Dynamic Predictions for Longitudinal and Event Time Outcomes with Applications in R

Dimitris Rizopoulos[1] **and Christos Thomadakis**[2]

[1]Erasmus University Medical Center, Rotterdam, the Netherlands
[2]Medical School, National and Kapodistrian University of Athens, Greece

45th Annual Conference of the International Society for Clinical Biostatistics

July 21, 2024, Thessaloniki, Greece

Contents

What is this Course About

- Often in follow-up studies different types of outcomes are collected

- Explicit outcomes
  - ▷ multiple longitudinal responses (e.g., markers, blood values)
  - ▷ time-to-event(s) of particular interest (e.g., death, relapse)

- Implicit outcomes
  - ▷ missing data
  - ▷ random visit times

What is this Course About (cont'd)

- Methods for the separate analysis of such outcomes are well established in the literature

- Survival data:
  - ▷ Cox model, accelerated failure time models, ...

- Longitudinal data
  - ▷ mixed effects models, GEE, marginal models, ...

What is this Course About (cont'd)

Purpose of this course is to present

Joint Modeling Techniques for Deriving Predictions

Learning Objectives

- After this course the participants will
  - ▷ be familiarized with joint modeling framework,
  - ▷ know how predictions are derived from joint models
  - ▷ know how to evaluate the accuracy of these predictions, and
  - ▷ be able to fit joint models in R and derive predictions

References

- Joint modeling sources*
  - ▷ Rizopoulos, D. (2012). *Joint Models for Longitudinal and Time-to-Event Data, with Applications in R*. Boca Raton: Chapman & Hall/CRC.
  - ▷ Fitzmaurice, G., Davidian, M., Verbeke, G. and Molenberghs, G. (2009). *Longitudinal Data Analysis*. Handbooks of Modern Statistical Methods. Boca Raton: Chapman & Hall/CRC, Chapter 15.
  - ▷ Wu, L. (2009). *Mixed Effects Models for Complex Data*. Boca Raton: Chapman & Hall/CRC, Chapter 8.
  - ▷ Ibrahim, J., Chen, M.-H. and Sinha, D. (2001). *Bayesian Survival Analysis*. New York: Springer-Verlag, Chapter 7.

* extra references of papers using joint modeling available at pp. 171–178.

References (cont'd)

- Useful material for package **JMbayes2**
  - ▷ a website with several examples: https://drizopoulos.github.io/JMbayes2/

- Useful material for package **JM** can be found in the web sites:
  - ▷ https://jmr.r-forge.r-project.org [R code used in the book]
  - ▷ https://www.drizopoulos.com/ → Software [additional R script files]

References (cont'd)

- Other software packages capable of fitting joint models
  - ▷ in **R**: **JMbayes** (by Rizopoulos), **joineR** (by Philipson et al.), **joineRML** (by Hickey et al.), function `stan_jm()` in **rstanarm** (by Brilleman), `jm_bamlss()` in **bamlss** (Koehler et al.), **lcmm** (by Proust-Lima et al.)
  - ▷ in **SAS**: **%JM** macro (by Garcia-Hernandez and Rizopoulos – http://www.jm-macro.com/), **%JMFit** macro (by Zhang et al.)
  - ▷ in **STATA**: **stjm** and **merlin** (by Crowther)

Part I

Introduction

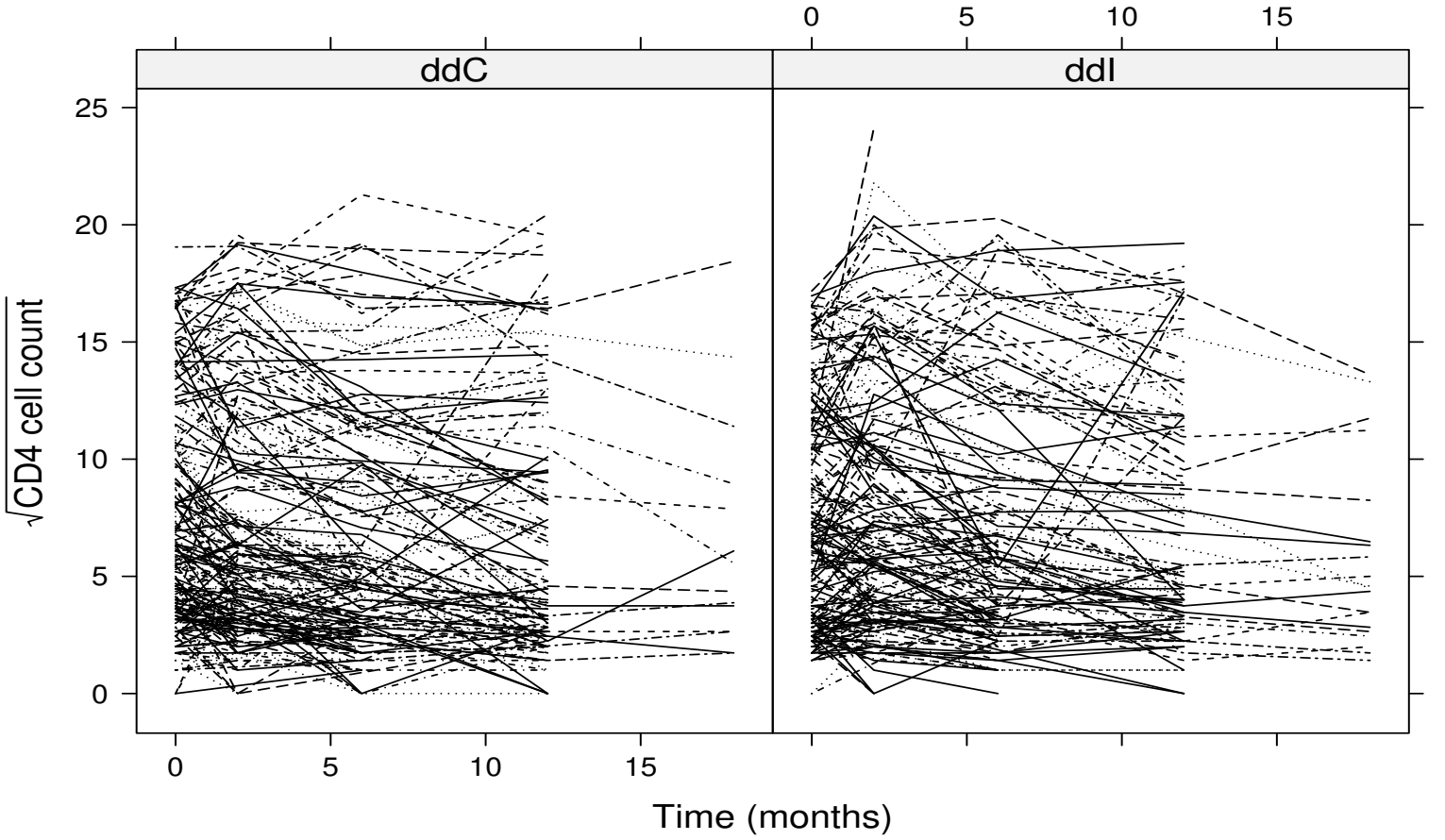
1.1 Motivating Longitudinal Studies

- AIDS: 467 HIV infected patients who had failed or were intolerant to zidovudine therapy (AZT) (Abrams et al., NEJM, 1994)

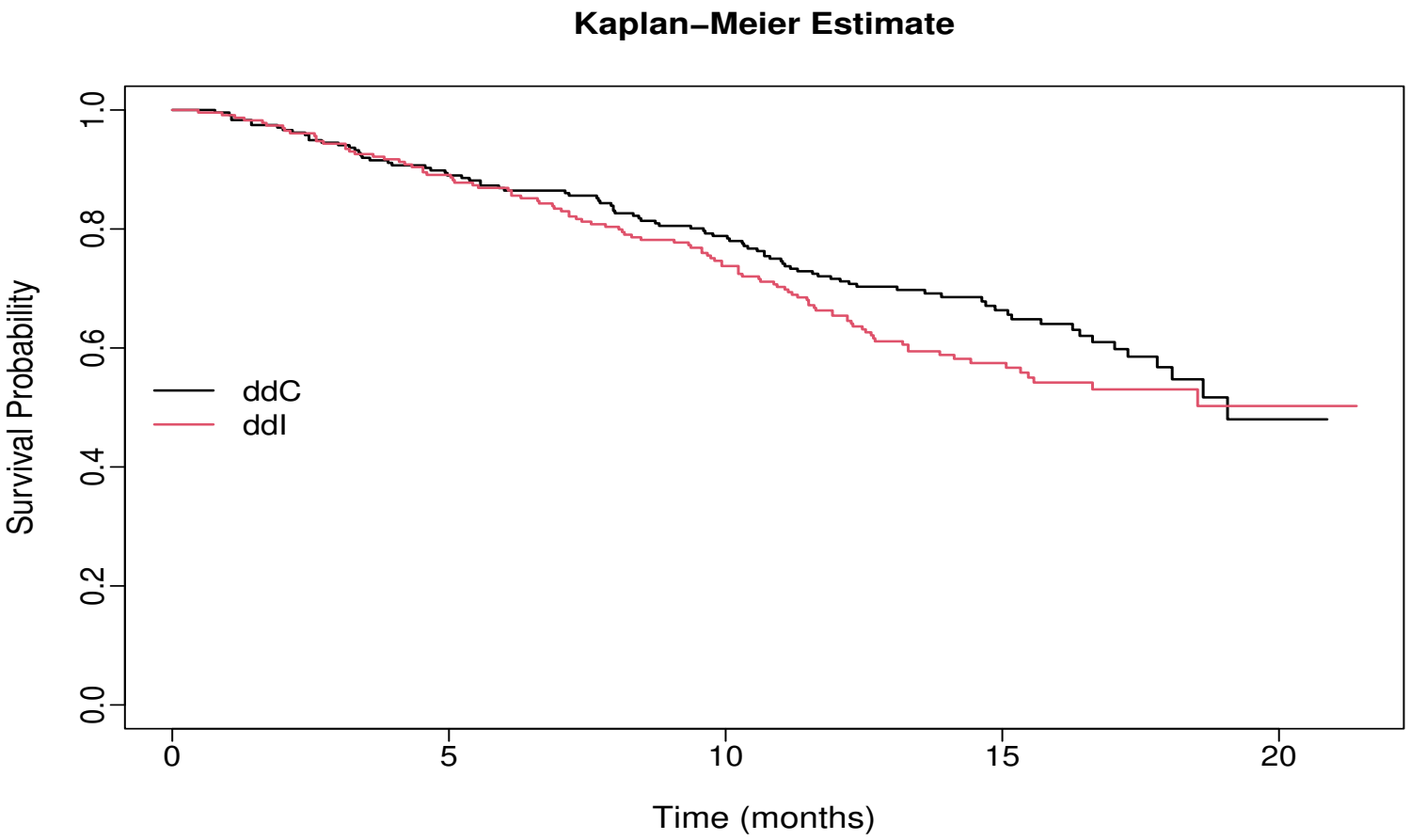
- The aim of this study was to compare the efficacy and safety of two alternative antiretroviral drugs, didanosine (ddI) and zalcitabine (ddC)

- Outcomes of interest:
  - ▷ time to death
  - ▷ randomized treatment: 230 patients ddI and 237 ddC
  - ▷ CD4 cell count measurements at baseline, 2, 6, 12 and 18 months

1.1 Motivating Longitudinal Studies (cont'd)



1.1 Motivating Longitudinal Studies (cont'd)



1.1 Motivating Longitudinal Studies (cont'd)

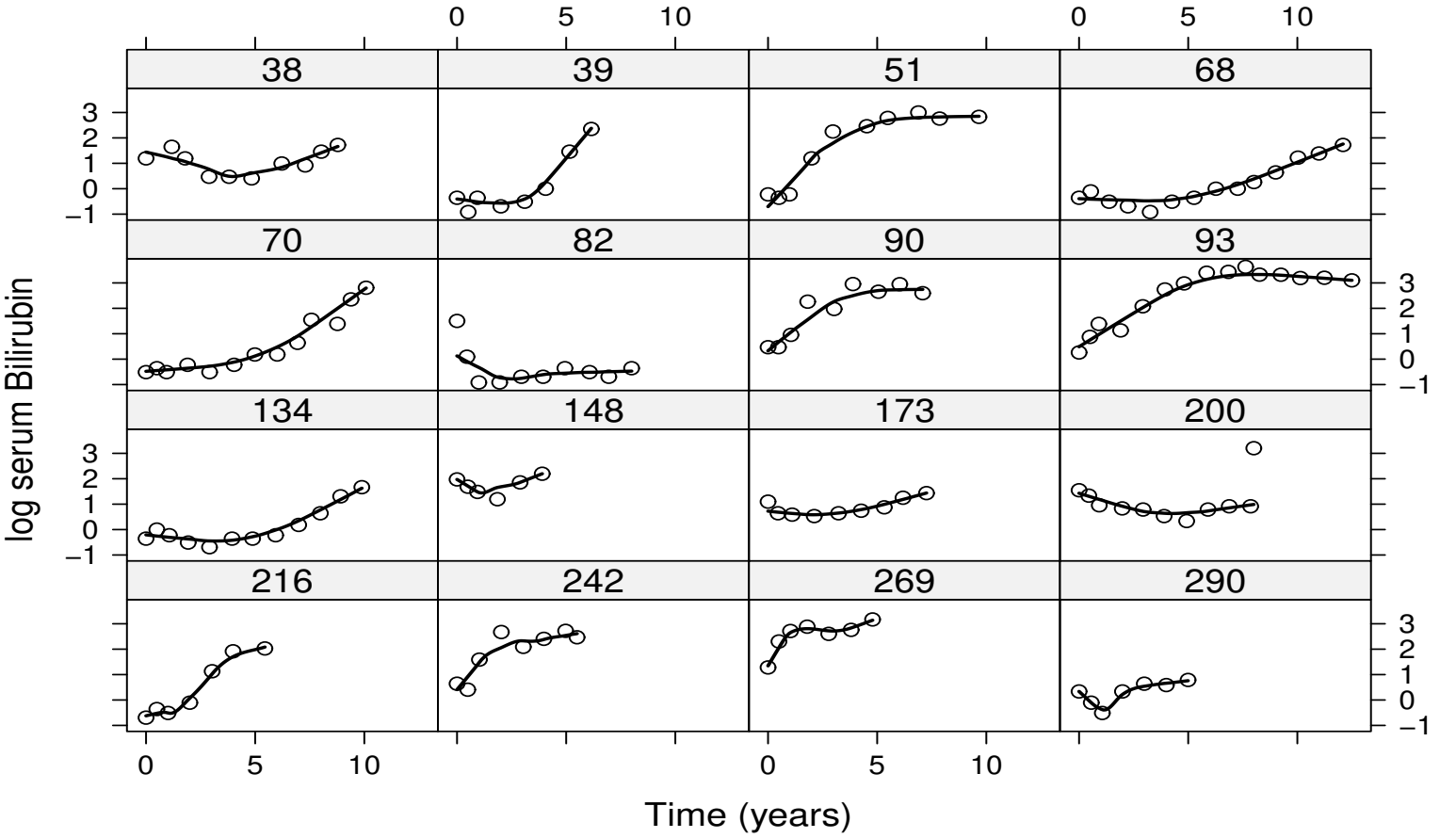
- Research Questions:
  - ▷ How strong is the association between CD4 cell count and the risk of death?
  - ▷ Is CD4 cell count a good biomarker?
    - * if treatment improves CD4 cell count, does it also improve survival?

1.1 Motivating Longitudinal Studies (cont'd)

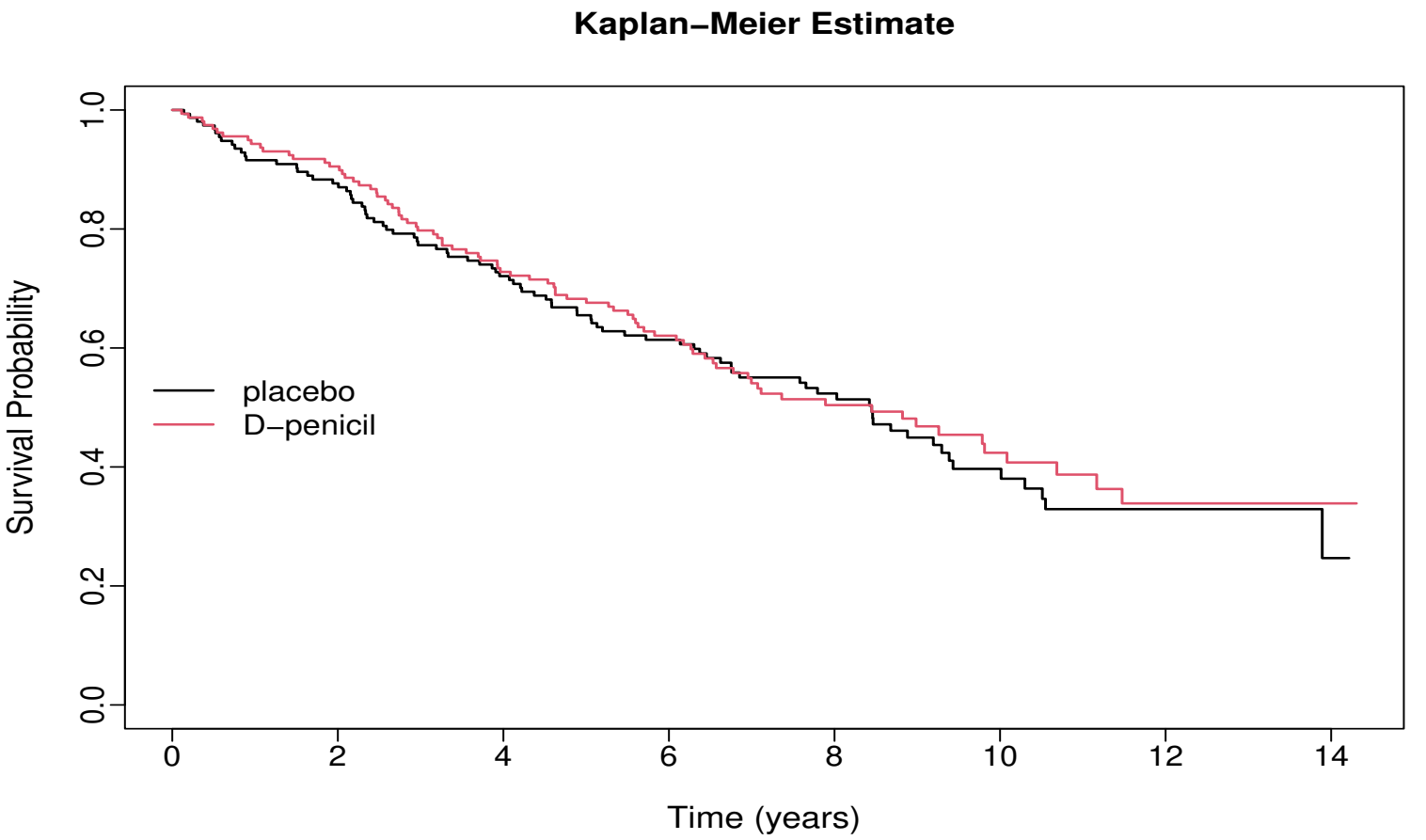
- PBC: Primary Biliary Cirrhosis:
  - ▷ a chronic, fatal but rare liver disease
  - ▷ characterized by inflammatory destruction of the small bile ducts within the liver

- Outcomes of interest:
  - ▷ time to death or liver transplantation
  - ▷ randomized treatment: 158 patients received D-penicillamine and 154 placebo
  - ▷ longitudinal bilirubin levels, cholesterol, prothrombin time (continuous)
  - ▷ longitudinal ascites, hepatomegaly, edema (categorical)

1.1 Motivating Longitudinal Studies (cont'd)



1.1 Motivating Longitudinal Studies (cont'd)



1.1 Motivating Longitudinal Studies (cont'd)

- Research Questions:
  - ▷ How strong is the association between bilirubin and the risk of death?
  - ▷ How the observed serum bilirubin levels could be utilized to provide predictions of survival probabilities?
  - ▷ Can bilirubin discriminate between patients of low and high risk?

1.2 Research Questions

- Depending on the questions of interest, different types of statistical analysis are required

- We will distinguish between two general types of analysis
  - ▷ separate analysis per outcome
  - ▷ joint analysis of outcomes

- Focus on each outcome separately
  - ▷ does treatment affect survival?
  - ▷ are the average longitudinal evolutions different between males and females?
  - ▷ ...

1.2 Research Questions (cont'd)

- Focus on multiple outcomes
  - ▷ **Complex effect estimation:** how strong is the association between the longitudinal evolution of CD4 cell counts and the hazard of death?
    - * *endogenous* vs. exogenous time-varying covariates
  - ▷ **Handling implicit outcomes:** focus on longitudinal outcomes but with dropout or random visit times
    - * *missing not at random* vs. missing at random

Part II

Review of Linear Mixed and Cox Models

2.1 Linear Mixed Models

- Repeated evaluations of the same outcome in each subject over time
  - ▷ CD4 cell count in HIV-infected patients
  - ▷ serum bilirubin in PBC patients

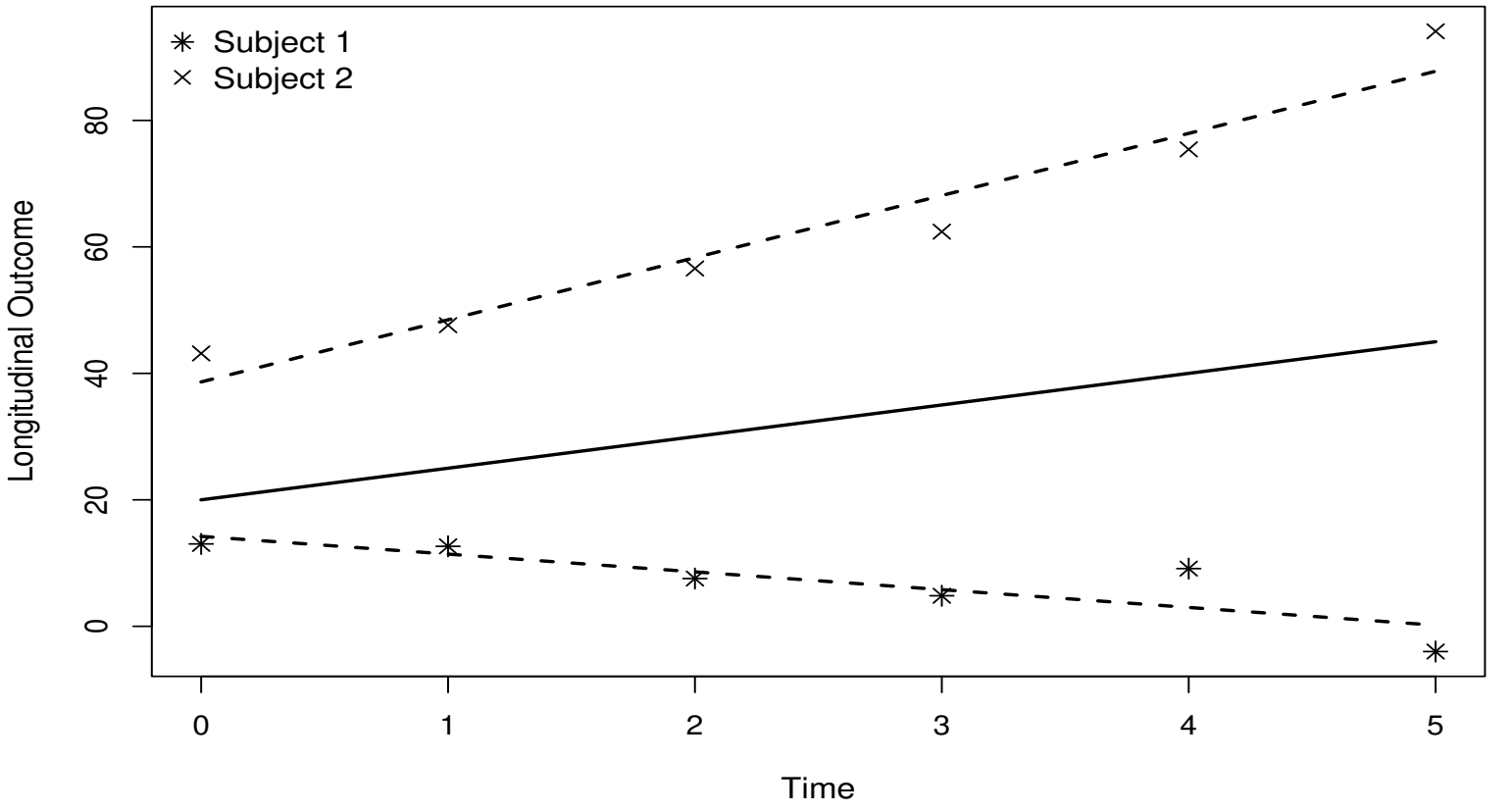
Measurements on the same subject are expected to be (positively) correlated

- Standard statistical tools, such as the t-test and linear regression that assume independent observations, not optimal for longitudinal data analysis

2.1 Linear Mixed Models (cont'd)

Random effects approach: Each subject in the population has her own subject-specific mean response profile over time

2.1 Linear Mixed Models (cont'd)



2.1 Linear Mixed Models (cont'd)

- The profile of each subject over time can be described by a linear model

$$y_{ij} = \tilde{\beta}_{i0} + \tilde{\beta}_{i1} t_{ij} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2),$$

  where

  - ▷ y_{ij} the jth response of the ith subject
  - ▷ $\tilde{\beta}_{i0}$ is the intercept and $\tilde{\beta}_{i1}$ the slope for subject i

- **Assumption:** Subjects are randomly sampled from a population $\Rightarrow$ subject-specific regression coefficients are also sampled from a population of regression coefficients

$$\tilde{\beta}_i \sim \mathcal{N}(\beta, D)$$

2.1 Linear Mixed Models (cont'd)

- We can reformulate the model as

$$y_{ij} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})t_{ij} + \varepsilon_{ij},$$

  where

  - ▷ βs are known as the *fixed effects*
  - ▷ b_is are known as the *random effects*

- In accordance for the random effects we assume

$$b_i = \begin{bmatrix} b_{i0} \\ b_{i1} \end{bmatrix} \sim \mathcal{N}(0, D)$$

2.1 Linear Mixed Models (cont'd)

- Put in a general form

$$\begin{cases} y_i = X_i\beta + Z_i b_i + \varepsilon_i, \\ \\ b_i \sim \mathcal{N}(0, D), \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2 \mathsf{I}_{n_i}), \end{cases}$$

with

- ▷ X design matrix for the fixed effects β
- ▷ Z design matrix for the random effects b_i
- ▷ $b_i \perp\!\!\!\perp \varepsilon_i$

2.1 Linear Mixed Models (cont'd)

- Interpretation:
  - ▷ β_j denotes the change in the average y_i when x_j is increased by one unit
  - ▷ b_i are interpreted in terms of how a subset of the regression parameters for the ith subject deviates from those in the population

- Advantageous feature: population + subject-specific predictions
  - ▷ β describes mean response changes in the population
  - ▷ $\beta + b_i$ describes individual response trajectories

2.2 Relative Risk Models

- The characteristic that distinguishes the analysis of time-to-event outcomes from other areas in statistics is **Censoring**
  - ▷ the event time of interest is not fully observed for all subjects under study

- Implications of censoring:
  - ▷ standard tools, such as the sample average, the t-test, and linear regression **cannot** be used
  - ▷ inferences may be sensitive to misspecification of the distribution of the event times

2.2 Relative Risk Models (cont'd)

- Notation (i denotes the subject)
  - ▷ T_i^* 'true' time-to-event
  - ▷ C_i the censoring time (e.g., the end of the study or a random censoring time)

- Available data for each subject
  - ▷ observed event time: $T_i = \min(T_i^*, C_i)$
  - ▷ event indicator: $\delta_i = 1$ if event; $\delta_i = 0$ if censored

Our aim is to make valid inferences for T_i^* but using only $\{T_i, \delta_i\}$

2.2 Relative Risk Models (cont'd)

- **Relative Risk Models** assume a multiplicative effect of covariates on the hazard scale, i.e.,

$$h_i(t) = h_0(t)\exp(\gamma_1 w_{i1} + \gamma_2 w_{i2} + \ldots + \gamma_p w_{ip}) \Rightarrow$$

$$\log h_i(t) = \log h_0(t) + \gamma_1 w_{i1} + \gamma_2 w_{i2} + \ldots + \gamma_p w_{ip},$$

where

  - ▷ $h_i(t)$ denotes the hazard of an event for patient i at time t
  - ▷ $h_0(t)$ denotes the baseline hazard
  - ▷ $w_{i1}, \ldots, w_{ip}$ a set of covariates

2.2 Relative Risk Models (cont'd)

- **Cox Model:** No assumptions for the baseline hazard function

- Parameter estimates and standard errors are based on the log partial likelihood function

$$p\ell(\gamma) = \sum_{i=1}^{n} \delta_i \Big[\gamma^\top w_i - \log \Big\{ \sum_{j:T_j \geq T_i} \exp(\gamma^\top w_j) \Big\} \Big],$$

  where only patients who had an event contribute

2.3 Time-Varying Covariates

- Often interest in the association between a time-varying covariate and the risk of an event
  - ▷ treatment changes with time (e.g., dose)
  - ▷ time-dependent exposure (e.g., smoking, diet)
  - ▷ markers of disease or patient condition (e.g., blood pressure, PSA levels)
  - ▷ ...

- Example: In the PBC study, are the longitudinal bilirubin measurements associated with the hazard of death?

2.3 Time-Varying Covariates (cont'd)

- There are two types of time-varying covariates

  (Kalbfleisch & Prentice, *The Stat. Anal. of Failure Time Data*, 2002)

  ▷ External (aka exogenous): the value of the covariate at time point t is not affected by the occurrence of an event at time point u, with $t > u$

  ▷ Internal (aka endogenous): not External

- This is a difficult concept and we will try to explain it with an example...

2.3 Time-Varying Covariates (cont'd)

- Example: A study on the time until an asthma attack for a group of patients

- We have two time-varying covariates: Pollution levels & a biomarker for asthma

- Say a patient had an asthma attack at a particular time point u
  - ▷ Pollution levels
    - * will the pollution levels at time $t > u$ be affected by the fact that the patient had an attack at u? $\Rightarrow$ **No**
  - ▷ Biomarker
    - * will the biomarker level at time $t > u$ be affected by the fact that the patient had an attack at u? $\Rightarrow$ **Yes**

2.3 Time-Varying Covariates (cont'd)

- It is **important** to distinguish between these two types of time-varying covariates, because the type of covariate dictates the appropriate type of analysis

- The extended Cox model is only valid for exogenous time-varying covariates

Treating endogenous covariates as exogenous may produce spurious results!

Part III

The Basic Joint Model

3.1 Joint Modeling Framework

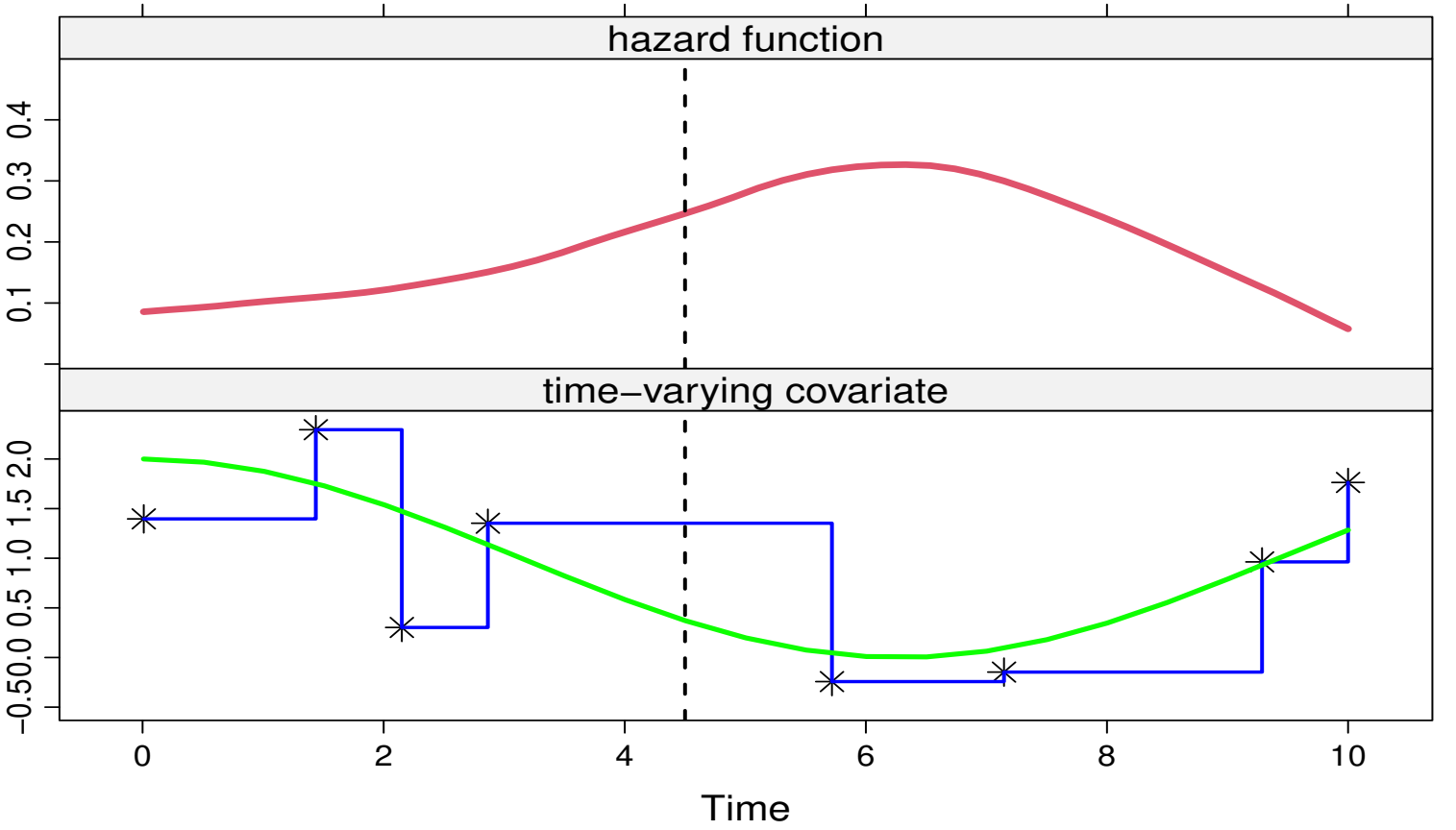
- To account for the special features of endogenous covariates a new class of models has been developed

Joint Models for Longitudinal and Time-to-Event Data

- Intuitive idea behind these models
  1. use an appropriate model to describe the evolution of the covariate/marker over time for each patient
  2. the estimated evolutions are then used in a Cox model

- Feature: covariate level's are **not** assumed constant between visits

3.1 Joint Modeling Framework (cont'd)



3.1 Joint Modeling Framework (cont'd)

- Some notation
  - ▷ T_i^*: True event time for patient i
  - ▷ T_i: Observed event time for patient i
  - ▷ δ_i: Event indicator, i.e., equals 1 for true events
  - ▷ y_i: Longitudinal covariate

- We will formulate the joint model in 3 steps – in particular, ...

3.1 Joint Modeling Framework (cont'd)

- Step 1: Let's assume that we know $m_i(t)$, i.e., the *true* & *unobserved* value of the covariate at time t

- Then, we can define a standard relative risk model

$$h_i(t \mid \mathcal{M}_i(t)) = h_0(t) \exp\{\gamma^\top w_i + \alpha m_i(t)\},$$

where

  - ▷ $\mathcal{M}_i(t) = \{m_i(s), 0 \leq s < t\}$ longitudinal history
  - ▷ α quantifies the association between the time-varying covariate and the risk of an event
  - ▷ w_i baseline covariates

3.1 Joint Modeling Framework (cont'd)

- Step 2: From the observed longitudinal data $y_i(t)$ reconstruct the covariate history for each subject

- Mixed effects model (we focus, for now, on continuous covariates)

$$\begin{aligned} y_i(t) &= m_i(t) + \varepsilon_i(t) \\ &= x_i^\top(t)\beta + z_i^\top(t)b_i + \varepsilon_i(t), \qquad \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2), \end{aligned}$$

where

  - ▷ $x_i(t)$ and β: Fixed-effects part
  - ▷ $z_i(t)$ and b_i: Random-effects part, $b_i \sim \mathcal{N}(0, D)$

3.1 Joint Modeling Framework (cont'd)

- Step 3: The two processes are associated $\Rightarrow$ define a model for their joint distribution

- Joint Models for such joint distributions are of the following form
  (Tsiatis & Davidian, Stat. Sinica, 2004)

$$p(y_i, T_i, \delta_i) = \int p(y_i \mid b_i) \left\{ h(T_i \mid b_i)^{\delta_i} S(T_i \mid b_i) \right\} p(b_i)\, db_i,$$

  where

  ▷ b_i a vector of random effects that explains the interdependencies

  ▷ $p(\cdot)$ density function; $S(\cdot)$ survival function

3.1 Joint Modeling Framework (cont'd)

- Key assumption: **Full Conditional Independence** $\Rightarrow$ random effects explain all interdependencies

  ▷ the longitudinal outcome is independent of the time-to-event outcome

  ▷ the repeated measurements in the longitudinal outcome are independent of each other

$$p(y_i, T_i, \delta_i \mid b_i) = p(y_i \mid b_i)\, p(T_i, \delta_i \mid b_i)$$

$$p(y_i \mid b_i) = \prod_j p(y_{ij} \mid b_i)$$

3.2 Bayesian Estimation

- Under the Bayesian paradigm, both θ and $\{b_i, i = 1, \ldots, n\}$ are regarded as parameters

- Inference via the full posterior distribution

$$p(\theta, b \mid T, \delta, y) = \frac{\prod_i p(T_i, \delta_i \mid b_i, \theta)\; p(y_i \mid b_i, \theta)\; p(b_i, \theta)\; p(\theta)}{\prod_i p(T_i, \delta_i, y_i)}$$

$$\propto \prod_{i=1}^{n} \Big\{ p(T_i, \delta_i \mid b_i, \theta)\; p(y_i \mid b_i, \theta)\; p(b_i, \theta) \Big\}\; p(\theta)$$

3.2 Bayesian Estimation (cont'd)

- Inference via sampling from the posterior
  - ▷ Markov Chain Monte Carlo
  - ▷ Hamiltonian Monte Carlo

- Model comparison: *Information Criteria for Predictive Accuracy*
  - ▷ Deviance information criterion (DIC)
  - ▷ Watanabe-Akaike information criterion (WAIC)
  - ▷ log pseudo-marginal likelihood (LPML)

3.3 A Comparison with the TD Cox

- Example: To illustrate the virtues of joint modeling, we compare it with the standard time-dependent Cox model for the AIDS data

$$\begin{cases} y_i(t) &= m_i(t) + \varepsilon_i(t) \\ &= \beta_0 + \beta_1 t + \beta_2 \{t \times \texttt{ddI}_i\} + b_{i0} + b_{i1} t + \varepsilon_i(t), \qquad \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2), \\ \\ h_i(t) &= h_0(t) \exp\{\gamma \texttt{ddI}_i + \alpha m_i(t)\}, \end{cases}$$

3.3 A Comparison with the TD Cox (cont'd)

	JM	Cox
	log HR (std.err)	log HR (std.err)
Treat	0.33 (0.2)	0.31 (0.15)
$CD4^{1/2}$	−0.29 (0.04)	−0.19 (0.02)

- Clearly, there is a considerable effect of ignoring the measurement error, especially for the CD4 cell counts

3.3 A Comparison with the TD Cox (cont'd)

- A unit decrease in $CD4^{1/2}$, results in a
  - ▷ **Joint Model**: 1.33-fold increase in risk (95% CI: 1.24; 1.43)
  - ▷ **Time-Dependent Cox**: 1.21-fold increase in risk (95% CI: 1.16; 1.27)

- Which one to believe?
  - ▷ a lot of theoretical and simulation work has shown that the Cox model underestimates the true association size of endogenous time-varying covariates

3.4 Joint Models in R

R> Joint models are fitted using function `jm()` from package **JMbayes2**, e.g.,

```
lmeFit <- lme(CD4 ~ obstime + obstime:drug, data = aids,
              random = ~ obstime | patient)

CoxFit <- coxph(Surv(Time, death) ~ drug, data = aids.id)

jointFit <- jm(CoxFit, lmeFit, time_var = "obstime")

summary(jointFit)
```

3.4 Joint Models in R (cont'd)

R> The data frame given in `lme()` should be in the long format, while the data frame given to `coxph()` should have one line per subject*

▷ the ordering of the subjects needs to be the same

R> The scale of the time variables in the mixed and Cox models need to be the same

▷ i.e., both in months, or both in years, etc.

R> Argument `time_var` specifies the time variable in the linear mixed model

* Unless you want to include exogenous time-varying covariates or handle competing risks

3.4 Joint Models in R (cont'd)

R> Useful functions

- ▷ `summary()`: summarizes the fitted model
- ▷ `compare_jm()`: compares fitted models using DIC and WAIC
- ▷ `coef()`, `fixef()`, `ranef()`: extract estimated coefficients and random effects
- ▷ `traceplot()` & `ggtraceplot`: produces traceplots
- ▷ `densplot()` & `ggdensityplot()`: produces density plots
- ▷ `predict()`: calculates predictions

Part IV

Joint Model Extensions

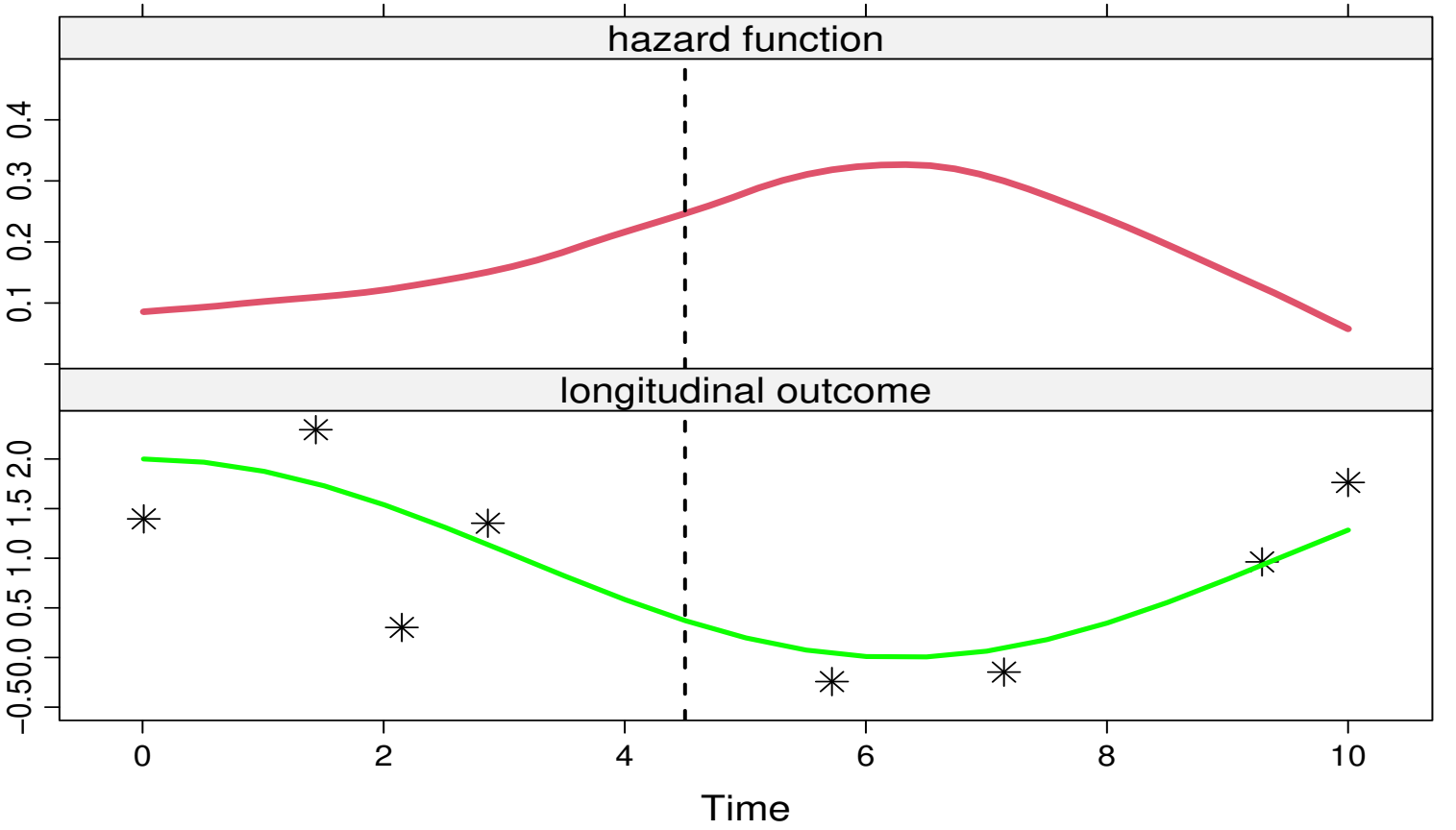
4.1 Functional Forms

- The standard joint model

$$\begin{cases} h_i(t \mid \mathcal{M}_i(t)) &= h_0(t)\exp\{\gamma^\top w_i + \alpha m_i(t)\}, \\ \\ y_i(t) &= m_i(t) + \varepsilon_i(t) \\ &= x_i^\top(t)\beta + z_i^\top(t)b_i + \varepsilon_i(t), \end{cases}$$

where $\mathcal{M}_i(t) = \{m_i(s), 0 \leq s < t\}$

4.1 Functional Forms (cont'd)



4.1 Functional Forms (cont'd)

- The standard joint model

$$\begin{cases} h_i(t \mid \mathcal{M}_i(t)) &= h_0(t)\exp\{\gamma^\top w_i + \alpha m_i(t)\}, \\ \\ y_i(t) &= m_i(t) + \varepsilon_i(t) \\ &= x_i^\top(t)\beta + z_i^\top(t)b_i + \varepsilon_i(t), \end{cases}$$

where $\mathcal{M}_i(t) = \{m_i(s), 0 \leq s < t\}$

Is this the only option? Is this the most optimal choice?

4.1 Functional Forms (cont'd)

- Note: Inappropriate modeling of time-varying covariates may result in surprising results

- Example: Cavender et al. (1992, J. Am. Coll. Cardiol.) conducted an analysis to test the effect of cigarette smoking on survival of patients who underwent coronary artery surgery
  - ▷ the estimated effect of current cigarette smoking was positive on survival although not significant (i.e., patient who smoked had higher probability of survival)
  - ▷ most of those who had died were smokers but many stopped smoking at the last follow-up before their death

4.1 Functional Forms (cont'd)

We need to carefully consider the functional form of time-varying covariates

- Let's see some possibilities...

4.1 Functional Forms (cont'd)

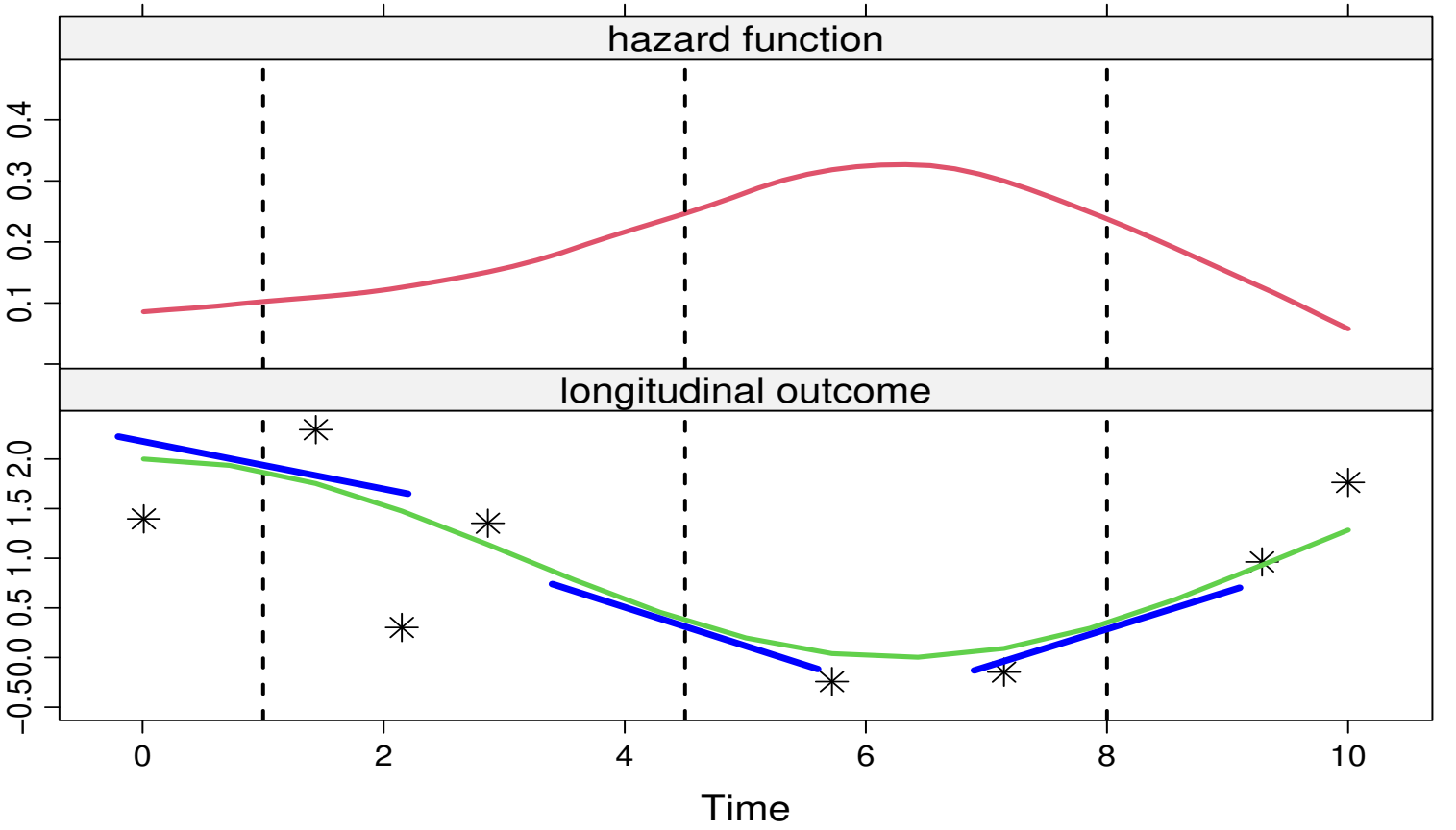
- *Time-dependent Slopes:* The hazard of an event at t is associated with both the current value and the slope of the trajectory at t (Ye et al., 2008, Biometrics):

$$h_i(t \mid \mathcal{M}_i(t)) = h_0(t) \exp\{\gamma^\top w_i + \alpha_1 m_i(t) + \alpha_2 m_i'(t)\},$$

where

$$m_i'(t) = \frac{d}{dt}\{x_i^\top(t)\beta + z_i^\top(t) b_i\}$$

4.1 Functional Forms (cont'd)



4.1 Functional Forms (cont'd)

- The definition of the slope is

$$m_i'(t) = \lim_{\epsilon \to 0} \frac{m_i(t+\epsilon) - m_i(t)}{\epsilon}$$

  the change in the longitudinal profile *as ϵ approaches zero*

- It can be challenging to interpret
  - ▷ it is the 'current' slope

4.1 Functional Forms (cont'd)

- *Time-dependent Slopes 2:* The hazard of an event at t is associated with the change of the trajectory the last year:

$$h_i(t \mid \mathcal{M}_i(t)) \;=\; h_0(t)\exp\{\gamma^\top w_i + \alpha\Delta m_i(t)\},$$

where

$$\Delta m_i(t) = m_i(t) - m_i(t-1)$$

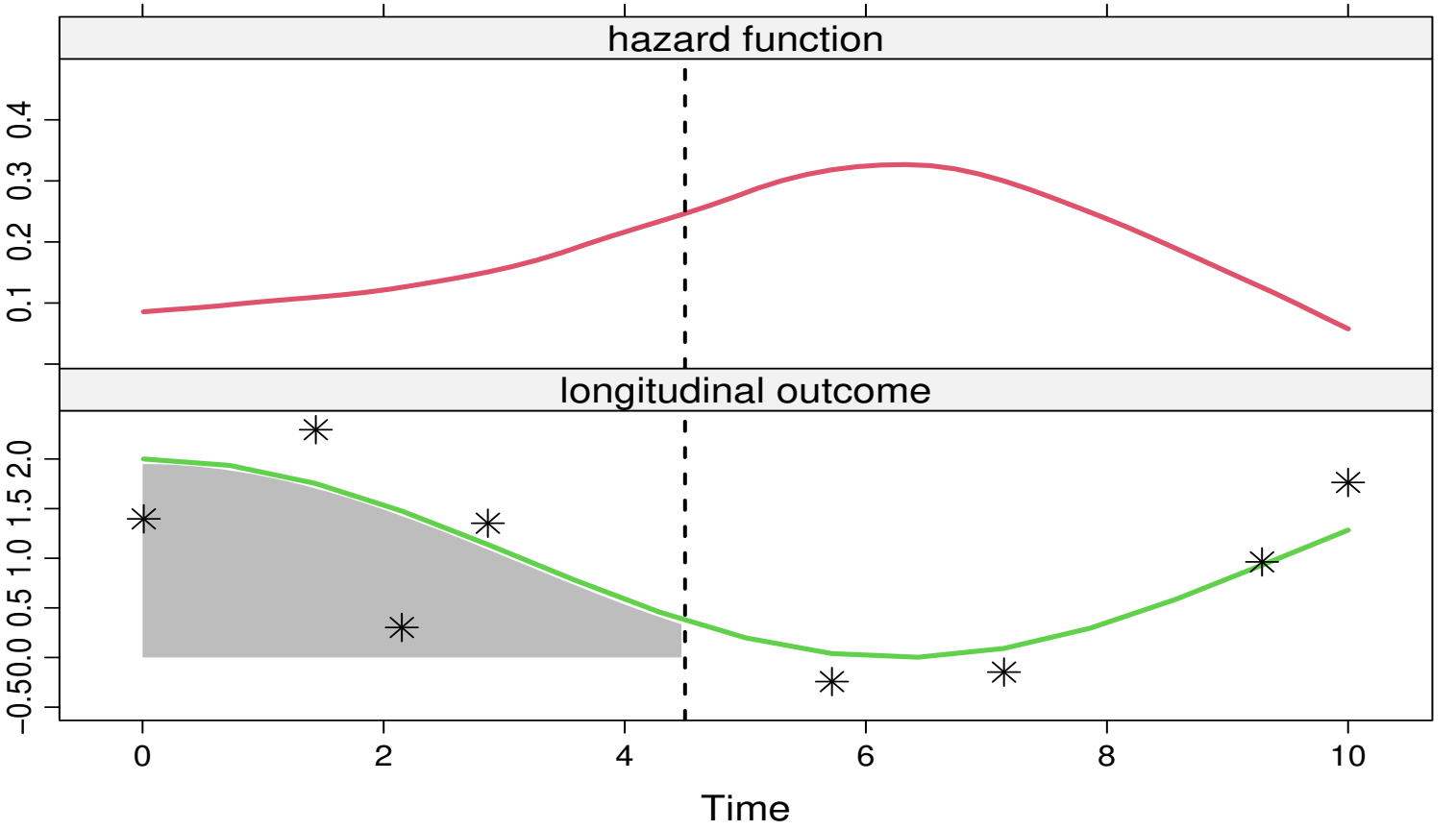
4.1 Functional Forms (cont'd)

- *Cumulative Effects:* The hazard of an event at t is associated with the whole area under the trajectory up to t:

$$h_i(t \mid \mathcal{M}_i(t)) \;=\; h_0(t)\exp\Big\{\gamma^\top w_i + \alpha \int_0^t m_i(s)\; ds\Big\}$$

- Area under the longitudinal trajectory taken as a summary of $\mathcal{M}_i(t)$

4.1 Functional Forms (cont'd)



4.1 Functional Forms (cont'd)

- *Cumulative Effects 2:* The hazard of an event at t is associated with the whole area under the trajectory up to t:

$$h_i(t \mid \mathcal{M}_i(t)) = h_0(t) \exp\left\{ \gamma^\top w_i + \alpha \frac{\int_0^t m_i(s)\, ds}{t} \right\}$$

- We account for the observation period

4.1 Functional Forms (cont'd)

R> In **JMbayes2** the specification of functional forms is done via the `functional_forms` argument

- ▷ e.g., the following code includes the area and slope in the linear predictor, and the interaction of the former with sex

```
jm(CoxFit, lmeFit, time_var = "time",
   functional_forms = ~ area(y) + value(y) + area(y):sex)
```

4.1 Functional Forms (cont'd)

R> The `area()` function calculates the *Cumulative Effects 2* functional form, where the integral is divide by the length of the period

R> The `slope()` function can be used for the *Time-dependent Slopes 2* functional form via

```
slope(..., eps = 1, direction = "back")
```

4.2 Multiple Longitudinal Markers

- So far we have concentrated on a single continuous longitudinal outcome

- But very often we may have several outcomes we wish to study, some of which could be categorical

- Example: In the PBC dataset we have used serum bilirubin as the most important marker, but during follow-up several other markers have been recorded
  - ▷ serum cholesterol (continuous)
  - ▷ edema (3 categories)
  - ▷ ascites (2 categories)
  - ▷ ...

4.2 Multiple Longitudinal Markers (cont'd)

We need to extend the basic joint model!

- To handle multiple longitudinal outcomes of different types we use Generalized Linear Mixed Models

  ▷ We assume $Y_{i1}, \ldots, Y_{iJ}$ for each subject, each one having a distribution in the exponential family, with expected value

$$m_{ij}(t) \;=\; E(y_{ij}(t) \mid b_{ij}) = g_j^{-1}\{x_{ij}^{\top}(t)\beta_j + z_{ij}^{\top}(t)b_{ij}\},$$

  with $g(\cdot)$ denoting a link function

4.2 Multiple Longitudinal Markers (cont'd)

- Correlation between the longitudinal outcomes is captured by assuming a multivariate normal distribution for the random effects

$$b_i = \begin{bmatrix} b_{i1} \\ \vdots \\ b_{iJ} \end{bmatrix} \sim \mathcal{N}(0, D)$$

4.2 Multiple Longitudinal Markers (cont'd)

- Two ways to include the longitudinal markers in the survival submodel
  - ▷ conditional expected value

$$h_i(t) = h_0(t)\exp\{\gamma^\top w_i + \sum_{j=1}^{J}\alpha_j m_{ij}(t)\}$$

  - ▷ or conditional linear predictor

$$\begin{cases} h_i(t) = h_0(t)\exp\{\gamma^\top w_i + \sum_{j=1}^{J}\alpha_j \eta_{ij}(t)\} \\ \\ \eta_{ij} = x_{ij}^\top(t)\beta_j + z_{ij}^\top(t)b_{ij} \end{cases}$$

4.2 Multiple Longitudinal Markers (cont'd)

- **Full Conditional Independence**: Given the random effects
  - ▷ the repeated measurements in each outcome are independent,
  - ▷ the longitudinal outcomes are independent of each other, and
  - ▷ longitudinal outcomes are independent of the time-to-event outcome

$$
\begin{aligned}
p(y_{ij} \mid b_{ij}) &= \prod_{k=1}^{n_{ij}} p(y_{ij,k} \mid b_{ij}) \\
p(y_i \mid b_i) &= \prod_{j} p(y_{ij} \mid b_{ij}) \\
p(y_i, T_i, \delta_i \mid b_i) &= \prod_{j} p(y_{ij} \mid b_{ij})\, p(T_i, \delta_i \mid b_i)
\end{aligned}
$$

4.2 Multiple Longitudinal Markers (cont'd)

- Features of multivariate joint models
  - ▷ using CI is straightforward to extend joint models to multiple longitudinal outcomes of different types
  - ▷ computationally much more intensive due to the high dimensional random effects

4.2 Multiple Longitudinal Markers (cont'd)

- Example: Multivariate joint model for the PBC dataset
  - ▷ log(ser Bilir): linear mixed-effects model
    - * fixed effects: intercept and linear time effect
    - * random effects: intercept and linear time effect
  - ▷ spiders: mixed-effects logistic regression model
    - * fixed effects: intercept and linear time effect
    - * random effects: intercept

4.2 Multiple Longitudinal Markers (cont'd)

- time-to-death: relative risk model
  - baseline covariates: drug and age
  - Analysis I: conditional linear predictor
  - Analysis II: conditional expected value

4.2 Multiple Longitudinal Markers (cont'd)

- Analysis I: conditional linear predictor

	Value	Std.Dev.	2.5%	97.5%
D-penicil	−0.071	0.234	−0.530	0.373
Age	0.064	0.009	0.046	0.082
value(logSB)	1.317	0.108	1.111	1.531
value(spiders)	0.070	0.048	−0.024	0.167

4.2 Multiple Longitudinal Markers (cont'd)

- Analysis II: conditional expected value

	Value	Std.Dev.	2.5%	97.5%
D-penicil	−0.080	0.234	−0.545	0.373
Age	0.063	0.009	0.045	0.081
value(logSB)	1.326	0.109	1.113	1.540
expit(value(spiders))	0.458	0.347	−0.228	1.134

4.2 Multiple Longitudinal Markers (cont'd)

R> To fit a multivariate joint model in **JMbayes2** we need first to fit a series of univariate mixed models.

▷ for non-Gaussian longitudinal data we use **GLMMadaptive**

```
mixed_model(spiders ~ year, data = pbc2,
            family = binomial(), random = ~ year | id)
```

- Arguments of `mixed_model()`
  - ▷ `fixed`: formula for the response outcome and fixed effects
  - ▷ `random`: formula for random effects
  - ▷ `family`: distribution of longitudinal outcome
  - ▷ `data`: dataset

4.2 Multiple Longitudinal Markers (cont'd)

R> To fit a multivariate joint model, we use `jm()` as before but we now provide a `list()` of mixed models

- ▷ an example for the PBC dataset using serum bilirubin (continuous) and spiders (binary)

```
lmmFit <- lme(log(serBilir) ~ year, data = pbc2, random = ~ year | id)

melrFit <- mixed_model(spiders ~ year, data = pbc2, family = binomial(),
                       random = ~ 1 | id)

CoxFit <- coxph(Surv(years, status2) ~ drug + age, data = pbc2.id)

jm(CoxFit, list(lmmFit, melrFit), time_var = "year")
```

4.2 Multiple Longitudinal Markers (cont'd)

R> The default in `jm()` is to include the conditional linear predictor $\eta_{ij}(t)$ in the survival submodel

▷ to include the conditional expected value, we can use the `functional_forms` argument, e.g.,

```
jm(CoxFit, list(lmmFit, melrFit), time_var = "year",
   functional_forms = ~ value(log(serBilir)) +
            vexpit(value(spiders)),
   n_iter = 20000L, n_burnin = 10000L)
```

4.2 Multiple Longitudinal Markers (cont'd)

R> Function `jm()` allows for various types of mixed models

▷ continuous: Student's t, beta, gamma, censored normal

▷ categorical: binomial, Poisson, negative binomial, beta binomial

For more info see
https://drizopoulos.github.io/JMbayes2/
→ Articles → Non-Gaussian Mixed Models

4.3 Competing Risks

- Often multiple failure times are recorded

- **Competing risks:** Occurrence of one event either
  - ▷ precludes the occurrence of other events or
  - ▷ substantially alters the probability of observing the other events

4.3 Competing Risks (cont'd)

- Example: In the PBC dataset $\Rightarrow$ competing risks
  - ▷ some patients received a liver transplantation
  - ▷ so far we have used the composite event, i.e. death or transplantation whatever comes first
  - ▷ when interest only is on one type of event, the other should be considered as a competing risk

- Example: In HIV studies
  - ▷ death while in care
  - ▷ disengagement from care

4.3 Competing Risks (cont'd)

- Example: Alzheimer's disease studies
  - ▷ dementia
  - ▷ death without dementia

4.3 Competing Risks (cont'd)

- In principle, competing-risk data can be analyzed through either
  - ▷ cause-specific hazards
  - ▷ cumulative incidence functions (CIFs)

4.3 Competing Risks (cont'd)

- Let
  - ▷ $T_i^* = \min(T_{i1}^*, \ldots, T_{iK}^*)$ be the survival time
  - ▷ $\delta_i^* \in \{1, \ldots, K\}$ be the failure cause

- **Cause-specific hazards**: the rate of failure from a particular cause at a specific time point given that the individual has survived up to that point:

$$h_{ik}(t) = \lim_{dt \to 0} \frac{P(t < T_i^* \leq t + dt, \delta_i^* = k \mid T_i^* > t)}{dt}$$

4.3 Competing Risks (cont'd)

- Proportional cause-specific hazards are usually applied in practice

$$h_{ik}(t) = h_{0k}(t) \exp(x_{ik}^\top \beta_k)$$

where

  - ▷ x_{ik} baseline covariates (possibly cause-specific)
  - ▷ β_k $\log$ cause-specific hazard ratios

4.3 Competing Risks (cont'd)

- If right-censoring occurs
  - ▷ $T_i = \min(T_{i1}^*, \ldots, T_{iK}^*, C_i)$, with C_i denoting the censoring time
  - ▷ $\delta_i \in \{0, 1, \ldots, K\}$, with 0 corresponding to censoring

- The likelihood becomes a product over failure causes

$$p(T_i, \delta_i) = \prod_{k=1}^{K} h_{ik}(T_i)^{I(\delta_i=k)} \exp\left\{ -\sum_{k=1}^{K} \int_0^{T_i} h_{ik}(u)du \right\}$$

Standard (e.g., Cox) models for each cause can be fitted separately by treating the other failure causes as non-informative right censoring!

4.3 Competing Risks (cont'd)

R> To fit cause-specific hazard models, e.g., through `coxph()`, we just treat events from other causes as right-censored

Death

```
CoxFitDeath <- coxph(Surv(years, status == "dead") ~ drug + age, data = pbc2.id)
  n= 312, number of events= 140

                   coef exp(coef)  se(coef)      z Pr(>|z|)
drugD-penicil -0.162071  0.850380  0.172501 -0.940    0.347
age            0.045718  1.046780  0.008487  5.387 7.16e-08 ***
```

4.3 Competing Risks (cont'd)

R> To fit cause-specific hazard models, e.g., through `coxph()`, we just treat events from other causes as right-censored

Transplantation

```
CoxFitTranspl <- coxph(Surv(years, status == "transplanted") ~ drug + age,
data = pbc2.id)
  n= 312, number of events= 29

                  coef exp(coef) se(coef)      z Pr(>|z|)
drugD-penicil -0.23680   0.78915  0.37723 -0.628     0.53
age           -0.09649   0.90802  0.02265 -4.259 2.05e-05 ***
```

The effect of age has an opposite direction!

4.3 Competing Risks (cont'd)

Cumulative incidence function (CIF) The probability of occurrence of a specific cause over time

$$F_{ik}(t) = \Pr(T_i^* \leq t, \delta_i^* = k) = \int_0^t h_{ik}(u) \exp\left\{ - \int_0^u \sum_{k=1}^{K} h_{ik}(s)\, ds \right\} du$$

- Complex function of cause-specific hazards

- Semi-parametric modeling of **sub-distribution hazards**, $\lambda_{ik}(t)$, proposed by Fine & Gray (1999) is typically performed for the event of interest as

$$F_{ik}(t) = 1 - \exp\left\{ - \int_0^t \lambda_{ik}(u) du \right\}$$

there is an **1-1 relationship** between $F_{ik}(t)$ and $\lambda_{ik}(t)$

4.3 Competing Risks (cont'd)

R> Proportional sub-distribution hazard models, for each event type, can be fitted through function `crr()` of package **cmprsk**

▷ For death

```
Call:
crr(ftime = pbc2.id$years, fstatus = pbc2.id$status,
    cov1 = mat, cengroup = "alive", failcode = "dead")

              coef exp(coef) se(coef)     z p-value
D-penicil -0.1915     0.826  0.16899 -1.13 2.6e-01
age        0.0479     1.049  0.00813  5.90 3.7e-09
```

4.3 Competing Risks (cont'd)

- Aetiological-type research questions $\rightarrow$ cause-specific hazards

- Prognosis of a disease and prediction purposes $\rightarrow$ CIF

4.3 Competing Risks (cont'd)

- Most of the research in joint modeling was initially focused on a single event

- Joint modeling of longitudinal data and competing-risk survival data has also gained attention

- Example: In the PBC dataset $\Rightarrow$ competing risks
  - ▷ death
  - ▷ liver transplantation

4.3 Competing Risks (cont'd)

- Joint models with competing risks:

$$\begin{cases} y_i(t) = m_i(t) + \varepsilon_i(t) = x_i^\top(t)\beta + z_i^\top(t)b_i + \varepsilon_i(t), \\ h_i^d(t) = h_0^d(t)\exp\{\gamma_d^\top w_i + \alpha_d m_i(t)\}, \\ h_i^{tr}(t) = h_0^{tr}(t)\exp\{\gamma_{tr}^\top w_i + \alpha_{tr} m_i(t)\}, \end{cases}$$

where

- ▷ $h_i^d(t)$ hazard function for death
- ▷ $h_i^{tr}(t)$ hazard function for transplantation

4.3 Competing Risks (cont'd)

- When two markers are used:

$$\begin{cases} y_{i1}(t) = m_{i1}(t) + \varepsilon_{i1}(t) = x_{i1}^{\top}(t)\beta_1 + z_{i1}^{\top}(t)b_{i1} + \varepsilon_{i1}(t), \\ y_{i2}(t) = m_{i2}(t) + \varepsilon_{i2}(t) = x_{i2}^{\top}(t)\beta_2 + z_{i2}^{\top}(t)b_{i2} + \varepsilon_{i2}(t), \\ h_i^d(t) = h_0^d(t)\exp\{\gamma_d^{\top} w_i + \alpha_{d1} m_{i1}(t) + \alpha_{d2} m_{i2}(t)\}, \\ h_i^{tr}(t) = h_0^{tr}(t)\exp\{\gamma_{tr}^{\top} w_i + \alpha_{tr1} m_{i1}(t) + \alpha_{tr2} m_{i2}(t)\}, \end{cases}$$

4.3 Competing Risks (cont'd)

- In the estimation, the only difference is in the construction of the likelihood part for the event process

$$p(T_i, \delta_i \mid b_i; \theta) = \prod_{k=1}^{K} \Big[h_{0k}(T_i) \exp\{\gamma_k^\top w_i + \alpha_k m_i(T_i)\} \Big]^{I(\delta_i = k)} \times \exp\left(- \sum_{k=1}^{K} \int_0^{T_i} h_{0k}(s) \exp\big\{ \gamma_k^\top w_i + \alpha_k m_i(s) \big\} \, ds \right),$$

with

- $T_i = \min(T_{i1}^*, \ldots, T_{iK}^*, C_i)$, with C_i denoting the censoring time
- $\delta_i \in \{0, 1, \ldots, K\}$, with 0 corresponding to censoring

4.3 Competing Risks (cont'd)

- This is different than in standard Cox models

We cannot fit a cause-specific hazard joint model by treating events from other causes as censored!

4.3 Competing Risks (cont'd)

- Example: Competing risks analysis for the PBC dataset
  - ▷ log(ser Bilir): linear mixed-effects model
    - * fixed effects: intercept, drug, linear time, interaction drug with time
    - * random effects: intercept and linear time
  - ▷ time to death or transplantation: relative risk model
    - * competing risks: transplantation and death
    - * baseline covariates: drug *different* per competing risk
    - * time-varying: current value log ser Bilir *different* per competing risk

4.3 Competing Risks (cont'd)

	Value	Std.Dev.	2.5%	97.5%
D-penicil	−0.439	0.522	−1.472	0.555
D-penicil:dead	0.528	0.529	−0.490	1.596
value(logSB)	1.266	0.180	0.941	1.615
value(logSB):dead	−0.014	0.183	−0.372	0.305

4.3 Competing Risks (cont'd)

R> Function `jm()` can fit joint models with competing risks

- ▷ First, the survival data have to be prepared in the competing risks long format using function `crisk_setup()`, e.g.,

```
pbc2.id[pbc2.id$id %in% c(1,2,5), c("id", "years", "status")]
```

```
  id     years       status
1  1  1.095170         dead
2  2 14.152338        alive
5  5  4.120578 transplanted
```

4.3 Competing Risks (cont'd)

```
pbc2.idCR <- crisk_setup(pbc2.id, statusVar = "status",
    censLevel = "alive", nameStrata = "CR")
```

```
pbc2.idCR[pbc2.idCR$id %in% c(1,2,5),
    c("id", "years", "status", "CR", "status2")]
```

```
    id     years       status           CR status2
1    1  1.095170         dead         dead       1
1.1  1  1.095170         dead transplanted       0
2    2 14.152338        alive         dead       0
2.1  2 14.152338        alive transplanted       0
5    5  4.120578 transplanted         dead       0
5.1  5  4.120578 transplanted transplanted       1
```

4.3 Competing Risks (cont'd)

R> To fit the joint model, we first fit the linear mixed and relative risk models as before

▷ for the latter we use the data in the competing risks long and put the event-type variable as strata

```
lmeFit_CR <- lme(log(serBilir) ~ drug * year, data = pbc2,
                 random = ~ year | id)

CoxFit_CR <- coxph(Surv(years, status2) ~ drug * strata(CR),
                   data = pbc2.idCR)
```

4.3 Competing Risks (cont'd)

R> Then the joint model is fitted with the code

```
jm(CoxFit_CR, lmeFit_CR, time_var = "year",
   functional_forms = ~ value(log(serBilir)):CR)
```

For more info see
https://drizopoulos.github.io/JMbayes2/
→ Articles → Competing Risks

4.3 Competing Risks (cont'd)

R> Function `jm()` can also fit joint models with multi-state processes

- ▷ this requires an analogous construction of a long dataset for multi-state models, and
- ▷ fitting a stratified Cox model

For more info see
https://drizopoulos.github.io/JMbayes2/
→ Articles → Multi-State Processes

Part V

Dynamic Predictions

5.1 Survival Probabilities

- Nowadays there is great interest for prognostic models and their application to personalized medicine

- Examples are numerous
  - ▷ cancer research, cardiovascular diseases, HIV research, ...

Physicians are interested in accurate prognostic tools to facilitate medical decision-making

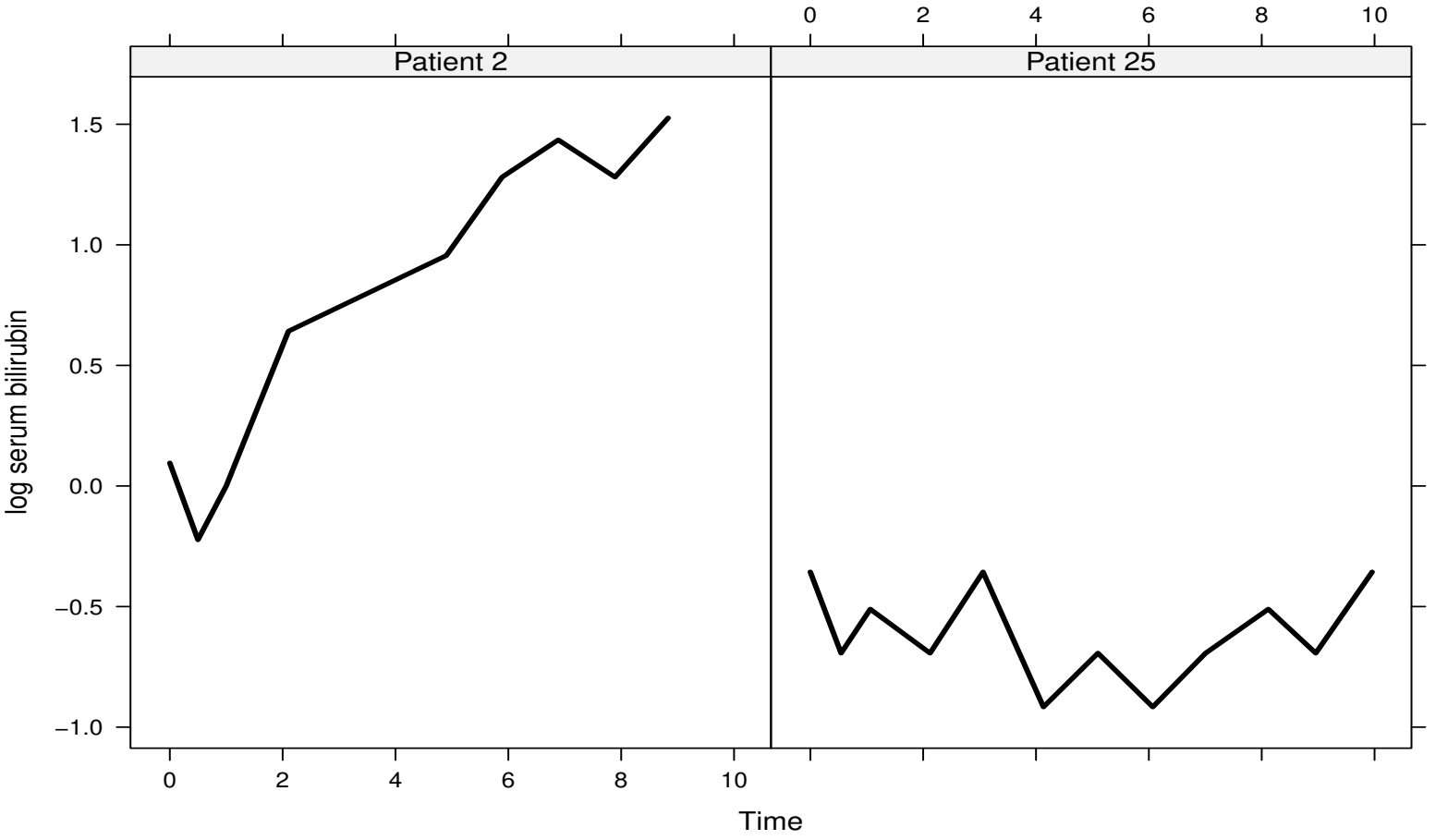
5.1 Survival Probabilities (cont'd)

- We want to obtain survival probabilities for a new patient j with longitudinal measurements up to time t

- Example: Patients 2 and 25 from the PBC dataset have 9 and 12 serum bilirubin measurements, respectively
  - ▷ **Dynamic Prediction** survival probabilities are dynamically updated as additional longitudinal information is recorded

- We need to account for the endogenous nature of the covariate
  - ▷ providing measurements up to time point $t \Rightarrow$ the patient was still alive at time t

5.1 Survival Probabilities (cont'd)



5.1 Survival Probabilities (cont'd)

- For a new subject j, we have available measurements up to t

$$\mathcal{Y}_j(t) = \{y_j(s), 0 \leq s \leq t\}$$

and we are interested in

$$\pi_j(u \mid t) = \mathsf{Pr}\{T_j^* \geq u \mid T_j^* > t, \mathcal{Y}_j(t), \mathcal{D}_n\},$$

where

- ▷ where $u > t$
- ▷ $\mathcal{D}_n$ denotes the training sample

5.1 Survival Probabilities (cont'd)

- We assume that the joint model has been fitted to the data at hand
- Based on the fitted model, we can estimate the conditional survival probabilities
  (Rizopoulos, 2011, Biometrics)

5.1 Survival Probabilities (cont'd)

- It is convenient to proceed using a Bayesian formulation of the problem $\Rightarrow$ $\pi_j(u \mid t)$ can be written as

$$\Pr\{T_j^* \geq u \mid T_j^* > t, \mathcal{Y}_j(t), \mathcal{D}_n\} = \int \Pr\{T_j^* \geq u \mid T_j^* > t, \mathcal{Y}_j(t), \theta\}\, p(\theta \mid \mathcal{D}_n)\, d\theta$$

- The first part of the integrand takes the form

$$\Pr\{T_j^* \geq u \mid T_j^* > t, \mathcal{Y}_j(t), \theta\} =$$

$$= \int \frac{S_j\{u \mid \mathcal{M}_j(u, b_j, \theta), \theta\}}{S_j\{t \mid \mathcal{M}_j(t, b_j, \theta), \theta\}}\, p(b_j \mid T_j^* > t, \mathcal{Y}_j(t), \theta)\, db_j$$

5.1 Survival Probabilities (cont'd)

- A Monte Carlo estimate of $\pi_j(u \mid t)$ can be obtained using the following simulation scheme:

Step 1. draw $\theta^{(\ell)} \sim [\theta \mid \mathcal{D}_n]$

Step 2. draw $b_j^{(\ell)} \sim [b_j \mid T_j^* > t, \mathcal{Y}_j(t), \theta^{(\ell)}]$

Step 3. compute $\pi_j^{(\ell)}(u \mid t) = S_j\{u \mid \mathcal{M}_j(u, b_j^{(\ell)}, \theta^{(\ell)}), \theta^{(\ell)}\} \Big/ S_j\{t \mid \mathcal{M}_j(t, b_j^{(\ell)}, \theta^{(\ell)}), \theta^{(\ell)}\}$

- Repeat Steps 1–3, $\ell = 1, \ldots, L$ times, where L denotes the number of Monte Carlo samples

5.1 Survival Probabilities (cont'd)

- Example: Dynamic predictions of survival probabilities for Patients 2 & 25 from the PBC dataset: We fit the joint model

- Longitudinal submodel
  - ▷ fixed effects: intercept & natural cubic splines of time with 3 d.f., sex, and interaction of the time effect with sex
  - ▷ random effects: intercept, natural cubic splines of time with 3 d.f.

- Survival submodel
  - ▷ sex effect + *underlying* serum bilirubin level

5.1 Survival Probabilities (cont'd)

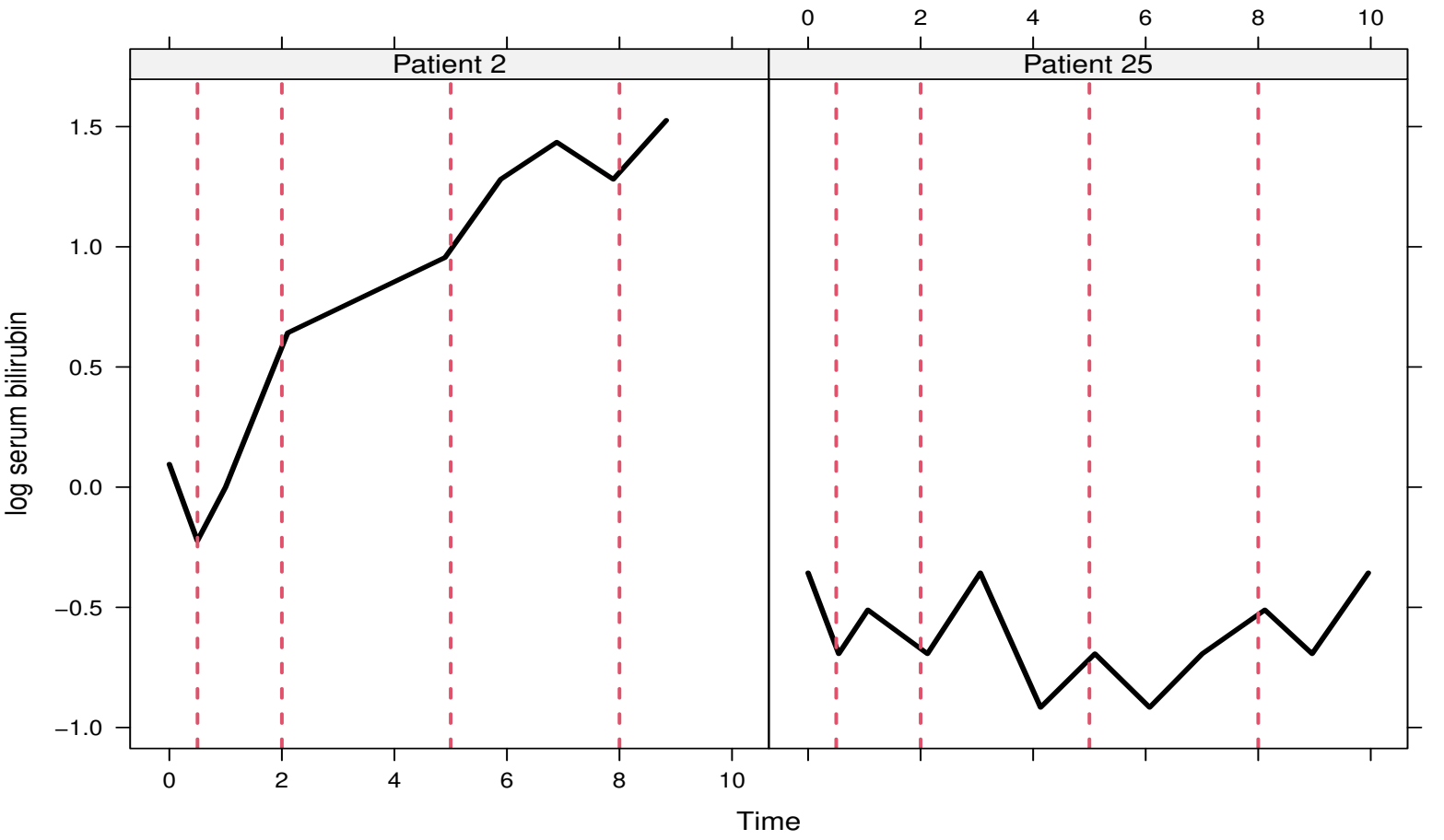
- Based on the fitted joint model we estimate $\pi_j(u \mid t)$ for Patients 2 and 25

- We use 500 Monte Carlo samples, and we took as estimate

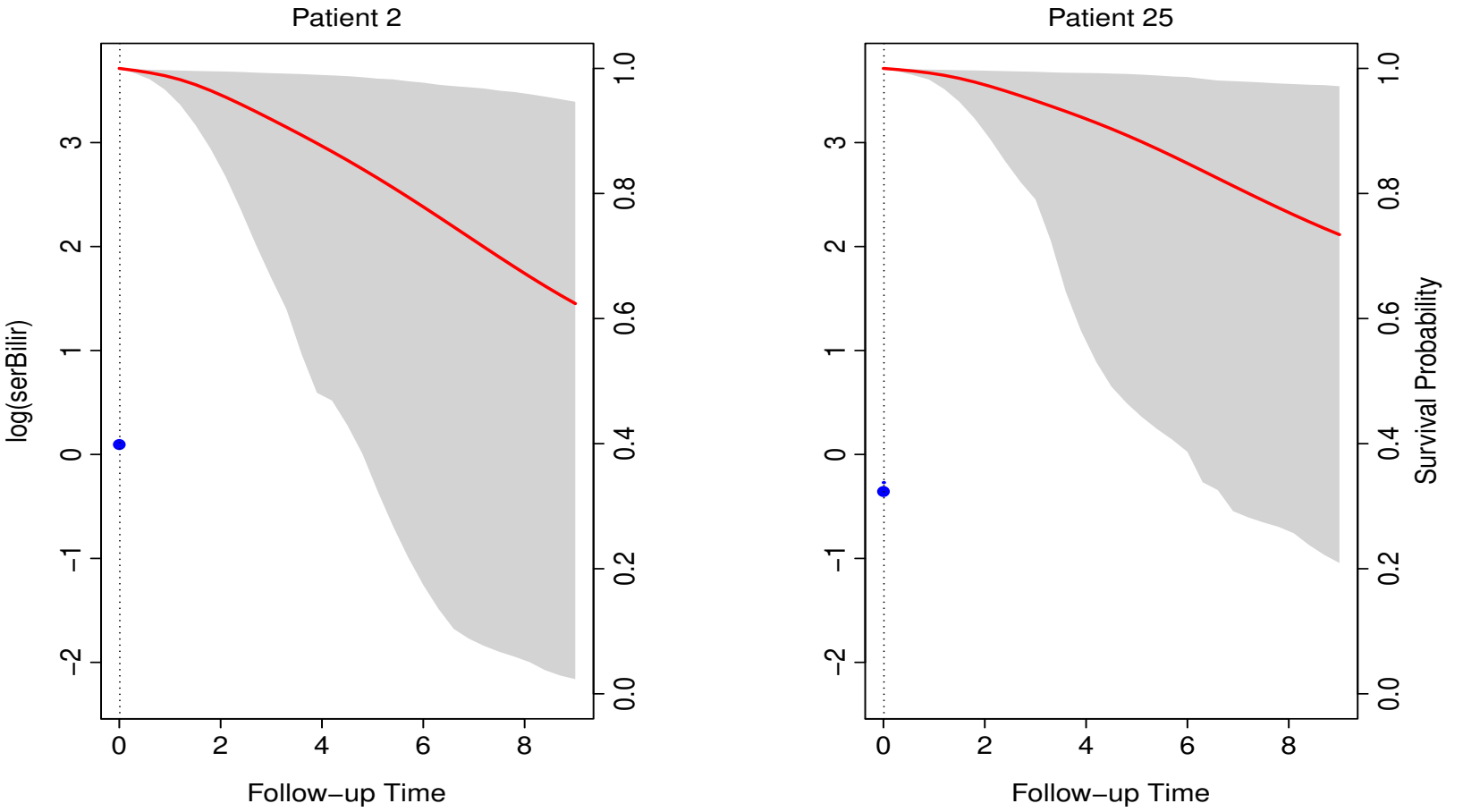
$$\hat{\pi}_j(u \mid t) = \text{mean}\{\pi_j^{(\ell)}(u \mid t), \ell = 1, \ldots, L\}$$

  and calculated a corresponding 95% pointwise CIs

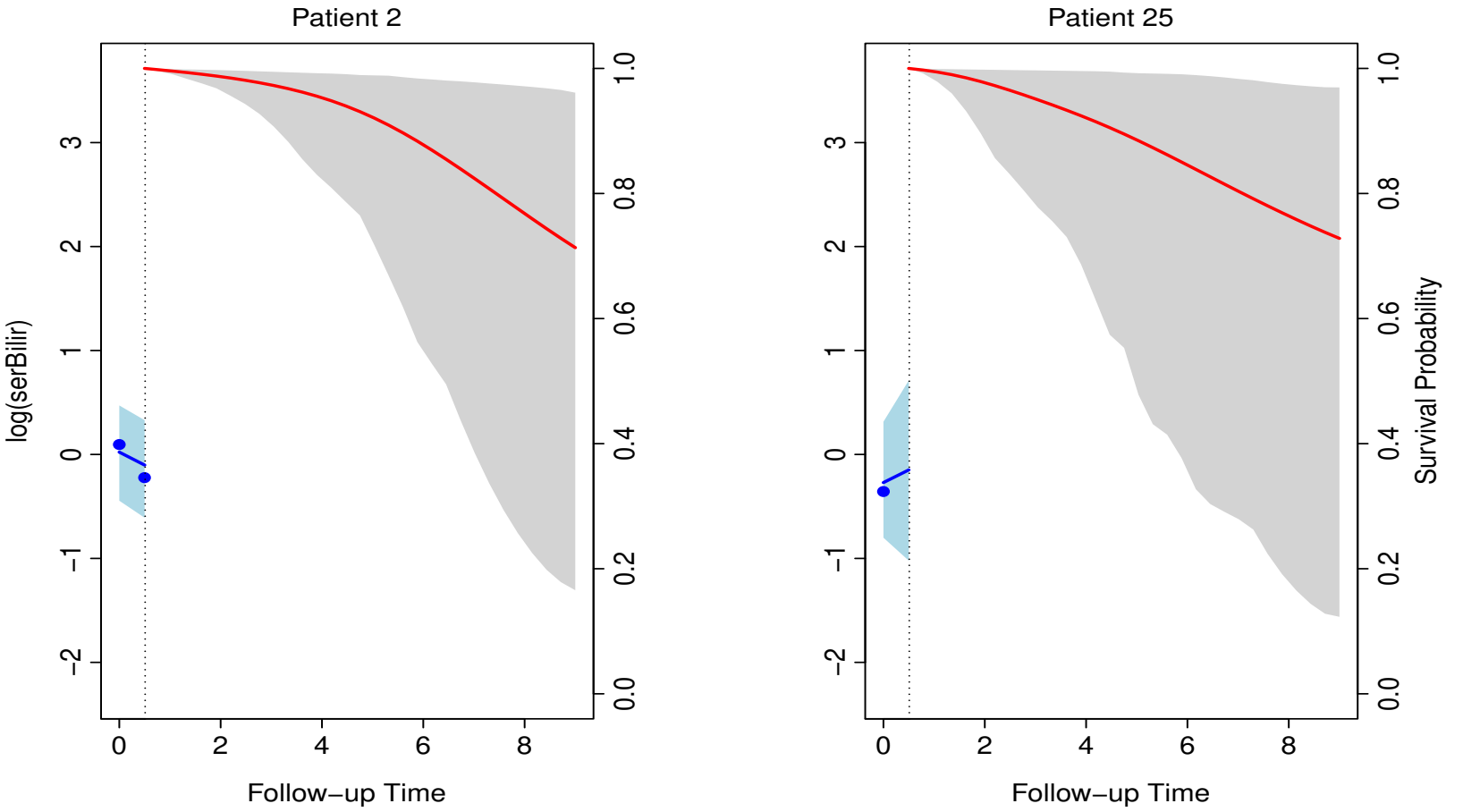
5.1 Survival Probabilities (cont'd)



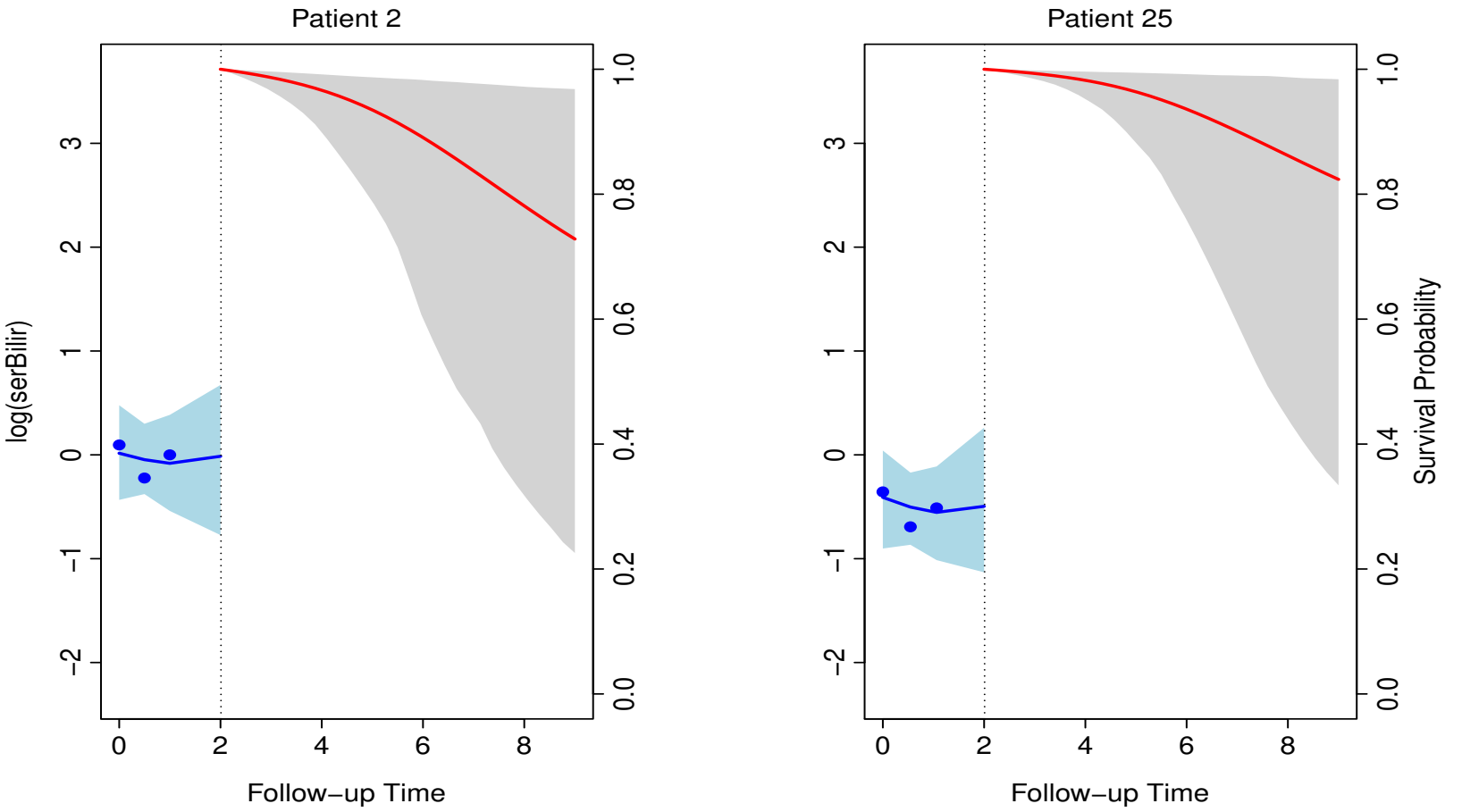
5.1 Survival Probabilities (cont'd)



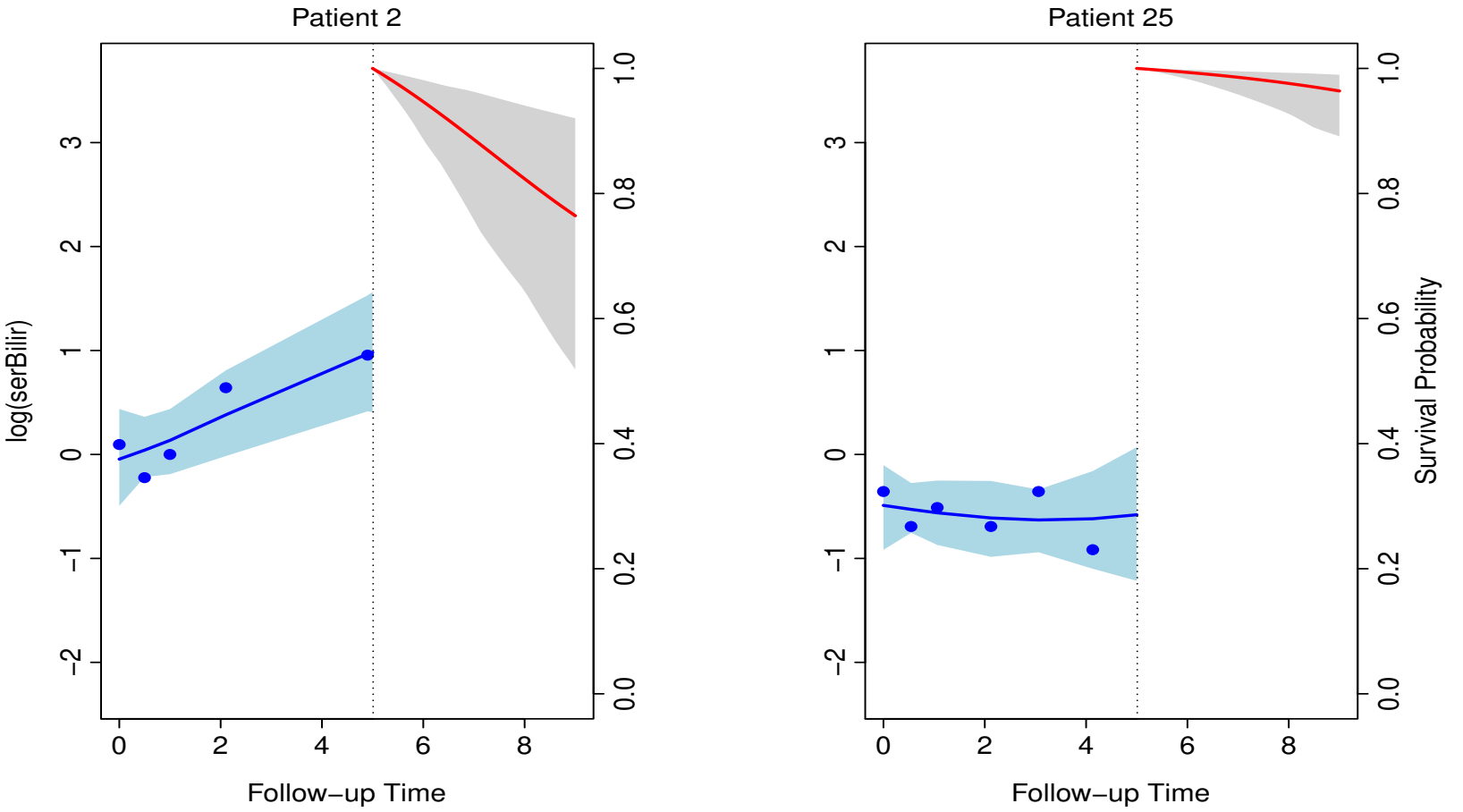
5.1 Survival Probabilities (cont'd)



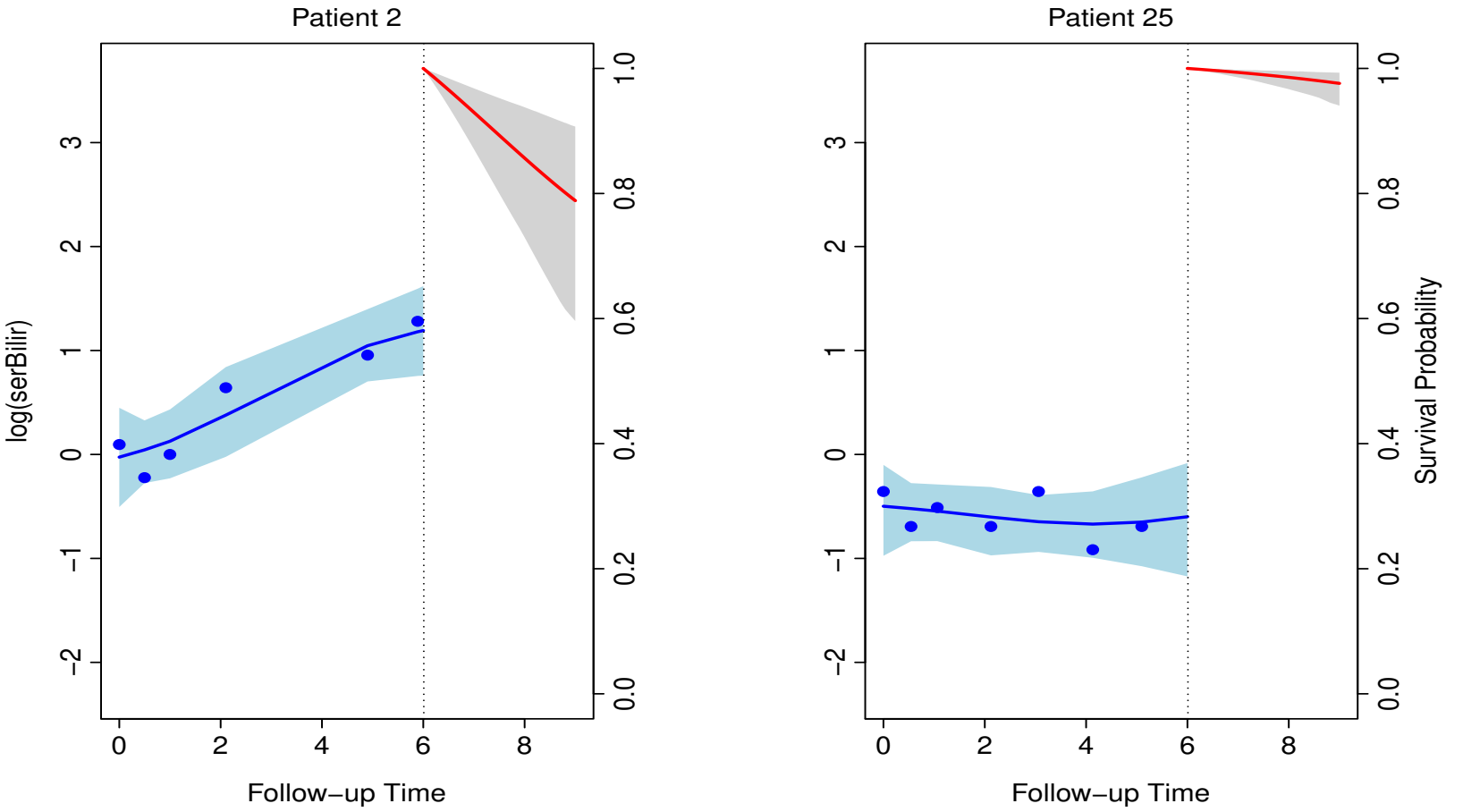
5.1 Survival Probabilities (cont'd)



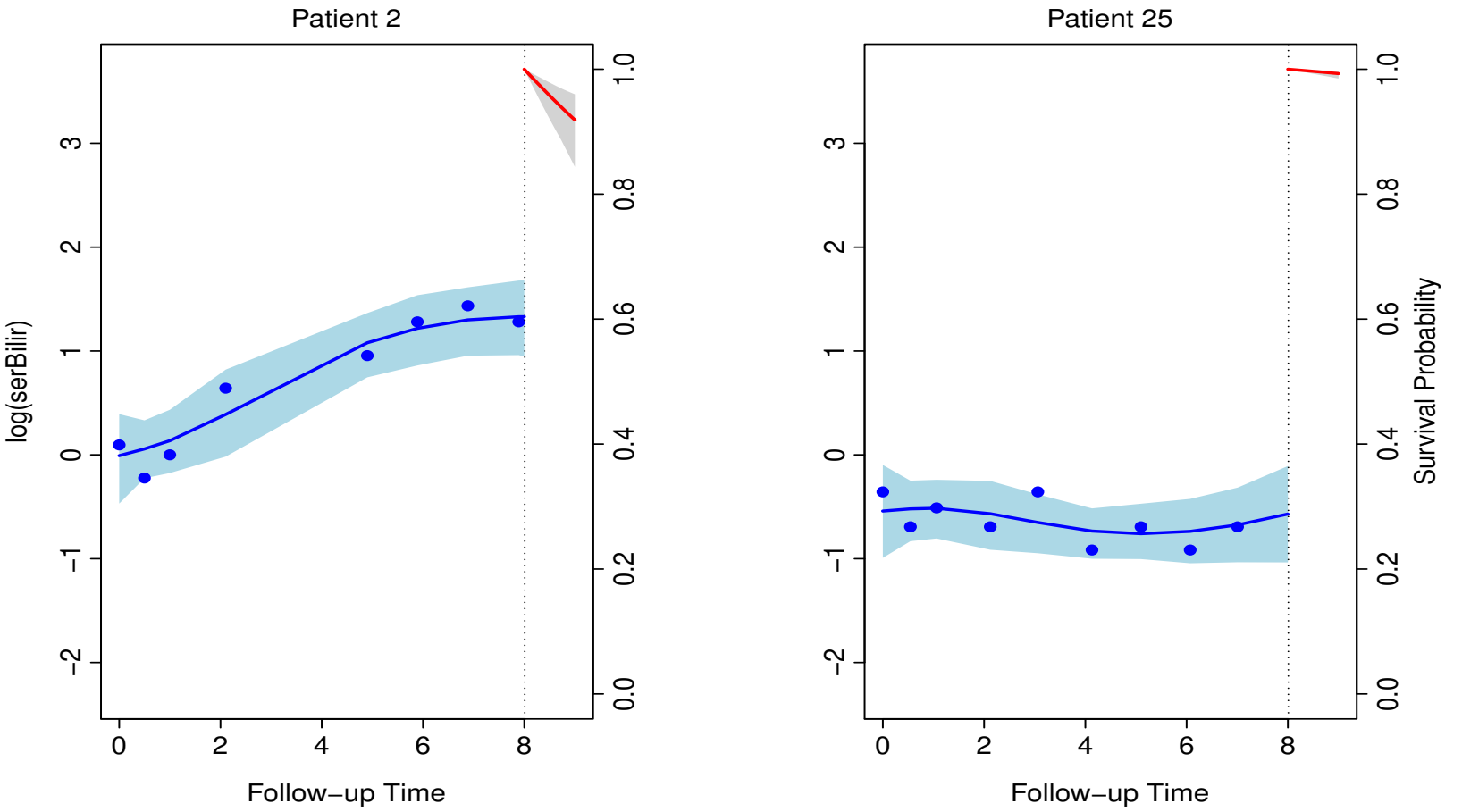
5.1 Survival Probabilities (cont'd)



5.1 Survival Probabilities (cont'd)



5.1 Survival Probabilities (cont'd)



5.1 Survival Probabilities (cont'd)

R> Individualized predictions of survival probabilities are computed by function `predict()` – for example, for Patient 2 from the PBC dataset we have

```
sfit <- predict(jointFit, newdata = pbc2[pbc2$id == "2", ],
                process = "event", return_newdata = TRUE)

sfit

plot(sfit)
```

5.2 Functional Forms

- All previous predictions were based on the standard joint model

$$\begin{cases} h_i(t \mid \mathcal{M}_i(t)) &= h_0(t)\exp\{\gamma^\top w_i + \alpha m_i(t)\}, \\ \\ y_i(t) &= m_i(t) + \varepsilon_i(t) \\ &= x_i^\top(t)\beta + z_i^\top(t)b_i + \varepsilon_i(t), \end{cases}$$

where $\mathcal{M}_i(t) = \{m_i(s), 0 \leq s < t\}$

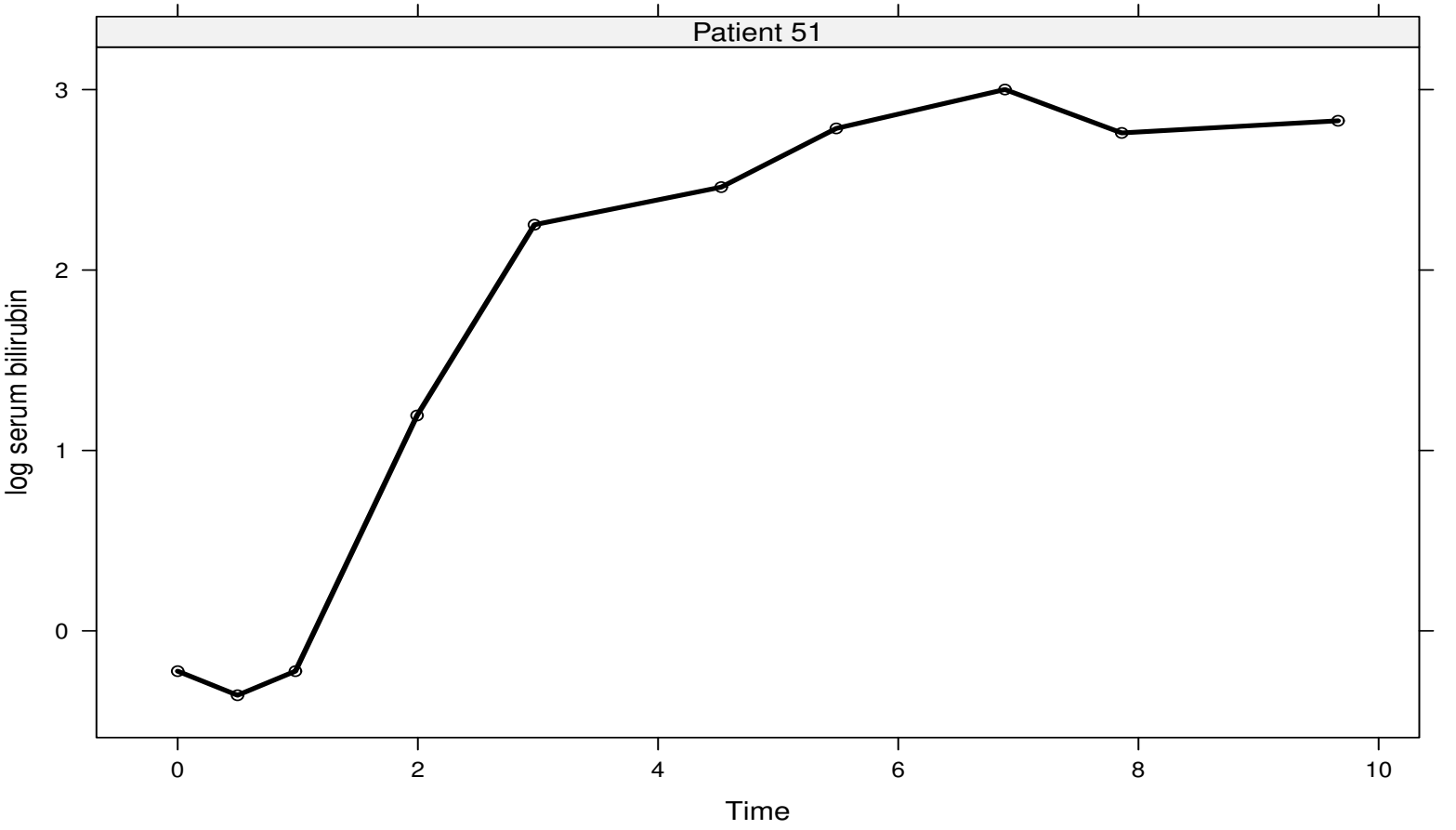
5.2 Functional Forms (cont'd)

- We have seen earlier that there are several alternative functional forms (see Section 5.1)

- Relevant questions:
  - ▷ Does the assumed functional form affect predictions?
  - ▷ Which functional form is the most optimal?

- Example: We compare predictions for the longitudinal and survival outcomes under different parameterizations for Patient 51 from the PBC study

5.2 Functional Forms (cont'd)



5.2 Functional Forms (cont'd)

- Predictions based on five joint models for the PBC dataset
  - ▷ the same longitudinal submodel as before, and
  - ▷ relative risk submodels:

$$h_i(t) = h_0(t)\exp\{\gamma \texttt{D-pnc}_i + \alpha_1 m_i(t)\},$$

$$h_i(t) = h_0(t)\exp\{\gamma \texttt{D-pnc}_i + \alpha_2 m_i'(t)\},$$

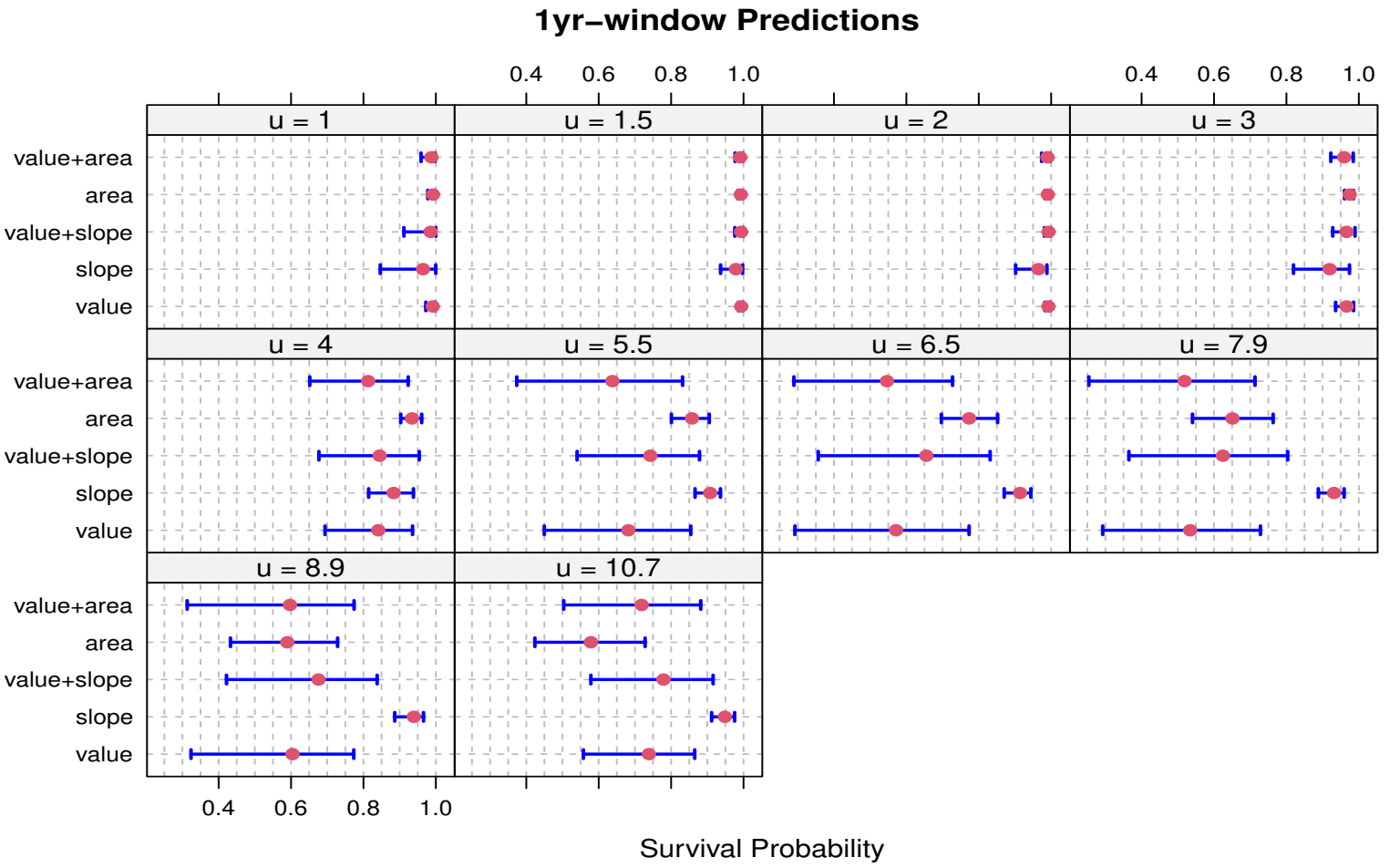
$$h_i(t) = h_0(t)\exp\{\gamma \texttt{D-pnc}_i + \alpha_1 m_i(t) + \alpha_2 m_i'(t)\}$$

5.2 Functional Forms (cont'd)

$$h_i(t) = h_0(t) \exp\left\{ \gamma \texttt{D-pnc}_i + \alpha_3 \frac{\int_0^t m_i(s)ds}{t} \right\},$$

$$h_i(t) = h_0(t) \exp\left\{ \gamma \texttt{D-pnc}_i + \alpha_1 m_i(t) + \alpha_3 \frac{\int_0^t m_i(s)ds}{t} \right\},$$

5.2 Functional Forms (cont'd)



5.2 Functional Forms (cont'd)

The chosen functional form can influence the derived predictions

5.2 Functional Forms (cont'd)

- We compare the models using the information criteria

	DIC	WAIC	LPML
area	4276.422	4568.705	−2713.276
value	4261.051	4574.446	−2763.496
value + area	4268.458	4604.367	−2639.927
value + slope	4274.964	4644.614	−2666.901
slope	4519.831	4891.027	−2896.365

- We continue with the area functional form

5.3 Discrimination

- We have seen how to calculate predictions of conditional survival probabilities
  - ▷ however, to use these predictions in practice we need to evaluate their accuracy

- Predictive accuracy measures
  - ▷ Discrimination: sensitivity, specificity, ROC and AUC
  - ▷ Calibration: comparison between predicted and observed probabilities
  - ▷ Overall: combination of discrimination and calibration

5.3 Discrimination (cont'd)

- To assess the discriminative power of the model, we assume the following setting
  - ▷ using the available longitudinal data up to time t,
  - ▷ we are interested in events occurring in a medically-relevant interval $(t, t + \Delta t]$

- Based on the fitted joint model and for a particular threshold value $c \in [0, 1]$, we can term subject j a **case** if

$$\pi_j(t + \Delta t \mid t) \leq c$$

5.3 Discrimination (cont'd)

- Following, we can define sensitivity

$$\mathsf{SN}_t^{\Delta t}(c) = \Pr\big\{\pi_j(t+\Delta t \mid t) \leq c \mid T_j^* \in (t, t+\Delta t]\big\},$$

specificity

$$\mathsf{SP}_t^{\Delta t}(c) = \Pr\big\{\pi_j(t+\Delta t \mid t) > c \mid T_j^* > t+\Delta t\big\},$$

and the corresponding AUC

$$\begin{aligned}&\mathsf{AUC}_t^{\Delta t}\\ &= \Pr\big[\pi_i(t+\Delta t \mid t) < \pi_j(t+\Delta t \mid t) \mid \{T_i^* \in (t, t+\Delta t]\} \cap \{T_j^* > t+\Delta t\}\big]\end{aligned}$$

5.3 Discrimination (cont'd)

- To estimate the sensitivity, specificity and the AUC, *we need to account for censoring*

- Two main approaches
  - ▷ model-based weights
  - ▷ inverse probability of censoring weighting (IPCW)
    (using Kaplan-Meier or other non-parametric estimators)

5.3 Discrimination (cont'd)

- IPCW
  - ▷ *Advantage:* it provides unbiased estimates even when the model is misspecified
  - ▷ *Disadvantage:* it requires that the model for the weights is correct
    - * in settings where joint models are used, challenging because censoring may depend on the longitudinal outcomes in a complex manner

5.3 Discrimination (cont'd)

- Model-based Weights
  - ▷ *Advantage:* it allows censoring to depend on the longitudinal history (in any possible manner)
  - ▷ *Disadvantage:* it requires that the model is well calibrated

5.3 Discrimination (cont'd)

Because censoring often depends on the longitudinal history,
we opt for model-based weights

5.3 Discrimination (cont'd)

- For the $\mathcal{R}(t)$ subjects at risk at time t (i.e., $T_i > t$), sensitivity is estimated as

$$\widehat{\mathsf{SN}}_t^{\Delta t}(c) = \frac{\sum_{i:T_i \geq t} I\{\hat{\pi}_i(t + \Delta t \mid t) \leq c\} \times \Omega_i}{\sum_{i:T_i \geq t} \Omega_i},$$

where

$$\Omega_i = \begin{cases} 1, & \text{if } T_i \leq t + \Delta t \text{ and } \delta_i = 1 \\ 1 - \hat{\pi}_i(t + \Delta t \mid T_i), & \text{if } T_i \leq t + \Delta t \text{ and } \delta_i = 0 \end{cases}$$

5.3 Discrimination (cont'd)

- And specificity as

$$\widehat{\mathsf{SP}}_t^{\Delta t}(c) = \frac{\sum_{i:T_i \geq t} I\{\hat{\pi}_i(t + \Delta t \mid t) > c\} \times \Phi_i}{\sum_{i:T_i \geq t} \Phi_i},$$

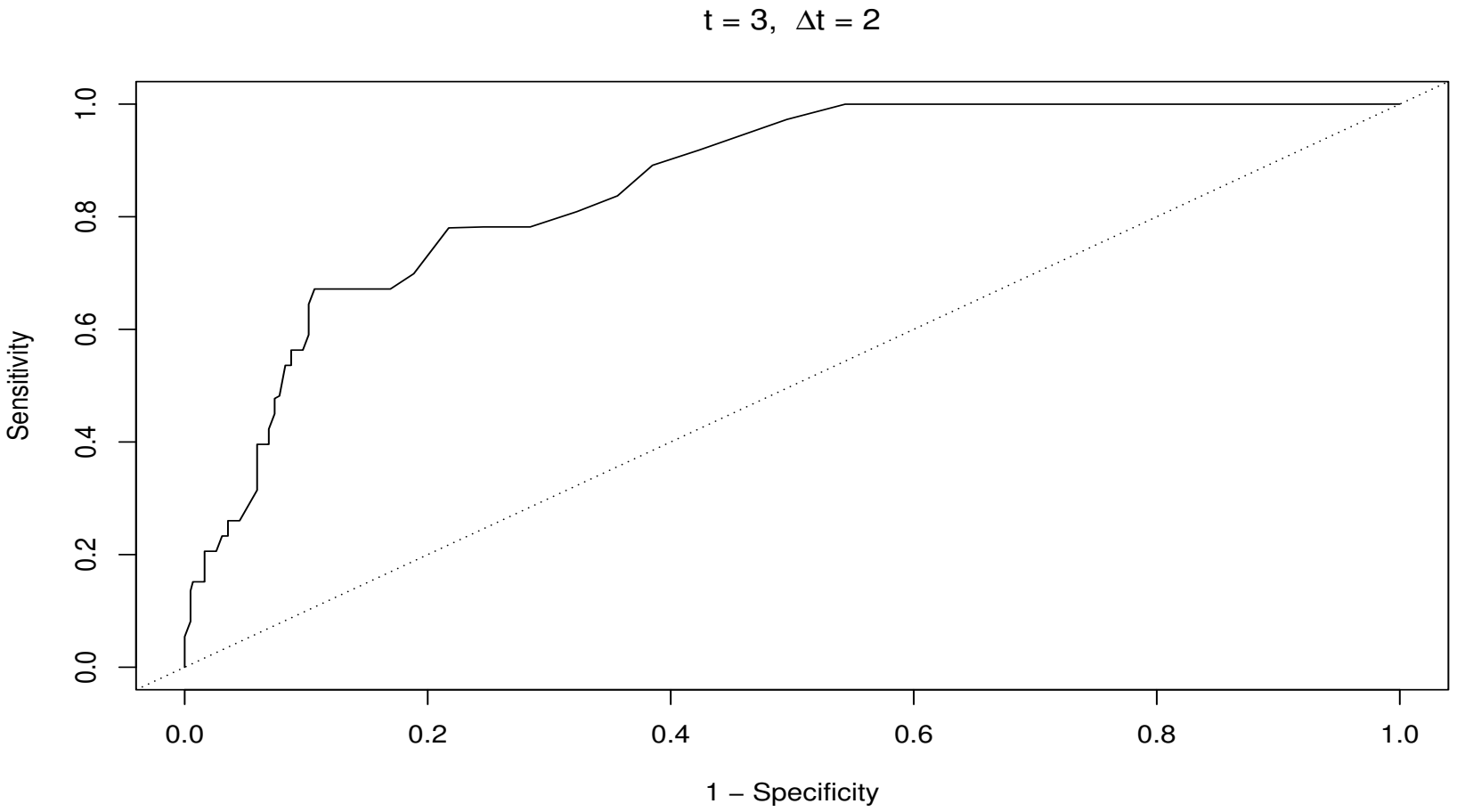
where

$$\Phi_i = \begin{cases} 1, & \text{if } T_i > t + \Delta t \\ \hat{\pi}_i(t + \Delta t \mid T_i), & \text{if } T_i \leq t + \Delta t \text{ and } \delta_i = 0 \end{cases}$$

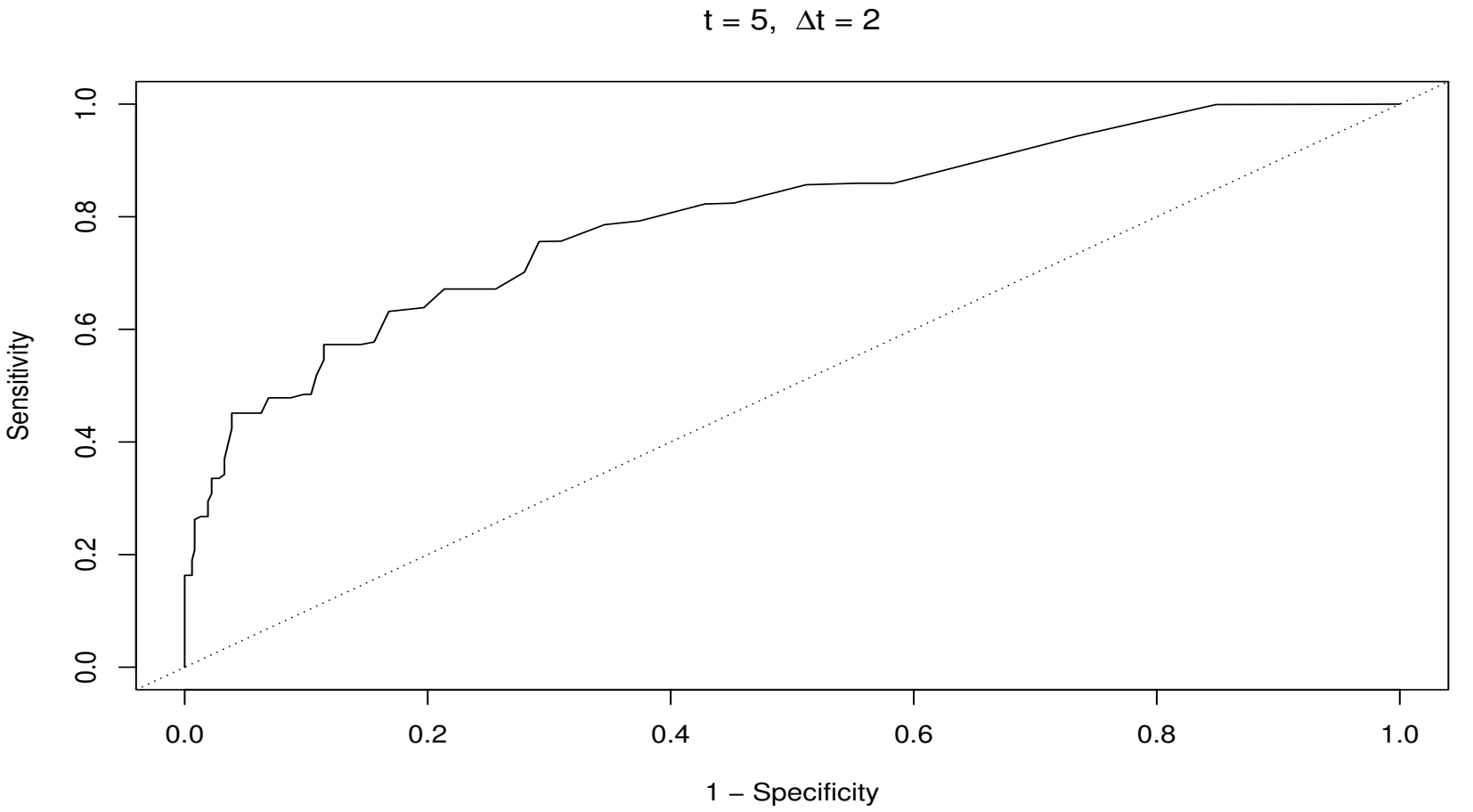
5.3 Discrimination (cont'd)

- Example: For the joint model fitted to the PBC dataset we have seen earlier
  - ▷ we estimate dynamic sensitivity, specificity and the ROC curve
  - ▷ at follow-up times $t = 3, 5,$ and 7
  - ▷ for $\Delta t = 2$

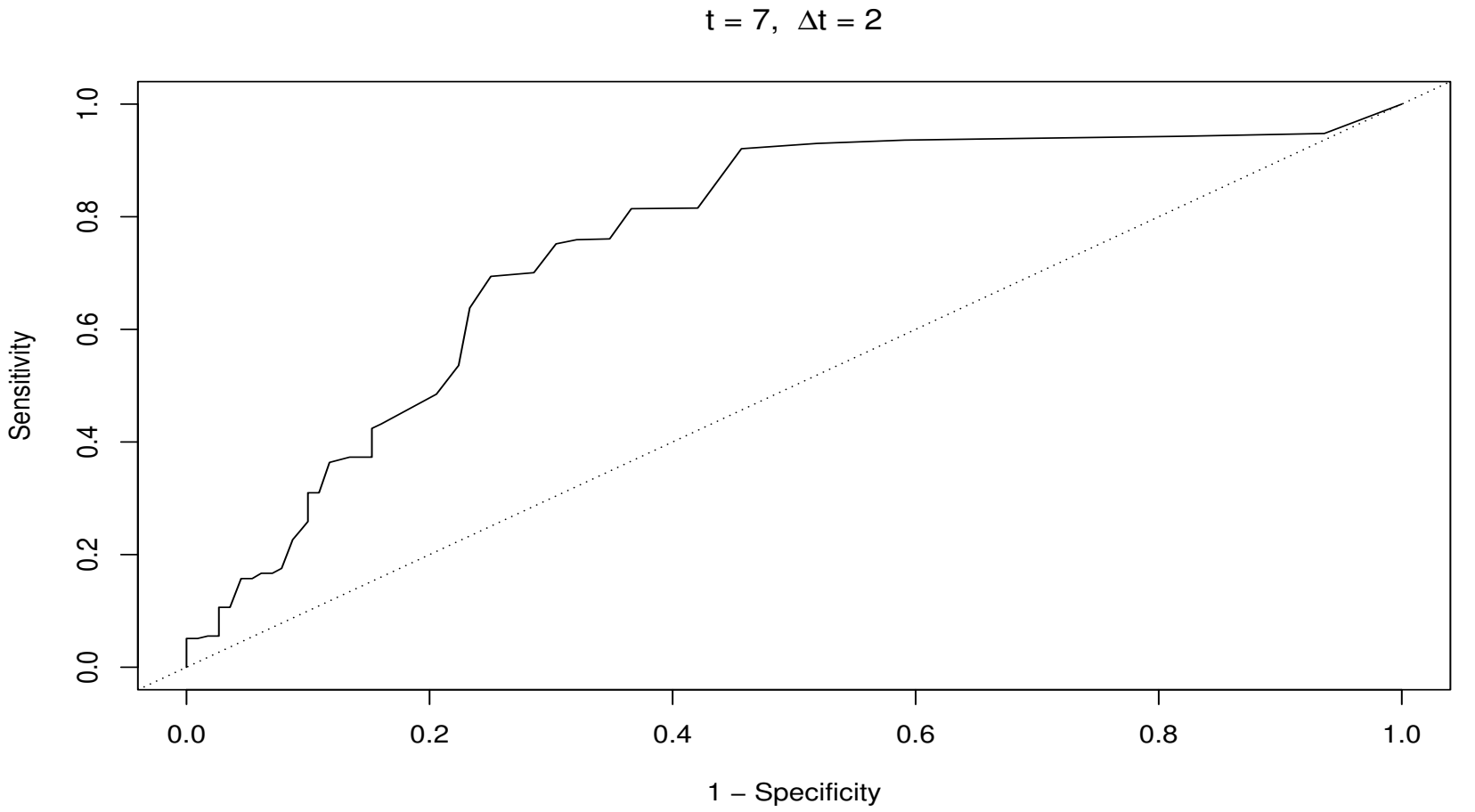
5.3 Discrimination (cont'd)



5.3 Discrimination (cont'd)



5.3 Discrimination (cont'd)



5.3 Discrimination (cont'd)

- The corresponding AUCs are

Time	AUC
t = 3	0.86
t = 5	0.80
t = 7	0.76

5.3 Discrimination (cont'd)

R> For a fitted joint model, we calculate the ROC curve and the corresponding AUC with the syntax

```
# model-based weights
roc <- tvROC(jointFit, newdata = pbc2, Tstart = 5, Dt = 2)

# Kaplan-Meier IPCW
roc <- tvROC(jointFit, newdata = pbc2, Tstart = 5, Dt = 2,
             type_weights = "IPCW")

roc

plot(roc)

tvAUC(roc)
```

5.4 Prediction Error

- We have covered *discrimination*
  - ▷ *calibration* assessed via calibration plots

- In standard survival analysis there are measures that combine the two concepts into one metric
  - ▷ the most-well know measure that achieves that is the *Brier score*

5.4 Prediction Error (cont'd)

- In the joint modeling framework, we need to take into account the dynamic nature of the longitudinal marker

- The expected quadratic error of prediction (Brier score) has the form

$$\text{PE}(t + \Delta t \mid t) = E\big[\{N_i(t + \Delta t) - \pi_i(t + \Delta t \mid t)\}^2\big]$$

  where

  ▷ $N_i(t) = I(T_i^* > t)$ is the "true" event status at time t

5.4 Prediction Error (cont'd)

- An estimator for $\text{PE}(t + \Delta t \mid t)$ that *accounts for censoring*

$$\widehat{\text{PE}}(t + \Delta t \mid t) = \{\mathcal{R}(t)\}^{-1} \sum_{i:T_i \geq t} I(t + \Delta t > u)\{1 - \hat{\pi}_i(t + \Delta t \mid t)\}^2$$

$$+ \delta_i I(T_i < t + \Delta t)\{0 - \hat{\pi}_i(t + \Delta t \mid t)\}^2$$

$$+ (1 - \delta_i) I(T_i < t + \Delta t) \Big[\hat{\pi}_i(t + \Delta t \mid T_i)\{1 - \hat{\pi}_i(t + \Delta t \mid t)\}^2$$

$$+ \{1 - \hat{\pi}_i(t + \Delta t \mid T_i)\}\{0 - \hat{\pi}_i(t + \Delta t \mid t)\}^2 \Big]$$

5.4 Prediction Error (cont'd)

where

- ▷ $\mathcal{R}(t)$ denotes the number of subjects at risk at t
- ▷ **red part**: subjects still event-free at $t + \Delta t$
- ▷ **blue part**: subjects who had the event before $t + \Delta t$
- ▷ **green part**: subject censored before $t + \Delta t$

- The weights used to account for censoring are model-based
  - ▷ censoring is allowed to depend on the longitudinal history in any possible manner
  - ▷ the model needs to be well specified

5.4 Prediction Error (cont'd)

- Example: For the joint model fitted to the PBC dataset we have seen earlier
  - ▷ we estimate the dynamic Brier score
  - ▷ at follow-up times $t = 3, 5,$ and 7
  - ▷ for $\Delta t = 2$

5.4 Prediction Error (cont'd)

- The estimated Brier scores are

Time	Brier Score
$t = 3$	0.10
$t = 5$	0.11
$t = 7$	0.12

5.4 Prediction Error (cont'd)

R> For a fitted joint model, we calculate the time-varying Brier score with the syntax

```
# model-based weights
predErr <- tvBrier(jointFit, newdata = pbc2, Tstart = 5, Dt = 2)

# Kaplan-Meier IPCW
predErr <- tvBrier(jointFit, newdata = pbc2, Tstart = 5, Dt = 2,
                   type_weights = "IPCW")

predErr
```

5.5 Cumulative Risk Probabilities

- We have presented dynamic predictions for a single longitudinal outcome and one event

- Extensions:
  - ▷ multiple longitudinal outcomes
  - ▷ competing risks

How can we account for the above?

5.5 Cumulative Risk Probabilities (cont'd)

- Suppose that for a new subject j, we have measurements from I multiple longitudinal outcomes up to time point t. The data for the ith marker

$$\mathcal{Y}_{ji}(t) = \{y_{ji}(t_{jik}); 0 \leq t_{jik} \leq t, k = 1, 2, \ldots, n_{ji}\}$$

with $\mathcal{Y}_j(t) = \{\mathcal{Y}_{j1}(t), \ldots, \mathcal{Y}_{jI}(t)\}$

- In the competing risk setting we are interested in predicting the cause-specific **cumulative incidence probabilities**

$$\Pr\left\{t < T_j^* \leq t + \Delta t, \delta_j^* = k \mid T_j^* > t, \mathcal{Y}_j(t), \mathcal{D}_n\right\}$$

for $k = 1, 2, \ldots, K$.

5.5 Cumulative Risk Probabilities (cont'd)

- Similarly to the single event case, to account for variability in the model parameters

$$\begin{aligned}&\Pr\left\{t < T_j^* \leq t + \Delta t, \delta_j^* = k \mid T_j^* > t, \mathcal{Y}_j(t), \mathcal{D}_n\right\} \\ &= \int \Pr\left\{t < T_j^* \leq t + \Delta t, \delta_j^* = k \mid T_j^* > t, \mathcal{Y}_j(t); \theta\right\} p(\theta \mid \mathcal{D}_n) d\theta\end{aligned}$$

- The first part of the integrand is the cumulative incidence (risk) for the kth event given that the individual is event-free at time t

$$\begin{aligned}&F_{kj}(t + \Delta t \mid t) = \Pr\left\{t < T_j^* \leq t + \Delta t, \delta_j^* = k \mid T_j^* > t, \mathcal{Y}_j(t); \theta\right\} = \\ &\int \Pr\left\{t < T_j^* \leq t + \Delta t, \delta_j^* = k \mid T_j^* > t, \mathcal{Y}_j(t), b_j; \theta\right\} p\{b_j \mid T_j^* > t, \mathcal{Y}_j(t); \theta\} d b_j\end{aligned}$$

  where b_j represents the random effects for all longitudinal markers.

5.5 Cumulative Risk Probabilities (cont'd)

- A Monte Carlo estimate of $F_{kj}(t + \Delta t \mid t)$ can be obtained using the following simulation scheme:

Step 1. draw $\theta^{(l)} \sim [\theta \mid \mathcal{D}_n]$

Step 2. draw $b_j^{(l)} \sim [b_j \mid T_j^* > t, \mathcal{Y}_j(t); \theta]$

Step 3. compute
$F_{kj}^{(l)}(t + \Delta t \mid t) = \Pr\left\{t < T_j^* \leq t + \Delta t, \delta_j^* = k \mid T_j^* > t, \mathcal{Y}_j(t), b_j^{(l)}; \theta^{(l)}\right\}$

- Repeat Steps 1-3, $l = 1, \ldots, L$ times, where L denotes the number of Monte Carlo samples.

5.5 Cumulative Risk Probabilities (cont'd)

- Example: Dynamic predictions of survival probabilities for Patient 2 from the PBC dataset

- Longitudinal submodels
  - ▷ log(ser Bilir)
    - * fixed effects: intercept, drug, linear and squared time, and interactions of linear and square time with drug
    - * random effects: intercept and linear and squared time
  - ▷ prothrombin
    - * fixed effects: intercept, drug, linear time, interaction of time with drug
    - * random effects: intercept and linear and squared time

5.5 Cumulative Risk Probabilities (cont'd)

- Example: Dynamic predictions of survival probabilities for Patient 2 from the PBC dataset

- time to death or transplantation: relative risk model
  - ▷ competing risks: transplantation and death
  - ▷ baseline covariates: drug and age *different* per competing risk
  - ▷ time-varying: current value log(ser Bilir) and prothrombin *different* per competing risk

5.5 Cumulative Risk Probabilities (cont'd)

R> Function jm() can fit joint models with multiple longitudinal outcomes and competing risks, with the survival data prepared in the competing risks long format using function crisk_setup(), e.g.,

```
pbc2.id[pbc2.id$id %in% c(1,2,5), c("id", "years", "status")]
```

```
  id     years       status
1  1  1.095170         dead
2  2 14.152338        alive
5  5  4.120578 transplanted
```

5.5 Cumulative Risk Probabilities (cont'd)

```
pbc2.idCR <- crisk_setup(pbc2.id, statusVar = "status",
    censLevel = "alive", nameStrata = "CR")
```

```
pbc2.idCR[pbc2.idCR$id %in% c(1,2,5),
    c("id", "years", "status", "CR", "status2")]
```

	id	years	status	CR	status2
1	1	1.095170	dead	dead	1
1.1	1	1.095170	dead	transplanted	0
2	2	14.152338	alive	dead	0
2.1	2	14.152338	alive	transplanted	0
5	5	4.120578	transplanted	dead	0
5.1	5	4.120578	transplanted	transplanted	1

5.5 Cumulative Risk Probabilities (cont'd)

- For the competing risk model we use the data in the competing risks long format and put the event-type variable log(CR) as strata

```
# Fit the competing risk cox model
CoxFit_CR <- coxph(Surv(years, status2) ~ (age + drug) * strata(CR),
                   data = pbc2.idCR)
```

5.5 Cumulative Risk Probabilities (cont'd)

R> We can create a `list()` for `functional_forms` for each longitudinal outcome to ensure an interaction with the event-type variable

```
# Functional forms
CR_forms <- list(
  "log(serBilir)" = ~ value(log(serBilir)):CR,
  "prothrombin" = ~ value(prothrombin):CR
)
```

5.5 Cumulative Risk Probabilities (cont'd)

R> We then fit two linear mixed models for log(ser Bilir) and prothrombin

```
fm1 <- lme(log(serBilir) ~ poly(year, 2) * drug, data = pbc2,
           random = ~ poly(year, 2) | id)

fm2 <- lme(prothrombin ~ year * drug, data = pbc2,
           random = ~ year | id)
```

5.5 Cumulative Risk Probabilities (cont'd)

R> The model is fitted using the code

```
# Fit the competing risk joint model
jFit_CR <- jm(CoxFit_CR, list(fm1, fm2), time_var = "year",
              functional_forms = CR_forms,
              n_iter = 25000L, n_burnin = 5000L, n_thin = 5L)
```

5.5 Cumulative Risk Probabilities (cont'd)

R> Individualized predictions of survival probabilities are computed by function `predict()`.

R> In contrast to the case with a single event, two datasets (with longitudinal and event information, respectively) are required in a named `list()`. For example, for Patient 2 from the PBC dataset we have

```
ND_long <- pbc2[pbc2$id == 2, ]
ND_event <- pbc2.idCR[pbc2.idCR$id == 2, ]

ND <- list(newdataL = ND_long, newdataE = ND_event)
```

5.5 Cumulative Risk Probabilities (cont'd)

R> In contrast to the case with a single event, two datasets (with longitudinal and event information, respectively) are required in a named `list()`. For example, for Patient 2 from the PBC dataset we have

R> `plot` can be used to visually the evolution of the multiple longitudinal outcomes and the cumulative risk probabilities of the competing risks.

```
predLong <- predict(jFit_CR, newdata = ND, return_newdata = TRUE,
                    times = seq(censTimes[i], 10, length = 25))

predEvent <- predict(jFit_CR, newdata = ND, return_newdata = TRUE,
                     process = "event",
                     times = seq(censTimes[i], 10, length = 25))
plot(predLong, predEvent, outcomes = 1:2)
```

5.5 Cumulative Risk Probabilities (cont'd)

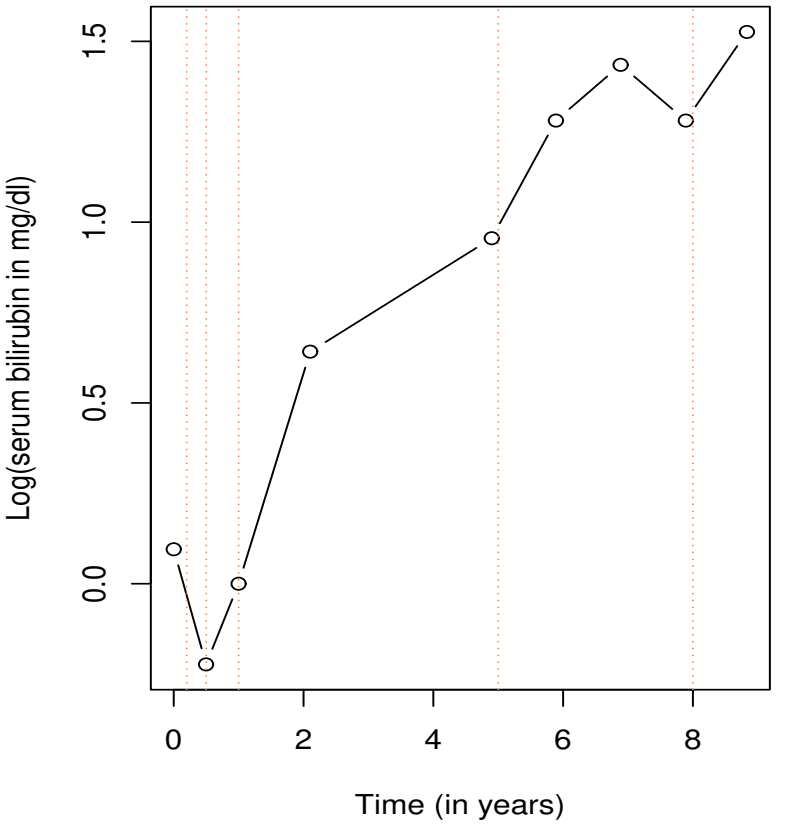
- Based on the fitted joint model we estimate $F_{kj}(t + \Delta t \mid t)$ for Patient 2

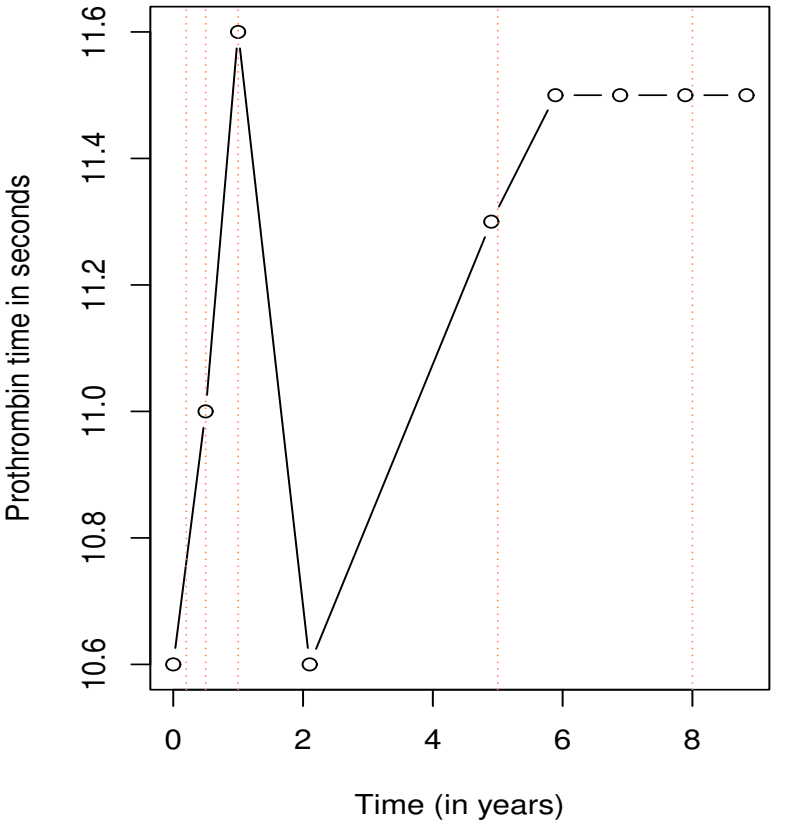
- We use 500 Monte Carlo samples, and we took as estimate

$$\hat{F}_{kj}(t + \Delta t \mid t) = \frac{1}{L} \sum_{l=1}^{L} F_{kj}^{(l)}(t + \Delta t \mid t)$$

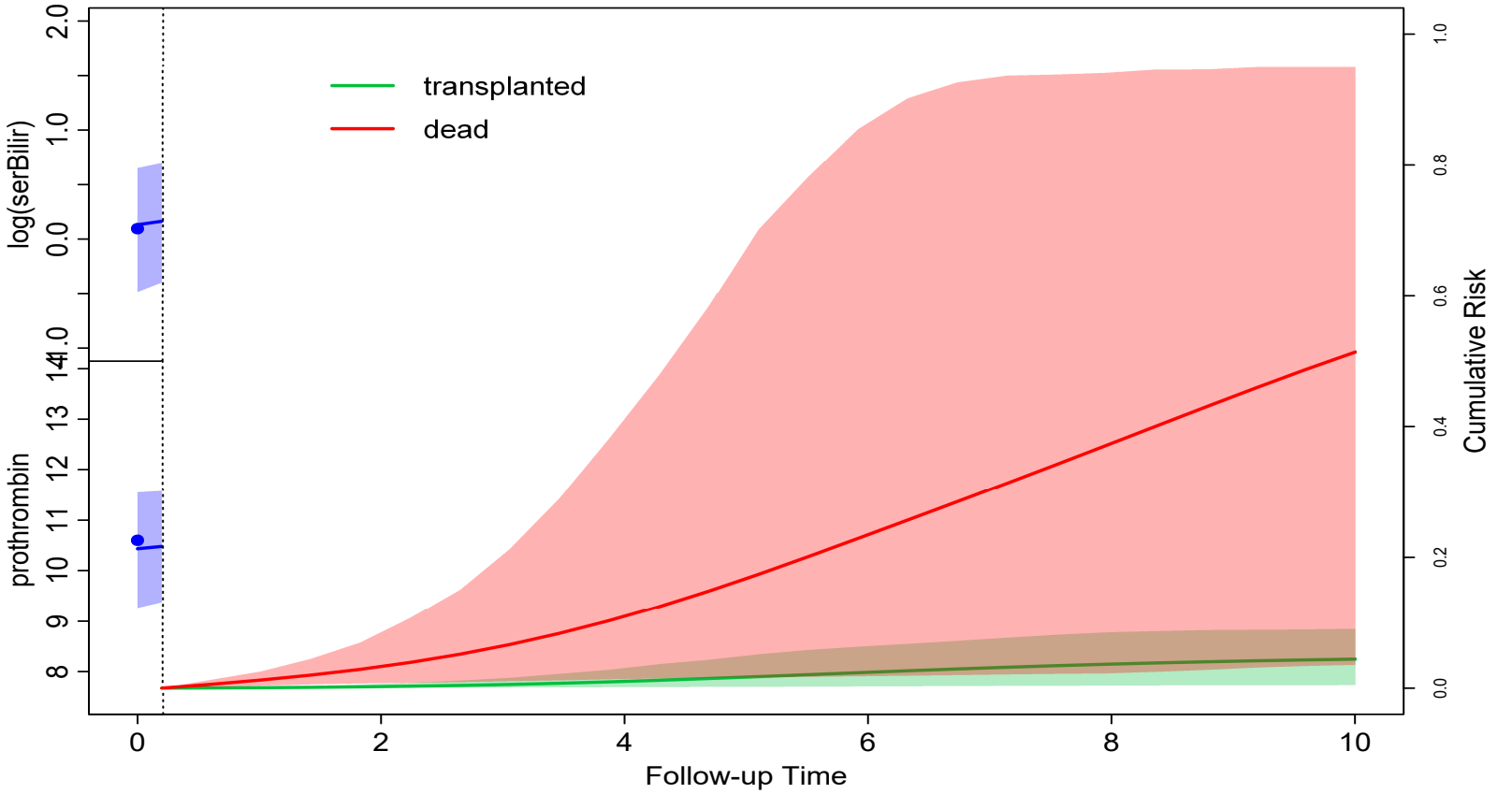
  and calculated a corresponding 95% pointwise CIs

5.5 Cumulative Risk Probabilities (cont'd)

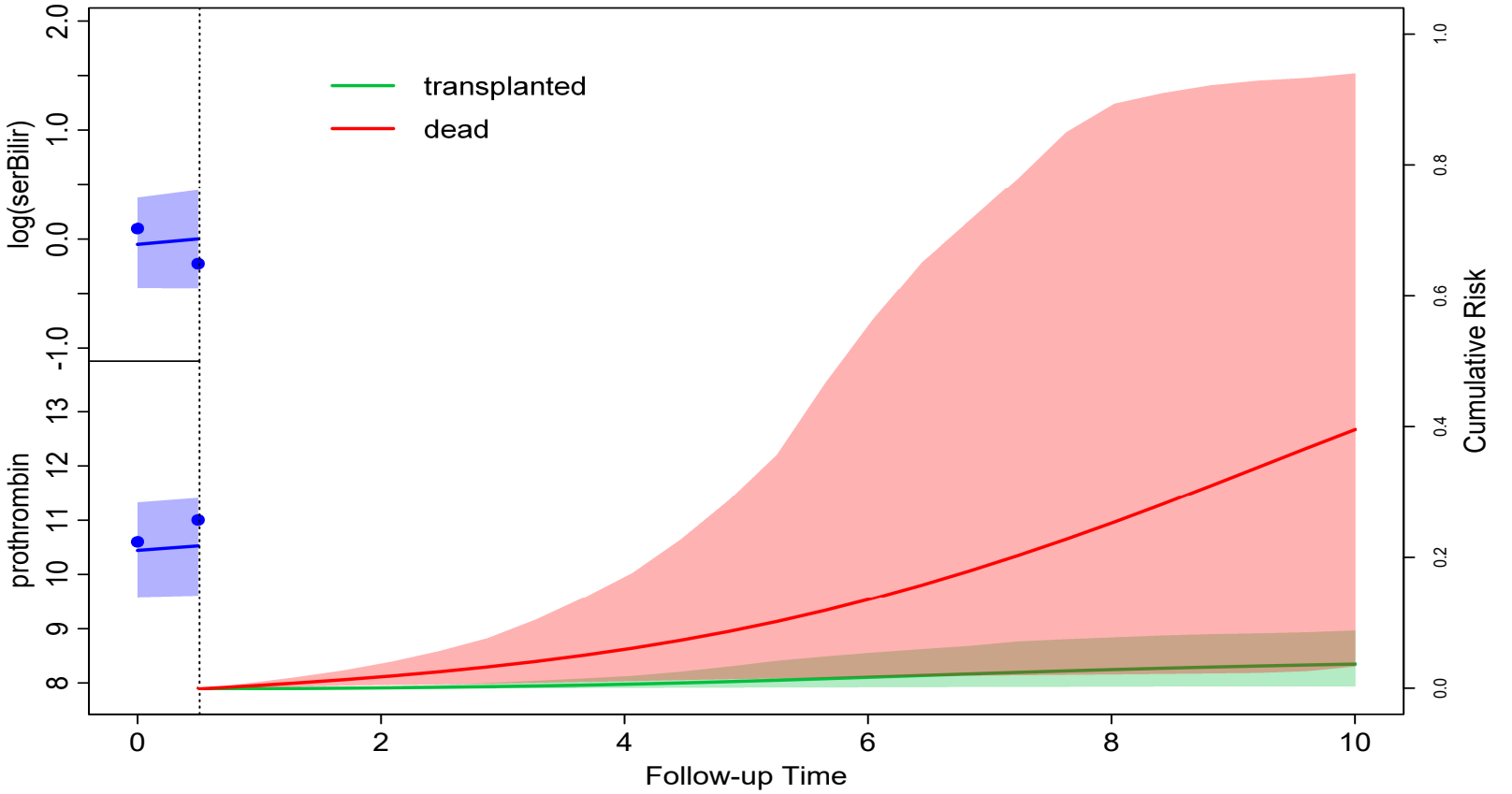




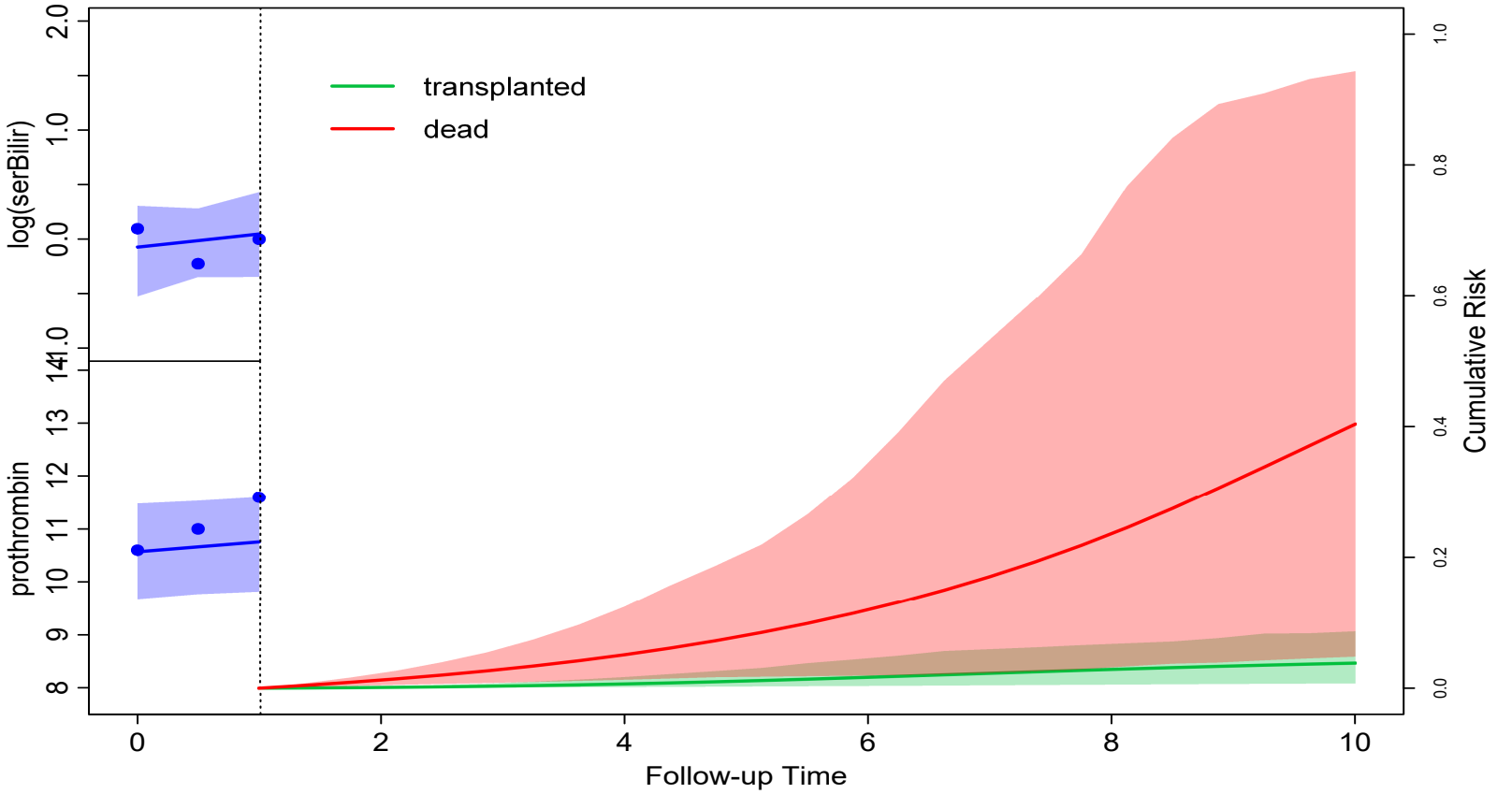
5.5 Cumulative Risk Probabilities (cont'd)



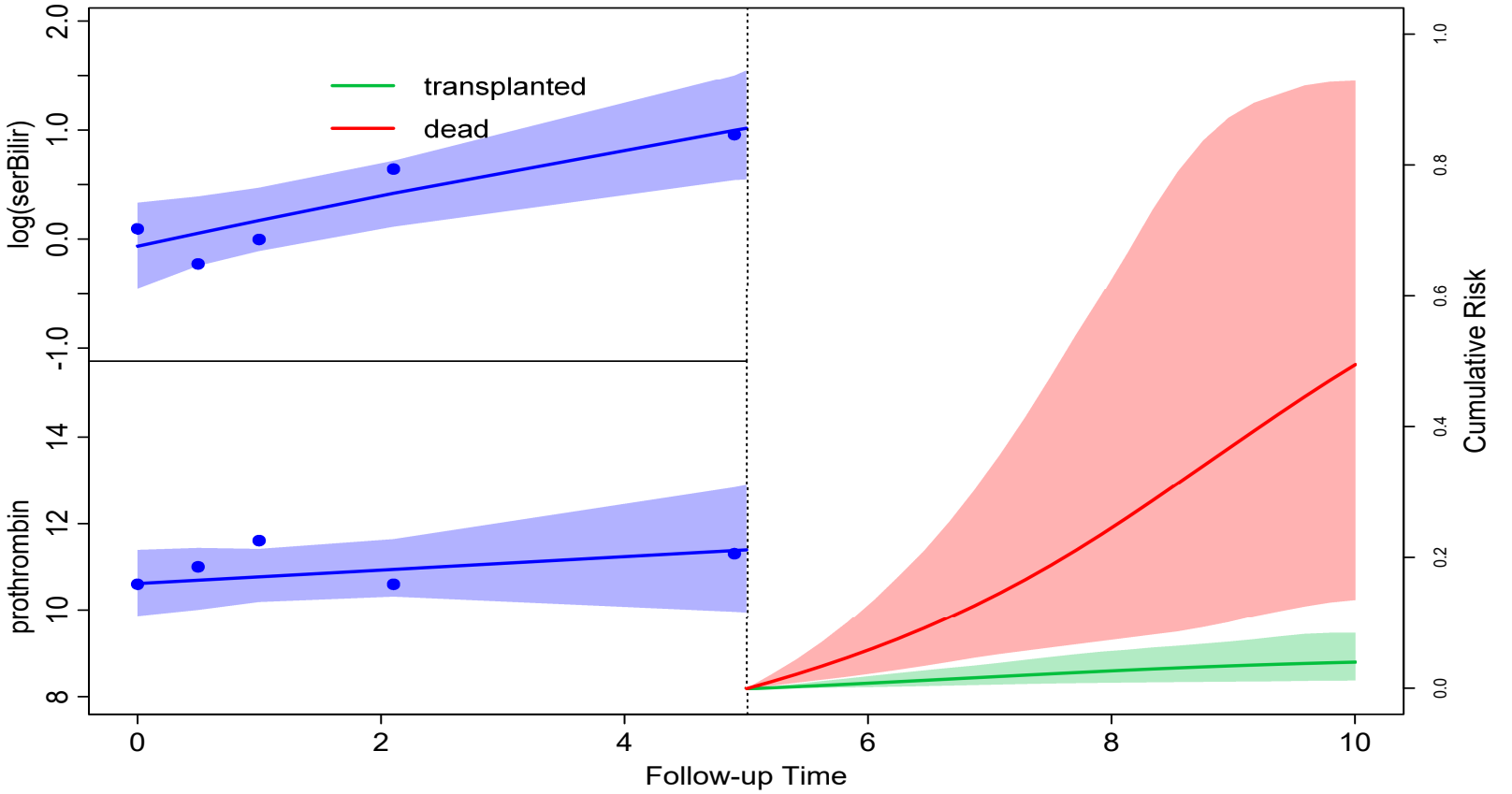
5.5 Cumulative Risk Probabilities (cont'd)



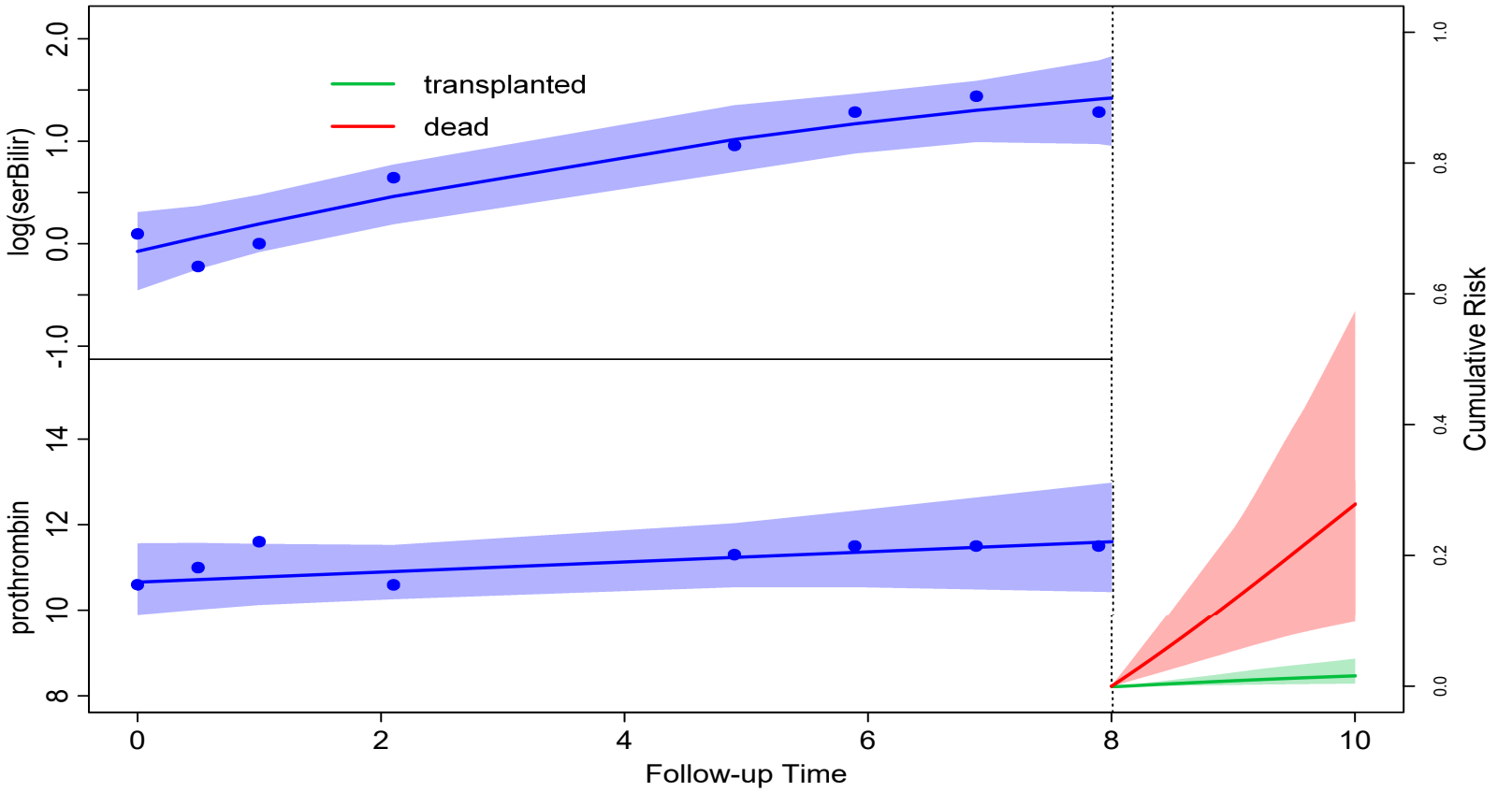
5.5 Cumulative Risk Probabilities (cont'd)



5.5 Cumulative Risk Probabilities (cont'd)



5.5 Cumulative Risk Probabilities (cont'd)



5.6 Discrimination with Competing Risks

- We have seen how to calculate conditional cumulative incidence functions $F_{kj}(t + \Delta t \mid t)$ on the basis of a competing risk joint model
  - ▷ Similarly to the single event case, their accuracy can be evaluated through appropriately defined measures

- Predictive accuracy measures
  - ▷ Discrimination: sensitivity, specificity, ROC and AUC
  - ▷ Calibration: comparison between predicted and observed probabilities
  - ▷ Overall: combination of discrimination and calibration

5.6 Discrimination with Competing Risks (cont'd)

- Without loss of generality, let us focus on the first event, $\delta_j^* = 1$ (main event)

- Definition of cases and controls is more challenging in the competing risk setting
  - ▷ Cases $\rightarrow \left\{T_j^* \in (t, t + \Delta t], \delta_j^* = 1\right\}$
  - ▷ Controls $\rightarrow$??

5.6 Discrimination with Competing Risks (cont'd)

- As in the case of a single event, to assess the discriminative power of the model, we assume the following setting

  ▷ available longitudinal information from multiple markers up to time t,

  ▷ we are interested in events occurring in a medically-relevant interval $(t, t + \Delta t]$

5.6 Discrimination with Competing Risks (cont'd)

- Based on a fitted joint model and for a specific threshold $c \in [0, 1]$, we can term a subject j a case if

$$F_{1j}(t + \Delta t \mid t) \geq c$$

- Definition of sensitivity should be clear

$$\mathsf{SN}_t^{\Delta t}(c) = \Pr\left\{F_{1j}(t + \Delta t \mid t) \geq c \mid T_j^* \in (t, t + \Delta t], \delta_j^* = 1\right\}$$

5.6 Discrimination with Competing Risks (cont'd)

- **Controls** may be defined via several ways, but here we define controls as subjects who are not cases, i.e.,
  - ▷ Event-free at $t + \Delta t$ or
  - ▷ Experienced a competing event within $(t, t + \Delta t]$

- Definition of specificity

$$\text{SP}_t^{\Delta t}(c) = \Pr\left[F_{1j}(t + \Delta t \mid t) \leq c \mid \left\{T_j^* > t + \Delta t\right\} \cup \left\{T_j^* \in (t, t + \Delta t], \delta_j^* \neq 1\right\}\right]$$

5.6 Discrimination with Competing Risks (cont'd)

- The previous definitions of sensitivity and specificity give rise to the following definition of the AUC

$$\text{AUC}_t^{\Delta t} = \Pr\Big[F_{1i}(t+\Delta t \mid t) \geq F_{1j}(t+\Delta t \mid t)$$
$$T_i^* \in (t, t+\Delta t], \delta_i^* = 1, \left\{T_j^* > t+\Delta t\right\} \cup \left\{T_j^* \in (t, t+\Delta t], \delta_j^* \neq 1\right\}\Big]$$

- The probability of observing a pair of subjects (i, j) where subject i has higher cumulative risk for the main event compared to subject j, given that subject i is a case and subject j a control

5.6 Discrimination with Competing Risks (cont'd)

- Subjects censored within $(t, t + \Delta t]$ have a **missing** status (cases or controls?)
- Blanche et al. (2013, 2014) derived IPCW estimators, accounting for missingness due to right censoring.
- Observed cases and controls were weighed by the probability of being observed

5.6 Discrimination with Competing Risks (cont'd)

- Sensitivity can be estimated by

$$\widehat{\mathsf{SN}}_t^{\Delta t}(c) = \frac{\sum_{i=1}^n I\left\{\hat{F}_{1j}(t+\Delta t \mid t) \geq c\right\} I\left\{T_i \in (t, t+\Delta t], \delta_i = 1\right\} \times \Omega_i}{\sum_{i=1}^n I\left\{T_i \in (t, t+\Delta t], \delta_i = 1\right\} \times \Omega_i}$$

- Let $\hat{G}()$ be the Kaplan-Meier estimator of the survival function of the censoring distribution

- $\Omega_i = \frac{\hat{G}(T_i)}{\hat{G}(t)}$ denotes the estimated conditional probability of not being censored at T_i conditional on being uncensored at t

- Recall that $T_i = \min(T_i^*, C_i)$ and $\delta_i = \delta_i^* I(T_i^* \leq C_i)$ represent the observed survival time and event type, respectively.

5.6 Discrimination with Competing Risks (cont'd)

- Subjects censored before t are only used to estimate the weights
- Blanche et al. (2014) derived similar estimators for the specificity and the AUC
- This procedure is different than the one we used before (**model-based weighting**)
- Model-based weights and IPCW have advantages and disadvantages (see our previous discussion)

5.6 Discrimination with Competing Risks (cont'd)

- As mentioned before, a metric that combines discrimination and calibration is the Brier score

- In competing risks, this is defined as

$$\text{PE}(t+\Delta t \mid t) = E\left[\{F_{1j}(t+\Delta t \mid t) - T_i^* \in (t, t+\Delta t], \delta_i^* = 1\}^2 \mid T_i^* > t\right]$$

- Blanche et al. (2014) derived a similar IPCW estimator based on the Kaplan-Meier distribution of censoring

5.6 Discrimination with Competing Risks (cont'd)

R> Not currently implemented in package **JMbayes2**

R> `tvAUC()` and `tvBrier()` will be extended soon to competing risks

Part VI

Closing

6.1 Concluding Remarks

- **When we need joint models for longitudinal and survival outcomes?**
  - ▷ to handle endogenous time-varying covariates in a survival analysis context
  - ▷ to account for nonrandom dropout in a longitudinal data analysis context

- **How joint models work?**
  - ▷ a mixed model for the longitudinal outcome
  - ▷ a relative risk model for the event process
  - ▷ explain interrelationships with shared random effects

6.1 Concluding Remarks (cont'd)

- **Where to pay attention when defining joint models?**
  - ▷ model flexibly the subject-specific evolutions for the longitudinal outcome
  - ▷ consider how to model the association structure between the two processes
    $\Rightarrow$ Functional Forms

- **Extensions**
  - ▷ under the full conditional independence assumption we can easily extend the basic joint model
  - ▷ multiple longitudinal outcomes and/or multiple failure times
  - ▷ though more computationally intensive

6.1 Concluding Remarks (cont'd)

- **Individualized predictions**
  - ▷ joint models can provide subject-specific predictions for the longitudinal and survival outcomes
  - ▷ these are dynamically updated as extra information is recorded for the subjects
  - ▷ joint models constitute an excellent tool for personalized medicine

The End!

6.2 Additional References

- Andrinopoulou, E.R., Rizopoulos, D., Takkenberg, J. and Lesaffre, E. (2014). Joint modeling of two longitudinal outcomes and competing risk data. *Statistics in Medicine*, to appear.

- Brown, E. and Ibrahim, J. (2003). A Bayesian semiparametric joint hierarchical model for longitudinal and survival data. *Biometrics* **59**, 221–228.

- Brown, E. Ibrahim, J. and DeGruttola, V. (2005). A flexible B-spline model for multiple longitudinal biomarkers and survival. *Biometrics* **61**, 64–73.

- Chi, Y.-Y. and Ibrahim, J. (2006). Joint models for multivariate longitudinal and multivariate survival data. *Biometrics* **62**, 432–445.

- DeGruttola, V. and Tu, X. (1994). Modeling progression of CD-4 lymphocyte count and its relationship to survival time. *Biometrics* **50**, 1003–1014.

- Elashoff, R., Li, G. and Li, N. (2008). A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics* **64**, 762–771.

6.2 Additional References (cont'd)

- Faucett, C. and Thomas, D. (1996). Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach. *Statistics in Medicine* **15**, 1663–1685.

- Gerds, T. and Schumacher, M. (2006). Consistent estimation of the expected Brier score in general survival models with right-censored event times. *Biometrical Journal* **48**, 1029–1040.

- Heagerty, P. and Zheng, Y. (2005). Survival model predictive accuracy and ROC curves. *Biometrics* **61**, 92–105.

- Henderson, R., Diggle, P. and Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics* **1**, 465–480.

- Henderson, R., Diggle, P. and Dobson, A. (2002). Identification and efficacy of longitudinal markers for survival. *Biostatistics* **3**, 33–50.

- Hsieh, F., Tseng, Y.-K. and Wang, J.-L. (2006). Joint modeling of survival and longitudinal data: Likelihood approach revisited. *Biometrics* **62**, 1037–1043.

6.2 Additional References (cont'd)

- Lin, H., Turnbull, B., McCulloch, C. and Slate, E. (2002). Latent class models for joint analysis of longitudinal biomarker and event process: Application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association* **97**, 53–65.

- Liu, L. and Huang, X. (2009). Joint analysis of correlated repeated measures and recurrent events processes in the presence of death, with application to a study on acquired immune deficiency syndrome. *Journal of the Royal Statistical Society, Series C* **58**, 65–81.

- Proust-Lima, C., Joly, P., Dartigues, J. and Jacqmin-Gadda, H. (2009). Joint modelling of multivariate longitudinal outcomes and a time-to-event: A nonlinear latent class approach. *Computational Statistics and Data Analysis* **53**, 1142–1154.

- Proust-Lima, C. and Taylor, J. (2009). Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: A joint modeling approach. *Biostatistics* **10**, 535–549.

- Rizopoulos, D. (2012). Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive Gaussian quadrature rule. *Computational Statistics and Data Analysis* **56**, 491–501.

- Rizopoulos, D. (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* **67**, 819–829.

6.2 Additional References (cont'd)

- 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. and Ghosh, P. (2011). A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Statistics in Medicine* **30**, 1366–1380.

- Rizopoulos, D., Hatfield, L.A., Carlin, B.P. and Takkenberg, J.J.M. (2014). Combining dynamic predictions from joint models for longitudinal and time-to-event data using Bayesian model averaging. *Journal of the American Statistical Association* **109**, 1385–1397.

- Rizopoulos, D., Murawska, M., Andrinopoulou, E.-R., Molenberghs, G., Takkenberg, J. and Lesaffre, E. (2013). Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Submitted*.

- Rizopoulos, D. and Lesaffre, E. (2014). Introduction to the special issue on joint modelling techniques. *Statistical Methods in Medical Research* **23**, 3–10.

- Rizopoulos, D., Verbeke, G. and Lesaffre, E. (2009). Fully exponential Laplace approximation for the joint modelling of survival and longitudinal data. *Journal of the Royal Statistical Society, Series B* **71**, 637–654.

6.2 Additional References (cont'd)

- Rizopoulos, D., Verbeke, G., Lesaffre, E. and Vanrenterghem, Y. (2008). A two-part joint model for the analysis of survival and longitudinal binary data with excess zeros. *Biometrics* **64**, 611–619.

- Rizopoulos, D., Verbeke, G. and Molenberghs, G. (2010). Multiple-imputation-based residuals and diagnostic plots for joint models of longitudinal and survival outcomes. *Biometrics* **66**, 20–29.

- Rizopoulos, D., Verbeke, G. and Molenberghs, G. (2008). Shared parameter models under random effects misspecification. *Biometrika* **95**, 63–74.

- Rubin, D. (1976). Inference and missing data. *Biometrika* **63**, 581–592.

- Song, X., Davidian, M. and Tsiatis, A. (2002). A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. *Biometrics* **58**, 742–753.

- Taylor, J., Park, Y., Ankerst, D., Proust-Lima, C., Williams, S., Kestin, L., Bae, K., Pickles, T., and Sandler, H. (2013). Real-time individual predictions of prostate cancer recurrence using joint models. *Biometrics*, **69**, 206–213.

- Tseng, Y.-K., Hsieh, F. and Wang, J.-L. (2005). Joint modelling of accelerated failure time and longitudinal data. *Biometrika* **92**, 587–603.

6.2 Additional References (cont'd)

- Tsiatis, A. and Davidian, M. (2001). A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika* **88**, 447–458.

- Tsiatis, A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: An overview. *Statistica Sinica* **14**, 809–834.

- Tsiatis, A., DeGruttola, V., and Wulfsohn, M. (1995). Modeling the relationship of survival to longitudinal data measured with error: Applications to survival and CD4 counts in patients with AIDS. *Journal of the American Statistical Association* **90**, 27–37.

- Viviani, S., Alfó, M. and Rizopoulos, D. (2014). Generalized linear mixed joint model for longitudinal and survival outcomes. *Statistics and Computing*, **24**, 417–427.

- Viviani, S., Rizopoulos, D. and Alfó, M. (2014). Local sensitivity of shared parameter models to nonignorability of dropout. *Statistical Modelling* **14**, 205–228.

- Wang, Y. and Taylor, J. (2001). Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association* **96**, 895–905.

6.2 Additional References (cont'd)

- Wu, M. and Bailey, K. (1988). Analysing changes in the presence of informative right censoring caused by death and withdrawal. *Statistics in Medicine* **7**, 337–346.

- Wu, M. and Bailey, K. (1989). Estimation and comparison of changes in the presence of informative right censoring: conditional linear model. *Biometrics* **45**, 939–955.

- Wu, M. and Carroll, R. (1988). Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics* **44**, 175–188.

- Wulfsohn, M. and Tsiatis, A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics* **53**, 330–339.

- Xu, C., Baines, P. and Wang, J.-L. (2014). Standard error estimation using the EM algorithm for the joint modeling of survival and longitudinal data. *Biostatistics*, to appear.

- Xu, J. and Zeger, S. (2001). Joint analysis of longitudinal data comprising repeated measures and times to events. *Applied Statistics* **50**, 375–387.

6.2 Additional References (cont'd)

- Ye, W., Lin, X., and Taylor, J. (2008). Semiparametric modeling of longitudinal measurements and time-to-event data - a two stage regression calibration approach. *Biometrics* **64**, 1238–1246.

- Yu, M., Law, N., Taylor, J., and Sandler, H. (2004). Joint longitudinal-survival-cure models and their application to prostate cancer. *Statistica Sinica* **14**, 835–862.

- Yu, M., Taylor, J. and Sandler, H. (2008). Individualized prediction in prostate cancer studies using a joint longitudinal-survival-cure model. *Journal of the American Statistical Association* **108**, 178–187.

- Zeng, D. and Cai, J. (2005). Asymptotic results for maximum likelihood estimators in joint analysis of repeated measurements and survival time. *The Annals of Statistics* **33**, 2132–2163.

- Zheng, Y. and Heagerty, P. (2007). Prospective accuracy for longitudinal markers. *Biometrics* **63**, 332–341.

6.3 Medical Papers with Joint Modeling

- Andrinopoulou, E.R., Rizopoulos, D., Jin, R., Bogers, A., Lesaffre, E. and Takkenberg, J. (2012). An introduction to mixed models and joint modeling: Analysis of valve function over time. *Annals of Thoracic Surgery* **93**, 1765–1772.

- Andrinopoulou, E.R., Rizopoulos, D., Geleijnse, M., Lesaffre, E., Bogers, A. and Takkenberg, J. (2015). Dynamic prediction of outcome for patients with severe aortic stenosis: Application of joint models for longitudinal and time-to-event data. *BMC Cardiovascular Disorders*, to appear.

- Daher Abdi, D.Z., Essig, M., Rizopoulos, D., Le Meur, Y., Premaud, A., Woillard, J.-B., Rerolle, J.-P., Marquet, P. and Rousseau, A. (2013). Impact of longitudinal exposure to mycophenolic acid on acute rejection in renal-transplant recipients using a joint modeling approach. *Pharmacological Research* **72**, 52–60.

- Ibrahim, J., Chu, H. and Chen, L.-M. (2010). Basic concepts and methods for joint models of longitudinal and survival data. *Journal of Clinical Oncology* **28**, 2796–2801.

- Nunez, J., Nunez, E., Rizopoulos, D., Minana, G., Bodi, V., Bondanza, L., Husser, O., Merlos, P., Santas, E., Pascual-Figal, D.,; Chorro, F. and Sanchis, J. (2014). Red blood cell distribution width is longitudinally associated with mortality and incident anemia in heart failure patients. *Circulation Journal* **78**, 410—418.

- Rizopoulos, D. and Takkenberg, J. (2014). Tools & Techniques: Dealing with time-varying covariates in survival analysis - joint models versus Cox models. *EuroIntervention* **10**, 285–288.

6.3 Medical Papers with Joint Modeling (cont'd)

- Thabut, G., Christie, J., Mal, H., Fournier, M., Brugiere, O., Leseche, G., Castier, Y. and Rizopoulos, D. (2013). Survival benefit of lung transplant for cystic fibrosis since lung-allocation-score implementation. *American Journal of Respiratory and Critical Care Medicine* **187**, 1335–1340.

- van der Linde, D., Roos-Hesselink, J., Rizopoulos, D., Heuvelman, H., Budts, W., van Dijk, A., Witsenburg, M., Yap, S., Bogers, A., Oxenius, A., Silversides, C., Oechslin, E. and Takkenberg, J. (2013). Surgical outcome of discrete subaortic stenosis in adults: A multicenter study. *Circulation* **127**, 1184–1191.

- van der Linde, D., Takkenberg, J., Rizopoulos, D., Heuvelman, H., Budts, W., van Dijk, A., Witsenburg, M., Yap, S., Bogers, A., Oxenius, A., Silversides, C., Oechslin, E. and Roos-Hesselink, J. (2013). Natural history of discrete subaortic stenosis in adults: A multicenter study. *European Heart Journal* **34**, 1548–1556.

Part VII

Practicals

7.1 R Practical: Dynamic Predictions

- We will work with the Liver Cirrhosis dataset
  - ▷ a placebo-controlled randomized trial on 488 liver cirrhosis patients

- Start R and load package **JMbayes2**, using `library("JMbayes2")`

- The longitudinal (long format) and survival information for the liver cirrhosis patients can be found in data frames `prothro` and `prothros`, respectively
  - ▷ the variables that we will need are:

7.1 R Practical: Dynamic Predictions (cont'd)

▷ `prothro`
  * `id`: patient id number
  * `pro`: prothrombin measurements
  * `time`: follow-up times in years
  * `treat`: randomized treatment

▷ `prothros`
  * `Time`: observed event times in years
  * `death`: event indicator with 0 = 'alive', and 1 = 'dead'
  * `treat`: randomized treatment

7.1 R Practical: Dynamic Predictions (cont'd)

- We will fit the following joint model to the Liver Cirrhosis dataset
  - ▷ longitudinal submodel: linear subject-specific random slopes for prothrombin levels allowing for different average evolutions in the two treatment groups

$$\begin{aligned} y_i(t) &= m_i(t) + \varepsilon_i(t) \\ m_i(t) &= \beta_0 + \beta_1 t + \beta_2 \{\texttt{Trt}_i \times t\} + b_{i0} + b_{i1} t \end{aligned}$$

  - ▷ survival submodel: treatment effect & *true* effect of prothrobin

$$h_i(t) = h_0(t) \exp\{\gamma \texttt{Trt}_i + \alpha m_i(t)\}$$

7.1 R Practical: Dynamic Predictions (cont'd)

- T1: Fit the linear mixed model using `lme()`, the Cox model using `coxph()`, and the corresponding joint model using `jm()` (see pp.41–43)

- We are interested in producing predictions of survival probabilities for Patient 155

- T2: Extract the data of Patient 155 using the code

```
dataP155 <- prothro[prothro$id == 155, ]
```

7.1 R Practical: Dynamic Predictions (cont'd)

- T3: Using the first measurement of Patient 155, and the fitted joint model calculate his conditional survival probabilities using function `predict()` and plot it using the `plot` method (see p.109)
  - ▷ set the `Time` variable equal to the time of the first measurement
  - ▷ set the `death` variable equal to 0

- T4: Combine the predictions in one plot
  - ▷ save as the object `Spred` the survival predictions, and `Lpred` the longitudinal ones
  - ▷ use `plot(Lpred, Spred)`

7.1 R Practical: Dynamic Predictions (cont'd)

- T5: Repeat the same procedure by including each time the next measurement of Patient 155 and see how his survival probabilities change over time as extra prothrombin measurements are recorded

  ▷ first using only the first measurement,

  ▷ and following update the predictions after each new longitudinal measurement has been recorded

  ▷ use a `for` loop to achieve this

7.1 R Practical: Dynamic Predictions (cont'd)

- T6: Calculate the ROC and the corresponding AUC under the postulated model at year 3 and with a 1-year window (see p.130)
  - ▷ using model-based weights and IPCW

- T7: Calculate the prediction error for the same period (see p.137)
  - ▷ using model-based weights and IPCW

7.2 R Practical: Dynamic Predictions CIFs

- We will work with the Mayo Clinic Primary Biliary Cirrhosis Data
  - ▷ A placebo-controlled randomized trial on 312 patients with primary biliary cirrhosis

- Start R and load package **JMbayes2**, using library("JMbayes2")

- The longitudinal (long format) and survival information for the primary biliary cirrhosis patients can be found in data frames pbc2 and pbc2.id, respectively
  - ▷ the variables that we will need are:

7.2 R Practical: Dynamic Predictions CIFs (cont'd)

- pbc2
  - ▷ id: patient id number
  - ▷ serBilir: serum bilirubin in mg/dl
  - ▷ prothrombin: prothrombin time in seconds
  - ▷ year: measurement times (in years)
  - ▷ drug: treatment group (placebo and D-penicil)
- pbc2.id
  - ▷ years: patient id number
  - ▷ status2: a factor with levels alive, transplanted and dead
  - ▷ drug: treatment group (placebo and D-penicil)
  - ▷ age: at baseline (in years)

7.2 R Practical: Dynamic Predictions CIFs (cont'd)

- We will fit the following joint model to the Mayo Clinic Primary Biliary Cirrhosis dataset
  - ▷ Longitudinal submodel for log(serBilir): linear and quadratic subject-specific random slopes for $\log$ bilirubin levels allowing for different average evolutions in the two treatment groups

$$\begin{aligned} y_{i1}(t) &= m_{i1}(t) + \epsilon_{i1}(t) \\ m_{i1}(t) &= \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 \{\texttt{Drug}_i \times t\} + \beta_4 \{\texttt{Drug}_i \times t^2\} \\ &+ b_{i0} + b_{i1} t + b_{i0} t^2 \end{aligned}$$

  - ▷ Longitudinal submodel for prothrombin: linear subject-specific random slopes for prothrombin levels allowing for different average evolutions in the two treatment groups

$$\begin{aligned} y_{i2}(t) &= m_{i2}(t) + \epsilon_{i2}(t) \\ m_{i2}(t) &= \beta_0 + \beta_1 t + \beta_2 \{\texttt{Drug}_i \times t\} + b_{i0} + b_{i1} t \end{aligned}$$

7.2 R Practical: Dynamic Predictions CIFs (cont'd)

- We will fit the following joint model to the Mayo Clinic Primary Biliary Cirrhosis dataset
  - ▷ Cause-specific hazard for death

$$h_i^d(t) = h_0^d(t) \exp\{\gamma_{d1}\texttt{Age}_i + \gamma_{d2}\texttt{Drug}_i + \alpha_{d1}m_{i1}(t) + \alpha_{d2}m_{i2}(t)\}$$

  - ▷ Cause-specific hazard for transplantation

$$h_i^{tr}(t) = h_0^{tr}(t) \exp\{\gamma_{tr1}\texttt{Age}_i + \gamma_{tr2}\texttt{Drug}_i + \alpha_{tr1}m_{i1}(t) + \alpha_{tr2}m_{i2}(t)\}$$

7.2 R Practical: Dynamic Predictions CIFs (cont'd)

- T1: Fit the longitudinal models for log(serBilir) and prothrombin using `lme()`
  - ▷ `poly(year, 2)` can automatically construct linear and quadratic slopes

- T2: Use `crisk_setup` to appropriately construct a competing risk format dataset
  - ▷ specify the event type variable, the level corresponding to right censoring and a name for the strata variable to be constructed

7.2 R Practical: Dynamic Predictions CIFs (cont'd)

- T3: Fit a `coxph()` model to the new dataset allowing for interaction with the event type

- Create a named `list()` for each longitudinal outcome to ensure an interaction with the event-type variable

```
CR_forms <- list(
  "log(serBilir)" = ~ value(log(serBilir)):CR,
  "prothrombin" = ~ value(prothrombin):CR
)
```

7.2 R Practical: Dynamic Predictions CIFs (cont'd)

- T4: Fit the competing risk joint model for the two longitudinal markers using `jm()` by providing the objects from `lme()` and `coxph()`
  - ▷ Use the argument `functional forms` to provide the `list()`

- T5: Extract the longitudinal and competing risk data of Patient 2 using the code

```
ND_long <- pbc2[pbc2$id == 2, ]
ND_event <- pbc2.idCR[pbc2.idCR$id == 2, ]
```

7.2 R Practical: Dynamic Predictions CIFs (cont'd)

- T6: Use the first observation in the longitudinal data
  - ▷ Set the years equal to 0.2
  - ▷ Set the status2 equal to 0 (event-free at 0.2 years)
  - ▷ Combine the datasets in a named `list()`

    ```
    ND <- list(newdataL = ND_long, newdataE = ND_event)
    ```

  - ▷ Use `predict()` to calculate predictions for the cumulative risk probabilities and future longitudinal values for the two markers up to 10 years since baseline
    - * Use `newdata = ND` in `predict()` and `process = "event"` for cumulative risk predictions

7.2 R Practical: Dynamic Predictions CIFs (cont'd)

- T7: Combine the predictions in one plot
  - ▷ Save as the object predEvent for the survival predictions, and predLong for the longitudinal ones
  - ▷ Use `plot(predLong, predEvent, outcomes = 1:2)`

- T8: Plot the predictions about future longitudinal outcomes for the two markers

```
par(mfrow = c(1,2))
plot(predLong, outcomes = 1)
plot(predLong, outcomes = 2)
```

7.2 R Practical: Dynamic Predictions CIFs (cont'd)

- Repeat the same procedure by keeping data of Patient 2 up to 0.2, 0.5, 1, 5, 8 years since baseline, respectively, and observe how their survival probabilities change over time as extra longitudinal measurements are recorded
  - ▷ first keep data up to 0.2 years,
  - ▷ and following update the predictions after new longitudinal information has been recorded
  - ▷ use a `for` loop to achieve this