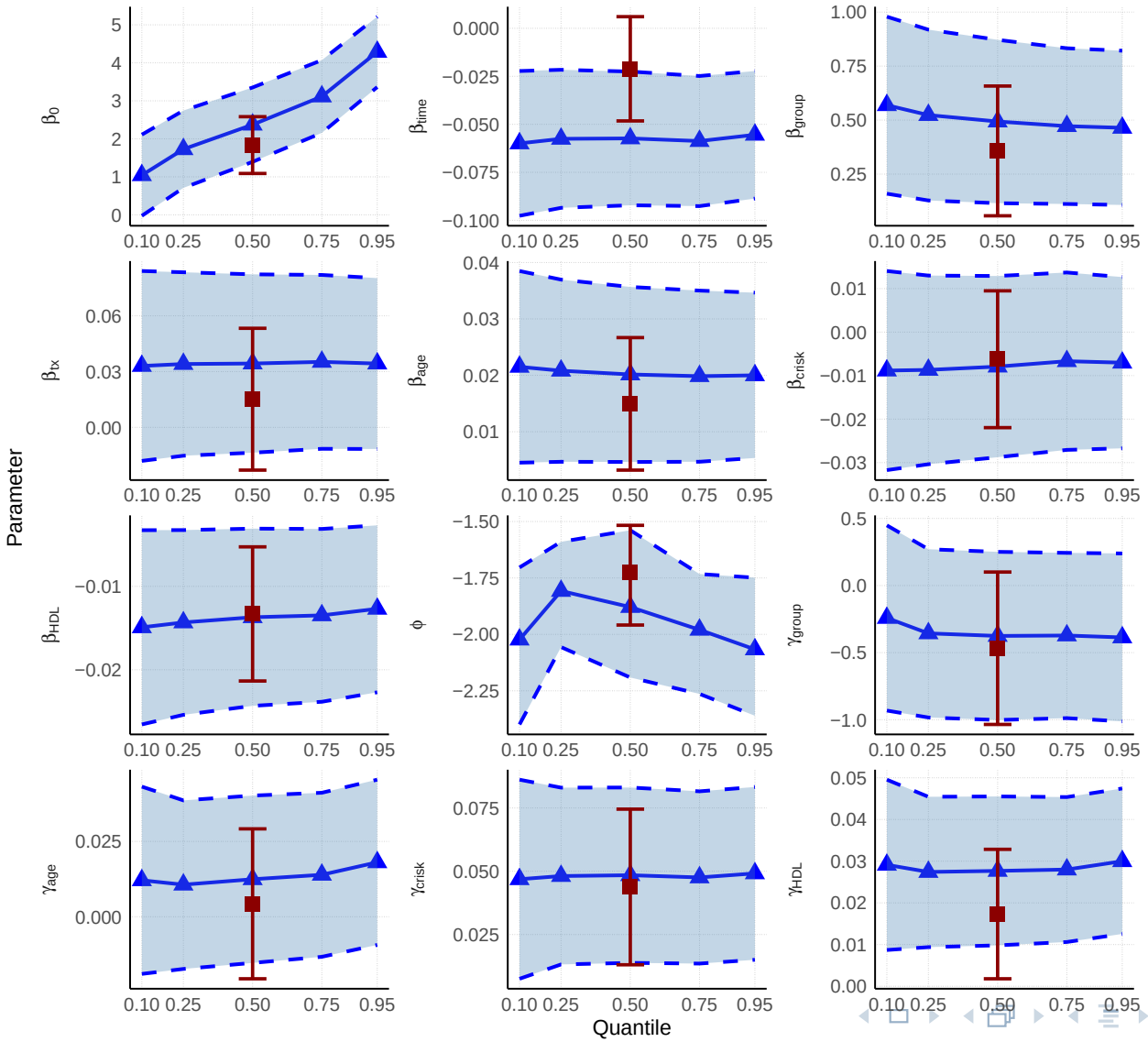
$$
\begin{aligned}
f\left(y_{ij}; \kappa_{q,i}\left(t_{ij}\right), \psi\right) &= \alpha_{ij_1}\alpha_2 y_{ij}^{\alpha_{ij_1}-1}\left(1-y_{ij}^{\alpha_{ij_1}}\right)^{\alpha_2-1}, \\
\kappa_{q,i}\left(t_{ij}\right) &= g^{-1}\left(\eta_{q,i}\left(t_{ij}\right)\right) \\
\eta_{q,i}\left(t_{ij}\right) &= \left(\beta_0 + b_{0_i}\right) + \left(\beta_{\text{time}} + b_{\text{time}_i} + \boldsymbol{z}_i'\boldsymbol{\beta}_{\text{tx}}\right) t_{ij} + \boldsymbol{x}_i'\boldsymbol{\beta}_{\text{cov}}
\end{aligned}
$$

with $\alpha_{ij_1} = \frac{\log\left(1-(1-q)^{\frac{1}{\alpha_2}}\right)}{\log\left(\kappa_{q,i}\left(t_{ij}\right)\right)}$ and $\alpha_2 = \frac{\log(1-q)}{\log\left(1-e^{-\psi}\right)}$.

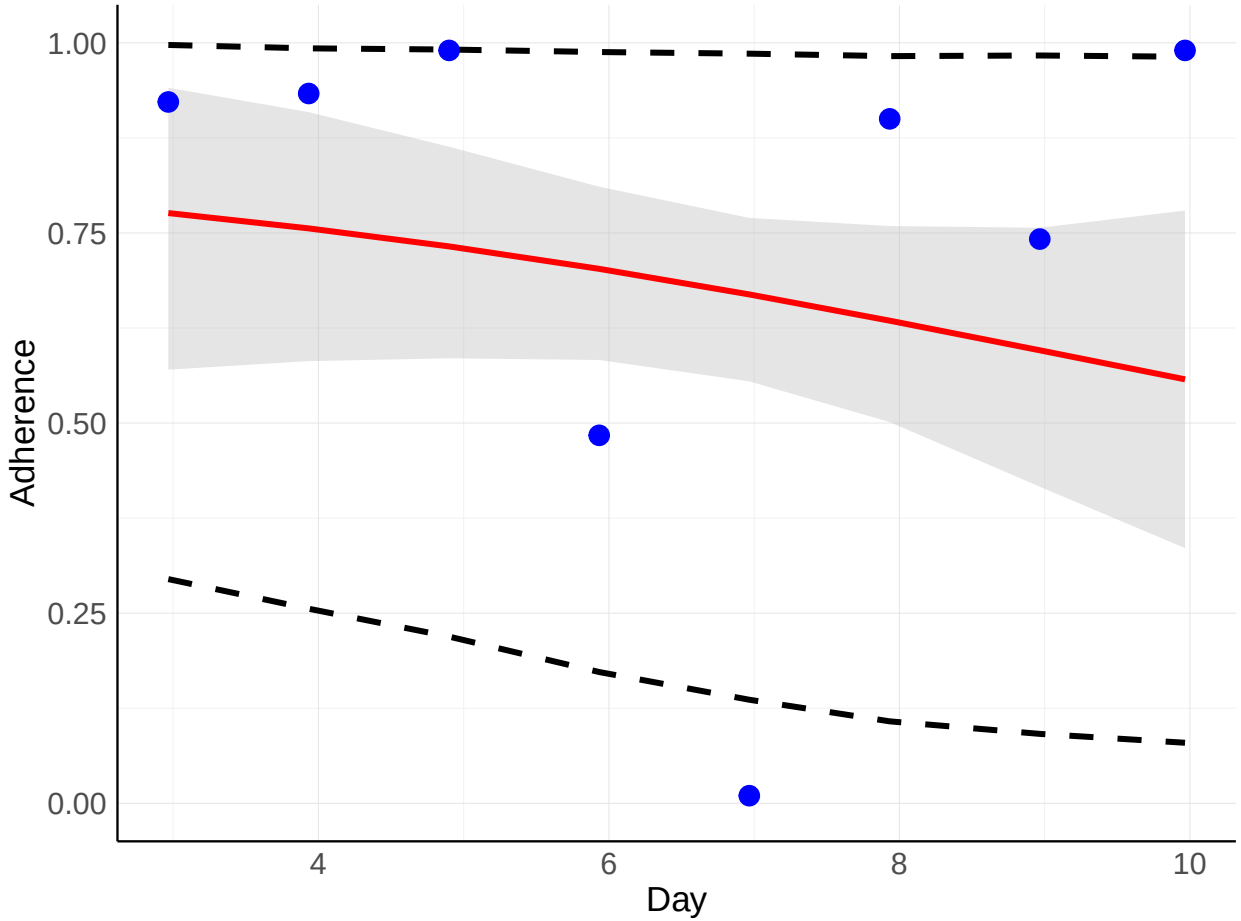
Survival model:

$$
h_i\left(t\right) = h_0\left(t\right)\exp\left(\phi\eta_i\left(t\right) + \boldsymbol{w}_i'\boldsymbol{\gamma}\right).
$$

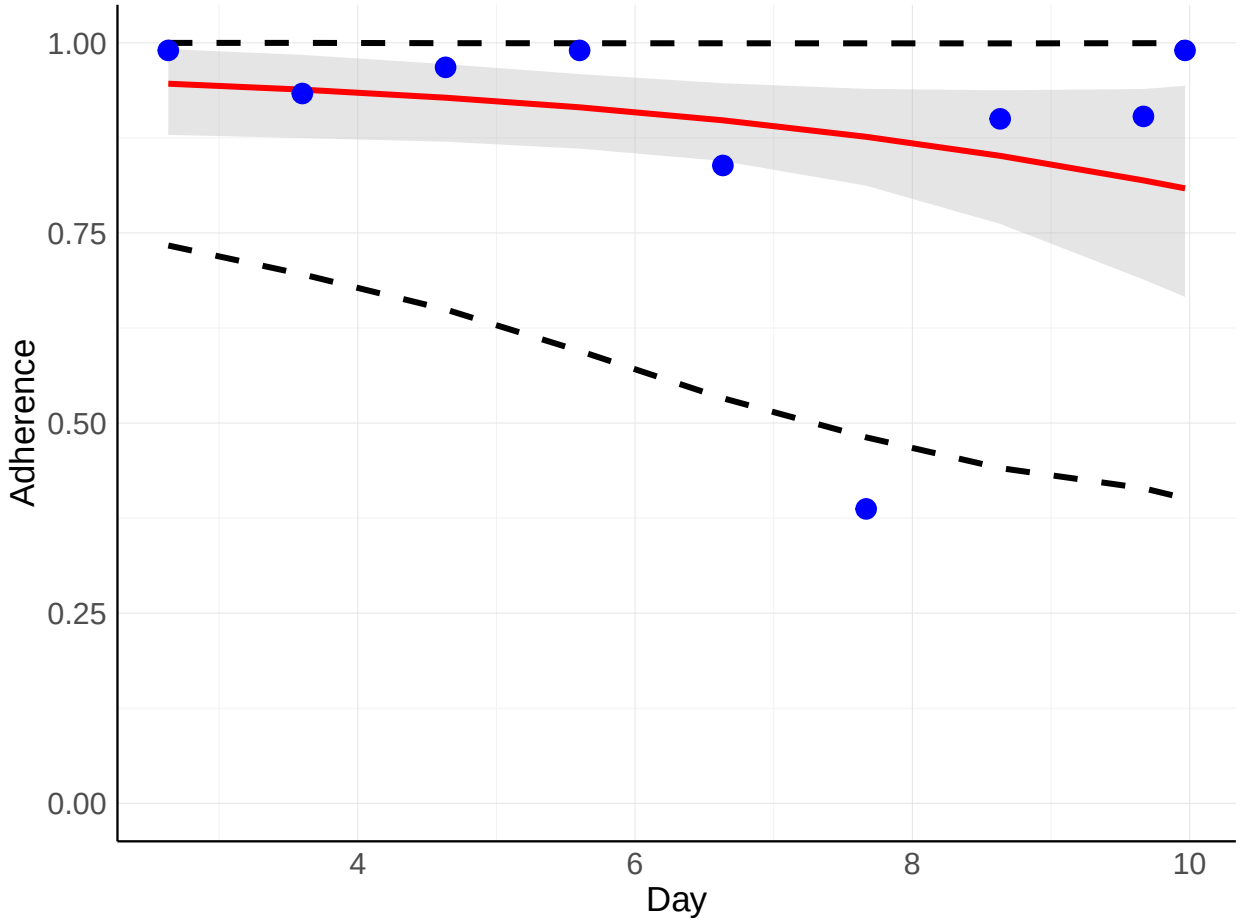
Posterior inference - fixed effects



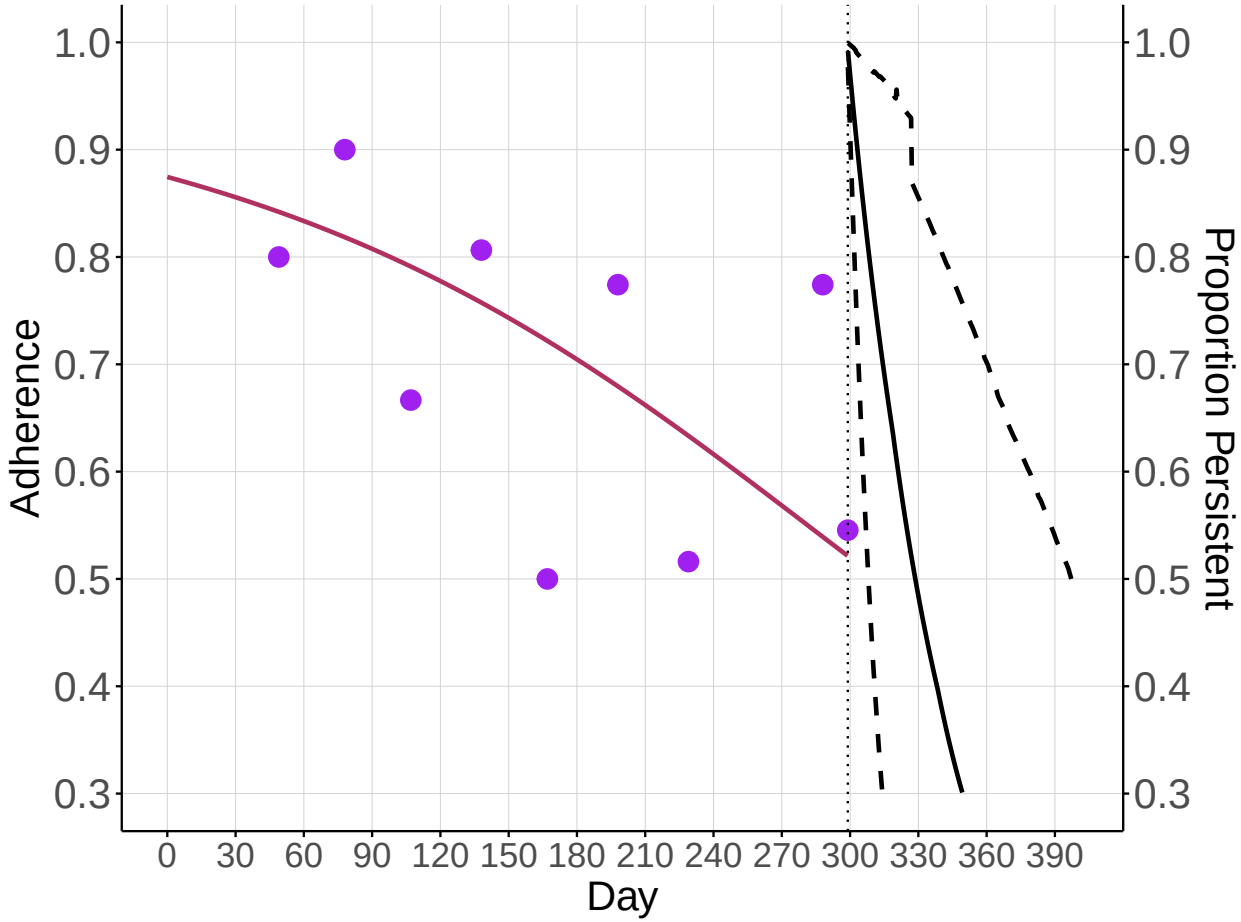
Posterior inference - Patient 46



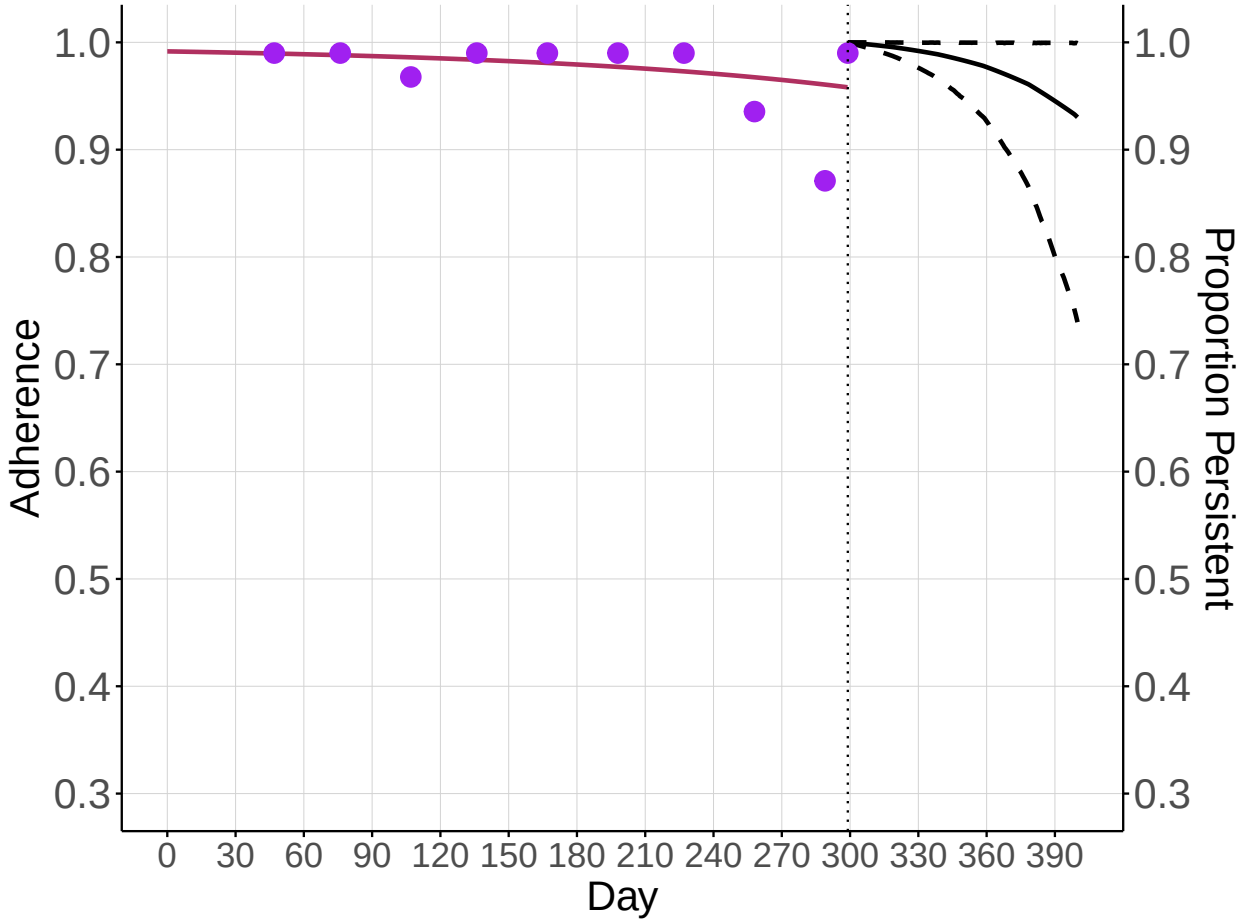
Posterior inference - Patient 56



Dynamic predictions - Patient 27



Dynamic predictions - Patient 10

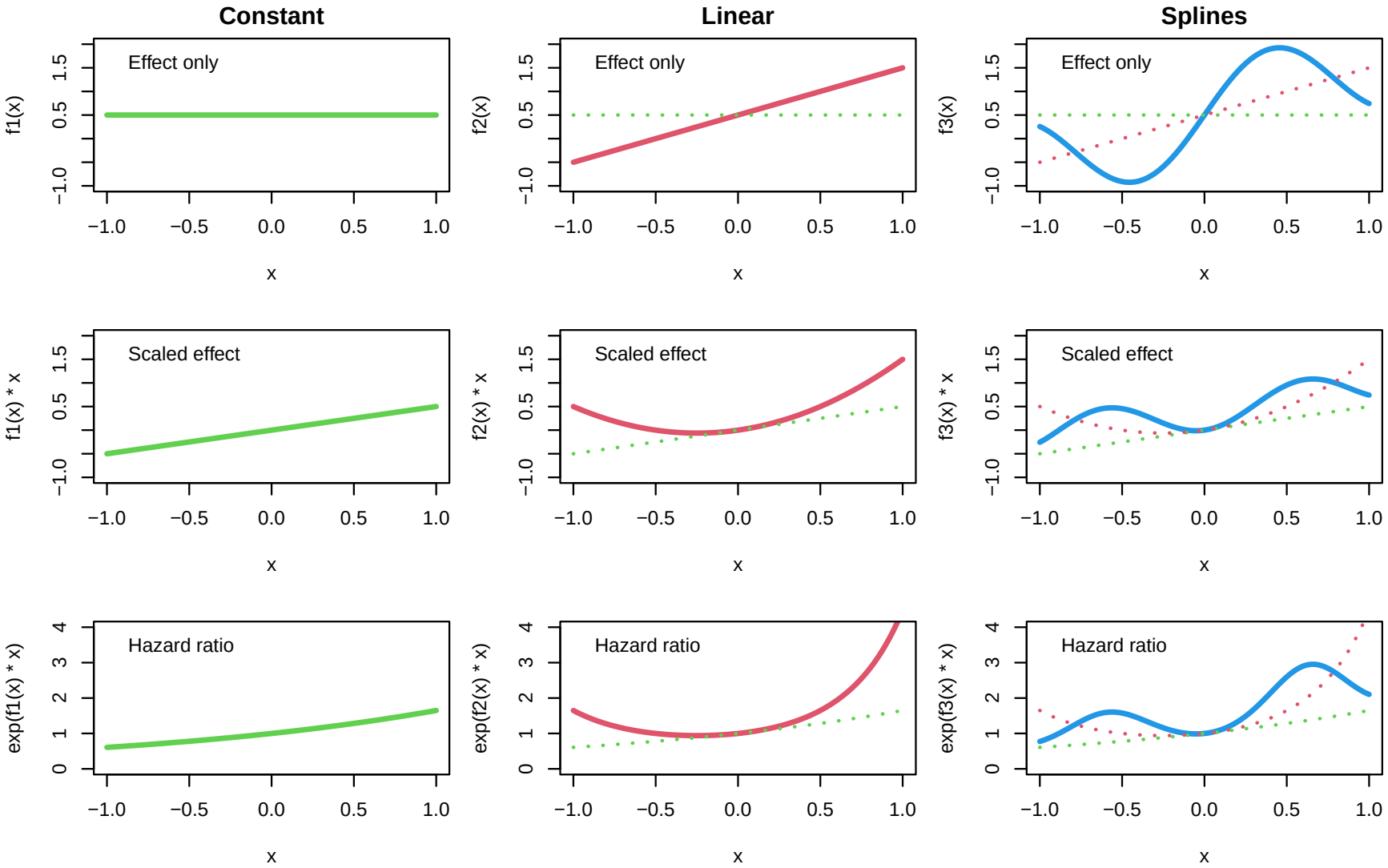


Summary

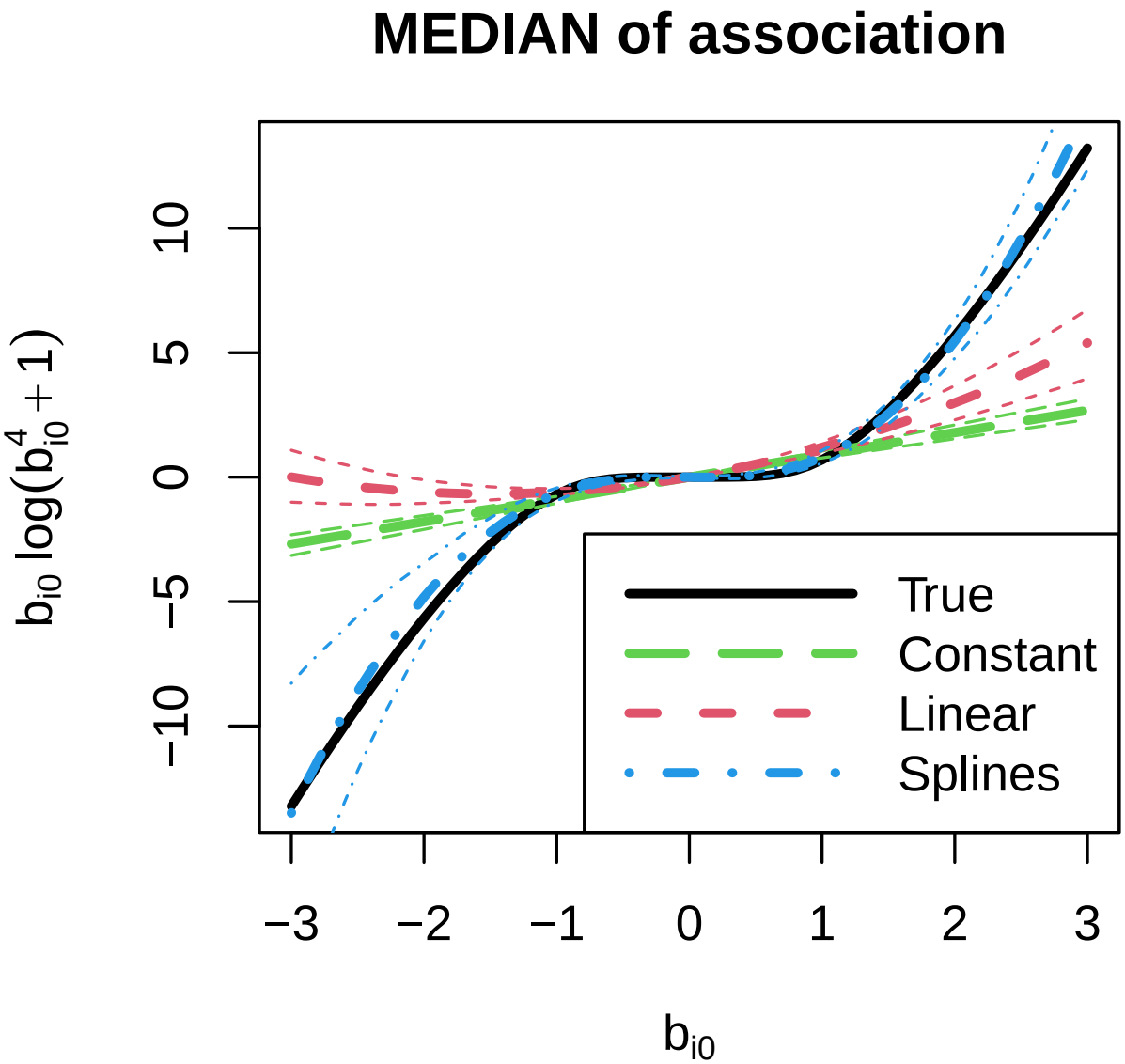
- **Most survival models** can be formulated as Latent Gaussian Models and thus **can be fitted with INLA**
- **INLA** provides an **efficient** and **reliable** approach to fit complex survival models.
- Scaling allows to go **beyond the limitations of other algorithms** (e.g., 24 correlated random effects).
- Its architecture allows to **tailor** the model to answer a given **question of interest**.
- **User-friendly interface** (INLAjoint) for longitudinal, survival and joint models
- **Spatial** and **spatio-temporal** random effects can be included in longitudinal and survival models - INLA, not INLAjoint.
- Dynamic development team, new features and improving efficiency

Current and future work

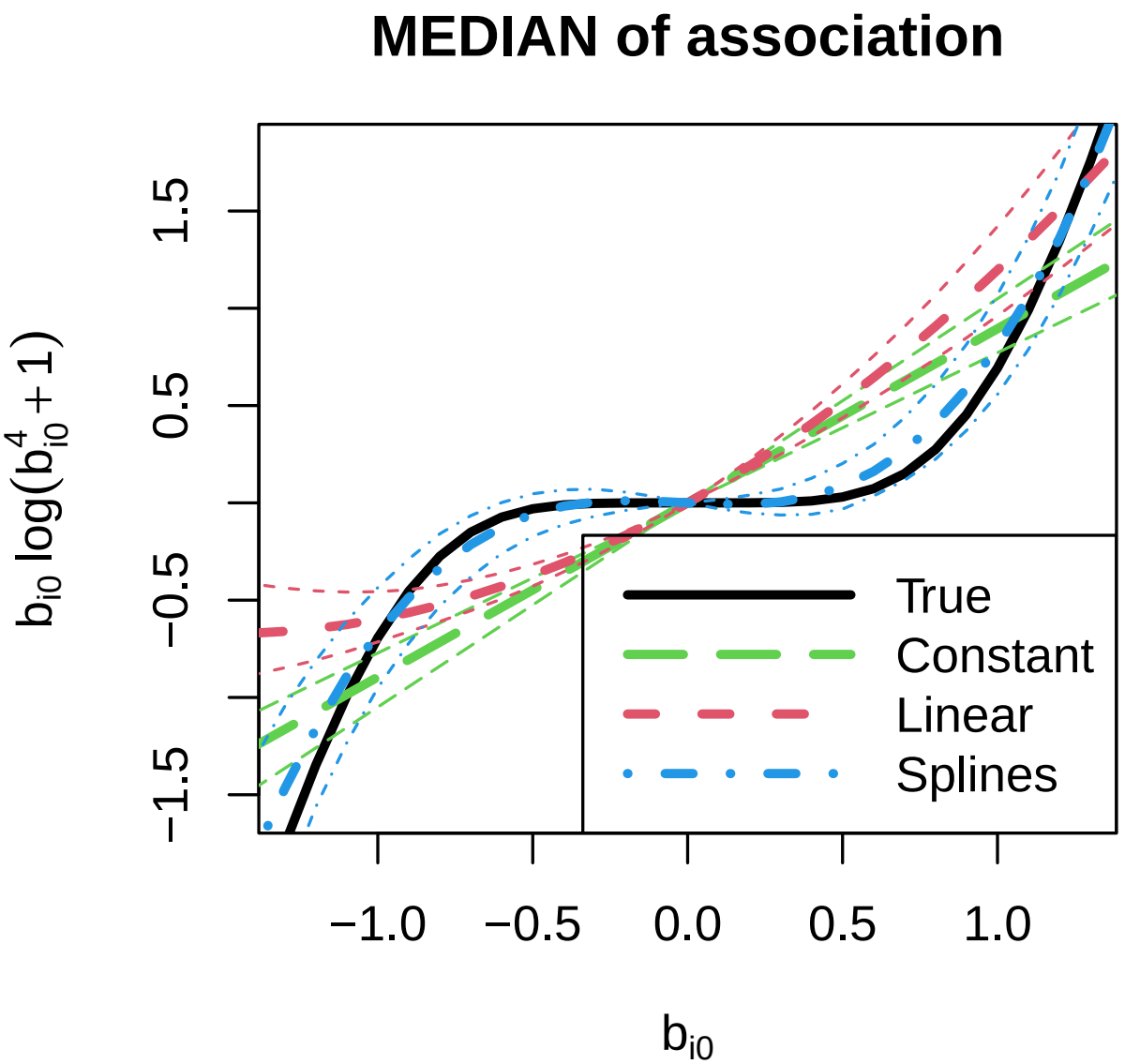
Non-parametric association



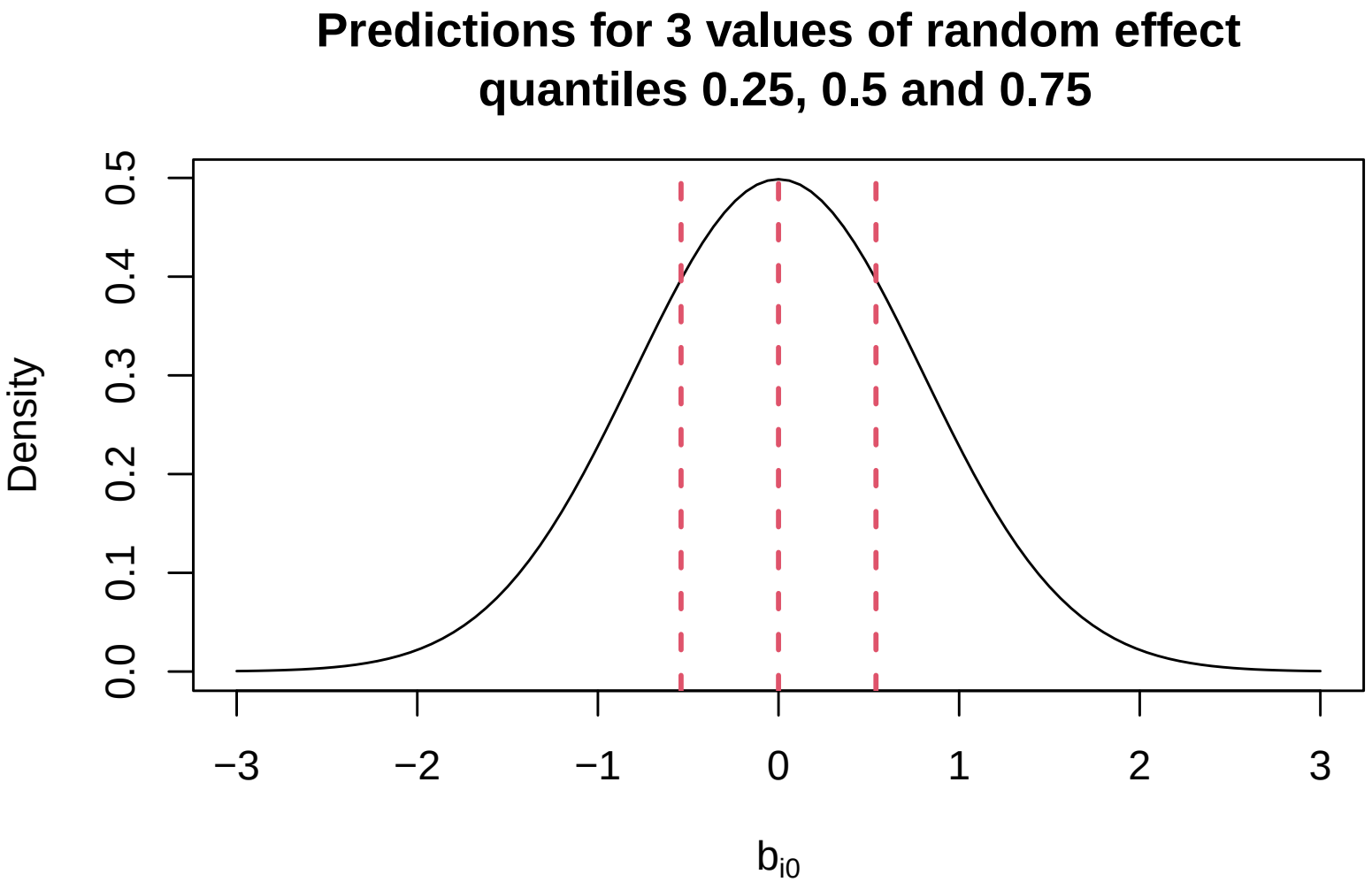
Simulations non-linear effect



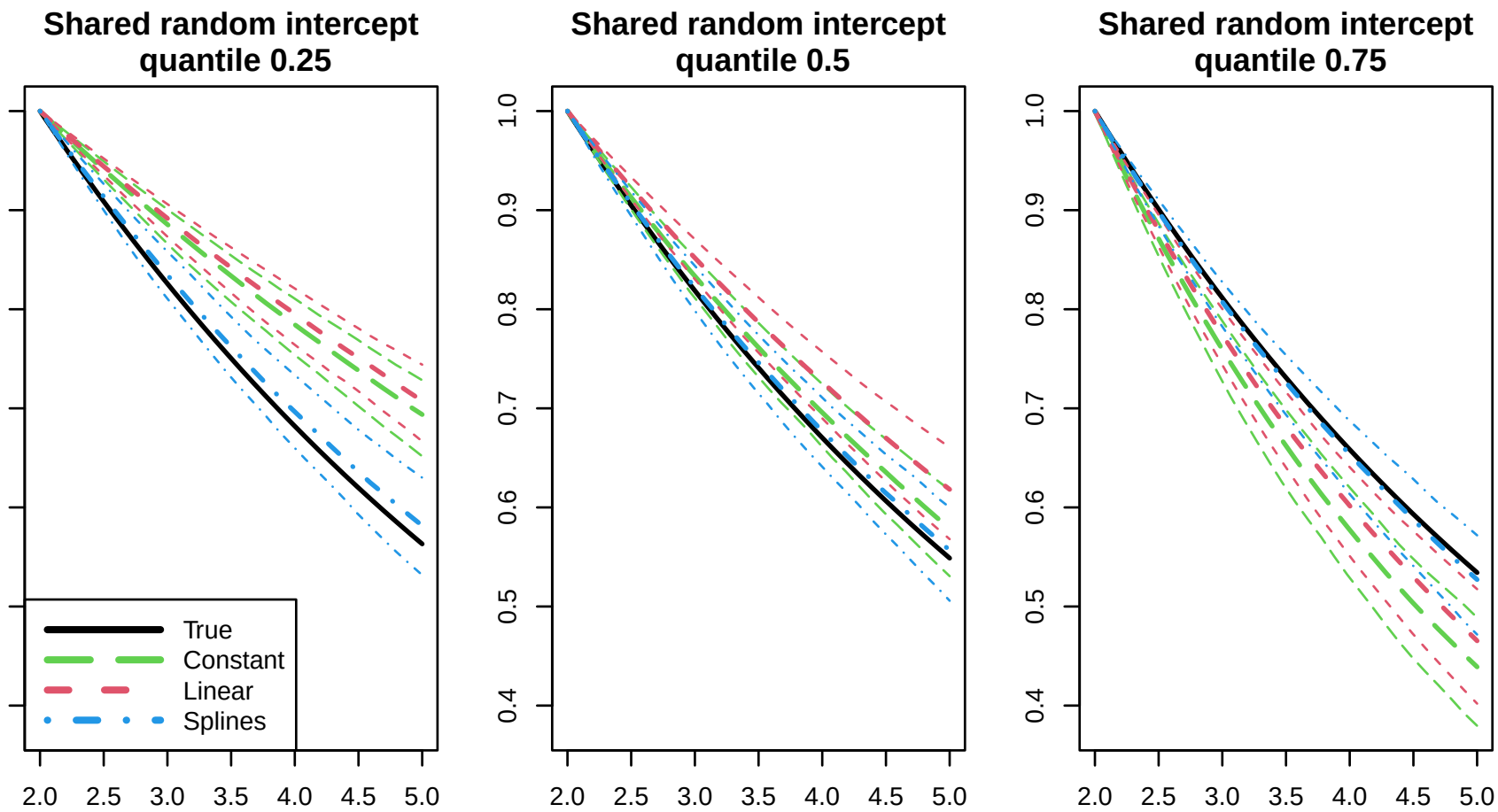
Simulations - scenario 1



Simulations - scenario 1

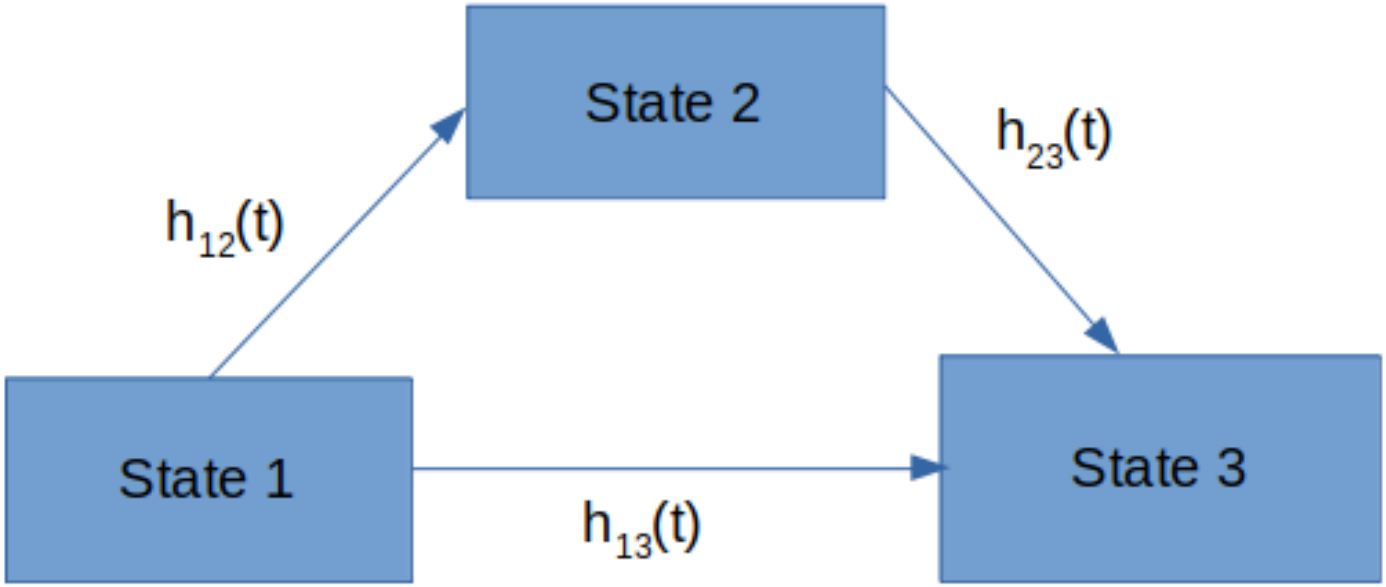


Simulations - scenario 1



Extra examples of other datasets

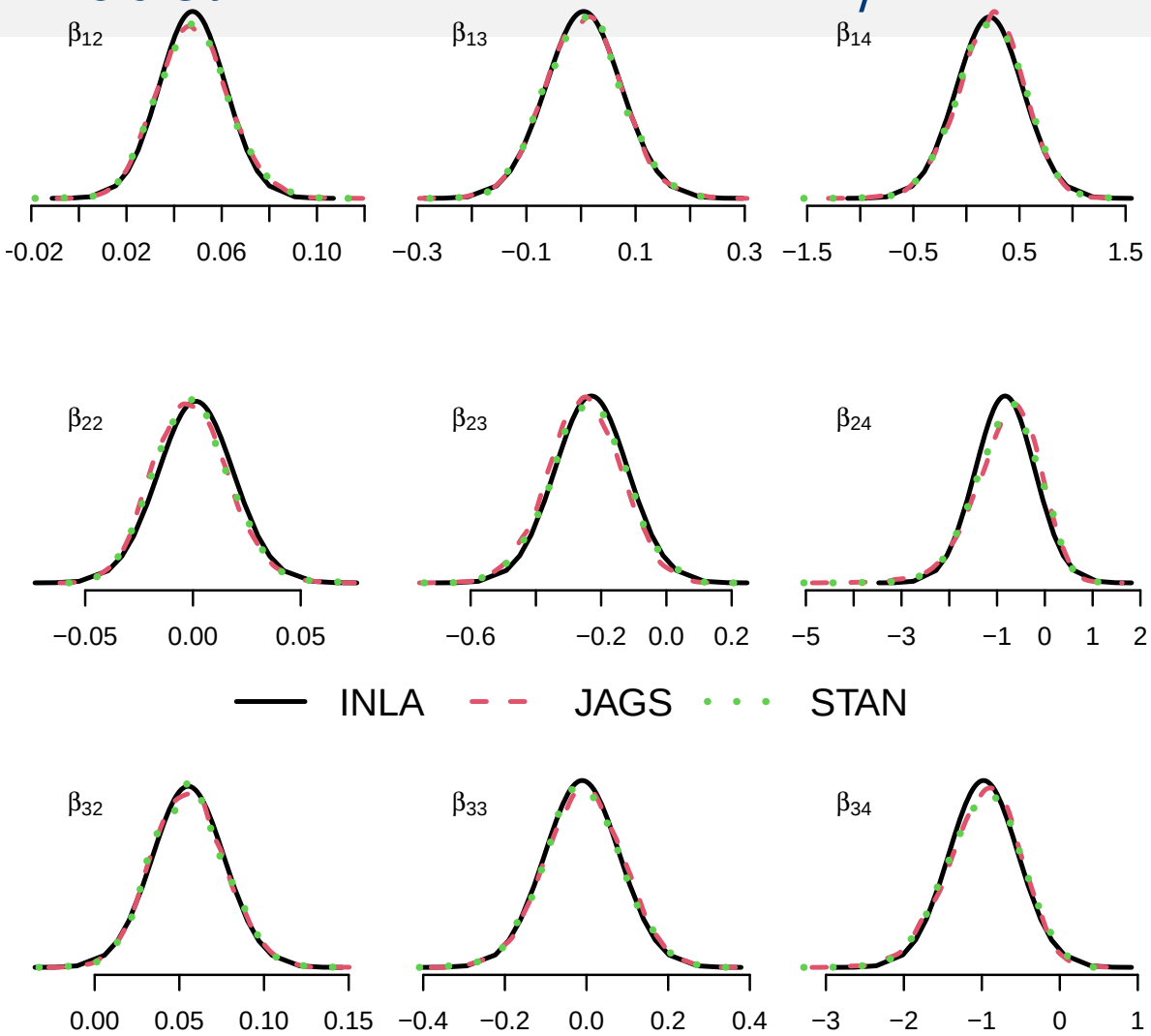
Multi-state model 1/2



$$\begin{cases} h_{i,12}(t) = h_{0,12}(t)\exp(\gamma_{12}X_i) \text{(S1)} \\ h_{i,13}(t) = h_{0,13}(t)\exp(\gamma_{13}X_i) \text{(S2)} \\ h_{i,23}(t) = h_{0,23}(t)\exp(\gamma_{23}X_i) \text{(S3)} \end{cases}$$

R: Example using data *SurvMS*

Multi-state model 2/2



Computation time: INLA = 1sec. ; JAGS = 155sec. ; STAN = 94sec.

Mixture cure model 1/3

bmt dataset (**smcure** R package)

The **bmt** dataset contains 91 patients with refractory acute lymphoblastic leukemia who participated in a bone marrow transplant study. Each patient has the following variables:

- **Time**: time to death (in days).
- **Status**: censoring indicator (1: if patient is uncensored; 0: otherwise).
- **TRT**: treatment group indicator (1: autologous; 0: allogeneic).

Mixture cure model 2/3

Model specification

Let Z be a binary random variable defined as 0 for susceptible and 1 for cured or immune individuals. The incidence submodel is given by:

$$Z \sim \text{Bernoulli}(\eta), \tag{4}$$

$$\text{logit}(\eta) = \beta_{C1} + \beta_{C2}\texttt{TRT}, \tag{5}$$

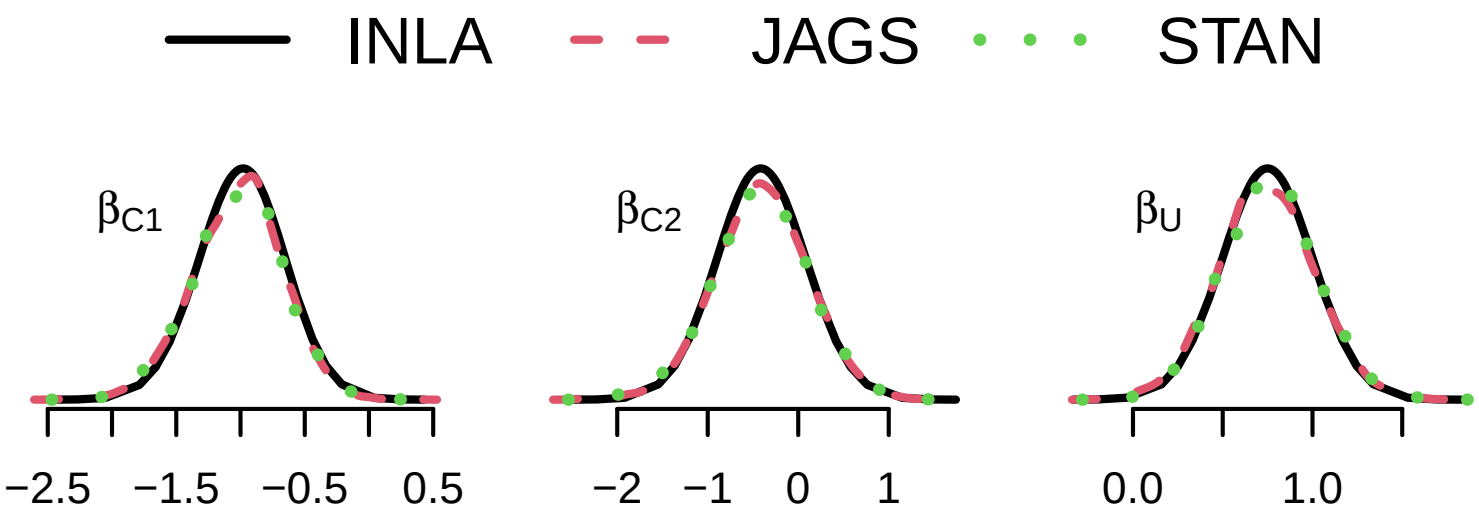
The latency submodel is expressed through a PH specification:

$$\lambda(t) = \lambda_0(t)\exp\left\{\beta_U \texttt{TRT}\right\}, \quad t > 0, \tag{6}$$

Note: Mixture cure is currently limited to parametric baseline.

R: Example with **INLAjoint**

Mixture cure model 3/3



Computation time: INLA = 0.6sec. ; JAGS = 62sec. ; STAN = 74sec.
JAGS: 4 chains, 1000 burn-in and 4000 iterations
STAN: 4 chains, 1000 burn-in and 1000 iterations

Shared frailty model

readmission dataset (**frailtypack** R package)
The **frailty model** can fit recurrent events with a gaussian distribution for the frailty term (i.e., lognormal frailty model).

$$\lambda_{i1}(t) = \lambda_{01}(t) \exp\left(\gamma_1 Chemo_i + \omega_i\right)$$

$$\omega_i \sim \mathcal{N}(\mu, \sigma^2)$$

R: Example with **INLAjoint**

Joint recurrent and terminal event

readmission dataset (**frailtypack** R package)
The **joint frailty model** deals with recurrent events with a frailty model and shares the frailty term in a proportional hazards model for a terminal event.
The parameter φ_1 is **scaling the frailty term in the terminal event submodel**.

$$\begin{cases} \lambda_{i1}(t) = \lambda_{01}(t)\exp\left(\gamma_{11}Chemo_i + \omega_i\right) \text{ (S1)} \\ \lambda_{i2}(t) = \lambda_{02}(t)\exp\left(\gamma_{12}Chemo_i + \varphi_1\omega_i\right) \text{(S2)} \end{cases}$$

$$\omega_i \sim log\mathcal{N}(\mu, \sigma^2)$$

Joint model longitudinal and multi-state

$$\begin{cases} Y_{ij} = \eta_i(t_{ij}) + \varepsilon_{ij} = \beta_0 + b_{i0} + (\beta_1 + b_{i1})time_{ij} + \beta_{12}X_i + \varepsilon_{ij} & \text{(L1)} \\ \lambda_{i,12}(t) = \lambda_{0,12}(t)\exp\left(\gamma_{12}X_i + \varphi_{12}\eta_i(t)\right) & \text{(S1)} \\ \lambda_{i,13}(t) = \lambda_{0,13}(t)\exp\left(\gamma_{13}X_i + \varphi_{13}\eta_i(t)\right) & \text{(S2)} \\ \lambda_{i,23}(t) = \lambda_{0,23}(t)\exp\left(\gamma_{23}X_i + \varphi_{23}\eta_i(t)\right) & \text{(S3)} \end{cases}$$

R: Example using data *SurvMS*

References

- Rustand D, van Niekerk J, Rue H, Tournigand C, Rondeau V, Briollais L. *Bayesian estimation of two-part joint models for a longitudinal semicontinuous biomarker and a terminal event with INLA: Interests for cancer clinical trial evaluation.* Biometrical Journal. 2023 Apr;65(4):2100322, 10.1002/bimj.202100322 .
- Rustand, D., van Niekerk, J., Krainski, E. T., Rue, H., Proust-Lima, C. *Fast and flexible inference for joint models of multivariate longitudinal and survival data using integrated nested Laplace approximations*, Biostatistics, 2024; 25(2), pp.429-448, 10.1093/biostatistics/kxad019.
- Danilo Alvares, Janet van Niekerk, Elias T. Krainski, Håvard Rue, and Denis Rustand. *Bayesian survival analysis with INLA.* Statistics in Medicine, 2024; 1-36, 10.1002/sim.10160.
- Denis Rustand, Janet van Niekerk, Elias T. Krainski, and Håvard Rue. *Joint Modeling of Multivariate Longitudinal and Survival Outcomes with the R package INLAjoint.* arXiv preprint arXiv:2402.08335 (2024).

Thank you · شكرا

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology