Joint Modeling of Longitudinal and Time-to-Event Data with Applications in R

Dimitris Rizopoulos
Department of Biostatistics, Erasmus University Medical Center

d.rizopoulos@erasmusmc.nl

ASA Lifetime Data Science Section

March 24, 2022

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 the state of the art in

Joint Modeling Techniques
for Longitudinal and Time-to-Event Data

Learning Objectives

- **Goals:** After this course participants will be able to
  - ▷ identify settings in which a joint modeling approach is required,
  - ▷ construct and fit an appropriate joint model, and
  - ▷ correctly interpret the obtained results

Agenda

- **Part I:** Introduction
  - ▷ Data sets that we will use throughout the course
  - ▷ Research questions

- **Part II:** (brief) Review of Linear Mixed Models
  - ▷ Features of repeated measurements data
  - ▷ Linear mixed models

Agenda (cont'd)

- **Part III:** (brief) Review of Relative Risk Models
  - ▷ Features of survival data
  - ▷ Relative risk models
  - ▷ Time-varying covariates

- **Part IV:** The Basic Joint Model
  - ▷ Definition
  - ▷ Estimation

Agenda (cont'd)

- **Part V:** Extensions of the Basic Joint Model
  - ▷ Functional forms
  - ▷ Multivariate joint models

- **Part VI:** Dynamic Predictions
  - ▷ Individualized predictions
  - ▷ Effect of the functional forms

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. 122–129.

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:
  - ▷ http://jmr.r-forge.r-project.org [R code used in the book]
  - ▷ http://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

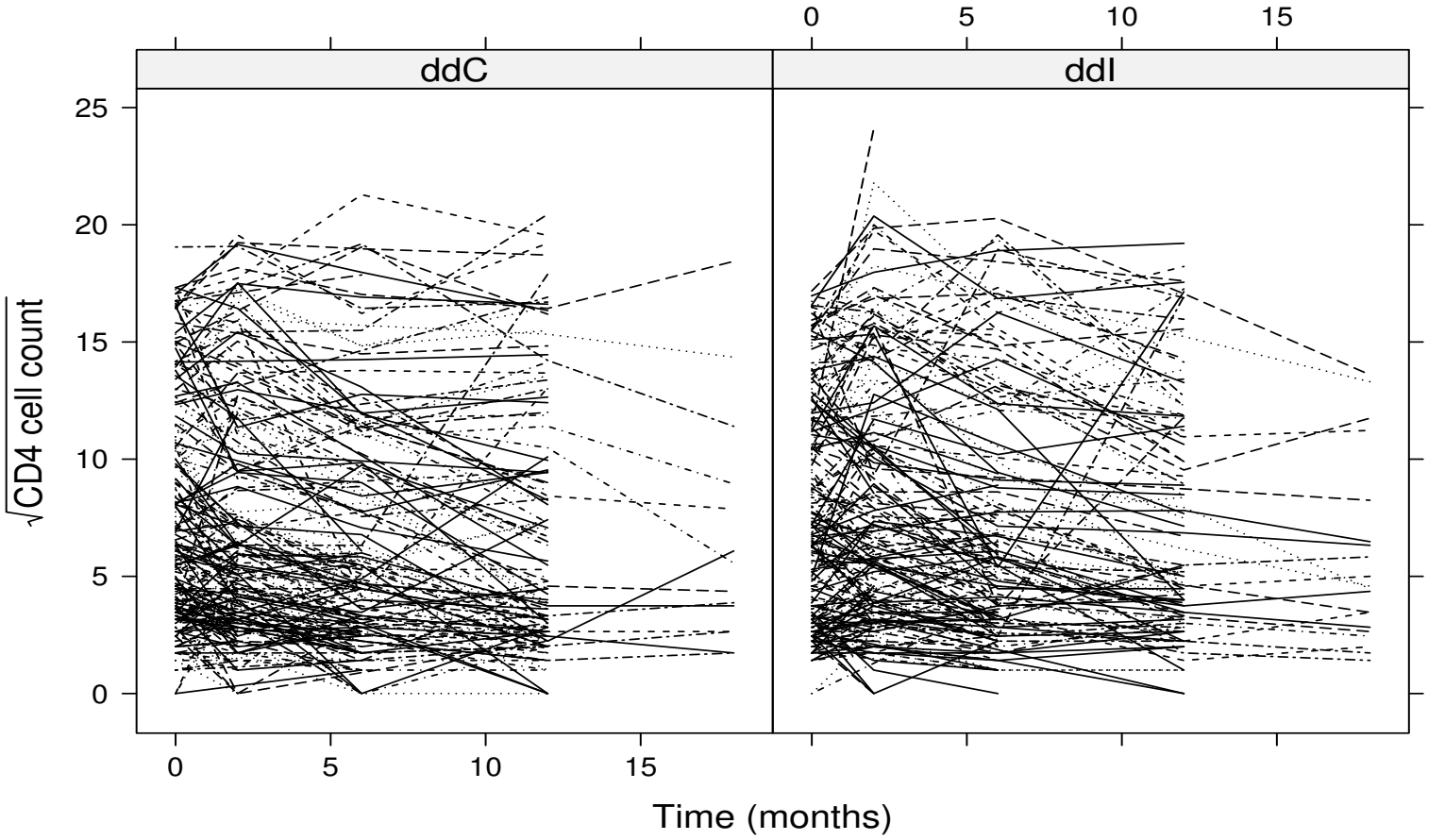
1.1 Motivating Longitudinal Studies

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

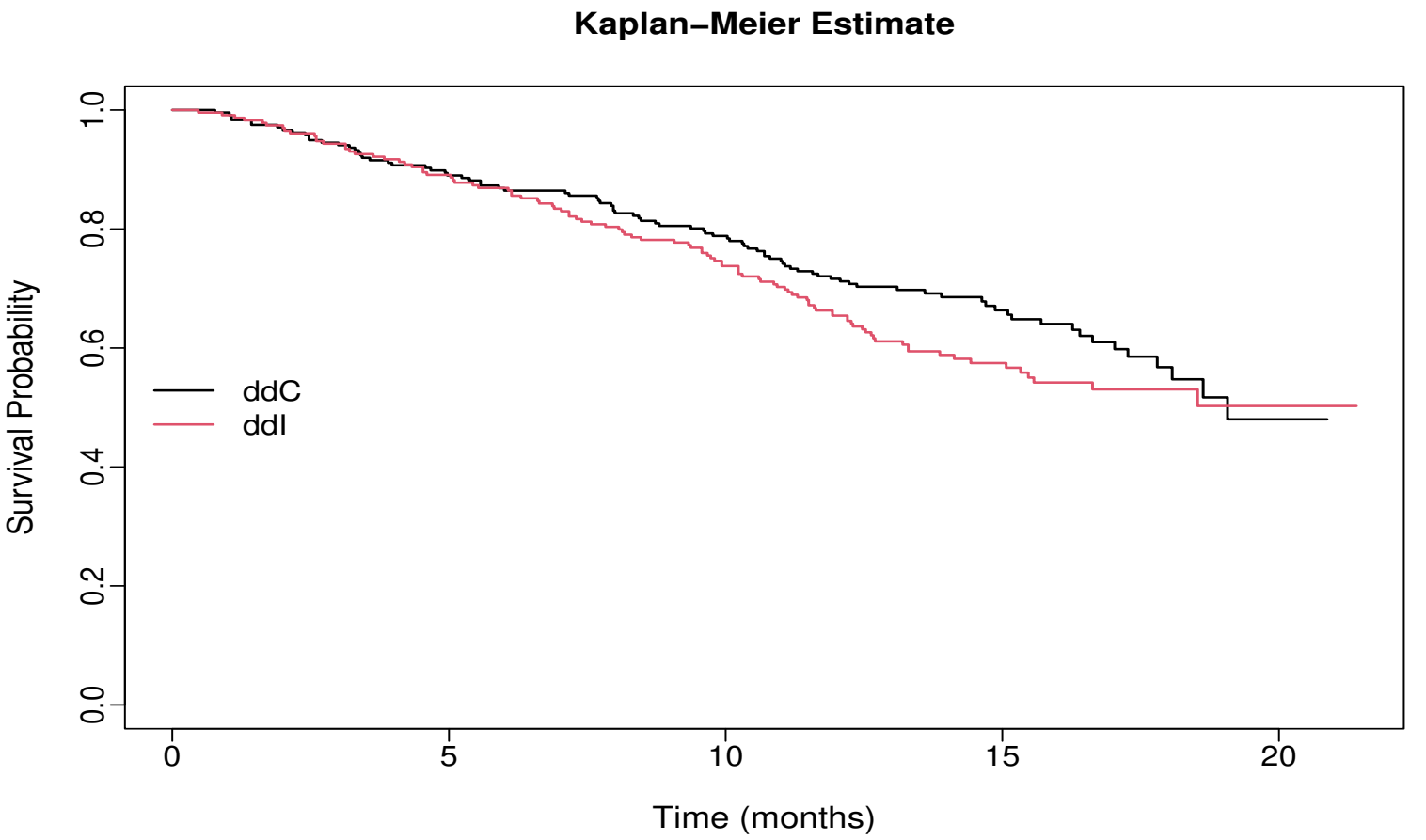
- The aim of this study was to compare the efficacy and safety of two alternative antiretroviral drugs, didanosine (ddI) and zalcitabine (ddC)

- 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)

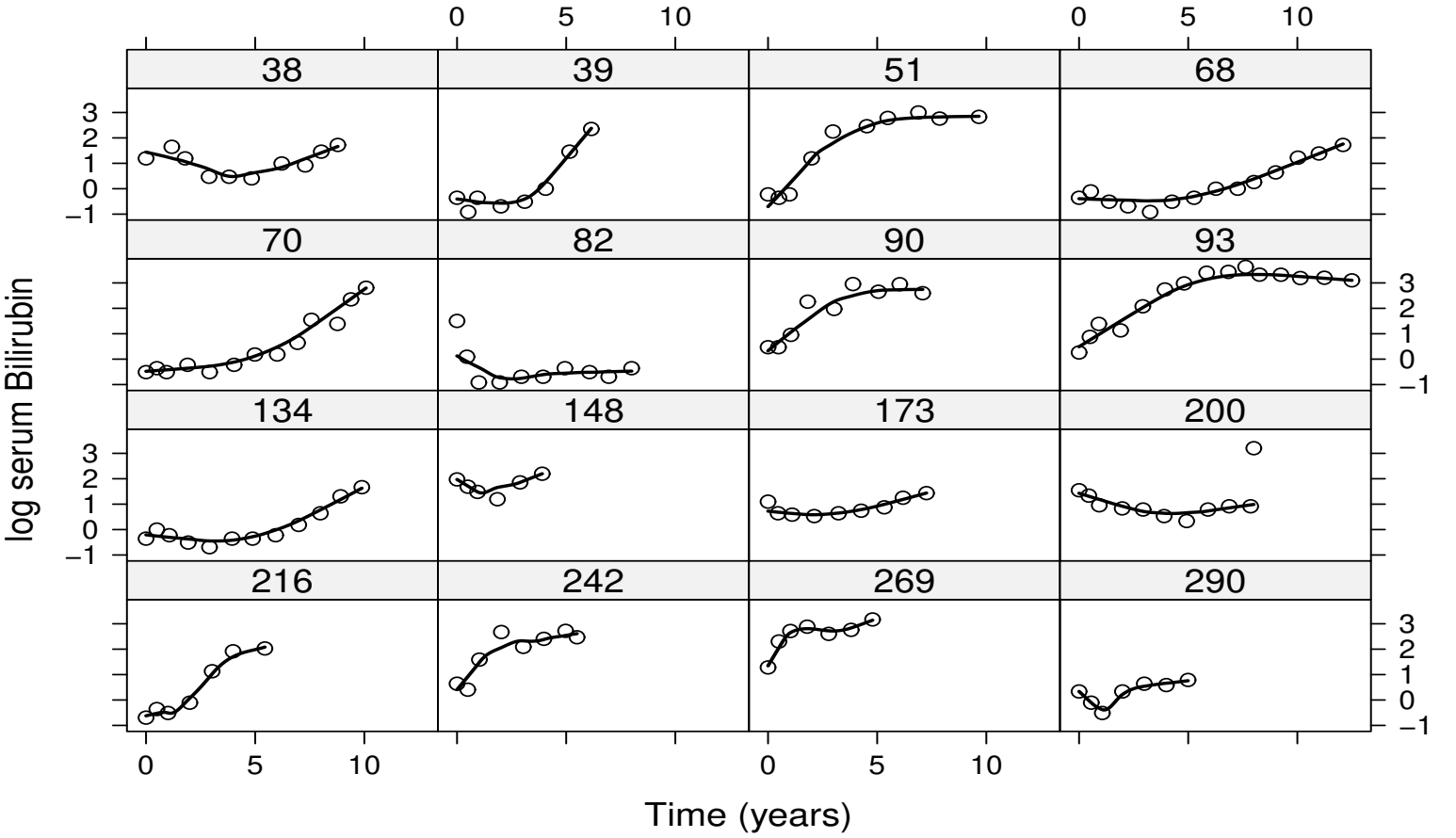
- Research Questions:
  - ▷ How strong is the association between CD4 cell count and the risk of death?
  - ▷ Is CD4 cell count a good biomarker?
    - * if treatment improves CD4 cell count, does it also improve survival?

1.1 Motivating Longitudinal Studies (cont'd)

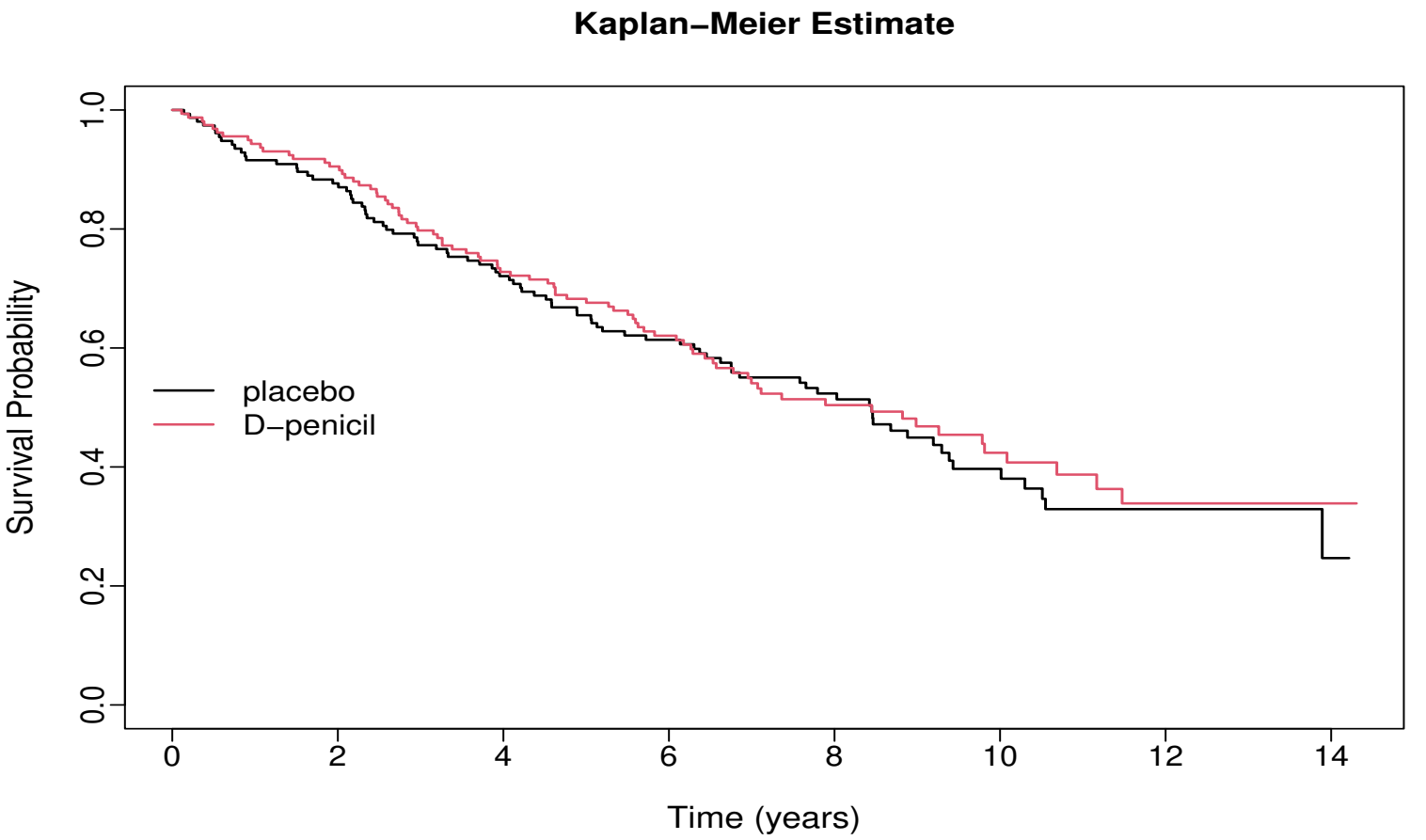
- PBC: Primary Biliary Cirrhosis:
  - ▷ a chronic, fatal but rare liver disease
  - ▷ characterized by inflammatory destruction of the small bile ducts within the liver

- 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?
  - ▷ **Handling implicit outcomes:** focus on longitudinal outcomes but with dropout or random visit times

Part II

Linear Mixed-Effects 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

- This implies that standard statistical tools, such as the t-test and simple linear regression that assume independent observations, are not optimal for longitudinal data analysis.

2.1 Linear Mixed Models (cont'd)

- The direct approach to model correlated data $\Rightarrow$ *multivariate regression*

$$y_i = X_i\beta + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, V_i),$$

where

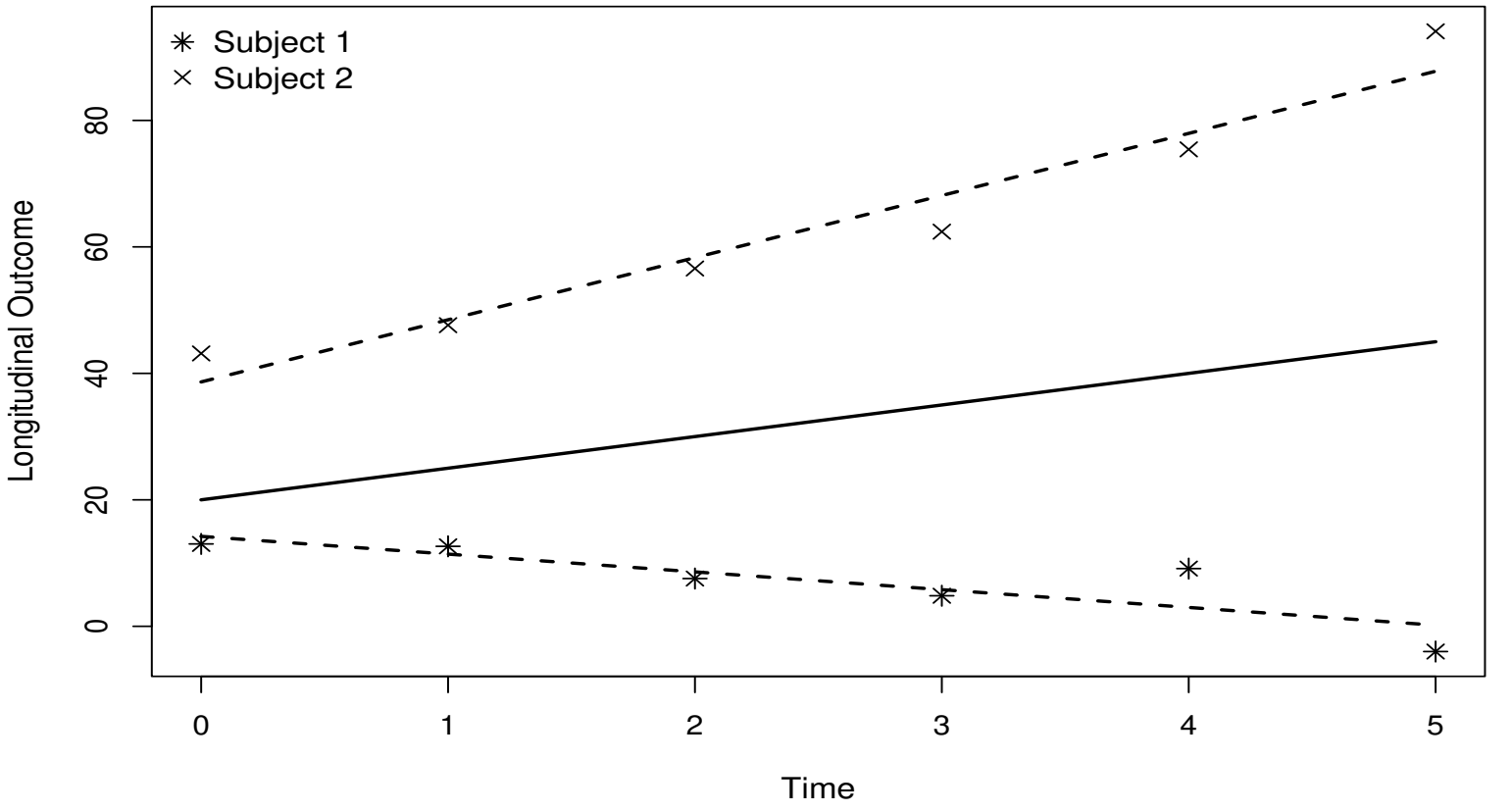
  - $\triangleright$ y_i the vector of responses for the ith subject
  - $\triangleright$ X_i design matrix describing structural component
  - $\triangleright$ V_i covariance matrix describing the correlation structure

- There are several options for modeling V_i, e.g., compound symmetry, autoregressive process, exponential spatial correlation, Gaussian spatial correlation, ...

2.1 Linear Mixed Models (cont'd)

- **Alternative intuitive 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

Part III

Relative Risk Models

3.1 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

3.1 Relative Risk Models (cont'd)

- Several types of censoring:
  - ▷ Location of the true event time wrt the censoring time: *right*, *left* & *interval*
  - ▷ Probabilistic relation between the true event time & the censoring time: *informative* & *non-informative*

Here we focus on non-informative right censoring

3.1 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 $\{T_i, \delta_i\}$

3.1 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

3.1 Relative Risk Models (cont'd)

- **Cox Model:** We make 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

3.2 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?

3.2 Time-Varying Covariates (cont'd)

- To answer our questions of interest we need to postulate a model that relates
  - ▷ the serum bilirubin with
  - ▷ the time-to-death

- The association between **baseline** marker levels and the risk of death can be estimated with standard statistical tools (e.g., Cox regression)

- When we want to study time-varying covariates, a more **careful consideration** is required

3.2 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...

3.2 Time-Varying Covariates (cont'd)

- Example: Consider a study on asthma, in particular on the time until an asthma attack for a group of patients

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

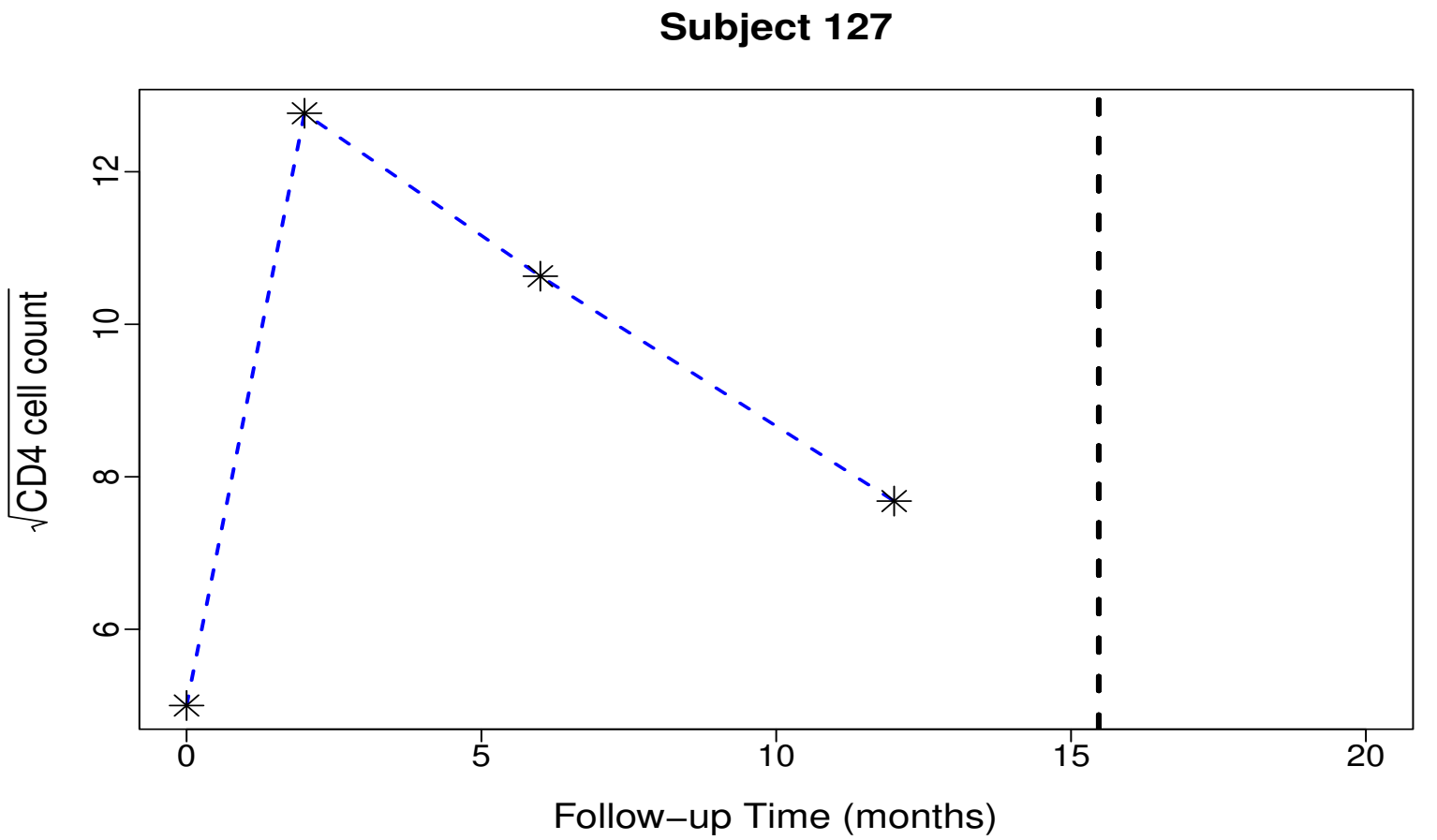
- Say a patient had an asthma attack at a particular time point u
  - ▷ Pollution levels
    - * will the pollution levels at time $t > u$ be affected by the fact that the patient had an attack at u? $\Rightarrow$ **No**
  - ▷ Biomarker
    - * will the biomarker level at time $t > u$ be affected by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

3.2 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

- In our motivating examples all time-varying covariates are **Biomarkers** $\Rightarrow$ These are always **endogenous** covariates
  - ▷ measured with error (i.e., biological variation)
  - ▷ the complete history is not available
  - ▷ existence directly related to failure status

3.2 Time-Varying Covariates (cont'd)



3.2 Time-Varying Covariates (cont'd)

- The Cox model presented earlier can be extended to handle time-varying covariates using the counting process formulation

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

where

  - ▷ $N_i(t)$ is a counting process which counts the number of events for subject i by time t,
  - ▷ $h_i(t)$ denotes the intensity process for $N_i(t)$,
  - ▷ $R_i(t)$ denotes the at risk process ('1' if subject i still at risk at t), and
  - ▷ $y_i(t)$ denotes the value of the time-varying covariate at t

3.2 Time-Varying Covariates (cont'd)

- Interpretation:

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

$\exp(\alpha)$ denotes the relative increase in the risk of an event at time t that results from one unit increase in $y_i(t)$ at the same time point

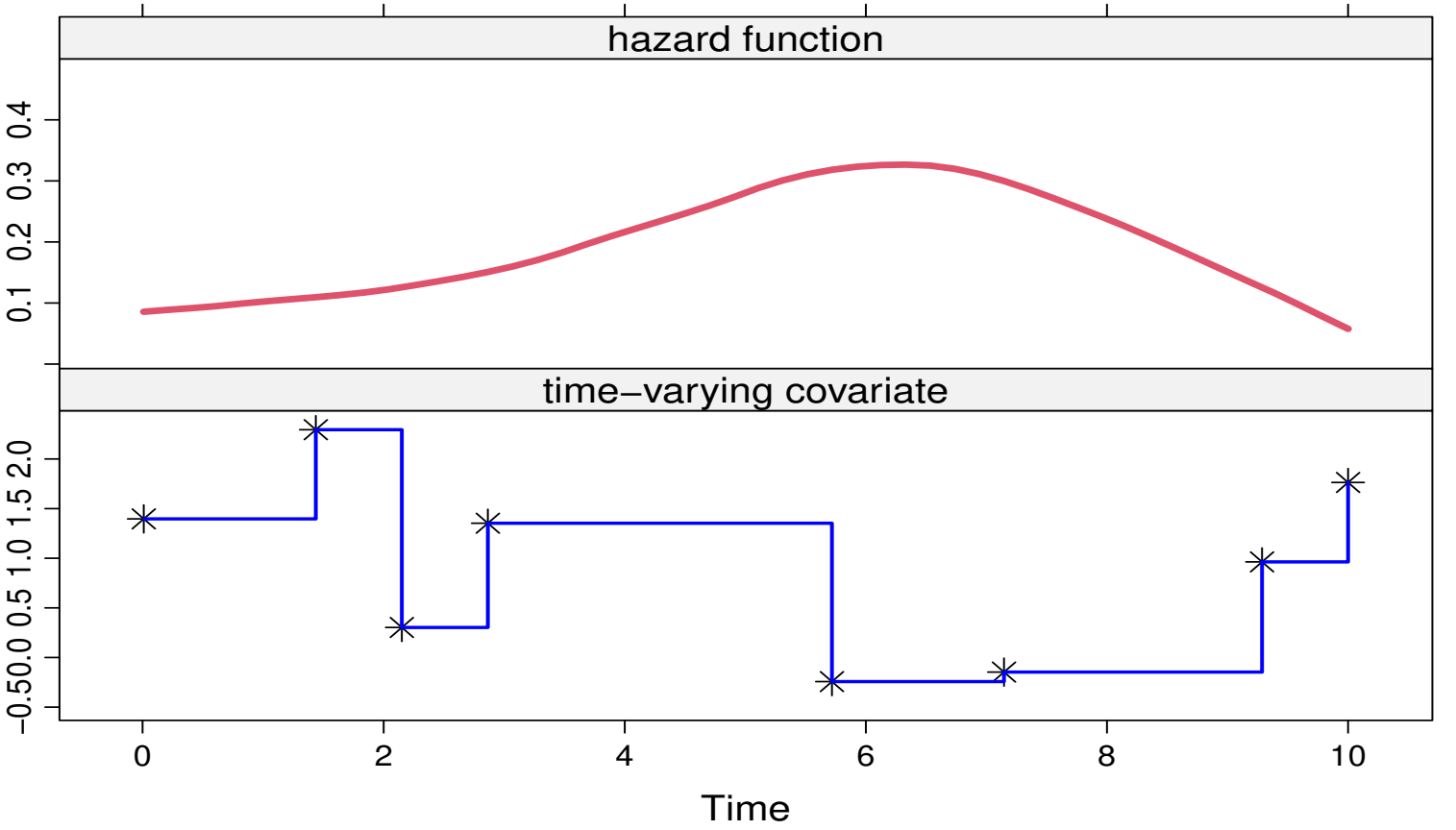
- Parameters are estimated based on the log-partial likelihood function

$$p\ell(\gamma, \alpha) = \sum_{i=1}^{n} \int_0^\infty \Big\{ R_i(t) \exp\{\gamma^\top w_i + \alpha y_i(t)\} - \log\Big[\sum_j R_j(t) \exp\{\gamma^\top w_j + \alpha y_j(t)\}\Big] \Big\} dN_i(t)$$

3.2 Time-Varying Covariates (cont'd)

- How does the extended Cox model handle time-varying covariates?
  - ▷ assumes no measurement error
  - ▷ step-function path
  - ▷ existence of the covariate is not related to failure status

3.2 Time-Varying Covariates (cont'd)



3.2 Time-Varying Covariates (cont'd)

- Therefore, the extended Cox model is only valid for exogenous time-varying covariates

Treating endogenous covariates as exogenous may produce spurious results!

Part IV

The Basic Joint Model

4.1 Joint Modeling Framework

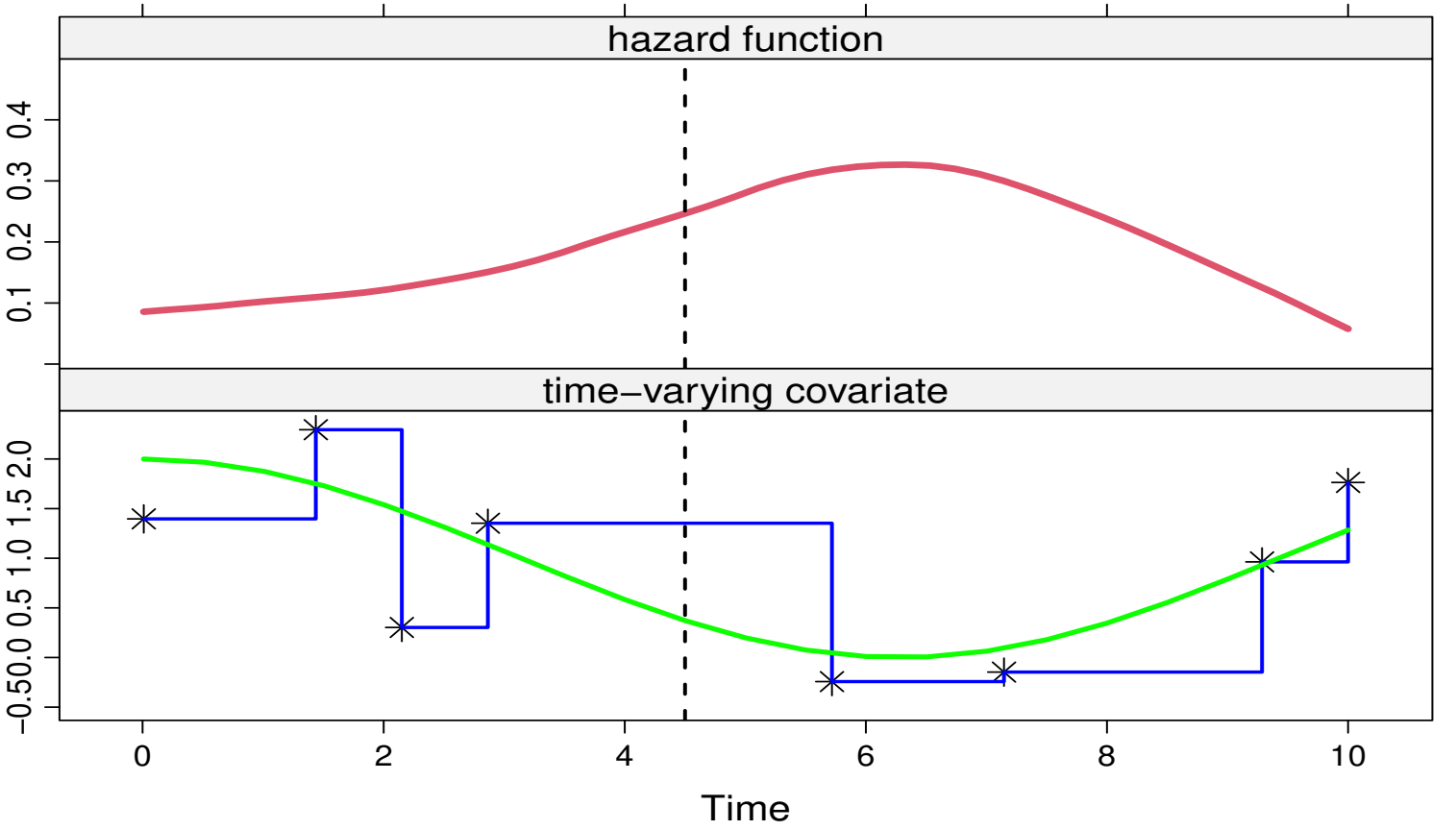
- To account for the special features of endogenous covariates a new class of models has been developed

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

4.1 Joint Modeling Framework (cont'd)



4.1 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, ...

4.1 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

4.1 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, $b_i \sim \mathcal{N}(0, D)$

4.1 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

4.1 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

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

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

Caveat: CI is difficult to test

4.1 Joint Modeling Framework (cont'd)

- The censoring and visiting* processes are assumed non-informative:

- Decision to withdraw from the study or appear for the next visit
  - ▷ **may depend** on observed past history (baseline covariates + observed longitudinal responses)
  - ▷ **no additional dependence** on underlying, latent subject characteristics associated with prognosis

*The visiting process is defined as the mechanism (stochastic or deterministic) that generates the time points at which longitudinal measurements are collected.

4.1 Joint Modeling Framework (cont'd)

- Joint models require a full specification of the joint distribution
  - ▷ **we need an assumption for the baseline hazard**

- General Advice: Use a parametric but flexible model for $h_0(t)$:

$$\log h_0(t) = \gamma_{h_0,0} + \sum_{q=1}^{Q} \gamma_{h_0,q} B_q(t, v),$$

where

- ▷ $B_q(t, v)$ denotes the q-th basis function of a B-spline with knots $v_1, \ldots, v_Q$
- ▷ γ_{h_0} a vector of spline coefficients

4.1 Joint Modeling Framework (cont'd)

- Penalize spline coefficients for smoothness

$$p(\gamma_{h_0} \mid \tau_h) \propto \tau_h^{\rho/2} \exp\left(-\frac{\tau_h}{2}\gamma_{h_0}^\top \Delta_r^\top \Delta_r \gamma_{h_0}\right),$$

where

  - τ_h smoothing parameter
  - Δ_r denotes r-th differences penalty matrix
  - ρ rank of $\Delta_r^\top \Delta_r$

4.2 Bayesian Estimation

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

- Inference is based on the full posterior distribution

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

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

4.2 Bayesian Estimation (cont'd)

- No closed-form solutions for the integrals in the normalizing constant
  $\Rightarrow$ MCMC or Hamiltonian Monte Carlo

- For MCMC estimation, combination of Gibbs and Metropolis-Hastings algorithm
  - ▷ Robbins-Monro adaptive optimal scaling

- To gain in efficiency, we can do block-updating for many of the parameters, i.e.,
  - ▷ fixed effects β
  - ▷ random effects b_i
  - ▷ baseline covariates in the survival submodel γ

4.2 Bayesian Estimation (cont'd)

- Inference then proceeds in the usual manner from the MCMC output, e.g.,
  - ▷ posterior means, variances, and standard errors
  - ▷ credible intervals
  - ▷ ...

4.2 Bayesian Estimation (cont'd)

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

- Two versions available
  - ▷ conditional on the random effects
  - ▷ marginalized over the random effects

Preferable is to work with the marginalized versions

4.3 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

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

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

	JM	Cox
	log HR (std.err)	log HR (std.err)
Treat	0.35 (0.21)	0.31 (0.15)
$CD4^{1/2}$	−0.28 (0.04)	−0.19 (0.02)

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

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

- A unit decrease in $CD4^{1/2}$, results in a
  - ▷ **Joint Model**: 1.32-fold increase in risk (95% CI: 1.23; 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

4.4 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)
```

4.4 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

4.4 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 V

Extensions of Joint Models

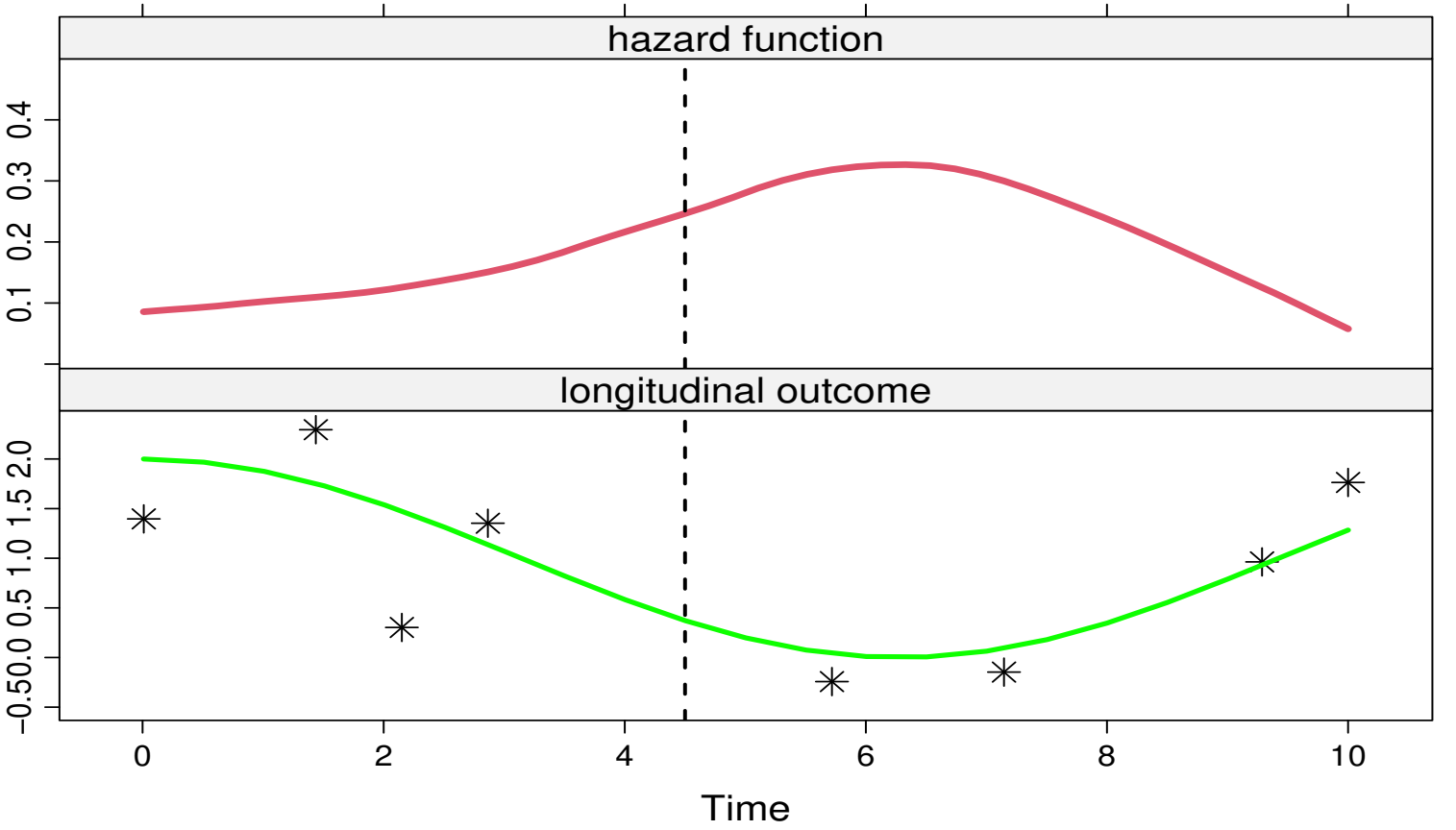
5.1 Functional Forms

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

5.1 Functional Forms (cont'd)



5.1 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?

5.1 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

5.1 Functional Forms (cont'd)

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

- Let's see some possibilities...

5.1 Functional Forms (cont'd)

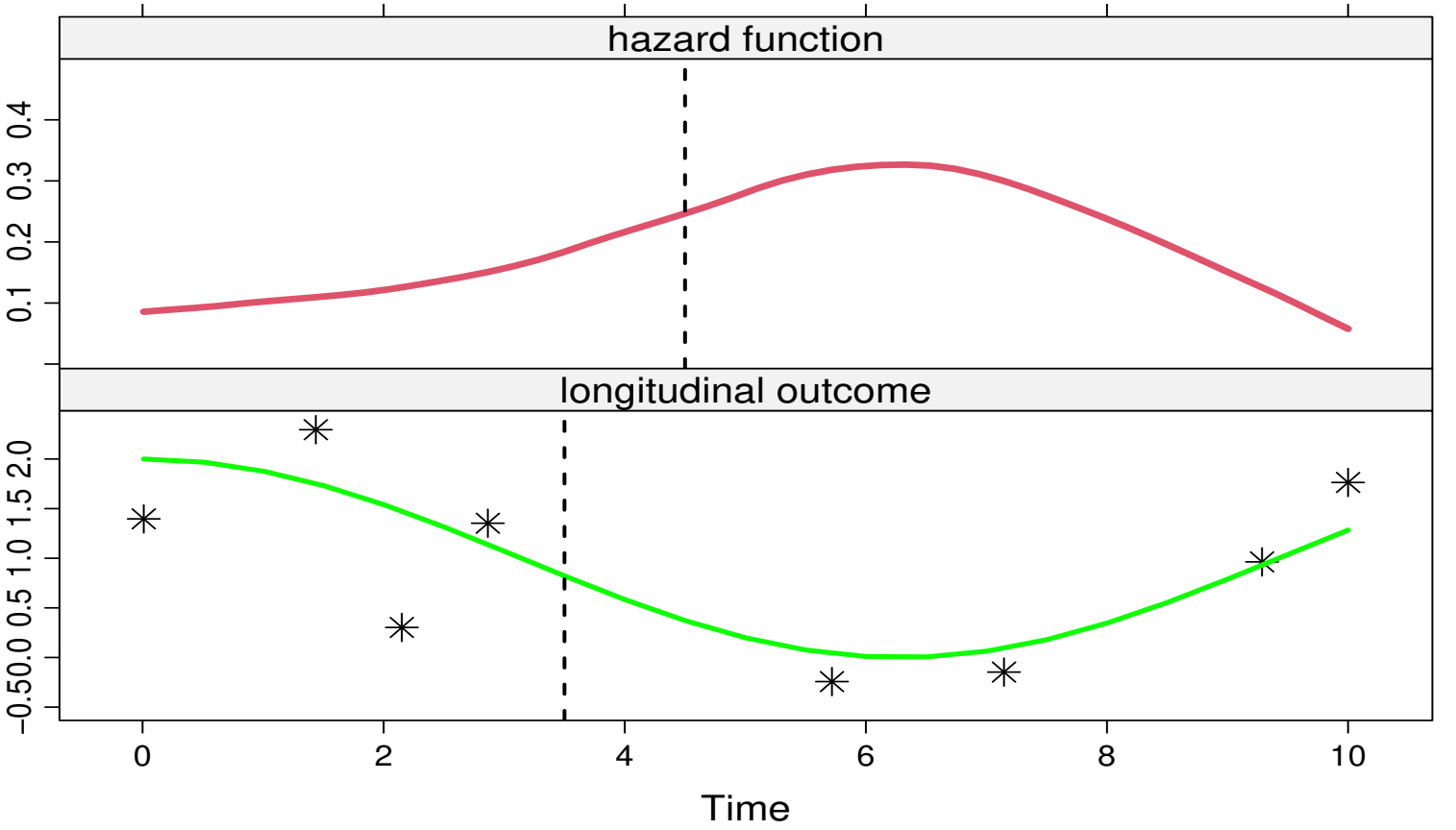
- *Lagged Effects:* The hazard of an event at t is associated with the level of the marker at a previous time point:

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

where

$$t_+^c = \max(t - c, 0)$$

5.1 Functional Forms (cont'd)



5.1 Functional Forms (cont'd)

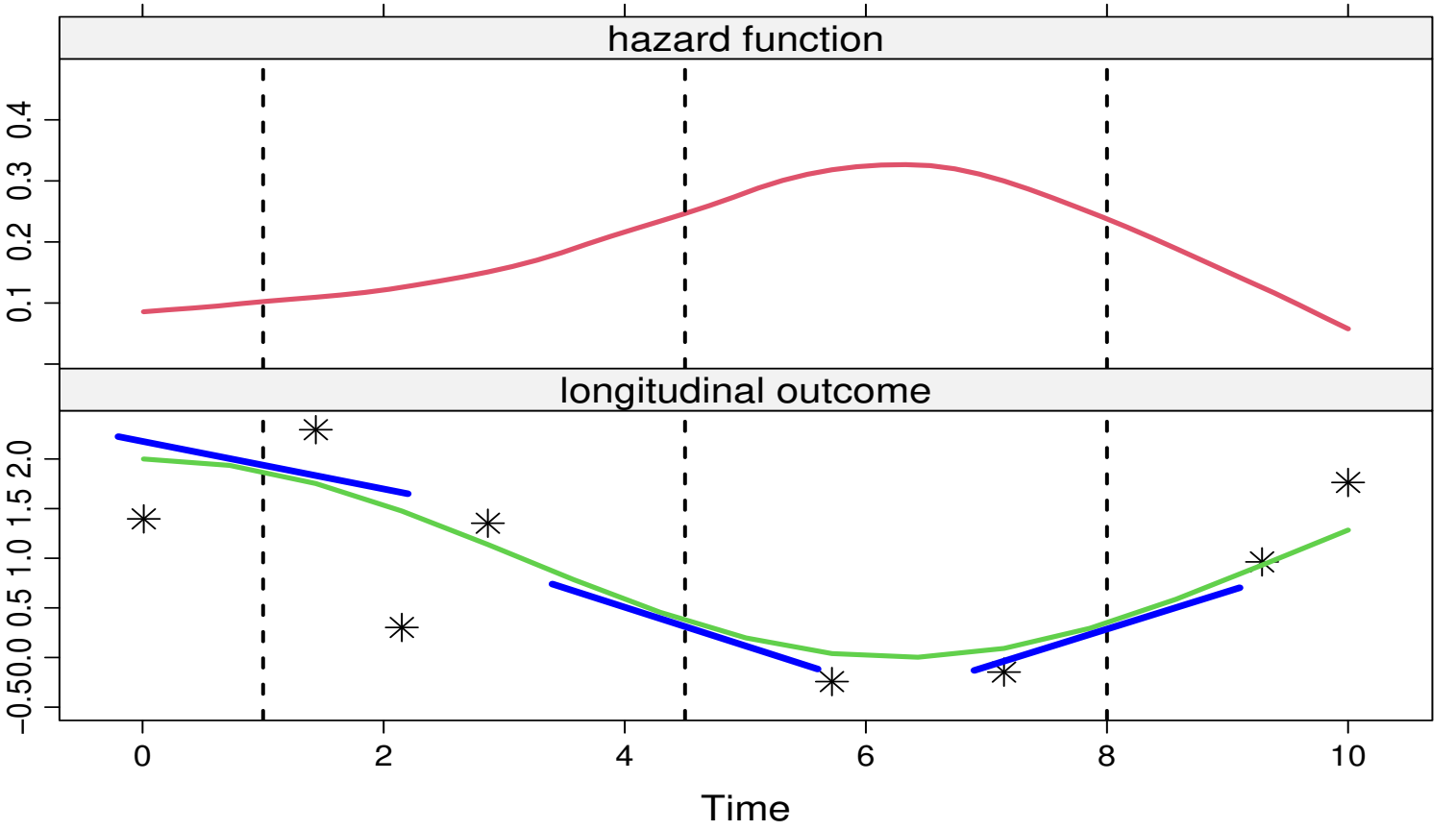
- *Time-dependent Slopes:* The hazard of an event at t is associated with both the current value and the slope of the trajectory at t (Ye et al., 2008, Biometrics):

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

where

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

5.1 Functional Forms (cont'd)



5.1 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

5.1 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

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

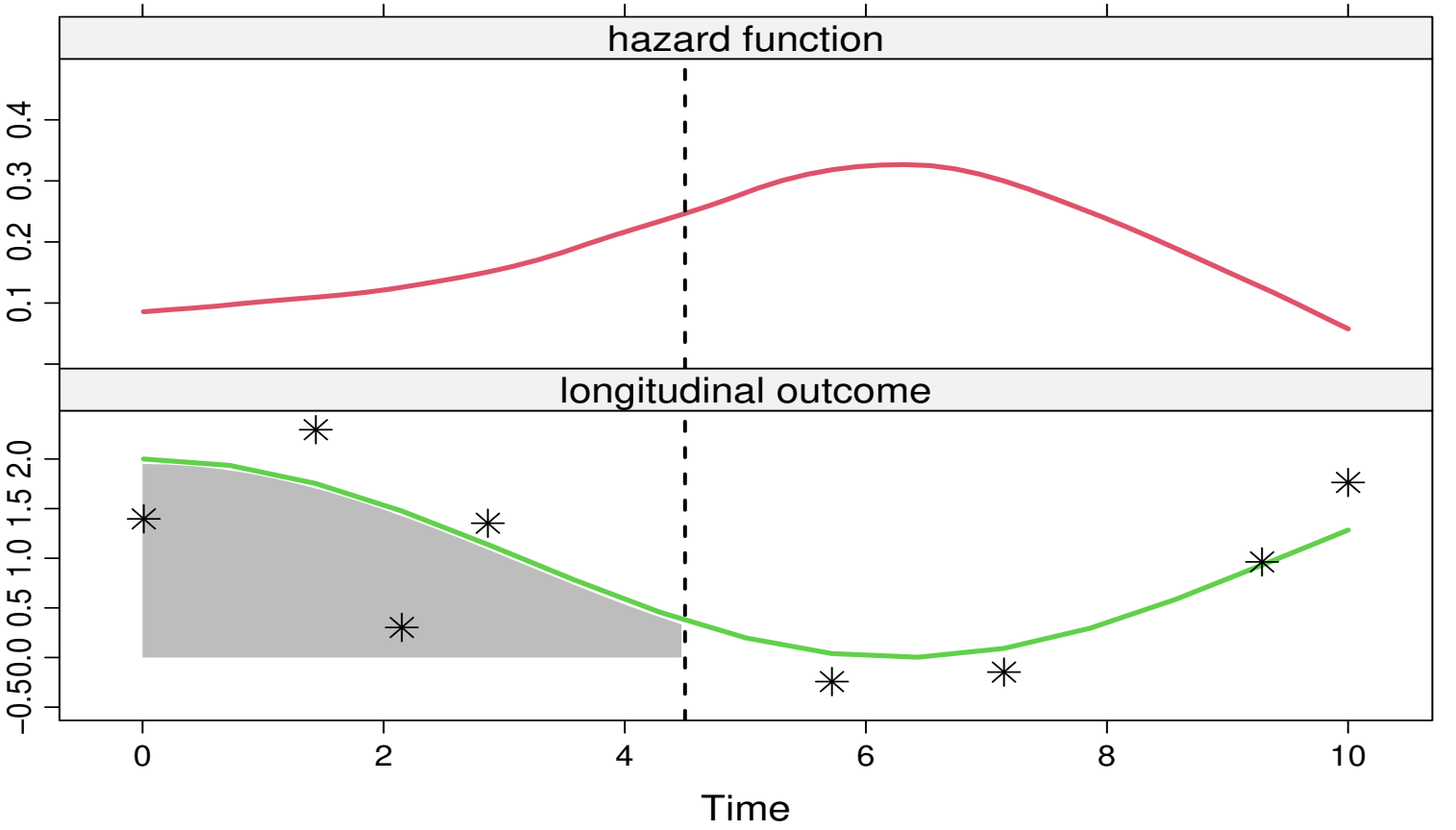
5.1 Functional Forms (cont'd)

- *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 $\mathcal{M}_i(t)$

5.1 Functional Forms (cont'd)



5.1 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

5.1 Functional Forms (cont'd)

- *Weighted Cumulative Effects (convolution):* The hazard of an event at t is associated with the area under the weighted trajectory up to t:

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

where $\varpi(\cdot)$ an appropriately chosen weight function, e.g.,

  - ▷ Gaussian density
  - ▷ Student's-t density
  - ▷ ...

5.1 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)
```

5.1 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")
```

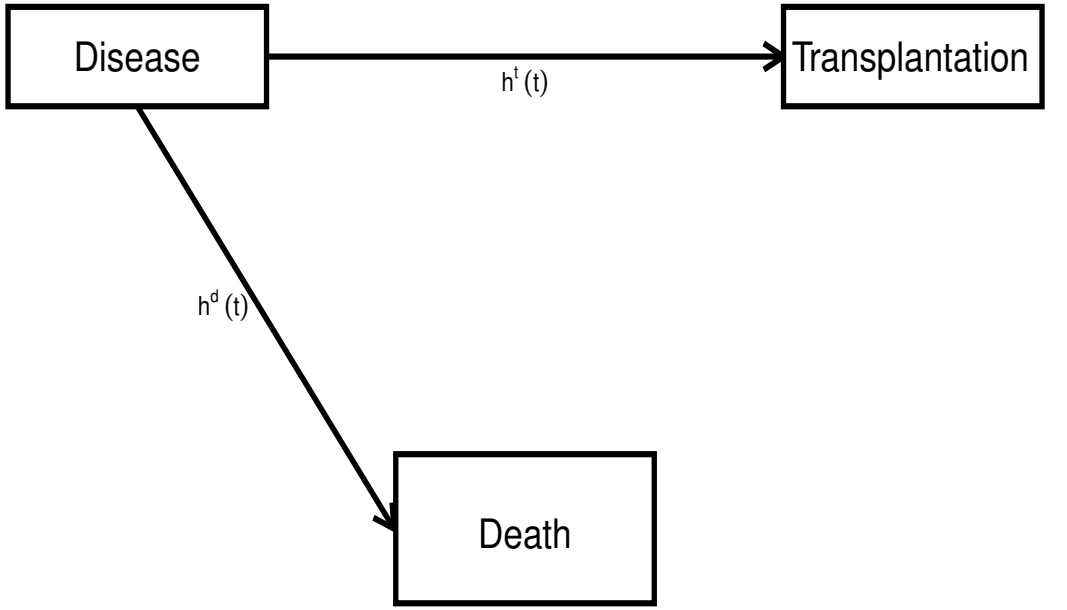
5.2 Multiple Failure Times

- Often multiple failure times are recorded
  - ▷ competing risks
  - ▷ transitions to multiple states
  - ▷ recurrent events

- Example: In the PBC dataset ⇒ 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

5.2 Multiple Failure Times (cont'd)

- Competing risks:
  - ▷ Death precludes the occurrence of transplantation
  - ▷ Transplantation modifies the risk of death



5.2 Multiple Failure Times (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

5.2 Multiple Failure Times (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

5.2 Multiple Failure Times (cont'd)

- This is different than in standard Cox models
  - ▷ i.e., we **cannot** fit a cause-specific hazard joint model by treating events from other causes as censored

5.2 Multiple Failure Times (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

5.2 Multiple Failure Times (cont'd)

	Value	Std.Dev.	2.5%	97.5%
D-penicil	−0.424	0.542	−1.430	0.644
D-penicil:dead	0.516	0.530	−0.515	1.522
value(logSB)	1.134	0.219	0.727	1.552
value(logSB):dead	0.107	0.228	−0.300	0.567

5.2 Multiple Failure Times (cont'd)

R> Function jm() can fit joint models with competing risks and multi-state processes; an example 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
```

5.2 Multiple Failure Times (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
```

5.2 Multiple Failure Times (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)
```

5.2 Multiple Failure Times (cont'd)

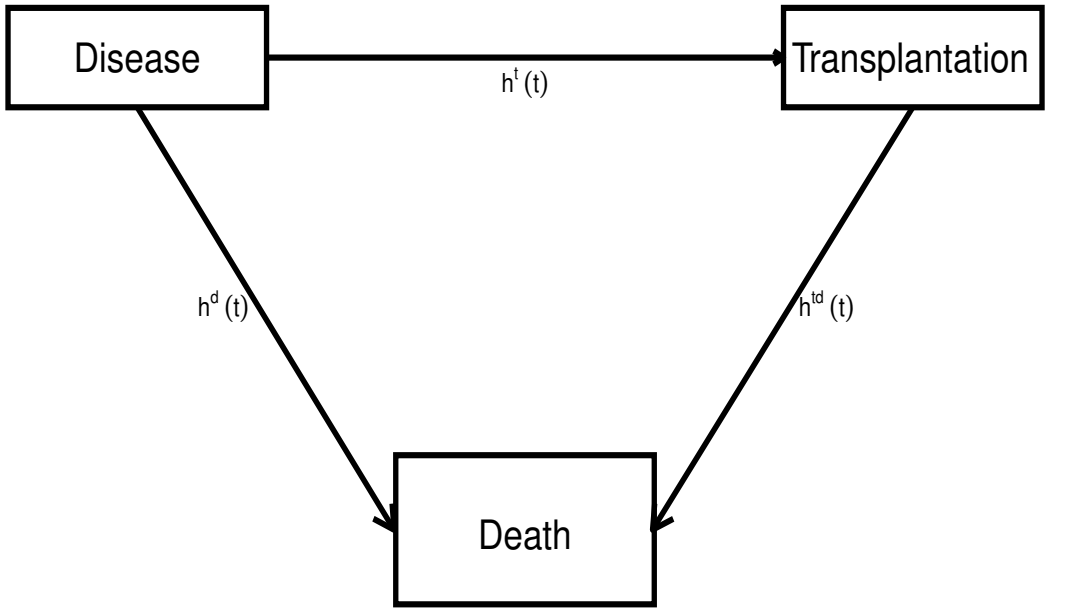
R> Then the joint model is fitted with the code

```
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

5.2 Multiple Failure Times (cont'd)

- Multi-state models:
  - ▷ Transition between transplantation and death is of interest
  - ▷ Effect of covariates and/or biomarkers can be **different for each transition**



5.2 Multiple Failure Times (cont'd)

- Joint models with multi-state processes:

$$\begin{cases} y_i(t) &= m_i(t) + \varepsilon_i(t) = x_i^\top(t)\beta + z_i^\top(t)b_i + \varepsilon_i(t), \\ h_i^d(t) &= h_0^d(t)\exp\left[w_i^{d\top}\gamma_d + \alpha^d m_i(t)\right], \\ h_i^t(t) &= h_0^t(t)\exp\left[w_i^{t\top}\gamma_t + \alpha^t m_i(t)\right], \\ h_i^{td}(t) &= h_0^{td}(t)\exp\left[w_i^{td\top}\gamma_{td} + \alpha^{td} m_i(t)\right], \end{cases}$$

where

- ▷ $h_i^d(t)$ transition intensity from disease to death
- ▷ $h_i^t(t)$ transition intensity from disease to transplantation
- ▷ $h_i^{td}(t)$ transition intensity from transplantation to death

5.2 Multiple Failure Times (cont'd)

Multi-state long-format **different** than the long format in Competing Risks

- General rule: 1 row per **possible** transition.
  - ▷ competing risks: always 2 rows per subject because **both transitions always possible** from starting state.
  - ▷ multi-state: unequal number of rows per subject because **not all transitions possible** from starting state)

5.2 Multiple Failure Times (cont'd)

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

5.2 Multiple Failure Times (cont'd)

- Multiple Failure Times: recurrent events

- Example: In the PBC dataset $\Rightarrow$ recurrent events
  - ▷ Patients showed irregular visiting patterns
  - ▷ So far, when we fitted the joint model we assumed that the visiting process is non-informative
  - ▷ If this assumption is violated, we should also model this process in order to obtain valid inferences

5.2 Multiple Failure Times (cont'd)

- Joint model with recurrent (visiting process) & terminal events

$$\begin{cases} y_i(t) &= m_i(t) + \varepsilon_i(t) = x_i^\top(t)\beta + z_i^\top(t)b_i + \varepsilon_i(t), \\ r_i(t) &= r_0(t)\exp\{\gamma_r^\top w_{ri} + \alpha_r m_i(t) + \mathsf{v}_i\}, \\ h_i(t) &= h_0(t)\exp\{\gamma_h^\top w_{hi} + \alpha_h m_i(t) + \zeta \mathsf{v}_i\}, \end{cases}$$

with

- ▷ $r_i(t)$ hazard function for the recurrent events
- ▷ $h_i(t)$ hazard function for the terminal event
- ▷ v_i frailty term accounting for the correlation in the recurrent events

5.2 Multiple Failure Times (cont'd)

- Conditional independence assumptions augmented
  - ▷ recurrent events are independent given v_i
  - ▷ longitudinal measurements are independent giver b_i
  - ▷ all three processes, namely
    * longitudinal process,
    * recurrent events process, and
    * terminating event process

    are independent given $\{b_i, \mathsf{v}_i\}$

- We need to postulate a distribution for the frailty terms
  - ▷ typical choices are log-Gamma or Gaussian

5.2 Multiple Failure Times (cont'd)

For more info see
https://drizopoulos.github.io/JMbayes2/
→ Articles → Recurrent Events

Part VI

Dynamic Predictions

6.1 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 that will inform them about the future prospect of a patient in order to adjust medical care

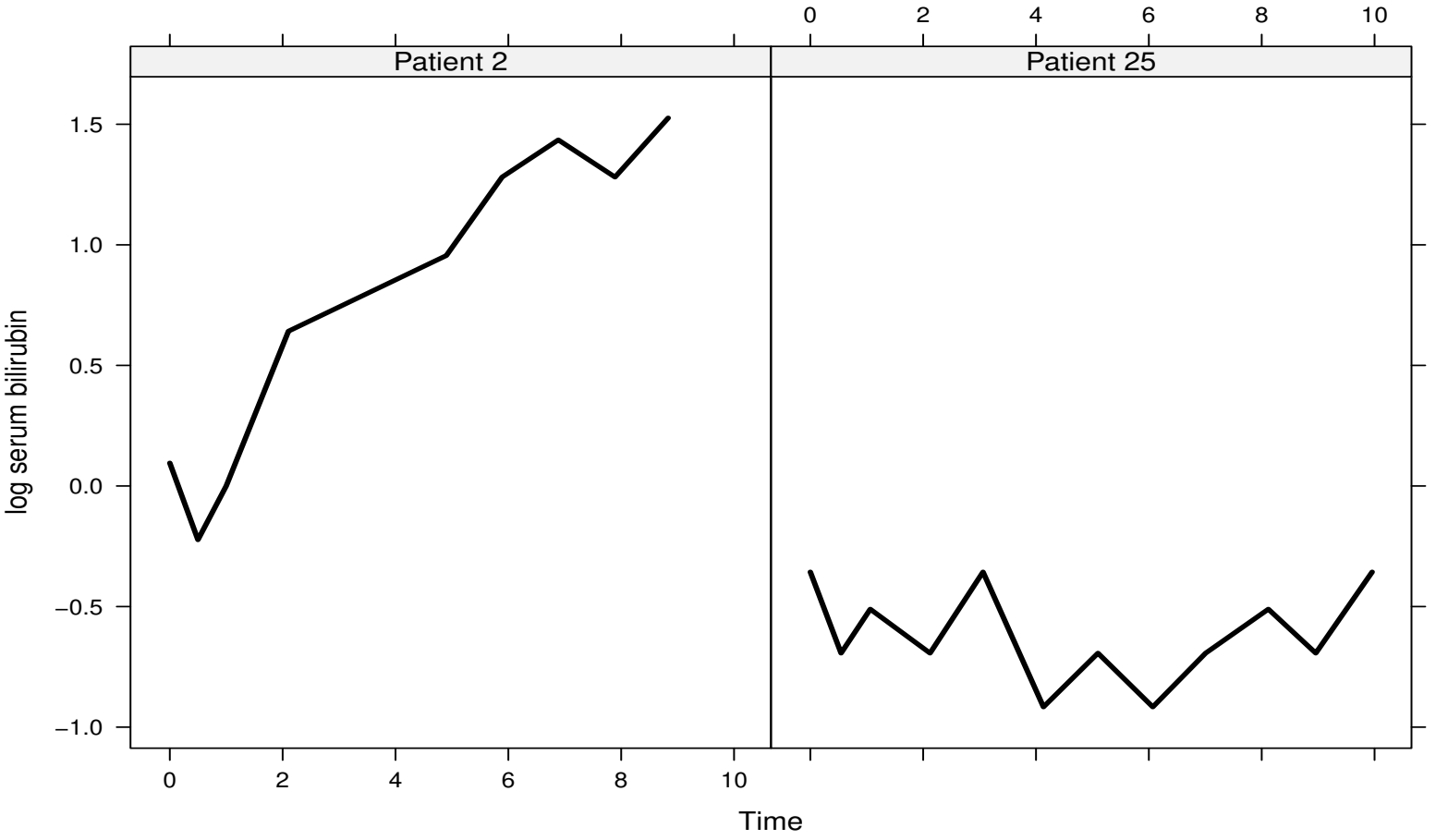
6.1 Survival Probabilities (cont'd)

- We are interested in predicting survival probabilities for a new patient j with serum bilirubin 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

6.1 Survival Probabilities (cont'd)



6.1 Survival Probabilities (cont'd)

- More formally, for a new subject j we have available measurements up to time point t

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

and we are interested in

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

where

▷ where $u > t$, and

▷ $\mathcal{D}_n$ denotes the sample on which the joint model was fitted

6.1 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)

6.1 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

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

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

6.1 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

6.1 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

6.1 Survival Probabilities (cont'd)

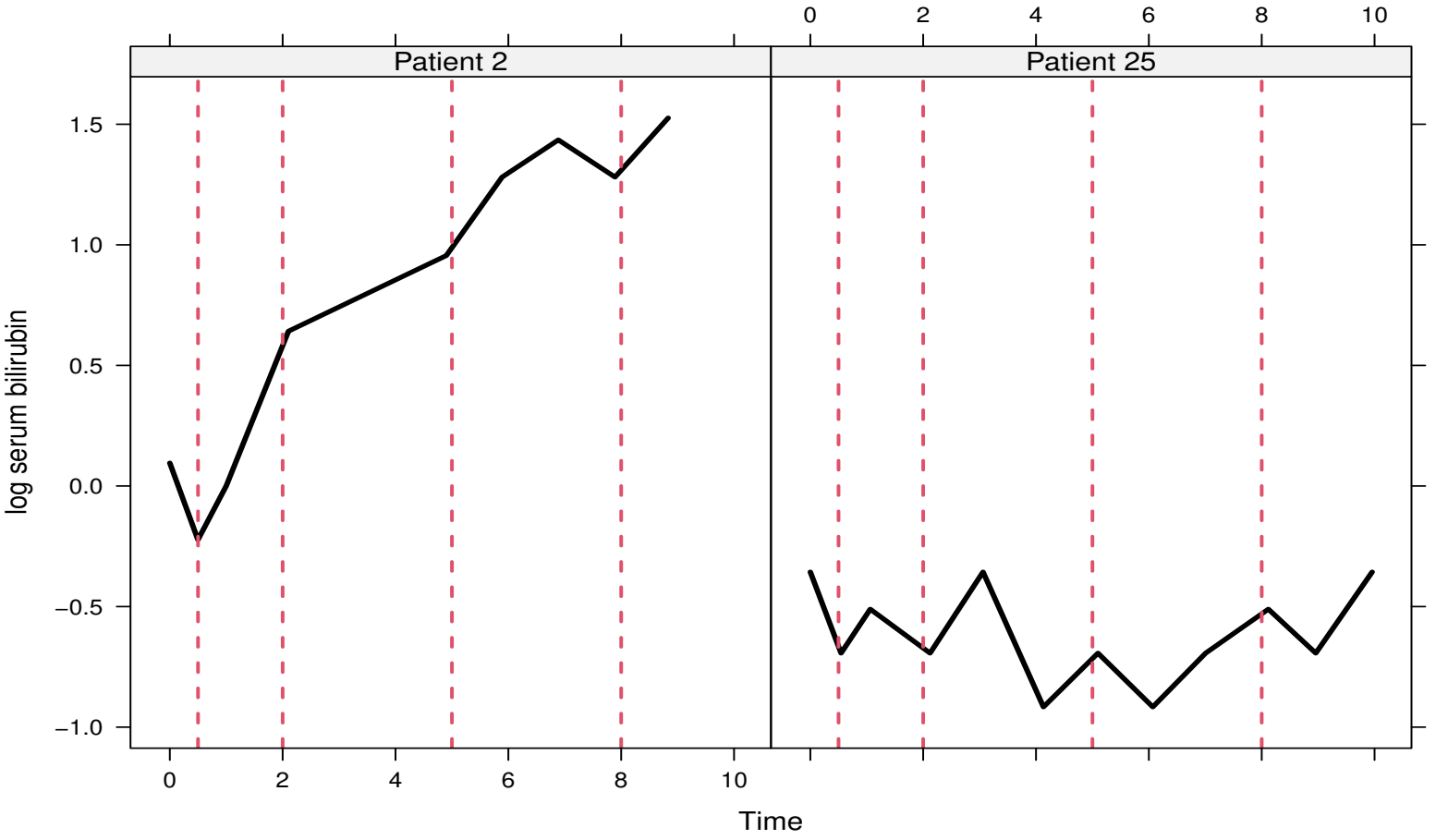
- Based on the fitted joint model we estimate $\pi_j(u \mid t)$ for Patients 2 and 25

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

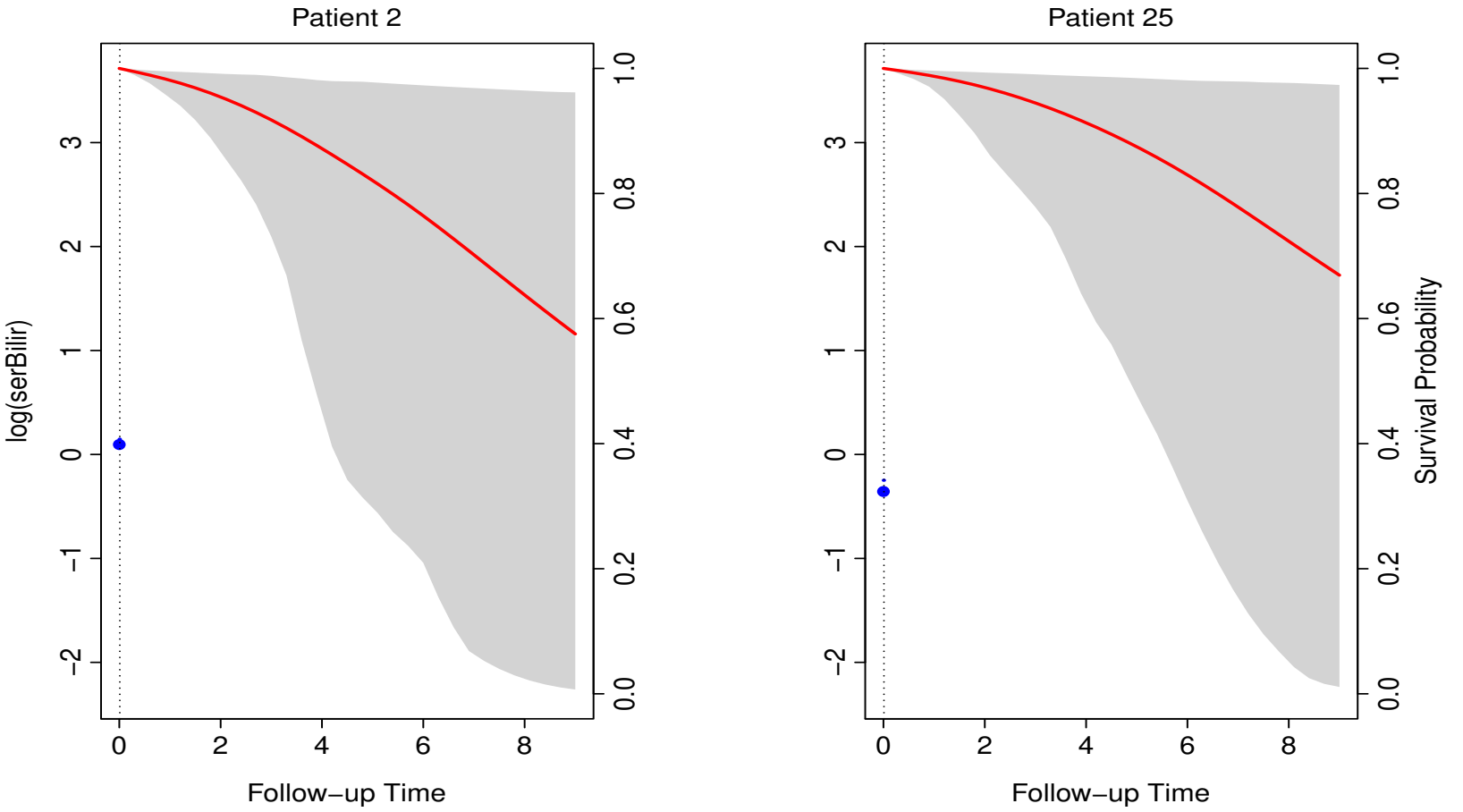
$$\hat{\pi}_j(u \mid t) = \text{mean}\{\pi_j^{(\ell)}(u \mid t), \ell = 1, \ldots, L\}$$

  and calculated a corresponding 95% pointwise CIs

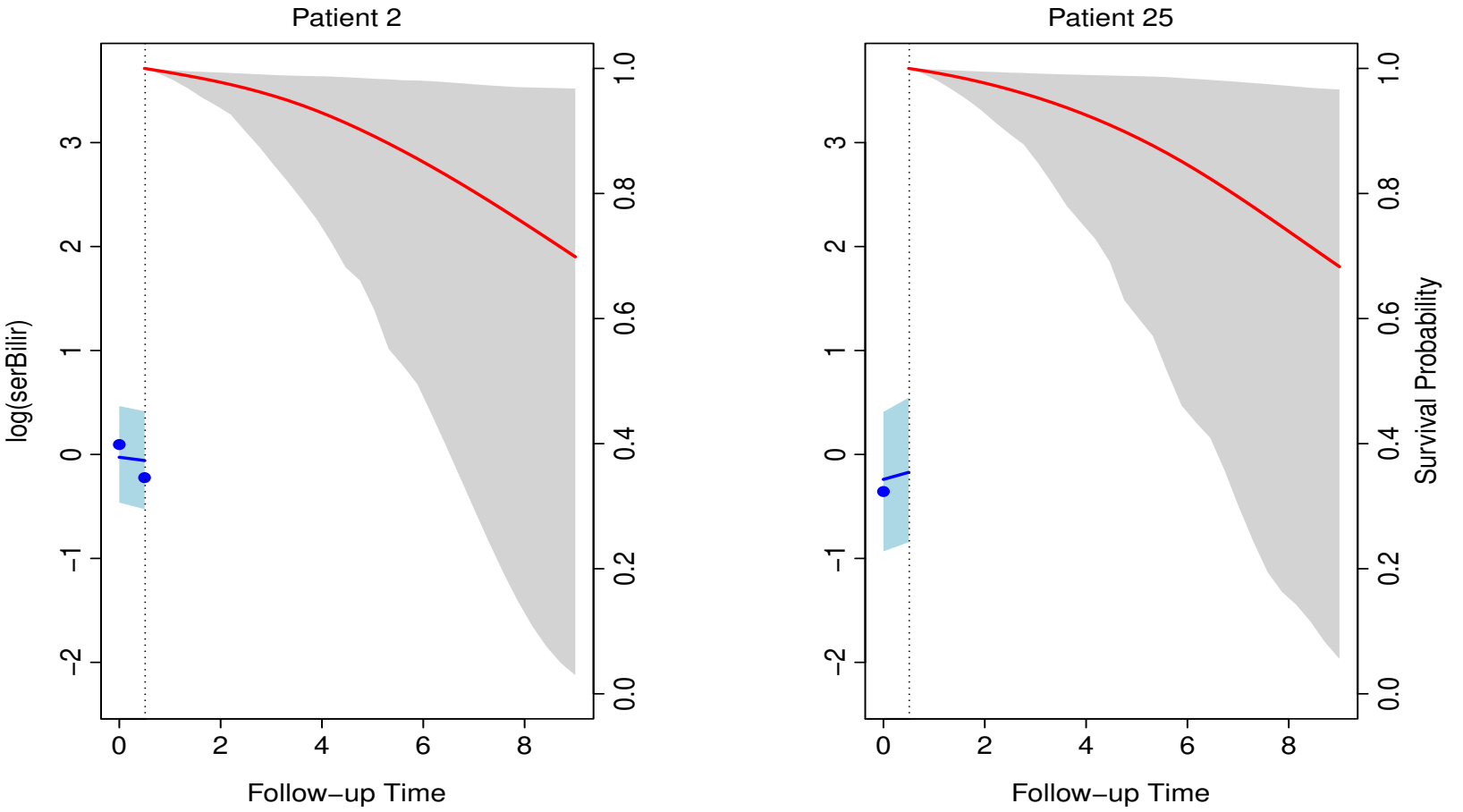
6.1 Survival Probabilities (cont'd)



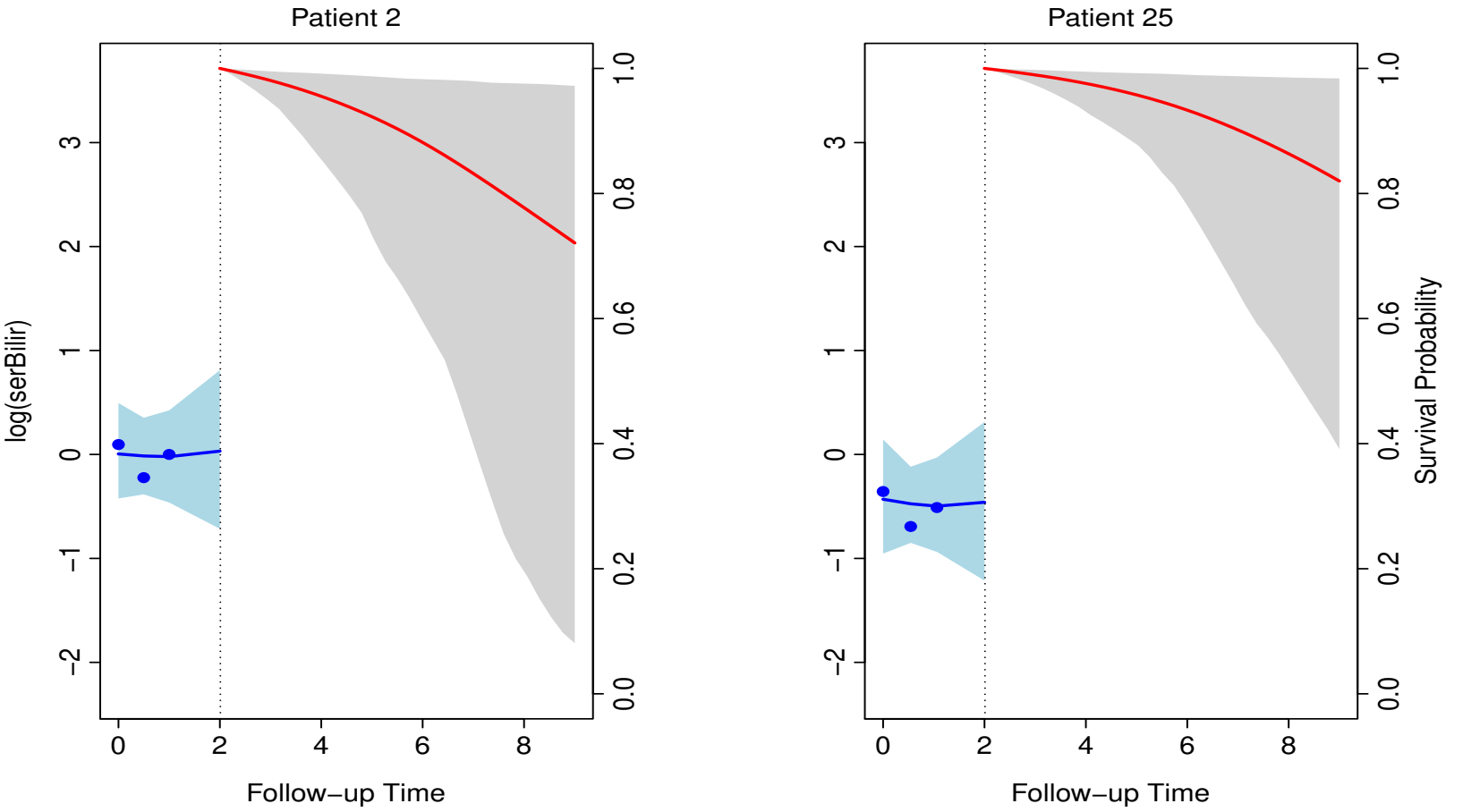
6.1 Survival Probabilities (cont'd)



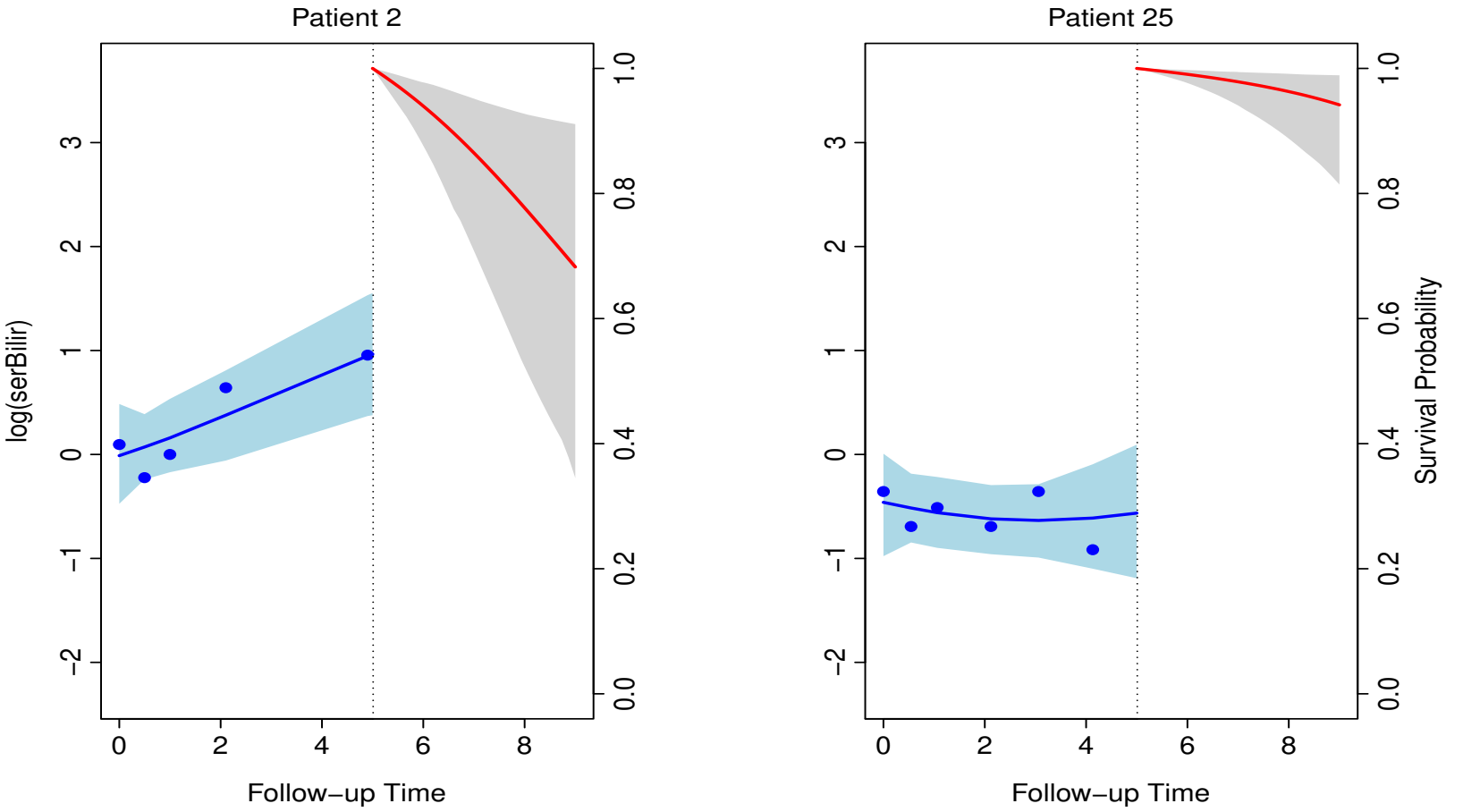
6.1 Survival Probabilities (cont'd)



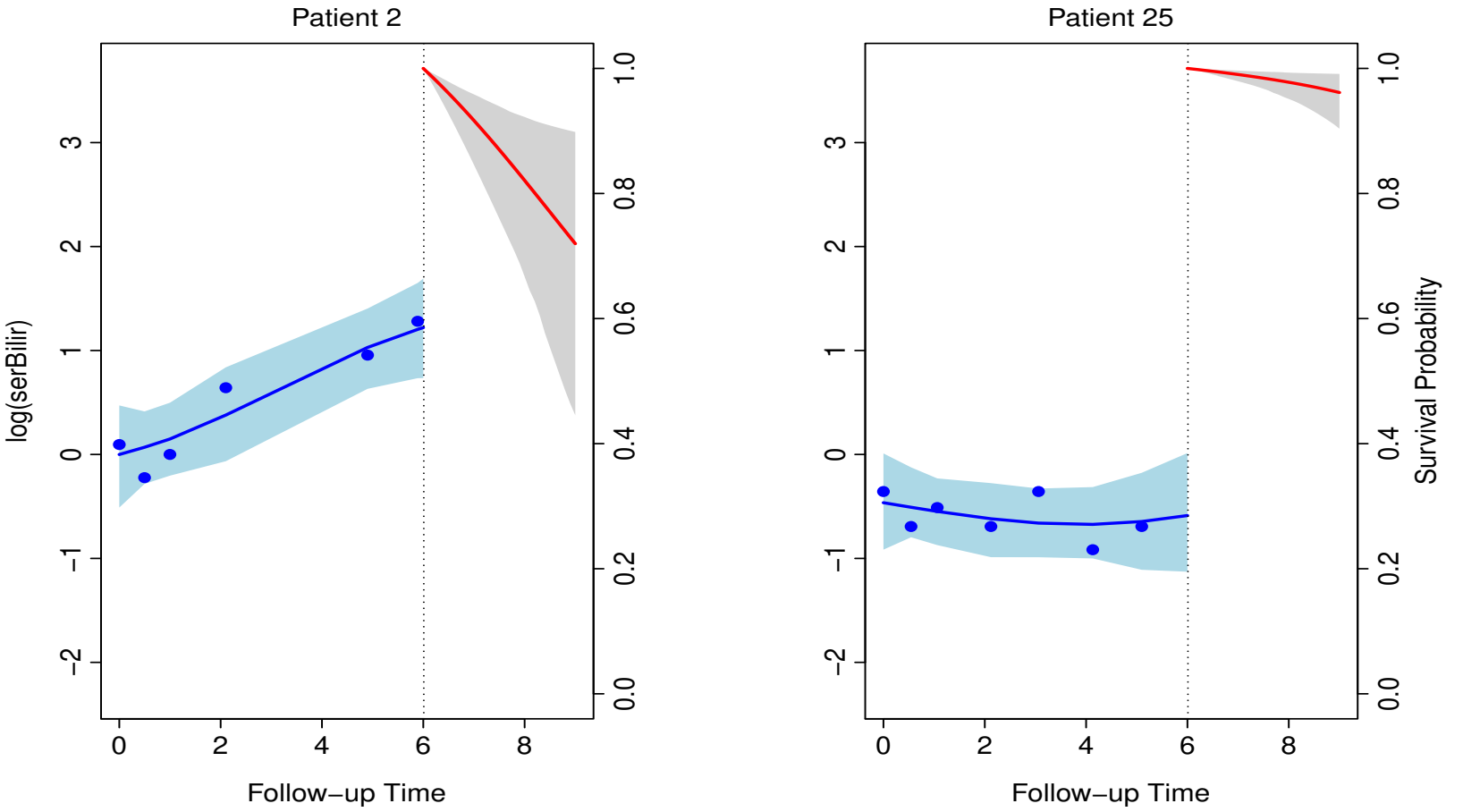
6.1 Survival Probabilities (cont'd)



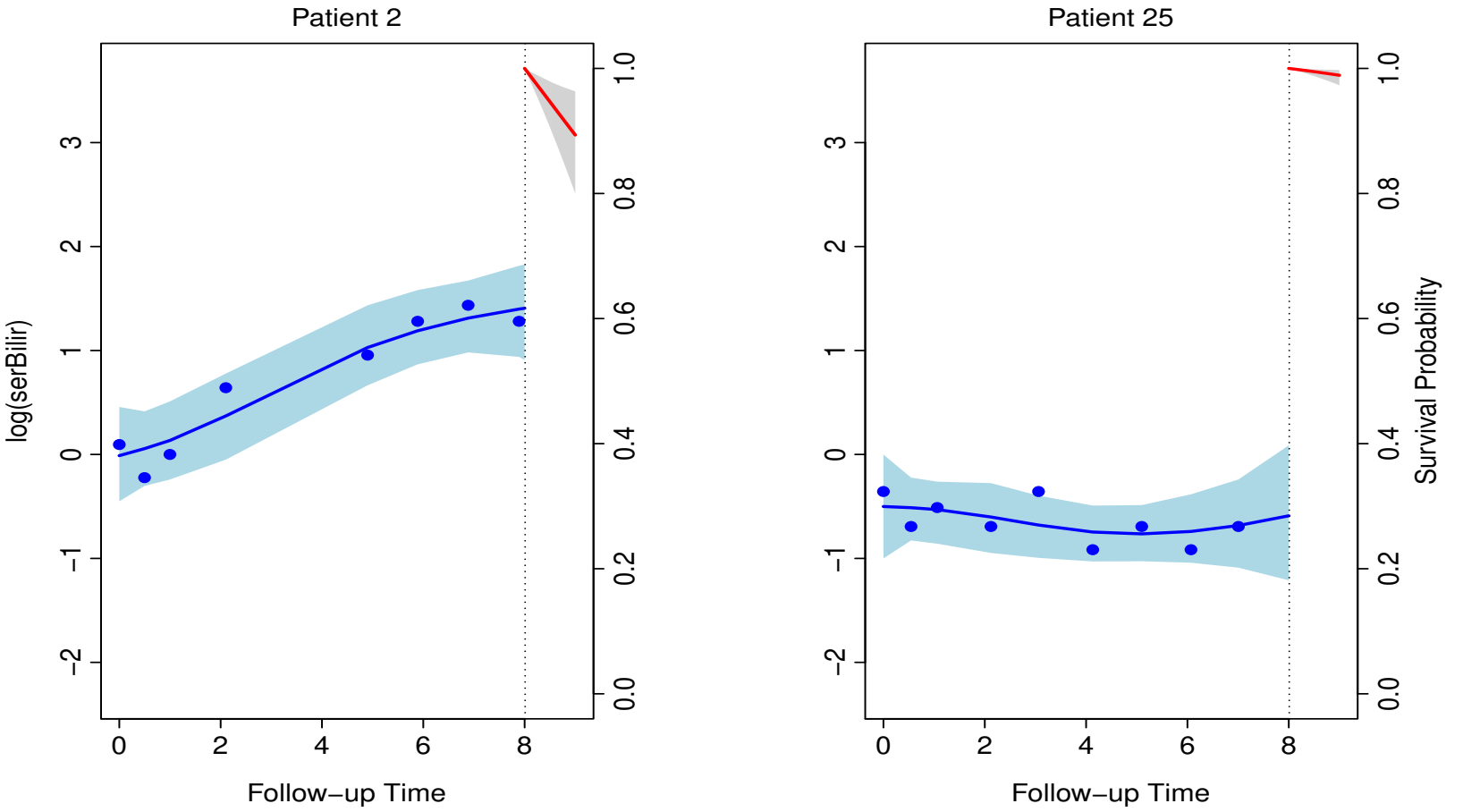
6.1 Survival Probabilities (cont'd)



6.1 Survival Probabilities (cont'd)



6.1 Survival Probabilities (cont'd)



6.1 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)
```

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

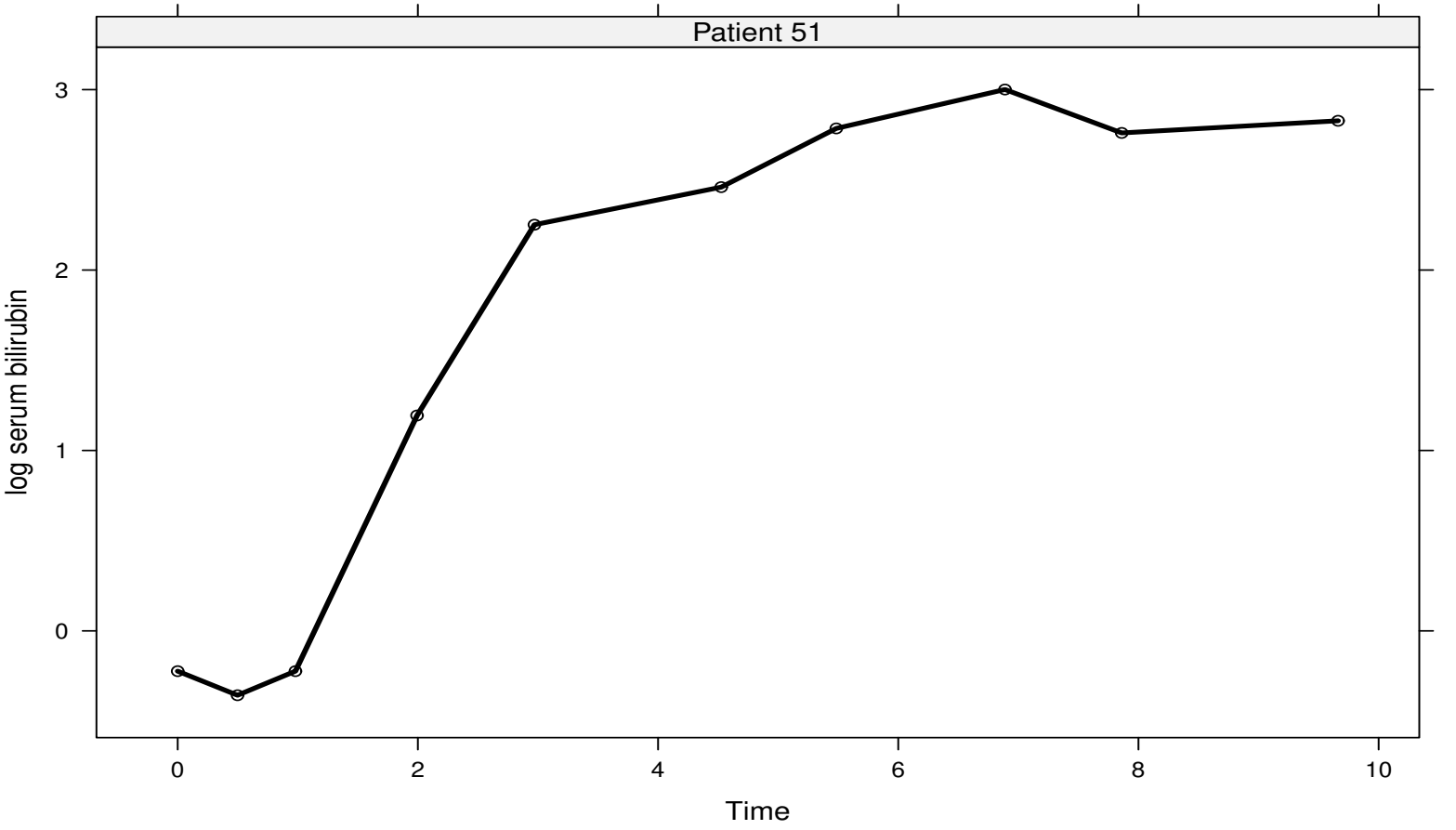
6.2 Functional Forms (cont'd)

- 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

6.2 Functional Forms (cont'd)



6.2 Functional Forms (cont'd)

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

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

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

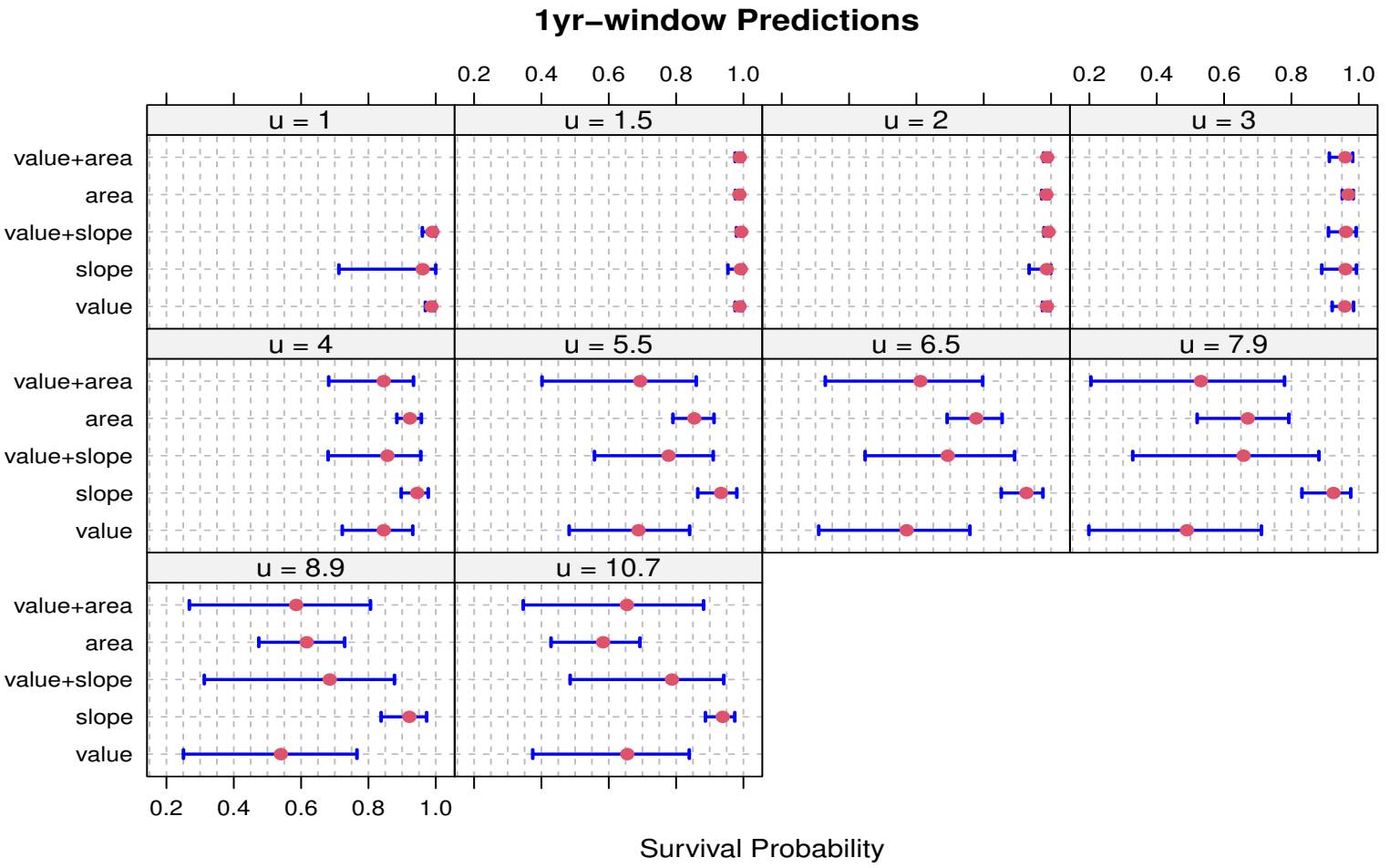
$$h_i(t) = h_0(t)\exp\{\gamma \texttt{D-pnc}_i + \alpha_1 m_i(t) + \alpha_2 m_i'(t)\}$$

6.2 Functional Forms (cont'd)

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

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

6.2 Functional Forms (cont'd)



6.2 Functional Forms (cont'd)

The chosen functional form can influence the derived predictions

6.2 Functional Forms (cont'd)

- We compare the models using the information criteria

	DIC	WAIC	LPML
value + slope	5322.683	22104.998	−5535.420
area	5346.029	23268.436	−5560.009
slope	5645.578	29600.396	−7353.621
value + area	5388.139	29840.361	−9110.958
value	5439.294	30513.206	−7230.238

- The value + slope model seems to be the 'best'

Part VII

Closing

7.1 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

7.1 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

7.1 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!

7.2 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.

7.2 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.

7.2 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.

7.2 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.

7.2 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.

7.2 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.

7.2 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.

7.2 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.

7.3 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.

7.3 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.