

An Introduction to the Joint Modeling of Longitudinal and Survival Data, with Applications in R

Dimitris Rizopoulos

Department of Biostatistics, Erasmus University Medical Center

<http://www.drizopoulos.com/>

ENAR Spring Meeting

March 25, 2018, Atlanta

Contents

1	Introduction	1
1.1	Motivating Longitudinal Studies	2
1.2	Research Questions	10
1.3	Recent Developments	13
1.4	Joint Models	15
2	Linear Mixed-Effects Models	18
2.1	Features of Longitudinal Data	19
2.2	The Linear Mixed Model	21

2.3 Mixed-Effects Models in R	30
2.4 Missing Data in Longitudinal Studies	36
2.5 Missing Data Mechanisms	40

3 Relative Risk Models 47

3.1 Features of Survival Data	48
3.2 Relative Risk Models	51
3.3 Relative Risk Models in R	55
3.4 Time Dependent Covariates	57
3.5 Extended Cox Model	62

4	The Basic Joint Model	68
4.1	Joint Modeling Framework	69
4.2	Estimation	80
4.3	Bayesian Estimation	86
4.4	A Comparison with the TD Cox	90
4.5	Joint Models in R	93
4.6	Connection with Missing Data	99
5	Extensions of Joint Models	108
5.1	Parameterizations	109
5.2	Multiple Longitudinal Markers	129

5.3 Multiple Failure Times	135
5.4 Extensions & Parameterizations	139
6 Dynamic Predictions, Discrimination & Calibration	143
6.1 Survival Probabilities: Definitions	144
6.2 Survival Probabilities: Estimation	148
6.3 Longitudinal Responses: Definitions*	159
6.4 Importance of the Parameterization	167
7 Closing	175
7.1 Concluding Remarks	176

7.2 Additional References	180
7.3 Medical Papers with Joint Modeling	189

Practicals	191
-------------------	------------

Practical 1: A Simple Joint Model	192
Practical 2: Dynamic Predictions	201

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 (e.g., dropout, intermittent missingness)
 - ▷ 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 Survival 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
- The course will be explanatory rather than mathematically rigorous
 - ▷ emphasis is given on sufficient detail in order for participants to obtain a clear view on the different joint modeling approaches, and how they should be used in practice

Agenda

- **Part I:** Introduction

- ▷ Data sets that we will use throughout the course
- ▷ Categorization of possible research questions

- **Part II:** (brief) Review of Linear Mixed Models

- ▷ Features of repeated measurements data
- ▷ Linear mixed models
- ▷ Missing data in longitudinal studies

Agenda (cont'd)

- **Part III:** (brief) Review of Relative Risk Models

- ▷ Features of survival data
- ▷ Relative risk models
- ▷ Time-dependent covariates

- **Part IV:** The Basic Joint Model

- ▷ Definition
- ▷ Estimation & Inference
- ▷ Connection with the missing data framework

Agenda (cont'd)

- **Part V:** Extensions of the Basic Joint Model
 - ▷ Parameterizations
 - ▷ Other extensions for the longitudinal and survival submodels (briefly)

- **Part VI:** Dynamic Predictions
 - ▷ Individualized predictions for the survival
 - ▷ Effect of the parameterization

Structure of the Course & Material

- Lectures & short software practicals using R package **JM** and/or **JMbayes**
- Material (also available in <http://www.drizopoulos.com/>):
 - ▷ Course Notes
 - ▷ R code in soft format
- Within the course notes there are several examples of R code which are denoted by the symbol '**R>**'

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. 180–188.

References (cont'd)

- 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]
 - ▷ <https://eur.academia.edu/DimitrisRizopoulos/Teaching-Documents> [additional R script files]
- Useful material for package **JMbayes**
 - ▷ a paper describing the current capabilities of the package is available on JSS <http://dx.doi.org/10.18637/jss.v072.i07>
- Blog about joint modeling <http://iprogn.blogspot.nl/>

References (cont'd)

- Other software packages capable of fitting joint models
 - ▷ in **R**: **joineR** (by Philipson et al.), **joineRML** (by Hickey et al.), **lcmm** (by Proust-Lima et al.), **bamlss** (by Umlauf 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** (by Crowther)

References (cont'd)

- Standard texts in survival analysis
 - ▷ Kalbfleisch, J. and Prentice, R. (2002). *The Statistical Analysis of Failure Time Data, 2nd Ed.*. New York: Wiley.
 - ▷ Therneau, T. and Grambsch, P. (2000). *Modeling Survival Data: Extending the Cox Model*. New York: Springer-Verlag.
 - ▷ Cox, D. and Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman & Hall.
 - ▷ Andersen, P., Borgan, O., Gill, R. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. New York: Springer-Verlag.
 - ▷ Klein, J. and Moeschberger, M. (2003). *Survival Analysis - Techniques for Censored and Truncated Data*. New York: Springer-Verlag.

References (cont'd)

- Standard texts in longitudinal data analysis
 - ▷ Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer-Verlag.
 - ▷ Molenberghs, G. and Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. New York: Springer-Verlag.
 - ▷ Fitzmaurice, G., Laird, N., and Ware, J. (2004). *Applied Longitudinal Analysis*. Hoboken: Wiley.
 - ▷ Diggle, P., Heagerty, P., Liang, K.-Y., and Zeger, S. (2002). *Analysis of Longitudinal Data*, 2nd edition. New York: Oxford University Press.

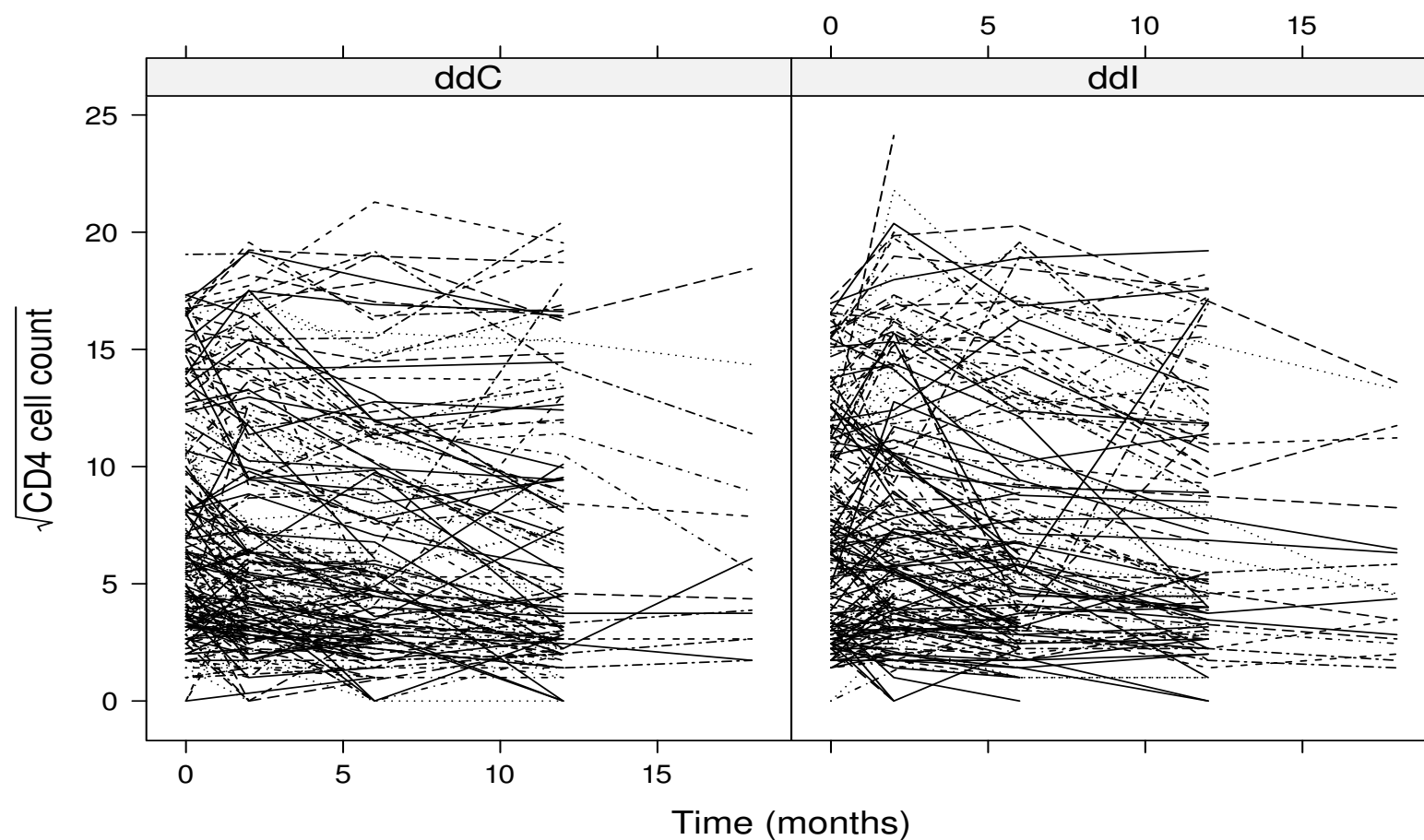
Chapter 1

Introduction

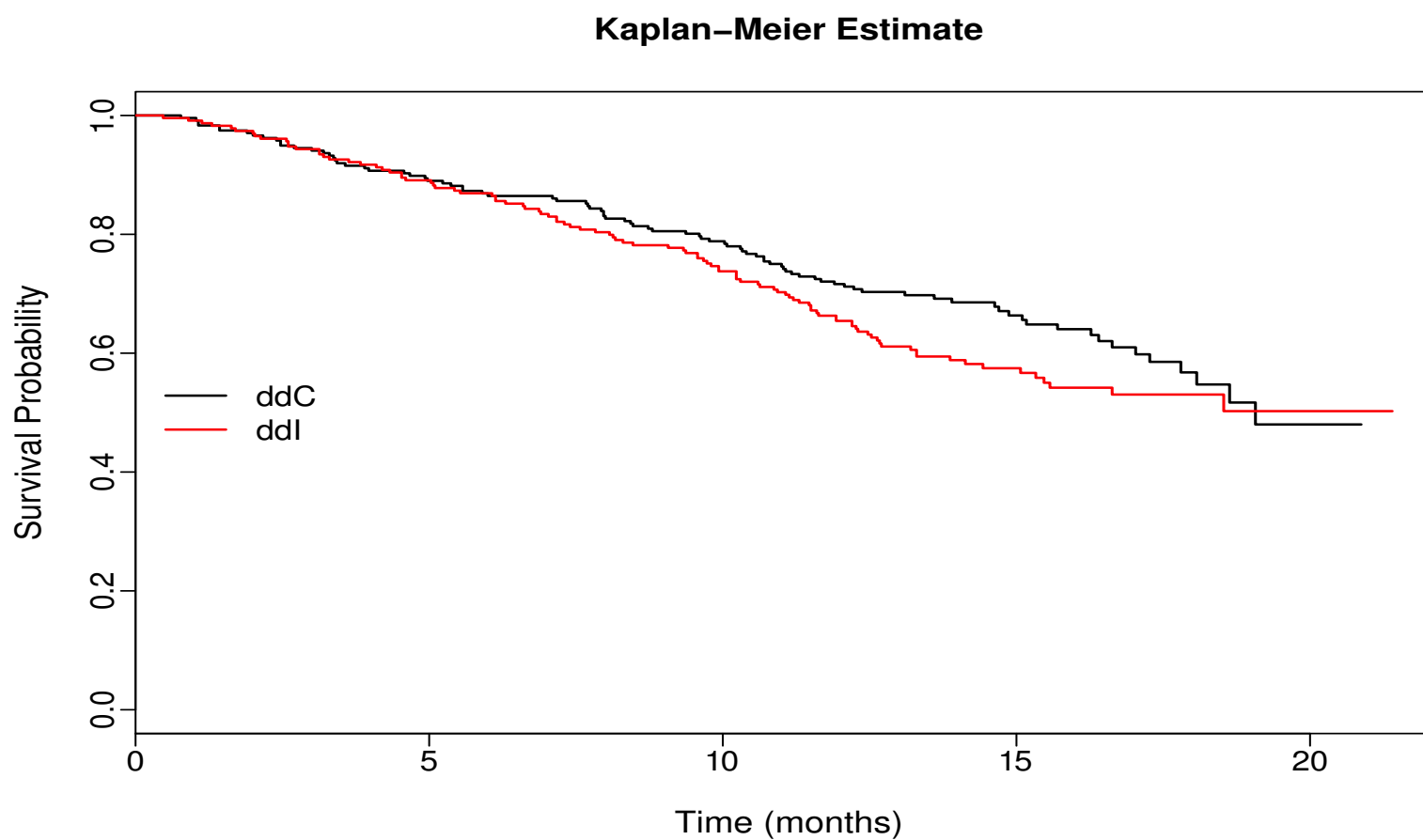
1.1 Motivating Longitudinal Studies

- **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 (ddl) and zalcitabine (ddC)
- Outcomes of interest:
 - ▷ time to death
 - ▷ randomized treatment: 230 patients ddl and 237 ddC
 - ▷ CD4 cell count measurements at baseline, 2, 6, 12 and 18 months
 - ▷ prevOI: previous opportunistic infections

1.1 Motivating Longitudinal Studies (cont'd)



1.1 Motivating Longitudinal Studies (cont'd)



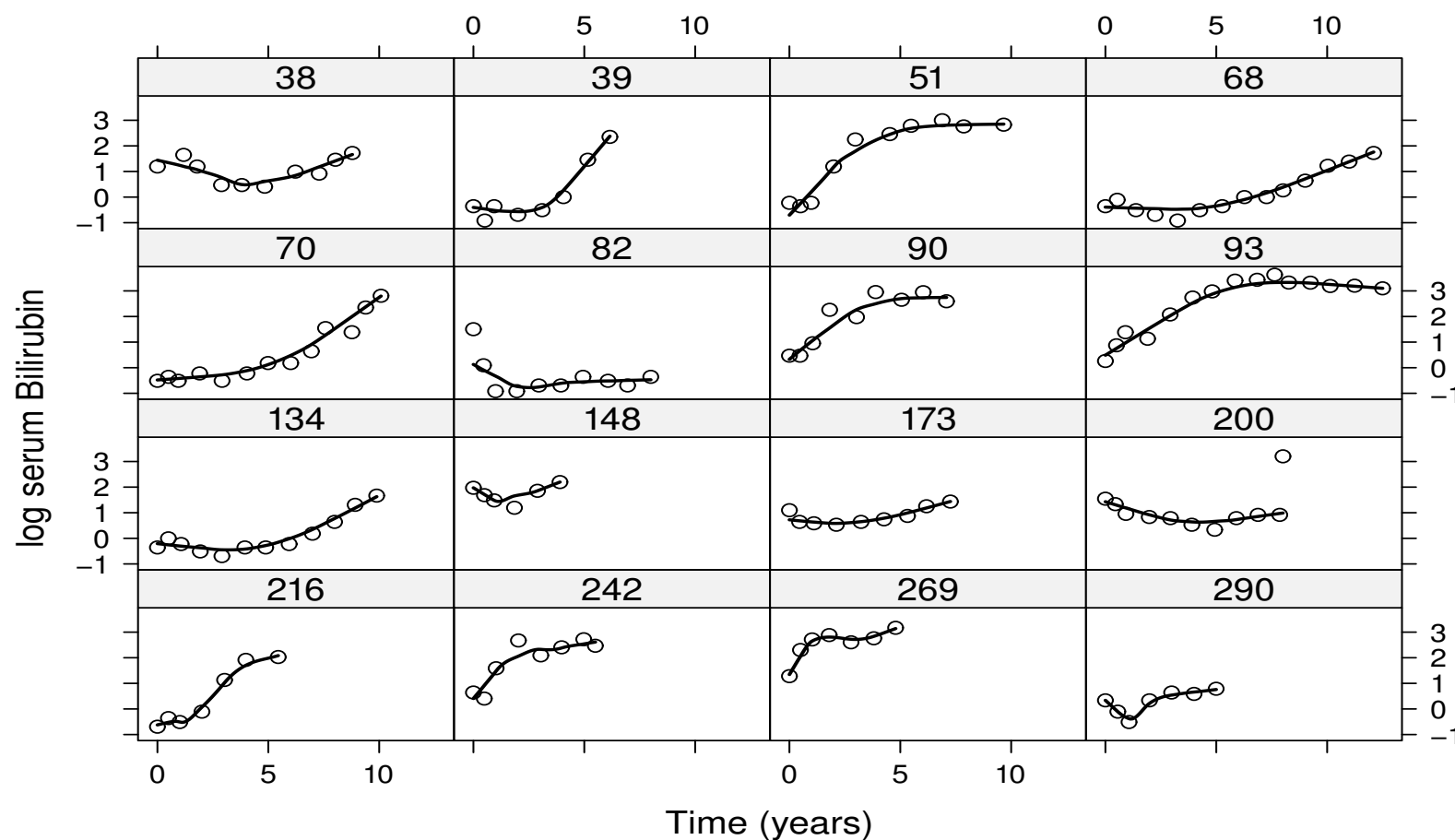
1.1 Motivating Longitudinal Studies (cont'd)

- 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?

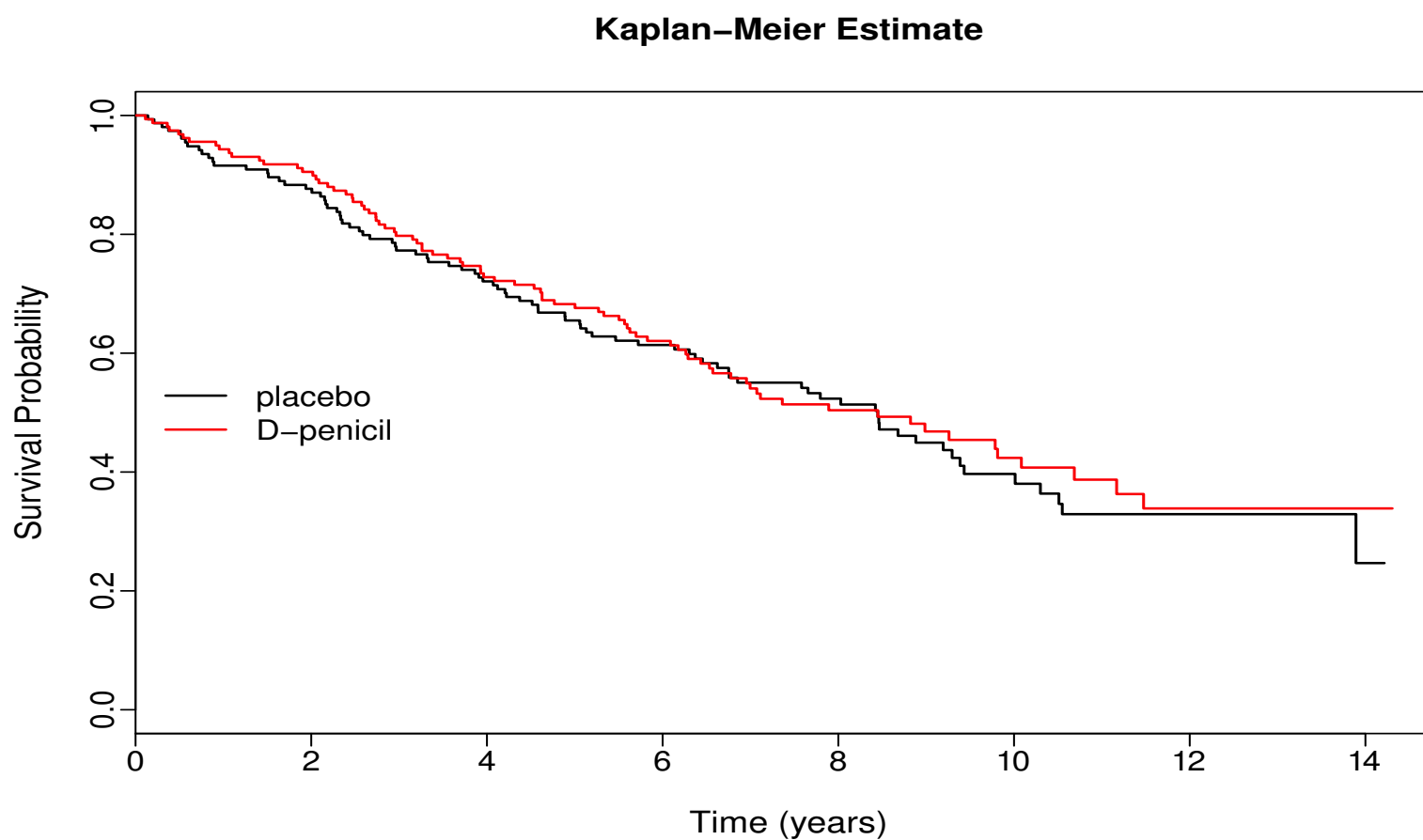
1.1 Motivating Longitudinal Studies (cont'd)

- **PBC:** Primary Biliary Cirrhosis:
 - ▷ a chronic, fatal but rare liver disease
 - ▷ characterized by inflammatory destruction of the small bile ducts within the liver
- Data collected by Mayo Clinic from 1974 to 1984 (Murtaugh et al., Hepatology, 1994)
- Outcomes of interest:
 - ▷ time to death and/or time to liver transplantation
 - ▷ randomized treatment: 158 patients received D-penicillamine and 154 placebo
 - ▷ longitudinal serum bilirubin levels

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 hypothesis testing: does treatment improve the average longitudinal profiles in all markers?
 - ▷ Complex effect estimation: how strong is the association between the longitudinal evolution of CD4 cell counts and the hazard rate of death?
 - ▷ Association structure among outcomes:
 - * how the association between markers evolves over time (evolution of the association)
 - * how marker-specific evolutions are related to each other (association of the evolutions)

1.2 Research Questions (cont'd)

- ▷ Prediction: can we improve prediction for the time to death by considering all markers simultaneously?
- ▷ Handling implicit outcomes: focus on a single longitudinal outcome but with dropout or random visit times

1.3 Recent Developments

- Up to now emphasis has been
 - ▷ *restricted* or *coerced* to separate analysis per outcome
 - ▷ or given to naive types of joint analysis (e.g., last observation carried forward)
- Main reasons
 - ▷ lack of appropriate statistical methodology
 - ▷ lack of efficient computational approaches & software

1.3 Recent Developments (cont'd)

- However, recently there has been an explosion in the statistics and biostatistics literature of joint modeling approaches
- Many different approaches have been proposed that
 - ▷ can handle different types of outcomes
 - ▷ can be utilized in pragmatic computing time
 - ▷ can be rather flexible
 - ▷ **most importantly:** can answer the questions of interest

1.4 Joint Models

- Let Y_1 and Y_2 two outcomes of interest measured on a number of subjects for which joint modeling is of scientific interest
 - ▷ both can be measured longitudinally
 - ▷ one longitudinal and one survival
- We have various possible approaches to construct a joint density $p(y_1, y_2)$ of $\{Y_1, Y_2\}$
 - ▷ Conditional models: $p(y_1, y_2) = p(y_1)p(y_2 \mid y_1)$
 - ▷ Copulas: $p(y_1, y_2) = c\{\mathcal{F}(y_1), \mathcal{F}(y_2)\}p(y_1)p(y_2)$

But **Random Effects Models** have (more or less) prevailed

1.4 Joint Models (cont'd)

- Random Effects Models specify

$$\begin{aligned} p(y_1, y_2) &= \int p(y_1, y_2 \mid b) p(b) db \\ &= \int p(y_1 \mid b) p(y_2 \mid b) p(b) db \end{aligned}$$

- ▷ Unobserved random effects b explain the association between Y_1 and Y_2
- ▷ Conditional Independence assumption

$$Y_1 \perp\!\!\!\perp Y_2 \mid b$$

1.4 Joint Models (cont'd)

- Features:
 - ▷ Y_1 and Y_2 can be of different type
 - * one continuous and one categorical
 - * one continuous and one survival
 - * ...
 - ▷ Extensions to more than two outcomes straightforward
 - ▷ Specific association structure between Y_1 and Y_2 is assumed
 - ▷ Computationally intensive (especially in high dimensions)

Chapter 2

Linear Mixed-Effects Models

2.1 Features of Longitudinal Data

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

- Longitudinal studies allow to investigate
 1. how treatment means differ at specific time points, e.g., at the end of the study (*cross-sectional effect*)
 2. how treatment means or differences between means of treatments change over time (*longitudinal effect*)

2.1 Features of Longitudinal Data (cont'd)

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.

2.2 The Linear Mixed Model

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

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

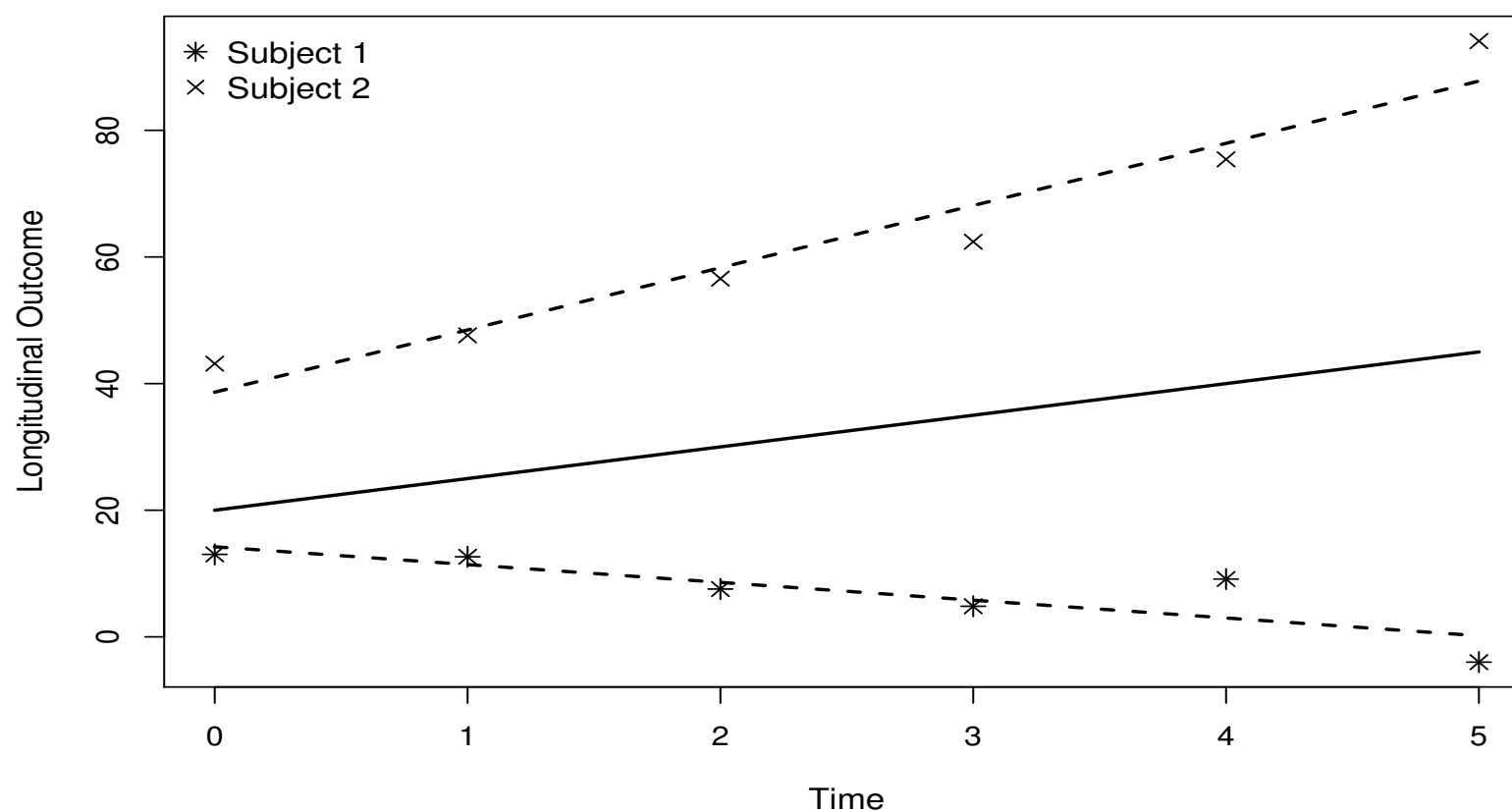
where

- ▷ y_i the vector of responses for the i th subject
 - ▷ X_i design matrix describing structural component
 - ▷ 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, . . .

2.2 The Linear Mixed Model (cont'd)

- **Alternative intuitive approach:** Each subject in the population has her own subject-specific mean response profile over time

2.2 The Linear Mixed Model (cont'd)



2.2 The Linear Mixed Model (cont'd)

- The evolution of each subject in 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 j th response of the i th 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

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

2.2 The Linear Mixed Model (cont'd)

- We can reformulate the model as

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

where

- ▷ β s are known as the *fixed effects*
 - ▷ b_i s are known as the *random effects*
- In accordance for the random effects we assume

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

2.2 The Linear Mixed Model (cont'd)

- Put in a general form

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

with

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

2.2 The Linear Mixed Model (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 i th 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 The Linear Mixed Model (cont'd)

- Example: We fit a linear mixed model for the AIDS dataset assuming
 - ▷ different average longitudinal evolutions per treatment group (**fixed part**)
 - ▷ random intercepts & random slopes (**random part**)

$$\left\{ \begin{array}{l} y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 \{\text{ddI}_i \times t_{ij}\} + b_{i0} + b_{i1} t_{ij} + \varepsilon_{ij}, \\ b_i \sim \mathcal{N}(0, D), \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2) \end{array} \right.$$

- Note: We did not include a main effect for treatment due to randomization

2.2 The Linear Mixed Model (cont'd)

	Value	Std.Err.	<i>t</i> -value	<i>p</i> -value
β_0	7.189	0.222	32.359	< 0.001
β_1	-0.163	0.021	-7.855	< 0.001
β_2	0.028	0.030	0.952	0.342

- No evidence of differences in the average longitudinal evolutions between the two treatments

2.3 Mixed-Effects Models in R

R> There are two primary packages in R for mixed models analysis:

▷ Package **nlme**

- * fits linear & nonlinear mixed effects models, and marginal models for normal data
- * allows for both random effects & correlated error terms
- * several options for covariances matrices and variance functions

▷ Package **lme4**

- * fits linear, nonlinear & generalized mixed effects models
- * uses only random effects
- * allows for nested and crossed random-effects designs

2.3 Mixed-Effects Models in R (cont'd)

- R> We will only use package **nlme** because package **JM** accepts as an argument a linear mixed model fitted by **nlme**
- R> The basic function to fit linear mixed models is `lme()` and has three basic arguments
- ▷ `fixed`: a formula specifying the response vector and the fixed-effects structure
 - ▷ `random`: a formula specifying the random-effects structure
 - ▷ `data`: a data frame containing all the variables

2.3 Mixed-Effects Models in R (cont'd)

R> The data frame that contains all variables should be in the *long format*

Subject	y	time	gender	age
1	5.1	0.0	male	45
1	6.3	1.1	male	45
2	5.9	0.1	female	38
2	6.9	0.9	female	38
2	7.1	1.2	female	38
2	7.3	1.5	female	38
⋮	⋮	⋮	⋮	⋮

2.3 Mixed-Effects Models in R (cont'd)

R> Using formulas in R

▷ CD4 = Time + Gender

⇒ `cd4 ~ time + gender`

▷ CD4 = Time + Gender + Time*Gender

⇒ `cd4 ~ time + gender + time:gender`

⇒ `cd4 ~ time*gender` (the same)

▷ CD4 = Time + Time²

⇒ `cd4 ~ time + I(time^2)`

R> Note: the intercept term is included by default

2.3 Mixed-Effects Models in R (cont'd)

R> The code used to fit the linear mixed model for the AIDS dataset (p. 28) is as follows

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

```
summary(lmeFit)
```

2.3 Mixed-Effects Models in R (cont'd)

R> The same fixed-effects structure but only random intercepts

```
lme(CD4 ~ obstime + obstime:drug, data = aids,  
    random = ~ 1 | patient)
```

R> The same fixed-effects structure, random intercepts & random slopes, with a diagonal covariance matrix (using the `pdDiag()` function)

```
lme(CD4 ~ obstime + obstime:drug, data = aids,  
    random = list(patient = pdDiag(form = ~ obstime)))
```

2.4 Missing Data in Longitudinal Studies

- A major challenge for the analysis of longitudinal data is the problem of missing data
 - ▷ studies are designed to collect data on every subject at a set of prespecified follow-up times
 - ▷ often subjects miss some of their planned measurements for a variety of reasons
- We can have different patterns of missing data

2.4 Missing Data in Longitudinal Studies (cont'd)

Subject	Visits				
	1	2	3	4	5
1	x	x	x	x	x
2	x	x	x	?	?
3	?	x	x	x	x
4	?	x	?	x	?

- ▷ Subject 1: Completer
- ▷ Subject 2: dropout
- ▷ Subject 3: late entry
- ▷ Subject 4: intermittent

2.4 Missing Data in Longitudinal Studies (cont'd)

- Implications of missingness:
 - ▷ we collect less data than originally planned \Rightarrow *loss of efficiency*
 - ▷ not all subjects have the same number of measurements \Rightarrow *unbalanced datasets*
 - ▷ missingness may depend on outcome \Rightarrow *potential bias*
- For the handling of missing data, we introduce the missing data indicator

$$r_{ij} = \begin{cases} 1 & \text{if } y_{ij} \text{ is observed} \\ 0 & \text{otherwise} \end{cases}$$

2.4 Missing Data in Longitudinal Studies (cont'd)

- We obtain a partition of the complete response vector y_i
 - ▷ observed data y_i^o , containing those y_{ij} for which $r_{ij} = 1$
 - ▷ missing data y_i^m , containing those y_{ij} for which $r_{ij} = 0$
- **For the remaining we will focus on dropout** \Rightarrow notation can be simplified
 - ▷ Discrete dropout time: $r_i^d = 1 + \sum_{j=1}^{n_i} r_{ij}$ (ordinal variable)
 - ▷ **Continuous time**: T_i^* denotes the time to dropout

2.5 Missing Data Mechanisms

- To describe the probabilistic relation between the measurement and missingness processes Rubin (1976, Biometrika) has introduced three mechanisms
- *Missing Completely At Random (MCAR)*: The probability that responses are missing is unrelated to both y_i^o and y_i^m

$$p(r_i \mid y_i^o, y_i^m) = p(r_i)$$

- Examples
 - ▷ subjects go out of the study after providing a pre-determined number of measurements
 - ▷ laboratory measurements are lost due to equipment malfunction

2.5 Missing Data Mechanisms (cont'd)

- Features of MCAR:
 - ▷ The observed data y_i^o can be considered a random sample of the complete data y_i
 - ▷ We can use any statistical procedure that is valid for complete data
 - * sample averages per time point
 - * linear regression, ignoring the correlation (consistent, but not efficient)
 - * t -test at the last time point
 - * ...

2.5 Missing Data Mechanisms (cont'd)

- *Missing At Random (MAR)*: The probability that responses are missing is related to y_i^o , but is unrelated to y_i^m

$$p(r_i \mid y_i^o, y_i^m) = p(r_i \mid y_i^o)$$

- Examples
 - ▷ study protocol requires patients whose response value exceeds a threshold to be removed from the study
 - ▷ physicians give rescue medication to patients who do not respond to treatment

2.5 Missing Data Mechanisms (cont'd)

- Features of MAR:
 - ▷ The observed data cannot be considered a random sample from the target population
 - ▷ Not all statistical procedures provide valid results

Not valid under MAR	Valid under MAR
sample marginal evolutions	sample subject-specific evolutions
methods based on moments, such as GEE	likelihood based inference
mixed models with misspecified correlation structure	mixed models with correctly specified correlation structure
marginal residuals	subject-specific residuals

2.5 Missing Data Mechanisms (cont'd)

- *Missing Not At Random (MNAR)*: The probability that responses are missing is related to y_i^m , and possibly also to y_i^o

$$p(r_i \mid y_i^m) \quad \text{or} \quad p(r_i \mid y_i^o, y_i^m)$$

- Examples
 - ▷ in studies on drug addicts, people who return to drugs are less likely than others to report their status
 - ▷ in longitudinal studies for quality-of-life, patients may fail to complete the questionnaire at occasions when their quality-of-life is compromised

2.5 Missing Data Mechanisms (cont'd)

- Features of MNAR
 - ▷ The observed data cannot be considered a random sample from the target population
 - ▷ Only procedures that explicitly model the joint distribution $\{y_i^o, y_i^m, r_i\}$ provide valid inferences \Rightarrow **analyses which are valid under MAR will not be valid under MNAR**

2.5 Missing Data Mechanisms (cont'd)

We cannot tell from the data at hand whether the missing data mechanism is MAR or MNAR

Note: We can distinguish between MCAR and MAR

Chapter 3

Relative Risk Models

3.1 Features of Survival Data

- The most important 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 Features of Survival Data (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* (similar to MNAR and MAR)

Here we focus on non-informative right censoring

- Note: Survival times may often be truncated; analysis of truncated samples requires similar calculations as censoring

3.1 Features of Survival Data (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 $\{T_i, \delta_i\}$

3.2 Relative Risk Models

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

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

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

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}, \dots, w_{ip} a set of covariates

3.2 Relative Risk Models (cont'd)

- Standard MLE can be applied based on the log-likelihood function

$$\ell(\theta) = \sum_{i=1}^n \delta_i \log p(T_i; \theta) + (1 - \delta_i) \log S_i(T_i; \theta),$$

which also can be re-expressed in terms of the hazard function

$$\ell(\theta) = \sum_{i=1}^n \delta_i \log h_i(T_i; \theta) - \int_0^{T_i} h_i(s; \theta) ds$$

Sensitivity to distributional assumptions due to censoring

3.2 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

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

where only patients who had an event contribute

3.2 Relative Risk Models (cont'd)

- **Example:** For the PBC dataset were interested in the treatment effect while correcting for sex and age effects

$$h_i(t) = h_0(t) \exp(\gamma_1 \text{D-penic}_i + \gamma_2 \text{Female}_i + \gamma_3 \text{Age}_i)$$

	Value	HR	Std.Err.	z-value	p-value
γ_1	-0.138	0.871	0.156	-0.882	0.378
γ_2	-0.493	0.611	0.207	-2.379	0.017
γ_3	0.021	1.022	0.008	2.784	0.005

3.3 Relative Risk Models in R

- R> The primary package in R for the analysis of survival data is the **survival** package
- R> A key function in this package that is used to specify the available event time information in a sample at hand is `Surv()`
- R> For right censored failure times (i.e., what we will see in this course) we need to provide the observed event times `time`, and the event indicator `status`, which equals 1 for true failure times and 0 for right censored times

`Surv(time, status)`

3.3 Relative Risk Models in R (cont'd)

R> Cox models are fitted using function `coxph()`. For instance, for the PBC data the following code fits the Cox model that contains the main effects of 'drug', 'sex' and 'age':

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

```
summary(coxFit)
```

R> The two main arguments are a formula specifying the design matrix of the model and a data frame containing all the variables

3.4 Time Dependent Covariates

- Often interest in the association between a time-dependent 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?

3.4 Time Dependent 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 move to the time-dependent setting, a more **careful consideration** is required

3.4 Time Dependent Covariates (cont'd)

- There are two types of time-dependent covariates

(Kalbfleisch and Prentice, 2002, Section 6.3)

- ▷ Exogenous (aka external): the future path of the covariate up to any time $t > s$ is not affected by the occurrence of an event at time point s , i.e.,

$$\Pr\{\mathcal{Y}_i(t) \mid \mathcal{Y}_i(s), T_i^* \geq s\} = \Pr\{\mathcal{Y}_i(t) \mid \mathcal{Y}_i(s), T_i^* = s\},$$

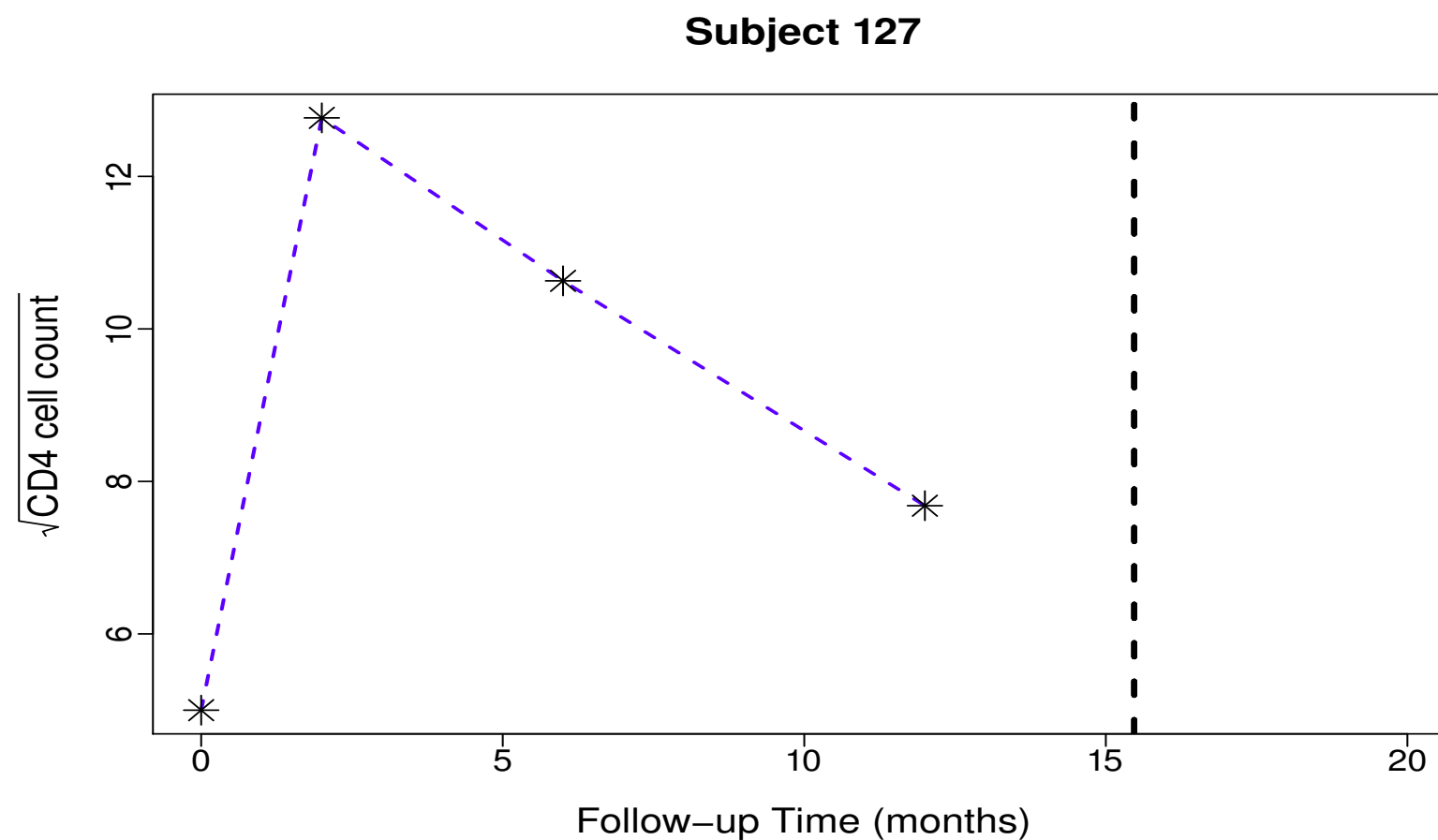
where $0 < s \leq t$ and $\mathcal{Y}_i(t) = \{y_i(s), 0 \leq s < t\}$

- ▷ Endogenous (aka internal): not Exogenous

3.4 Time Dependent Covariates (cont'd)

- It is **very important** to distinguish between these two types of time-dependent 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

3.4 Time Dependent Covariates (cont'd)



3.5 Extended Cox Model

- The Cox model presented earlier can be extended to handle time-dependent 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

3.5 Extended Cox Model (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

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

3.5 Extended Cox Model (cont'd)

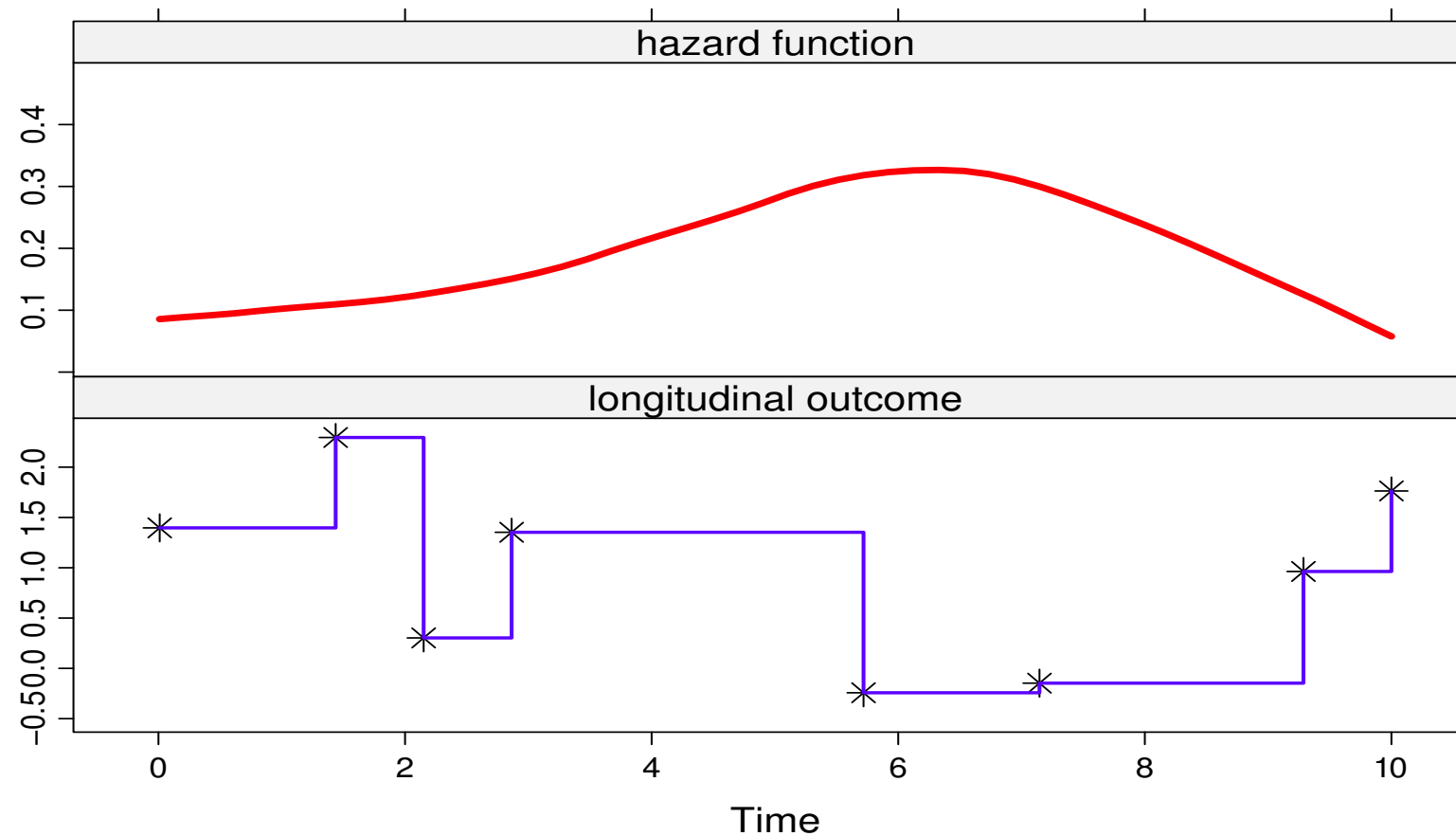
- Typically, data must be organized in the long format

Patient	Start	Stop	Event	$y_i(t)$	Age
1	0	135	1	5.5	45
2	0	65	0	2.2	38
2	65	120	0	3.1	38
2	120	155	1	4.1	38
3	0	115	0	2.5	29
3	115	202	0	2.9	29
⋮	⋮	⋮	⋮	⋮	⋮

3.5 Extended Cox Model (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

3.5 Extended Cox Model (cont'd)



3.5 Extended Cox Model (cont'd)

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

**Treating endogenous covariates as exogenous may
produce spurious results!**

Chapter 4

The Basic Joint Model

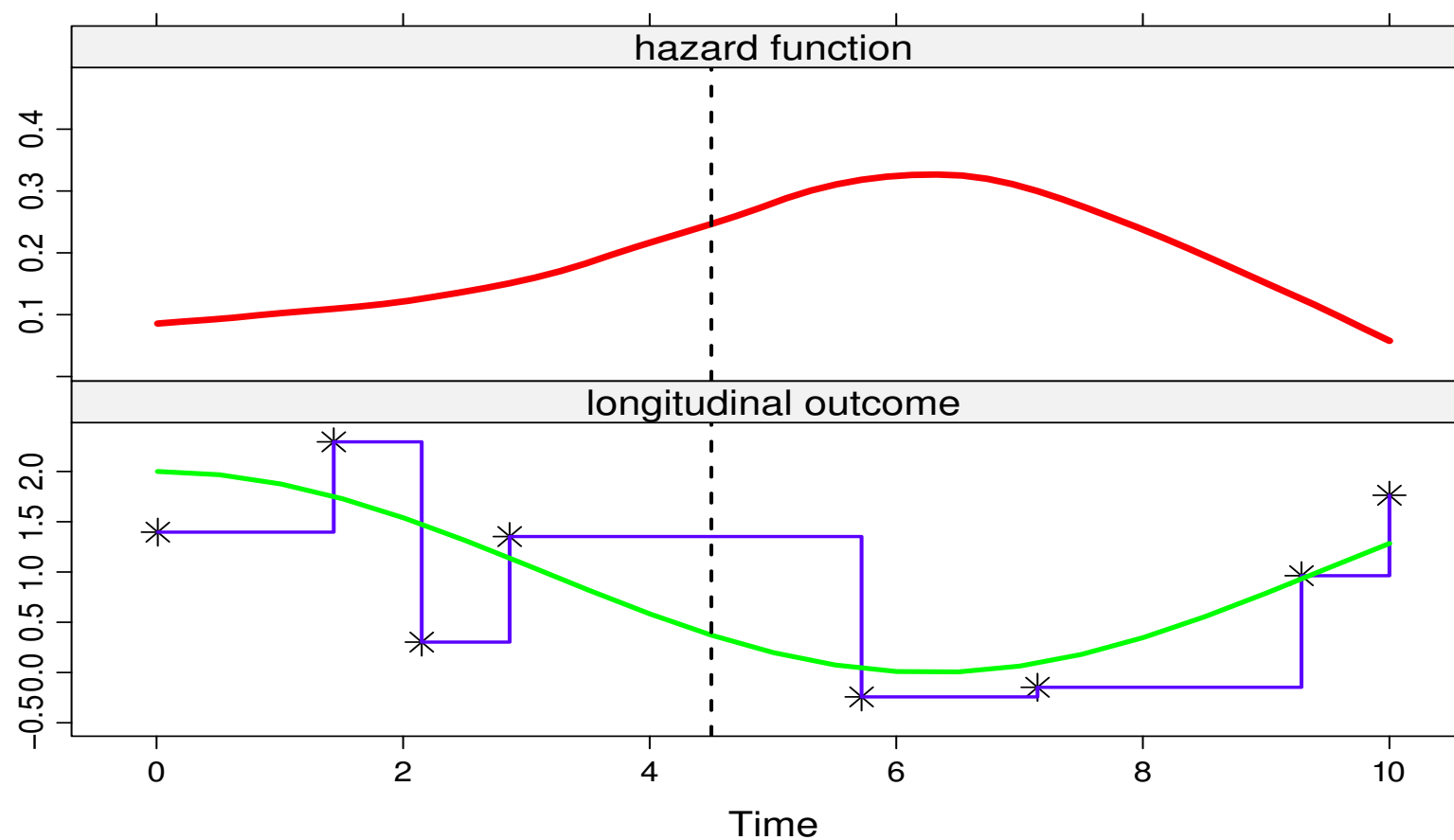
4.1 Joint Modeling Framework

- 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 marker in time for each patient
 2. the estimated evolutions are then used in a Cox model
- Feature: Marker 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 responses
- 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 marker 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 strength of the association between the marker and the risk of an event
- ▷ w_i baseline covariates

4.1 Joint Modeling Framework (cont'd)

- Step 2: From the observed longitudinal response $y_i(t)$ reconstruct the covariate history for each subject
- Mixed effects model (we focus, for now, on continuous markers)

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

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)

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

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 be tested

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)

- The survival function, which is a part of the likelihood of the model, depends on the whole longitudinal history

$$S_i(t \mid b_i) = \exp\left(-\int_0^t h_0(s) \exp\{\gamma^\top w_i + \alpha m_i(s)\} ds\right)$$

- Therefore, care in the definition of the design matrices of the mixed model
 - ▷ when subjects have nonlinear profiles \Rightarrow
 - ▷ use splines or polynomials to model them flexibly

4.1 Joint Modeling Framework (cont'd)

- Assumptions for the baseline hazard function $h_0(t)$
 - ▷ parametric \Rightarrow possibly restrictive
 - ▷ unspecified \Rightarrow within JM framework underestimates standard errors
- It is advisable to use parametric but flexible models for $h_0(t)$
 - ▷ splines

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

where

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

4.1 Joint Modeling Framework (cont'd)

- It is advisable to use parametric but flexible models for $h_0(t)$
 - ▷ step-functions: piecewise-constant baseline hazard often works satisfactorily

$$h_0(t) = \sum_{q=1}^Q \xi_q I(v_{q-1} < t \leq v_q),$$

where $0 = v_0 < v_1 < \dots < v_Q$ denotes a split of the time scale

4.2 Estimation

- Mainly maximum likelihood but also Bayesian approaches
- The log-likelihood contribution for subject i :

$$\ell_i(\theta) = \log \int \left\{ \prod_{j=1}^{n_i} p(y_{ij} \mid b_i; \theta) \right\} \left\{ h(T_i \mid b_i; \theta)^{\delta_i} S_i(T_i \mid b_i; \theta) \right\} p(b_i; \theta) db_i,$$

where

$$S_i(t \mid b_i; \theta) = \exp \left(- \int_0^t h_0(s; \theta) \exp \{ \gamma^\top w_i + \alpha m_i(s) \} ds \right)$$

4.2 Estimation (cont'd)

- Both integrals do not have, in general, a closed-form solution \Rightarrow need to be approximated numerically
- Standard numerical integration algorithms
 - ▷ Gaussian quadrature
 - ▷ Monte Carlo
 - ▷ ...
- More difficult is the integral with respect to b_i because it can be of high dimension
 - ▷ Laplace approximations
 - ▷ pseudo-adaptive Gaussian quadrature rules

4.2 Estimation (cont'd)

- To maximize the approximated log-likelihood

$$\ell(\theta) = \sum_{i=1}^n \log \int p(y_i | b_i; \theta) \{h(T_i | b_i; \theta)^{\delta_i} S_i(T_i | b_i; \theta)\} p(b_i; \theta) db_i,$$

we need to employ an optimization algorithm

- Standard choices
 - ▷ EM (treating b_i as missing data)
 - ▷ Newton-type
 - ▷ hybrids (start with EM and continue with quasi-Newton)

4.2 Estimation (cont'd)

- Standard errors: Standard asymptotic MLE

$$\text{var}(\hat{\theta}) = \left\{ - \sum_{i=1}^n \frac{\partial^2 \log p(y_i, T_i, \delta_i; \theta)}{\partial \theta^\top \partial \theta} \bigg|_{\theta=\hat{\theta}} \right\}^{-1}$$

- Standard asymptotic tests + information criteria
 - ▷ likelihood ratio test
 - ▷ score test
 - ▷ Wald test
 - ▷ AIC, BIC, ...

4.2 Estimation (cont'd)

- Based on a fitted joint model, estimates for the random effects are based on the posterior distribution:

$$p(b_i \mid T_i, \delta_i, y_i; \theta) = \frac{p(T_i, \delta_i \mid b_i; \theta) p(y_i \mid b_i; \theta) p(b_i; \theta)}{p(T_i, \delta_i, y_i; \theta)}$$

$$\propto p(T_i, \delta_i \mid b_i; \theta) p(y_i \mid b_i; \theta) p(b_i; \theta),$$

in which θ is replaced by its MLE $\hat{\theta}$

4.2 Estimation (cont'd)

- Measures of location

$$\begin{cases} \bar{b}_i = \int b_i p(b_i | T_i, \delta_i, y_i; \hat{\theta}) db_i \\ \hat{b}_i = \operatorname{argmax}_b \{\log p(b | T_i, \delta_i, y_i; \hat{\theta})\} \end{cases}$$

- Measures of dispersion

$$\begin{cases} \operatorname{var}(b_i) = \int (b_i - \bar{b}_i)(b_i - \bar{b}_i)^\top p(b_i | T_i, \delta_i, y_i; \hat{\theta}) db_i \\ H_i = \left\{ -\frac{\partial^2 \log p(b | T_i, \delta_i, y_i; \hat{\theta})}{\partial b^\top \partial b} \Big|_{b=\hat{b}_i} \right\}^{-1} \end{cases}$$

4.3 Bayesian Estimation

- Bayesian estimation
 - ▷ under the Bayesian paradigm both θ and $\{b_i, i = 1, \dots, n\}$ are regarded as parameters
- Inference is based on the full posterior distribution

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

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

4.3 Bayesian Estimation (cont'd)

- No closed-form solutions for the integrals in the normalizing constant \Rightarrow **MCMC**
- For the standard joint model we have defined thus far, the majority of the parameters can be updated using Gibbs sampling (or slice sampling)
 - ▷ when no closed-form posterior conditionals are available, we can use the Metropolis-Hastings algorithm
- 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

4.3 Bayesian Estimation (cont'd)

- Good proposal distributions can be obtained from the separate fits of the two submodels
- Not directly programmable in WinBUGS, INLA, etc., due to the integral in the definition of the survival function

$$S_i(t \mid b_i; \theta) = \exp\left(-\int_0^t h_0(s; \theta) \exp\{\gamma^\top w_i + \alpha m_i(s)\} ds\right)$$

extra steps required...

4.3 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
 - ▷ Bayes factors
 - ▷ DIC, CPO
 - ▷ ...

4.4 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

$$\left\{ \begin{array}{l} y_i(t) = m_i(t) + \varepsilon_i(t) \\ \quad = \beta_0 + \beta_1 t + \beta_2 \{t \times \text{ddI}_i\} + b_{i0} + b_{i1} t + \varepsilon_i(t), \quad \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2), \\ h_i(t) = h_0(t) \exp\{\gamma \text{ddI}_i + \alpha m_i(t)\}, \end{array} \right.$$

where

▷ $h_0(t)$ is assumed piecewise-constant

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

	JM	Cox
	log HR (std.err)	log HR (std.err)
Treat	0.33 (0.16)	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

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

- A unit decrease in $CD4^{1/2}$, results in a
 - ▷ **Joint Model**: 1.3-fold increase in risk (95% CI: 1.24; 1.43)
 - ▷ **Time-Dependent Cox**: 1.2-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 markers

4.5 Joint Models in R

R> Joint models are fitted using function `jointModel()` from package **JM**. This function accepts as main arguments a linear mixed model and a Cox PH model based on which it fits the corresponding joint model

```
lmeFit <- lme(CD4 ~ obstime + obstime:drug,  
             random = ~ obstime | patient, data = aids)  
  
coxFit <- coxph(Surv(Time, death) ~ drug, data = aids.id, x = TRUE)  
  
jointFit <- jointModel(lmeFit, coxFit, timeVar = "obstime",  
                      method = "piecewise-PH-aGH")  
  
summary(jointFit)
```


4.5 Joint Models in R (cont'd)

- R>** As before, 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>** In the call to `coxph()` you need to set `x = TRUE` (or `model = TRUE`) such that the design matrix used in the Cox model is returned in the object fit
- R>** Argument `timeVar` specifies the time variable in the linear mixed model

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

4.5 Joint Models in R (cont'd)

R> Argument `method` specifies the type of relative risk model and the type of numerical integration algorithm – the syntax is as follows:

`<baseline hazard>-<parameterization>-<numerical integration>`

Available options are:

- ▷ `"piecewise-PH-GH"`: PH model with piecewise-constant baseline hazard
- ▷ `"spline-PH-GH"`: PH model with B-spline-approximated log baseline hazard
- ▷ `"weibull-PH-GH"`: PH model with Weibull baseline hazard
- ▷ `"weibull-AFT-GH"`: AFT model with Weibull baseline hazard
- ▷ `"Cox-PH-GH"`: PH model with unspecified baseline hazard

`GH` stands for standard Gauss-Hermite; using `aGH` invokes the pseudo-adaptive Gauss-Hermite rule

4.5 Joint Models in R (cont'd)

R> Joint models under the Bayesian approach are fitted using function `jointModelBayes()` from package **JMbayes**. This function works in a very similar manner as function `jointModel()`, e.g.,

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

```
coxFit <- coxph(Surv(Time, death) ~ drug, data = aids.id, x = TRUE)
```

```
jointFitBayes <- jointModelBayes(lmeFit, coxFit, timeVar = "obstime")
```

```
summary(jointFitBayes)
```

4.5 Joint Models in R (cont'd)

- R> JMbayes** is more flexible (in some respects):
- ▷ directly implements the MCMC
 - ▷ allows for categorical longitudinal data as well
 - ▷ allows for general transformation functions
 - ▷ penalized B-splines for the baseline hazard function
 - ▷ ...

4.5 Joint Models in R (cont'd)

R> In both packages methods are available for the majority of the standard generic functions + extras

- ▷ `summary()`, `anova()`, `vcov()`, `logLik()`
- ▷ `coef()`, `fixef()`, `ranef()`
- ▷ `fitted()`, `residuals()`
- ▷ `plot()`
- ▷ `xtable()` (you need to load package **xtable** first)

4.6 Connection with Missing Data

- So far we have attacked the problem from the survival point of view
- However, often, we may be also interested on the longitudinal outcome
- **Issue:** When patients experience the event, they dropout from the study
 - ▷ a direct connection with the missing data field

Dropout must be taken into account when deriving inferences for the longitudinal outcome

4.6 Connection with Missing Data (cont'd)

- To show this connection more clearly
 - ▷ T_i^* : true time-to-event
 - ▷ y_i^o : longitudinal measurements before T_i^*
 - ▷ y_i^m : longitudinal measurements after T_i^*
- **Important to realize** that the model we postulate for the longitudinal responses is for the complete vector $\{y_i^o, y_i^m\}$
 - ▷ implicit assumptions about missingness

4.6 Connection with Missing Data (cont'd)

- Missing data mechanism:

$$p(T_i^* \mid y_i^o, y_i^m) = \int p(T_i^* \mid b_i) p(b_i \mid y_i^o, y_i^m) db_i$$

still depends on y_i^m , which corresponds to nonrandom dropout

Intuitive interpretation: Patients who dropout show different longitudinal evolutions than patients who do not

4.6 Connection with Missing Data (cont'd)

- Implications of nonrandom dropout
 - ▷ observed data do not constitute a random sample from the target population
- This feature complicates the validation of the joint model's assumptions using standard residual plots
 - ▷ **what is the problem:** Residual plots may show systematic behavior due to dropout and not because of model misfit

4.6 Connection with Missing Data (cont'd)

- What about censoring?
 - ▷ censoring also corresponds to a discontinuation of the data collection process for the longitudinal outcome
- Likelihood-based inferences for joint models provide valid inferences when censoring is MAR
 - ▷ a patient relocates to another country (MCAR)
 - ▷ a patient is excluded from the study when her longitudinal response exceeds a prespecified threshold (MAR)
 - ▷ censoring depends on random effects (MNAR)

4.6 Connection with Missing Data (cont'd)

- Joint models belong to the class of *Shared Parameter Models*

$$p(y_i^o, y_i^m, T_i^*) = \int p(y_i^o, y_i^m \mid b_i) p(T_i^* \mid b_i) p(b_i) db_i$$

the association between the longitudinal and missingness processes is explained by the *shared* random effects b_i

4.6 Connection with Missing Data (cont'd)

- The other two well-known frameworks for MNAR data are
 - ▷ Selection models

$$p(y_i^o, y_i^m, T_i^*) = p(y_i^o, y_i^m) p(T_i^* \mid y_i^o, y_i^m)$$

- ▷ Pattern mixture models:

$$p(y_i^o, y_i^m, T_i^*) = p(y_i^o, y_i^m \mid T_i^*) p(T_i^*)$$

- These two model families are primarily applied with discrete dropout times and cannot be easily extended to continuous time

4.6 Connection with Missing Data (cont'd)

- Example: In the AIDS data the association parameter α was highly significant, suggesting nonrandom dropout
- A comparison between
 - ▷ linear mixed-effects model \Rightarrow MAR
 - ▷ joint model \Rightarrow MNARis warranted
- MAR assumes that missingness depends only on the observed data

$$p(T_i^* \mid y_i^o, y_i^m) = p(T_i^* \mid y_i^o)$$

4.6 Connection with Missing Data (cont'd)

	LMM (MAR) value (s.e.)	JM (MNAR) value (s.e.)
Inter	7.19 (0.22)	7.22 (0.22)
Time	−0.16 (0.02)	−0.19 (0.02)
Treat:Time	0.03 (0.03)	0.01 (0.03)

- Minimal sensitivity in parameter estimates & standard errors
 ⇒ **Warning:** This does not mean that this is always the case!

Chapter 5

Extensions of Joint Models

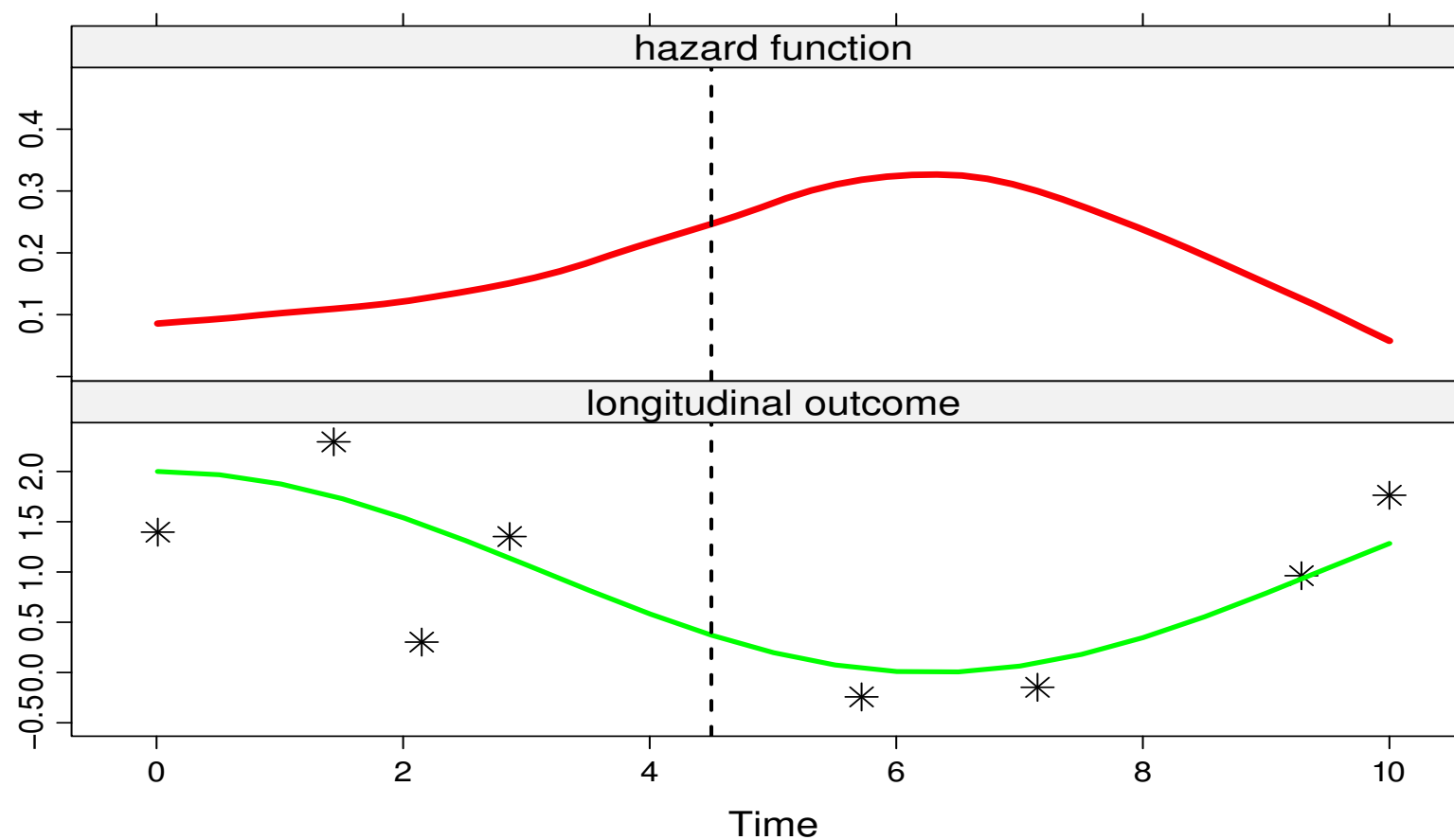
5.1 Parameterizations

- The standard joint model

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

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

5.1 Parameterizations (cont'd)



5.1 Parameterizations (cont'd)

- The standard joint model

$$\left\{ \begin{array}{l} h_i(t \mid \mathcal{M}_i(t)) = h_0(t) \exp\{\gamma^\top w_i + \alpha m_i(t)\}, \\ y_i(t) = m_i(t) + \varepsilon_i(t) \\ \quad = x_i^\top(t)\beta + z_i^\top(t)b_i + \varepsilon_i(t), \end{array} \right.$$

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 Parameterizations (cont'd)

- Note: Inappropriate modeling of time-dependent 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., patients 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 Parameterizations (cont'd)

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

- Let's see some possibilities. . .

5.1 Parameterizations (cont'd)

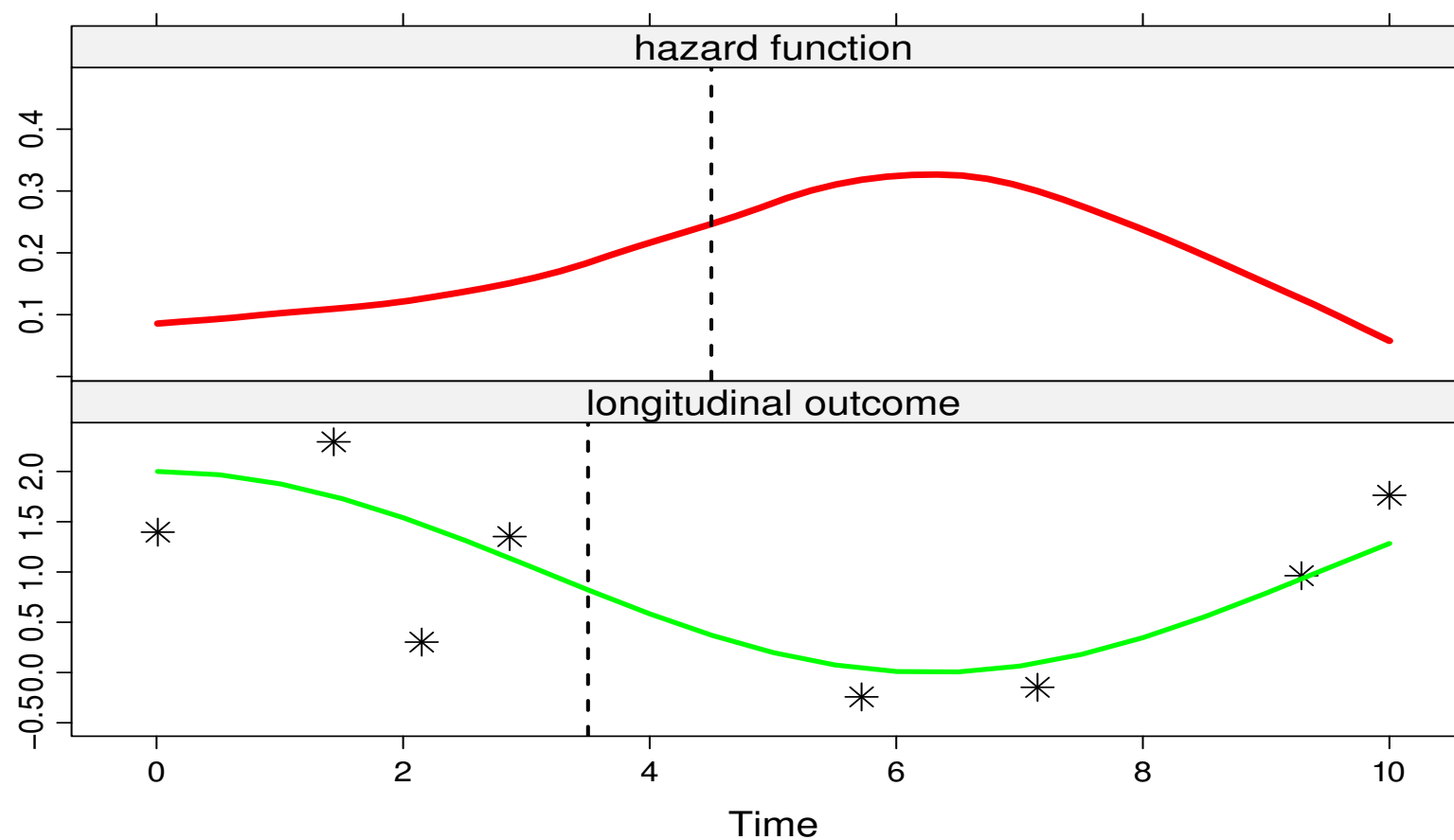
- *Lagged Effects*: The hazard of an event at t is associated with the level of the marker at a previous time point:

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

where

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

5.1 Parameterizations (cont'd)



5.1 Parameterizations (cont'd)

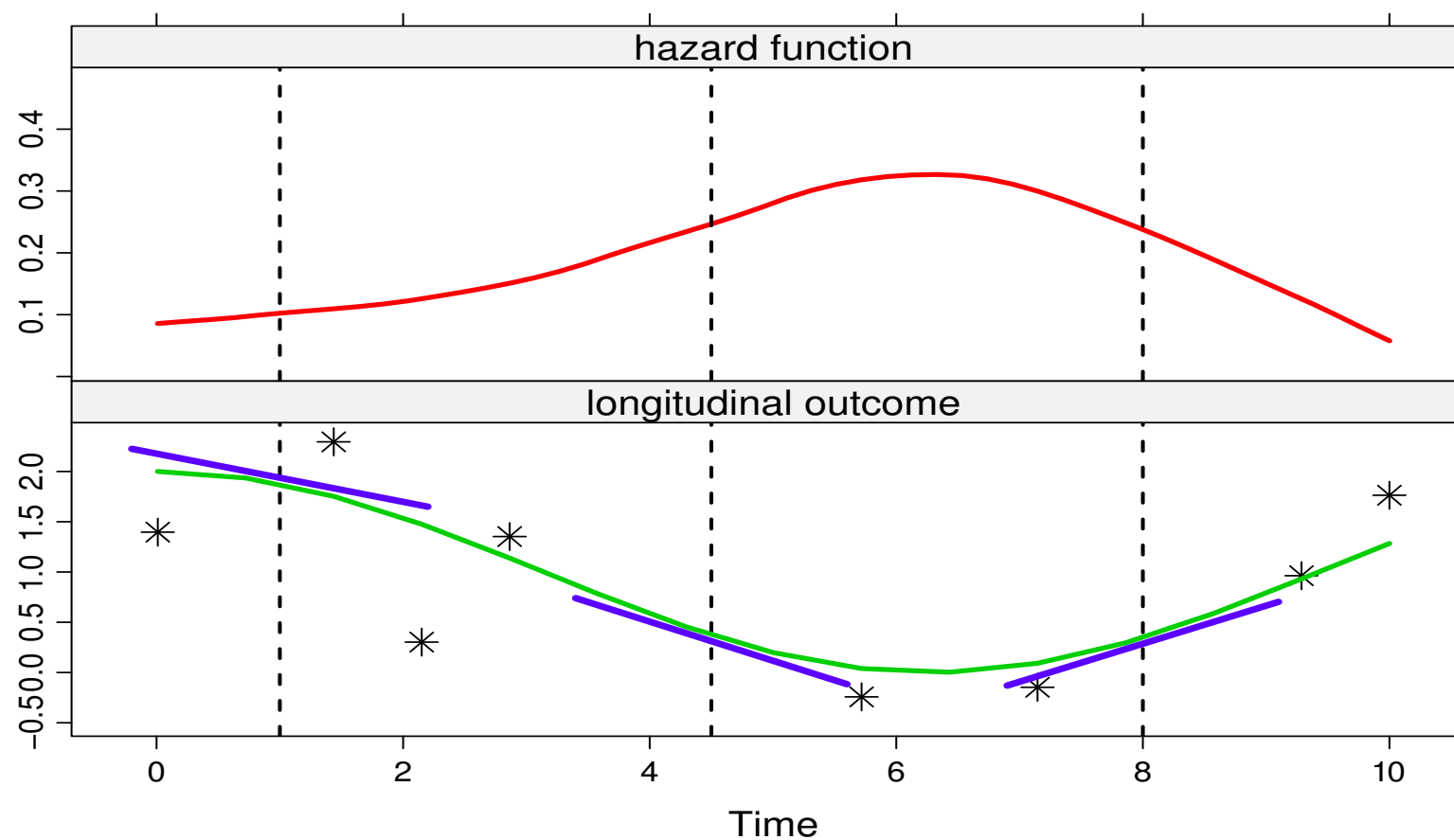
- *Time-dependent Slopes*: The hazard of an event at t is associated with both the current value and the slope of the trajectory at t (Ye et al., 2008, Biometrics):

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

where

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

5.1 Parameterizations (cont'd)



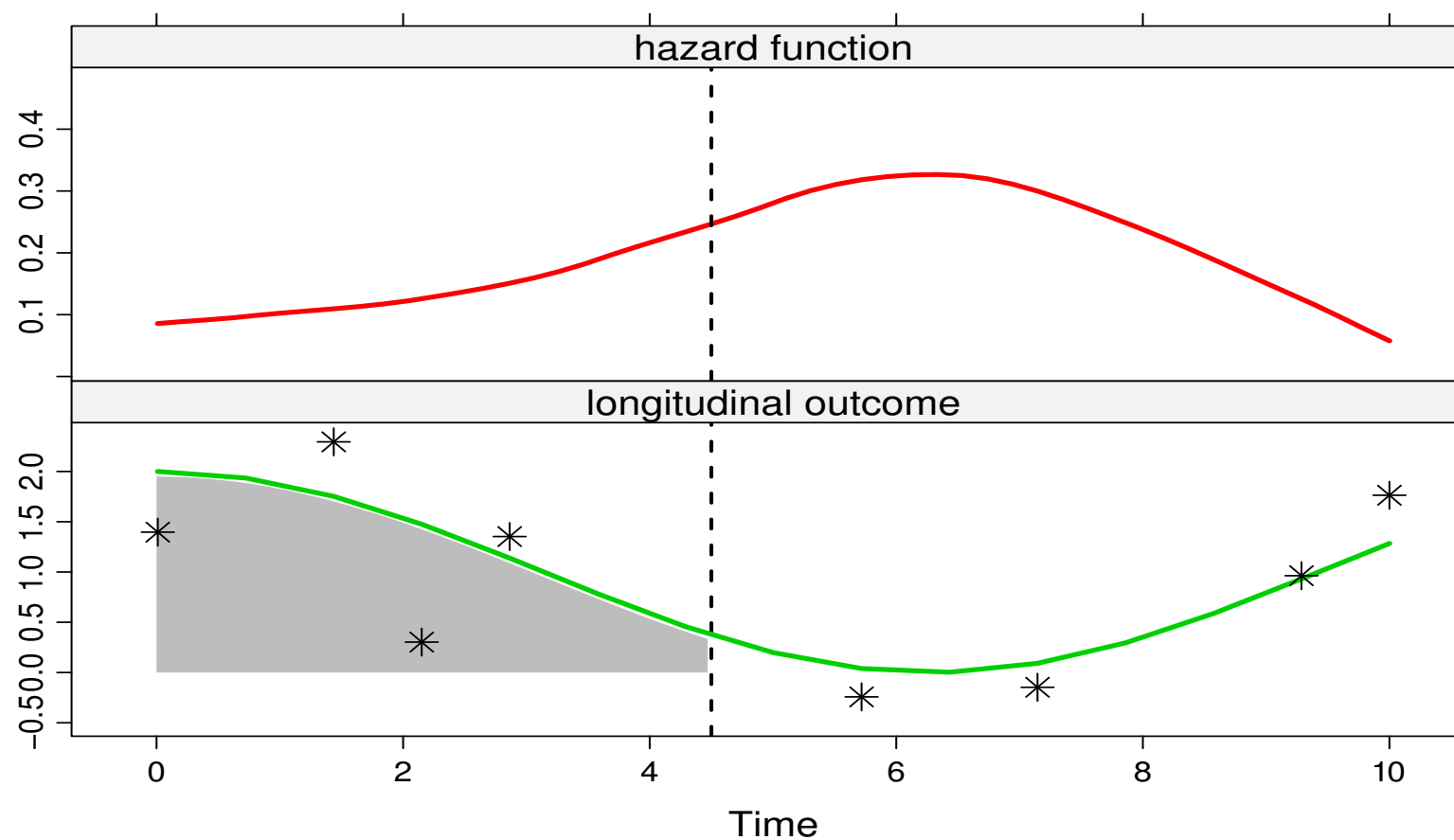
5.1 Parameterizations (cont'd)

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

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

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

5.1 Parameterizations (cont'd)



5.1 Parameterizations (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 :

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

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

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

5.1 Parameterizations (cont'd)

- *Random Effects*: The hazard of an event at t is associated only with the random effects of the longitudinal model:

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

- Features:
 - ▷ avoids numerical integration for the survival function
 - ▷ interpretation of α more difficult, especially in high-dimensional random-effects settings

5.1 Parameterizations (cont'd)

- Example: Sensitivity of inferences for the longitudinal process to the choice of the parameterization for the AIDS data
- We use the same mixed model as before, i.e.,

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

and the following four survival submodels

5.1 Parameterizations (cont'd)

- Model I (current value)

$$h_i(t) = h_0(t) \exp\{\gamma \text{ddI}_i + \alpha_1 m_i(t)\}$$

- Model II (current value + current slope)

$$h_i(t) = h_0(t) \exp\{\gamma \text{ddI}_i + \alpha_1 m_i(t) + \alpha_2 m'_i(t)\},$$

where

$$\triangleright m'_i(t) = \beta_1 + \beta_2 \text{ddI}_i + b_{i1}$$

5.1 Parameterizations (cont'd)

- Model III (random slope)

$$h_i(t) = h_0(t) \exp\{\gamma d dI_i + \alpha_3 b_{i1}\}$$

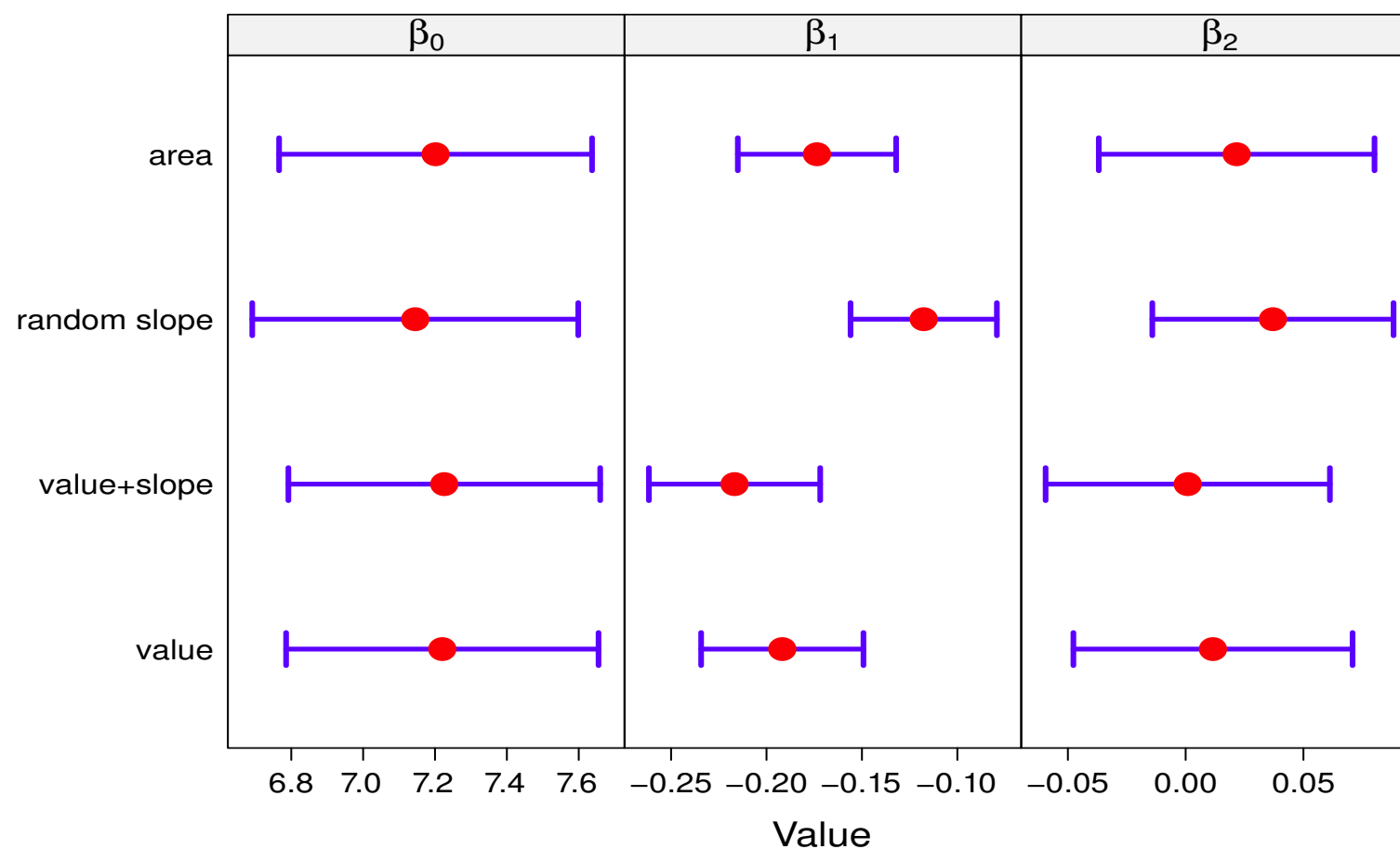
- Model IV (area)

$$h_i(t) = h_0(t) \exp\left\{\gamma d dI_i + \alpha_4 \int_0^t m_i(s) ds\right\},$$

where

$$\triangleright \int_0^t m_i(s) ds = \beta_0 t + \frac{\beta_1}{2} t^2 + \frac{\beta_2}{2} \{t^2 \times d dI_i\} + b_{i0} t + \frac{b_{i1}}{2} t^2$$

5.1 Parameterizations (cont'd)



5.1 Parameterizations (cont'd)

- There are noticeable differences between the parameterizations
 - ▷ especially in the slope parameters
- Therefore, a sensitivity analysis should not stop at the standard joint model parameterization but also consider alternative association structures

5.1 Parameterizations (cont'd)

- R>** Lagged effects can be fitted using the `lag` argument of `jointModel()`. For example, the following code fits a joint model for the PBC dataset with
- ▷ random intercepts and random slopes for log serum bilirubin, and
 - ▷ a relative risk model with piecewise-constant baseline hazard and the *true effect at the previous year*

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

```
coxFit <- coxph(Surv(years, status2) ~ 1, data = pbc2.id, x = TRUE)
```

```
jointFit <- jointModel(lmeFit, coxFit, timeVar = "year",  
  method = "piecewise-PH-aGH", lag = 1)
```

```
summary(jointFit)
```

5.1 Parameterizations (cont'd)

R> For the time-dependent slopes and cumulative effects parameterizations, arguments `parameterization` and `derivForm` of `jointModel()` should be used

▷ the first one just specifies whether we want to include a single or two terms involving $m_i(t)$ in the linear predictor of the survival submodel, options are

- * `parameterization = "value"`
- * `parameterization = "slope"`
- * `parameterization = "both"`

▷ the second one requires a few extra steps to specify – we will see an example in the practical

5.2 Multiple Longitudinal Markers

- So far we have concentrated on a single continuous marker
- But very often we may have several markers 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)
 - ▷ ...

5.2 Multiple Longitudinal Markers (cont'd)

We need to extend the basic joint model!

- To handle multiple longitudinal markers of different types we use Generalized Linear Mixed Models
 - ▷ We assume Y_{i1}, \dots, 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

5.2 Multiple Longitudinal Markers (cont'd)

- ▷ Correlation between the outcomes is built by assuming a multivariate normal distribution for the random effects

$$b_i = (b_{i1}^\top, \dots, b_{iJ}^\top)^\top \sim \mathcal{N}(0, D)$$

- The expected value of each longitudinal marker is incorporated in the linear predictor of the survival submodel

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

5.2 Multiple Longitudinal Markers (cont'd)

R> Joint models for multiple longitudinal outcomes can be fitted with function `mvJointModelBayes()` from package **JMbayes**

- The use of this function mimics the one of `jointModelBayes()` but with some small differences, namely
 - ▷ we fit a *multivariate* mixed model using `mvglmer()`,
 - ▷ following we fit a Cox model using `coxph()`, and
 - ▷ and we give the resulting objects as input in `mvJointModelBayes()`

5.2 Multiple Longitudinal Markers (cont'd)

R> An example for the PBC dataset using serum bilirubin (continuous) and spiders (binary)

```
multMixedFit <- mvglmer(list(log(serBilir) ~ year + (year | id),  
                             spiders ~ year + (1 | id)), data = pbc2,  
                        families = list(gaussian, binomial))
```

```
CoxFit <- coxph(Surv(Time, event) ~ drug + age, data = pbc2.id,  
               model = TRUE)
```

```
multJMFit <- mvJointModelBayes(multMixedFit, CoxFit, timeVar = "year")  
summary(multJMFit)
```


5.2 Multiple Longitudinal Markers (cont'd)

- R> Function `mvJointModelBayes()` also allows for
- ▷ right, left, interval censored data
 - ▷ left truncated data
 - ▷ exogenous time-varying covariates

More info and vignettes in
<http://www.drizopoulos.com/> → Software

5.3 Multiple Failure Times

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

5.3 Multiple Failure Times (cont'd)

- Joint models with competing risks:

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

where

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

5.3 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

5.3 Multiple Failure Times (cont'd)

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

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

with

- ▷ $r_i(t)$ hazard function for the recurrent events
- ▷ $h_i(t)$ hazard function for the terminal event
- ▷ \mathbf{v}_i frailty term accounting for the correlation in the recurrent events

5.4 Extensions & Parameterizations

- Note: In the previous extensions of joint models, i.e.,
 - ▷ multiple longitudinal markers
 - ▷ multiple failure times

we used the default parameterization that includes the current value term $m_i(t)$ in the linear predictor of the survival submodel(s)

Nonetheless, all the other parameterizations we have seen earlier are also applicable

5.4 Extensions & Parameterizations (cont'd)

- For example in the case of multiple longitudinal outcomes

$$g_j[E\{y_{ij}(t) \mid b_{ij}\}] = m_{ij}(t) = x_{ij}^\top(t)\beta_j + z_{ij}^\top(t)b_{ij}$$

$$h_i(t) = h_0(t) \exp\left\{\gamma^\top w_i + \sum_{j=1}^J \sum_{l=1}^L f_{jl}(\mathcal{H}_{ij}(t), \alpha_{jl})\right\}$$

5.4 Extensions & Parameterizations (cont'd)

- In this case we face a challenging model selection problem
- Different possible solutions
 - ▷ lasso
 - ▷ ridge
 - ▷ horseshoe
 - ▷ ...

5.4 Extensions & Parameterizations (cont'd)

- R> Function `mvJointModelBayes()` also allows to consider multiple parameterization per outcome in a similar manner as `jointModelBayes()` does
- R> It also implements a global-local ridge-type prior for the association parameters

$$\alpha_{jl} \sim \mathcal{N}(0, \tau\psi_{jl})$$

$$\tau^{-1} \sim \text{Gamma}(0.1, 0.1)$$

$$\psi_{jl}^{-1} \sim \text{Gamma}(1, 0.01)$$

Chapter 6

Dynamic Predictions, Discrimination & Calibration

6.1 Survival Probabilities: Definitions

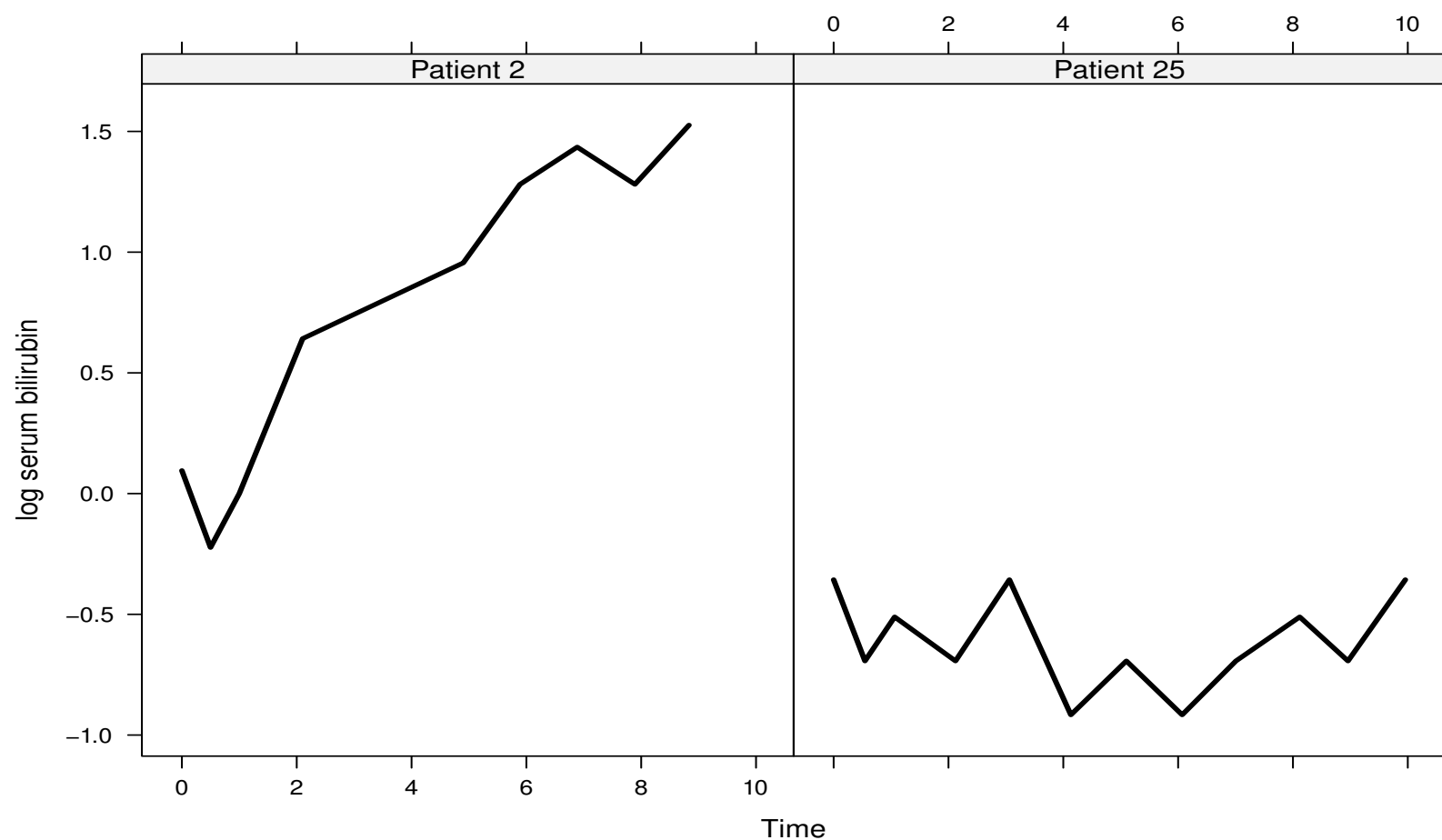
- Nowadays there is great interest for prognostic models and their application to personalized medicine
- Examples are numerous
 - ▷ cancer research, cardiovascular diseases, HIV research, ...

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: Definitions (cont'd)

- We are interested in predicting survival probabilities for a new patient j that has provided a set of serum bilirubin measurements up to a specific time point t
- Example: We consider Patients 2 and 25 from the PBC dataset that have provided us with 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 marker
 - ▷ providing measurements up to time point $t \Rightarrow$ the patient was still alive at time t

6.1 Survival Probabilities: Definitions (cont'd)



6.1 Survival Probabilities: Definitions (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

6.2 Survival Probabilities: Estimation

- 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)

6.2 Survival Probabilities: Estimation (cont'd)

- $\pi_j(u \mid t)$ can be rewritten as

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

- A naive estimator for $\pi_j(u \mid t)$ can be constructed by plugging-in the MLEs and the Empirical Bayes estimates

$$\tilde{\pi}_j(u \mid t) = \frac{S_j\{u \mid \mathcal{M}_j(u, \hat{b}_j, \hat{\theta}); \hat{\theta}\}}{S_j\{t \mid \mathcal{M}_j(t, \hat{b}_j, \hat{\theta}); \hat{\theta}\}}$$

- ▷ this works relatively well in practice, but
- ▷ standard errors are difficult to compute

6.2 Survival Probabilities: Estimation (cont'd)

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

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

- We have already seen the first part of the integrand

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

6.2 Survival Probabilities: Estimation (cont'd)

- Provided that the sample size is sufficiently large, we can approximate the posterior of the parameters by

$$\{\theta \mid \mathcal{D}_n\} \sim \mathcal{N}(\hat{\theta}, \hat{\mathcal{H}}),$$

where

- ▷ $\hat{\theta}$ are the MLEs, and
- ▷ $\hat{\mathcal{H}}$ their asymptotic covariance matrix

6.2 Survival Probabilities: Estimation (cont'd)

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

Step 1. draw $\theta^{(\ell)} \sim \mathcal{N}(\hat{\theta}, \hat{\mathcal{H}})$

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

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

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

6.2 Survival Probabilities: Estimation (cont'd)

- Steps 1 and 3 are straightforward
- In Step 2 we need to sample from $\{b_j \mid T_j^* > t, \mathcal{Y}_j(t), \theta^{(\ell)}\}$, which is nonstandard
 - ▷ as n_i increases, this posterior converges to a multivariate normal distribution (Rizopoulos et al., Biometrika, 2008)
 - ▷ we use a Metropolis-Hastings algorithm with multivariate t proposals

6.2 Survival Probabilities: Estimation (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: Linear & quadratic time, treatment and their interaction
 - ▷ random effects: Intercept, linear & quadratic time effects
- Survival submodel
 - ▷ treatment effect + *underlying* serum bilirubin level
 - ▷ piecewise-constant baseline hazard in 7 intervals

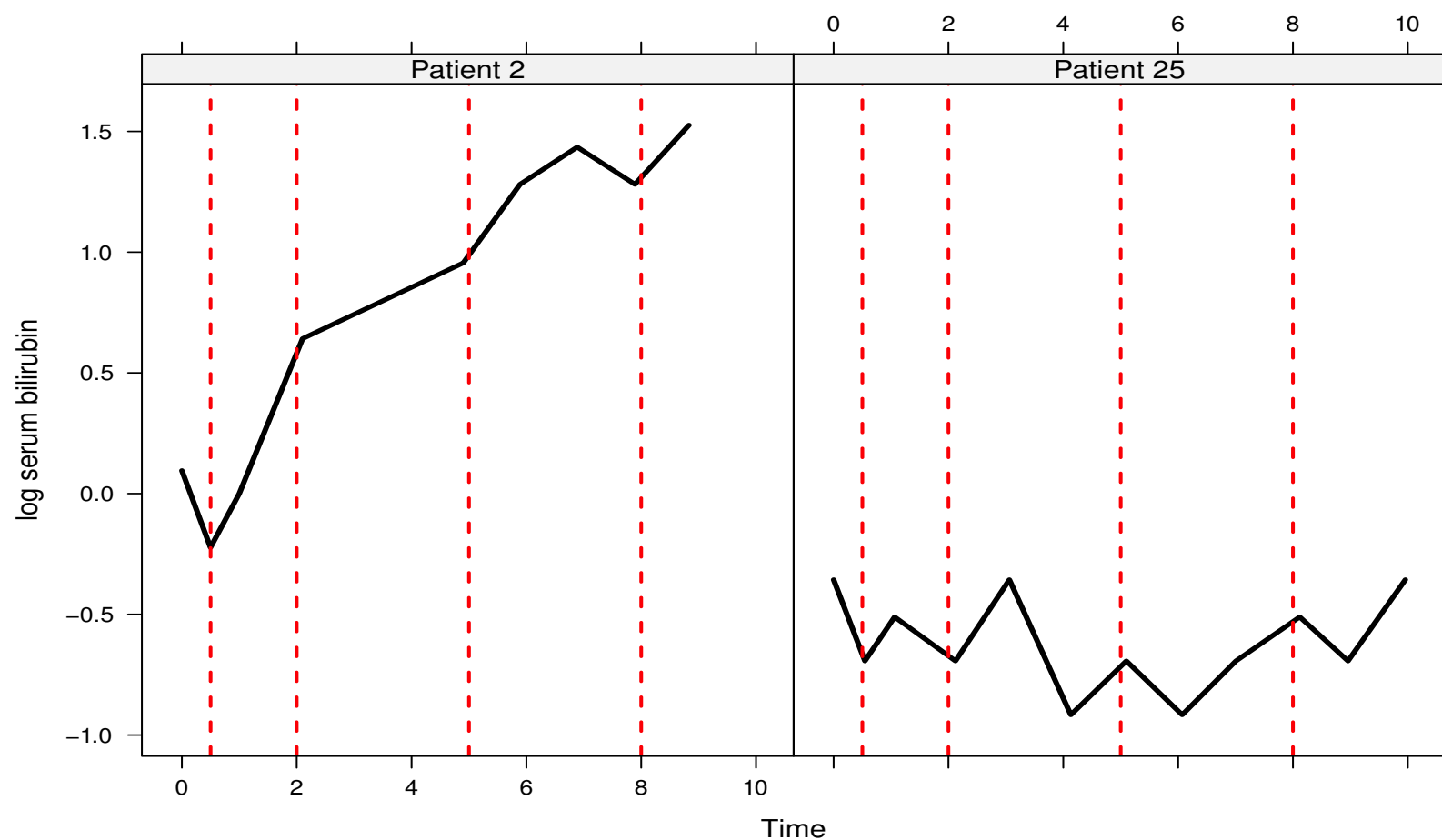
6.2 Survival Probabilities: Estimation (cont'd)

- 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

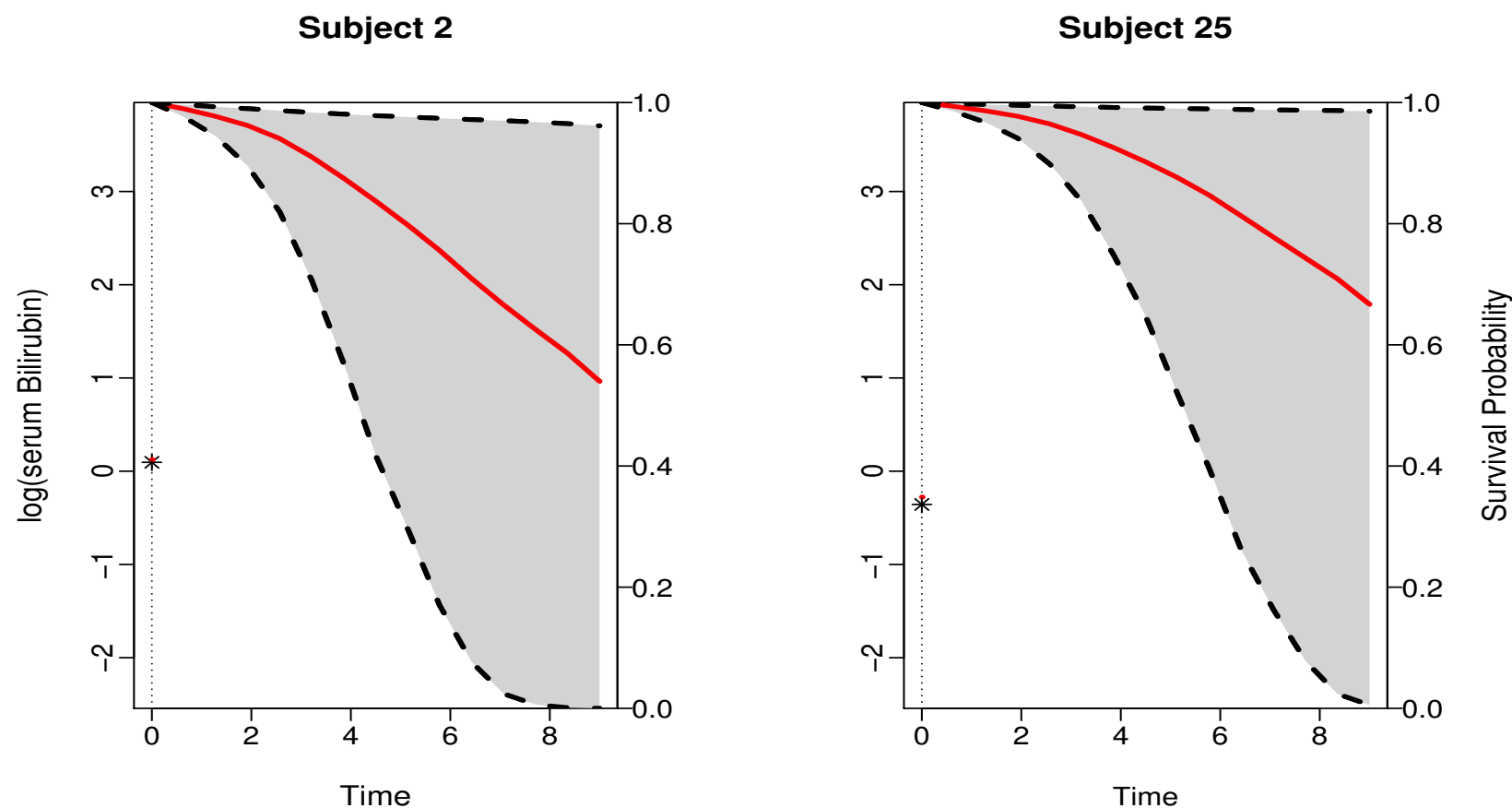
$$\hat{\pi}_j(u \mid t) = \text{median}\{\pi_j^{(\ell)}(u \mid t), \ell = 1, \dots, L\}$$

and calculated a corresponding 95% pointwise CIs

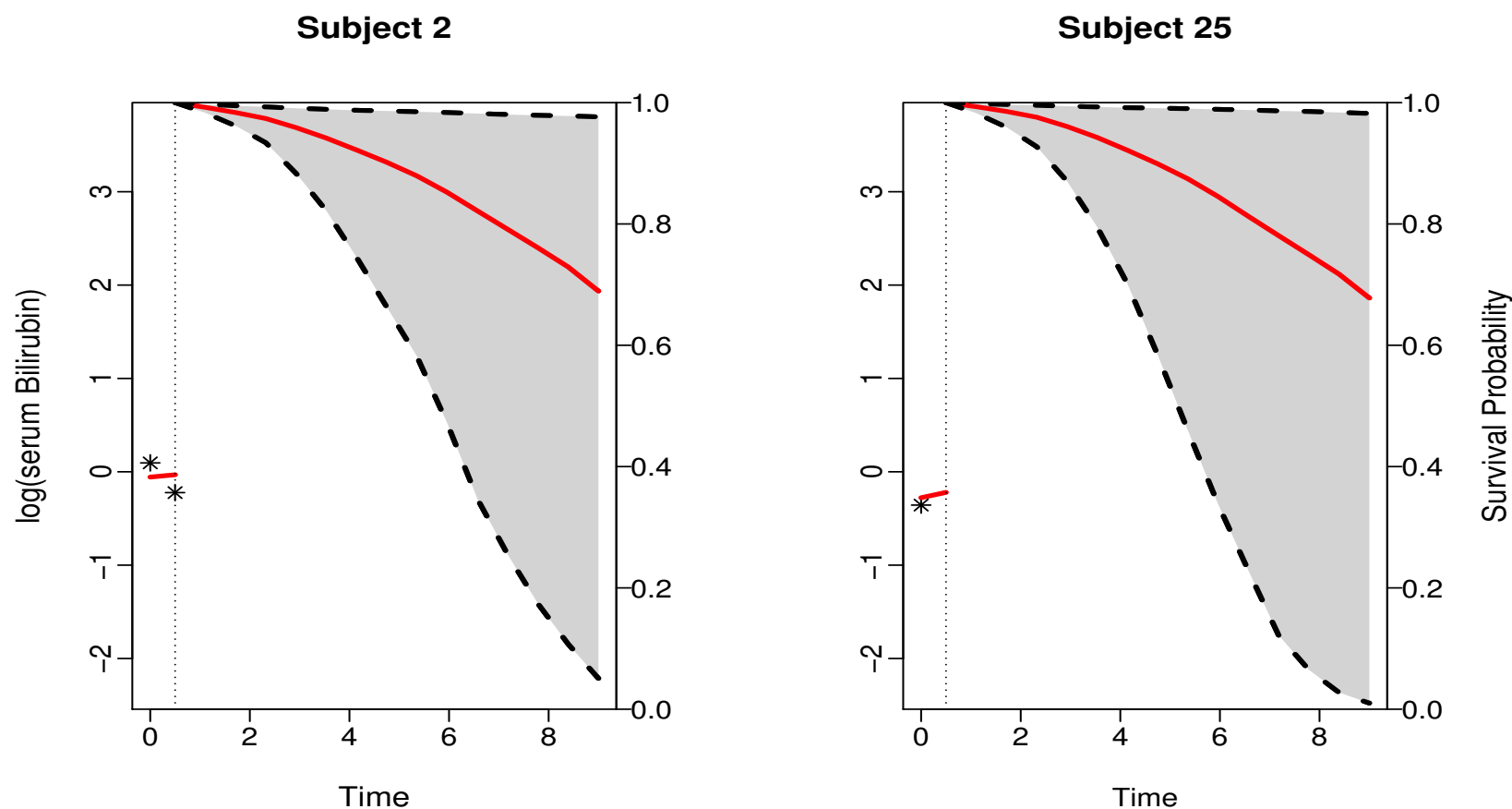
6.2 Survival Probabilities: Estimation (cont'd)



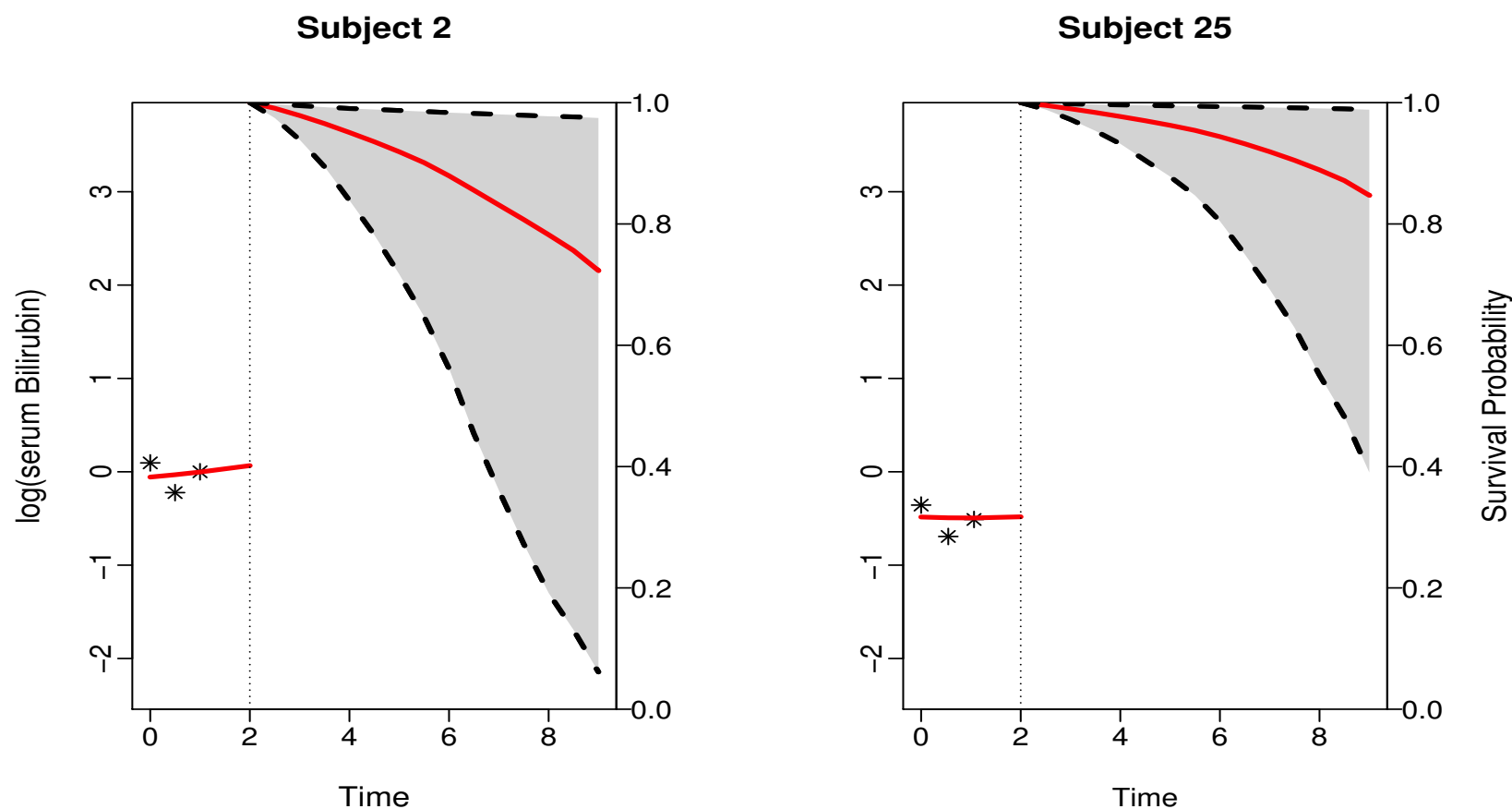
6.2 Survival Probabilities: Estimation (cont'd)



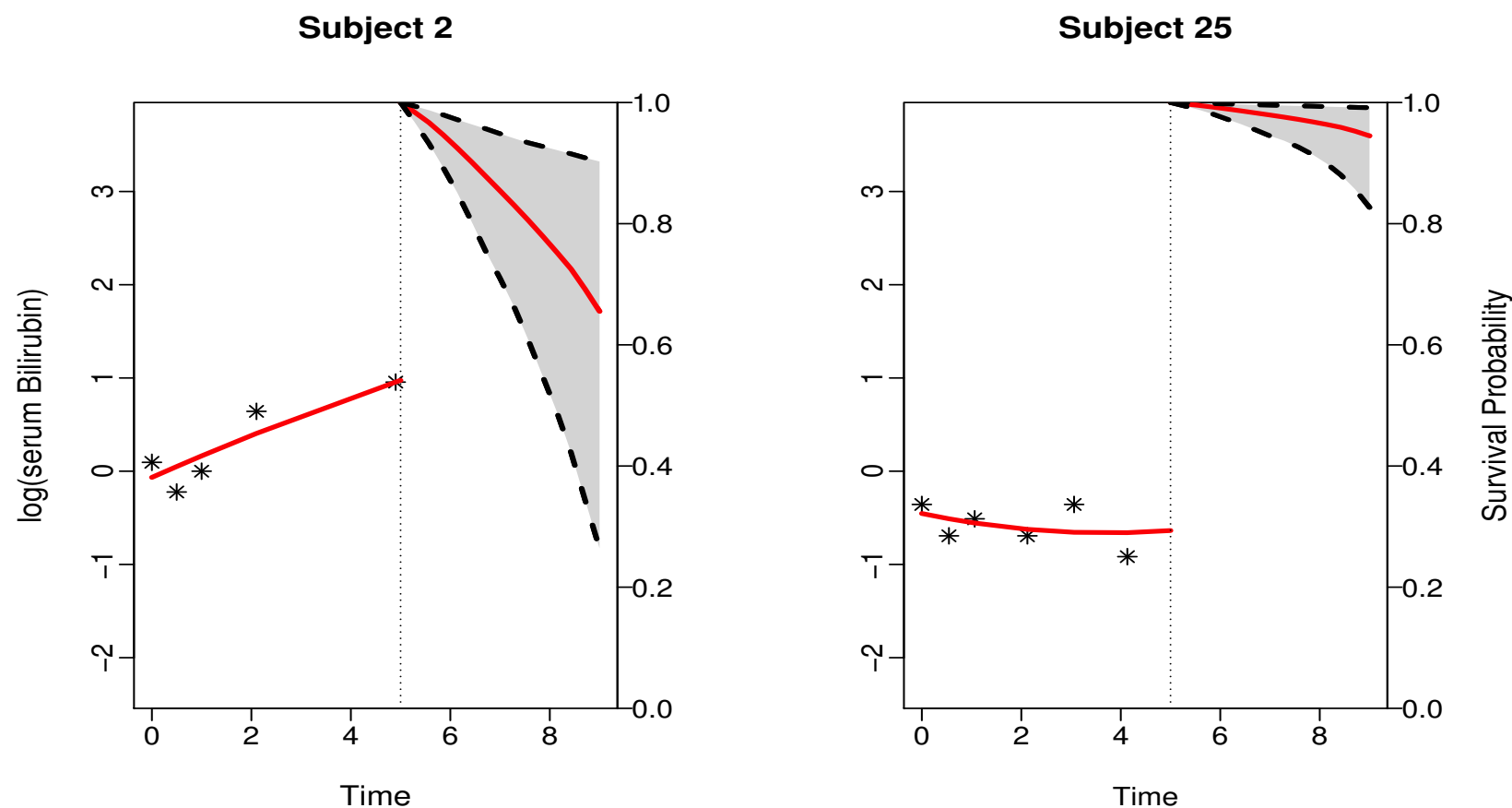
6.2 Survival Probabilities: Estimation (cont'd)



6.2 Survival Probabilities: Estimation (cont'd)

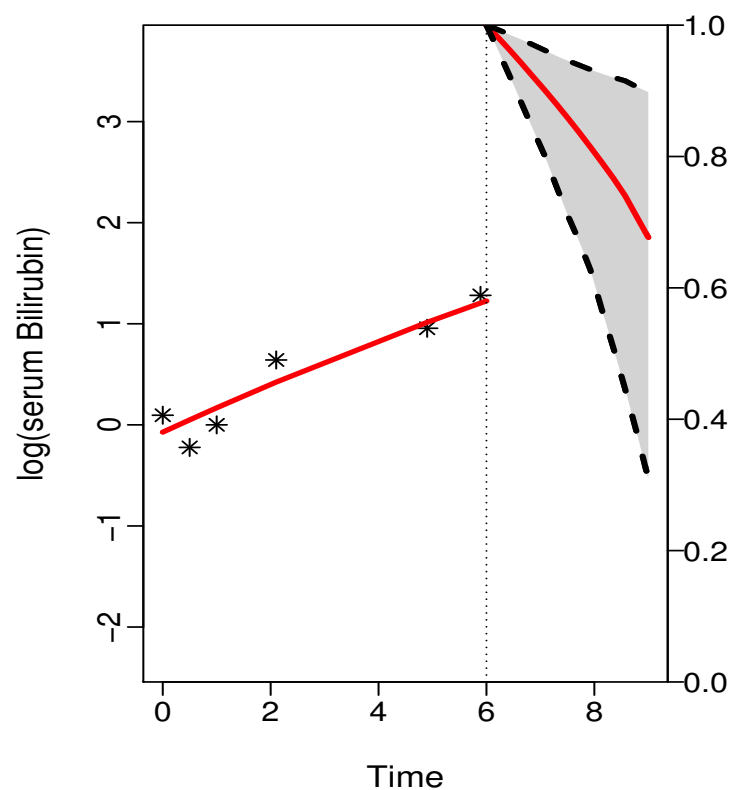


6.2 Survival Probabilities: Estimation (cont'd)

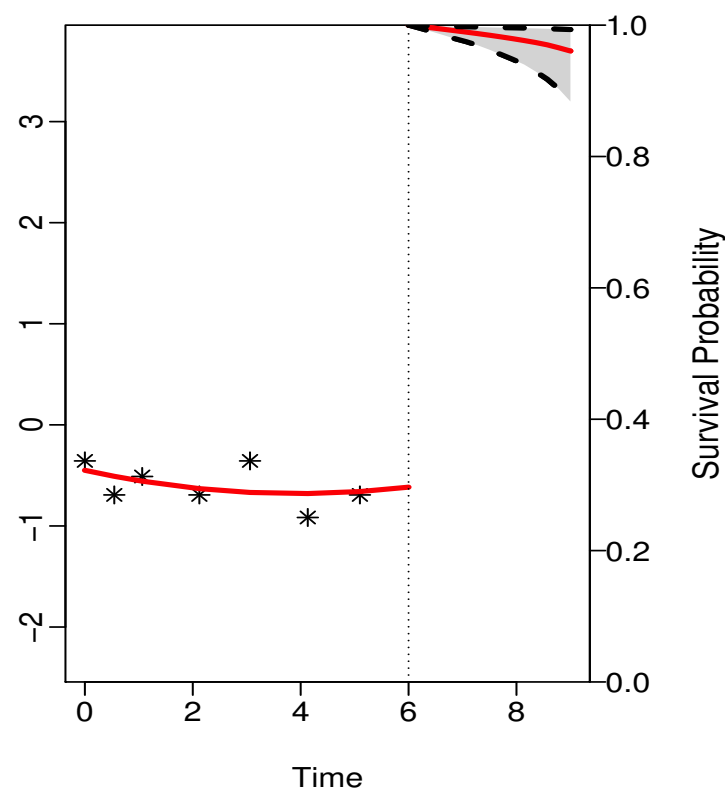


6.2 Survival Probabilities: Estimation (cont'd)

Subject 2

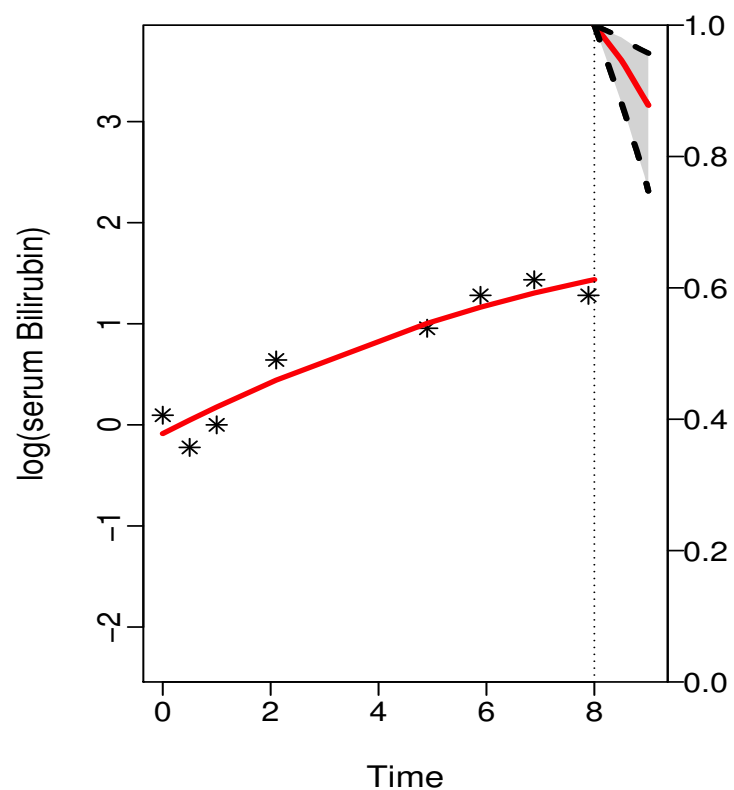


Subject 25

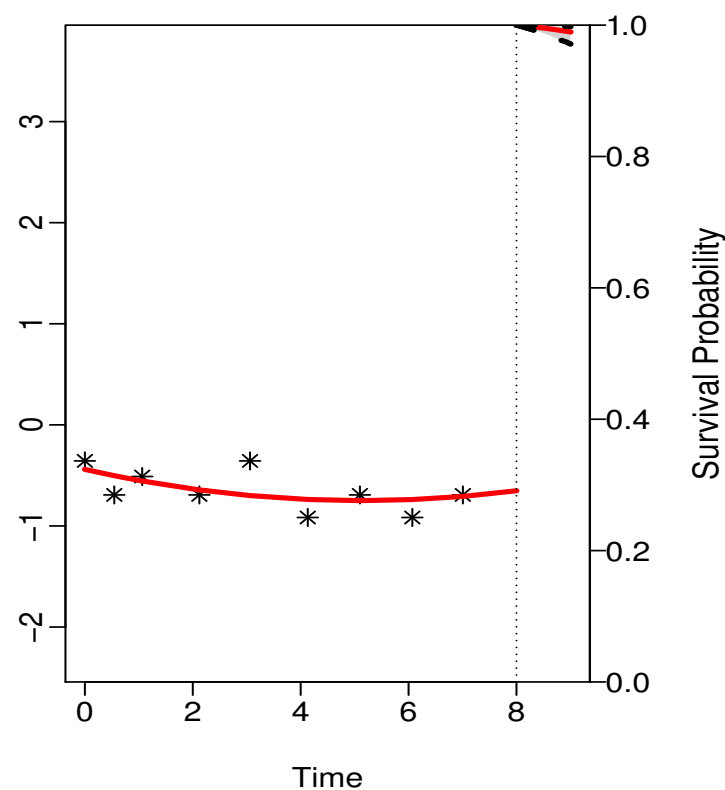


6.2 Survival Probabilities: Estimation (cont'd)

Subject 2



Subject 25



6.2 Survival Probabilities: Estimation (cont'd)

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

```
sfit <- survfitJM(jointFit, newdata = pbc2[pbc2$id == "2", ])
```

```
sfit
```

```
plot(sfit)
```

```
plot(sfit, include.y = TRUE)
```

6.3 Longitudinal Responses: Definitions*

- In some occasions it may be also of interest to predict the longitudinal outcome
- We can proceed in the same manner as for the survival probabilities: 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

$$\omega_j(u \mid t) = E\{y_j(u) \mid T_j^* > t, \mathcal{Y}_j(t), \mathcal{D}_n\}, \quad u > t$$

6.3 Longitudinal Responses: Definitions* (cont'd)

- To estimate $\omega_j(u | t)$ we can follow a similar approach as for $\pi_j(u | t)$ – Namely, $\omega_j(u | t)$ is written as:

$$E\{y_j(u) | T_j^* > t, \mathcal{Y}_j(t), \mathcal{D}_n\} = \int E\{y_j(u) | T_j^* > t, \mathcal{Y}_j(t), \mathcal{D}_n; \theta\} p(\theta | \mathcal{D}_n) d\theta$$

- With the first part of the integrand given by:

$$\begin{aligned} E\{y_j(u) | T_j^* > t, \mathcal{Y}_j(t), \mathcal{D}_n; \theta\} &= \\ &= \int \{x_j^\top(u)\beta + z_j^\top(u)b_j\} p(b_j | T_j^* > t, \mathcal{Y}_j(t); \theta) db_j \end{aligned}$$

6.3 Longitudinal Responses: Estimation* (cont'd)

- A similar Monte Carlo simulation scheme:

Step 1. draw $\theta^{(\ell)} \sim \mathcal{N}(\hat{\theta}, \hat{\mathcal{H}})$

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

Step 3. compute $\omega_j^{(\ell)}(u \mid t) = x_j^\top(u)\beta^{(\ell)} + z_j^\top(u)b_j^{(\ell)}$

- Note: Prediction intervals can be easily computed by replacing Step 3 with a draw from:

$$\omega_j^{(\ell)}(u \mid t) \sim \mathcal{N}\left\{x_j^\top(u)\beta^{(\ell)} + z_j^\top(u)b_j^{(\ell)}, \quad [\sigma^2]^{(\ell)}\right\}$$

6.3 Longitudinal Responses: Estimation* (cont'd)

- Example: Dynamic predictions of serum bilirubin for Patients 2 & 25 from the PBC dataset: We fit the joint model
- Longitudinal submodel
 - ▷ fixed effects: Linear & quadratic time, treatment and their interaction
 - ▷ random effects: Intercept, linear & quadratic time effects
- Survival submodel
 - ▷ treatment effect + *underlying* serum bilirubin level
 - ▷ piecewise-constant baseline hazard in 7 intervals

6.3 Longitudinal Responses: Estimation* (cont'd)

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

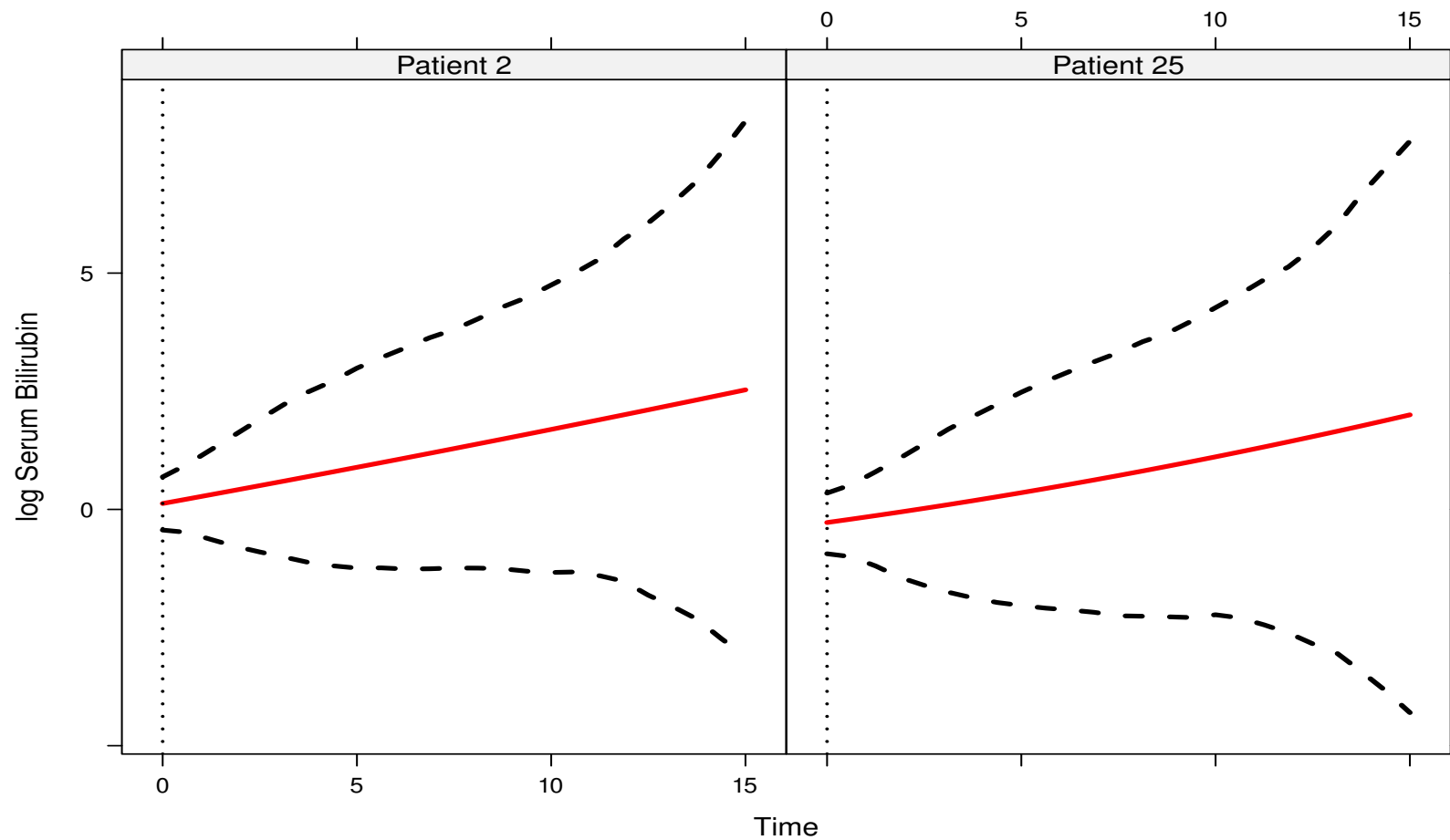
- Point estimates

$$\hat{\omega}_j(u | t) = x_j^\top(u) \hat{\beta} + z_j^\top(u) \hat{b}_j,$$

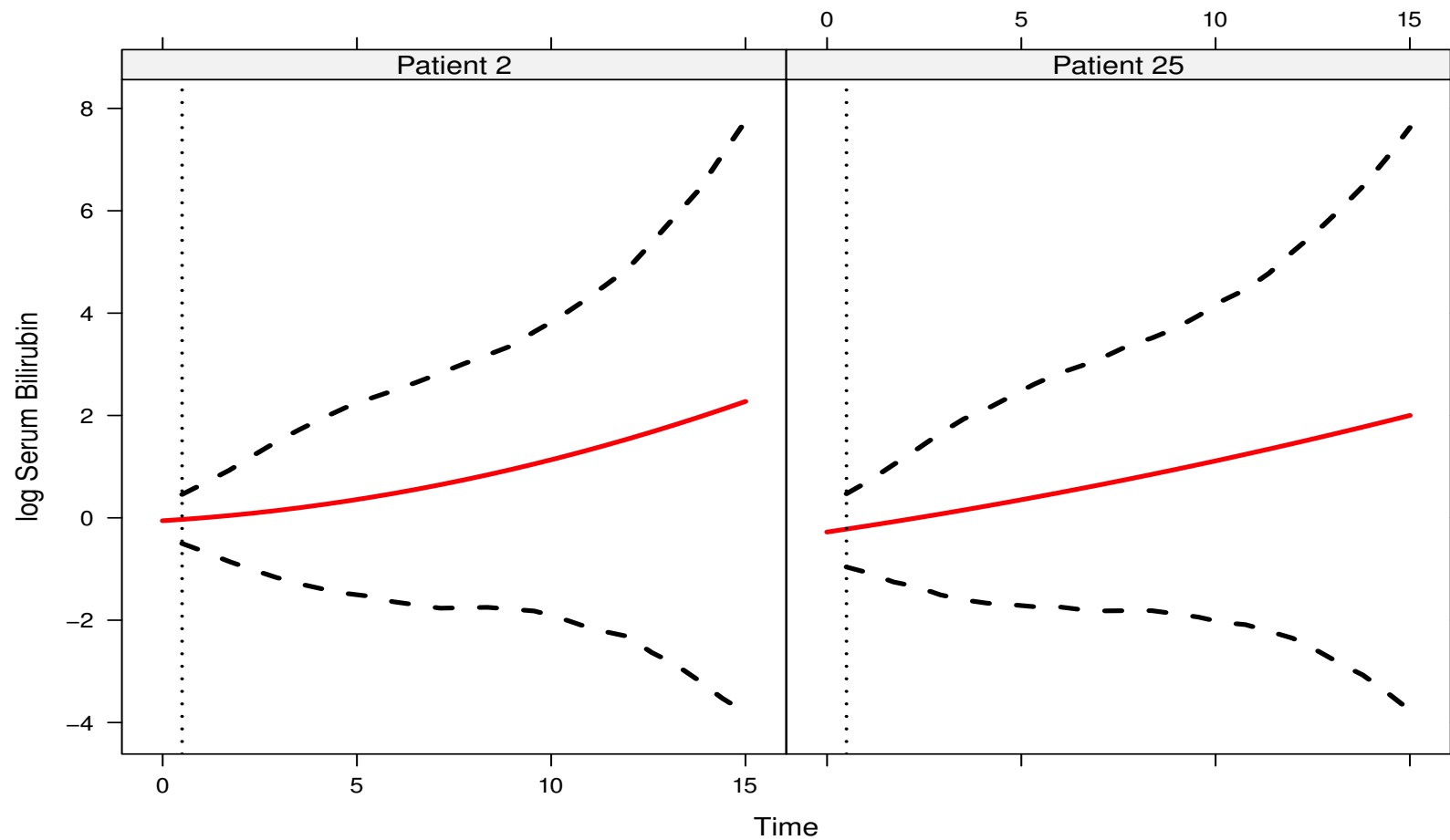
where $\hat{\beta}$: MLEs & \hat{b}_j : empirical Bayes estimates

- 95% pointwise CIs
 - ▷ simulation scheme: 2.5% and 97.5% percentiles of 500 Monte Carlo samples of $\omega_j^{(\ell)}(u | t)$

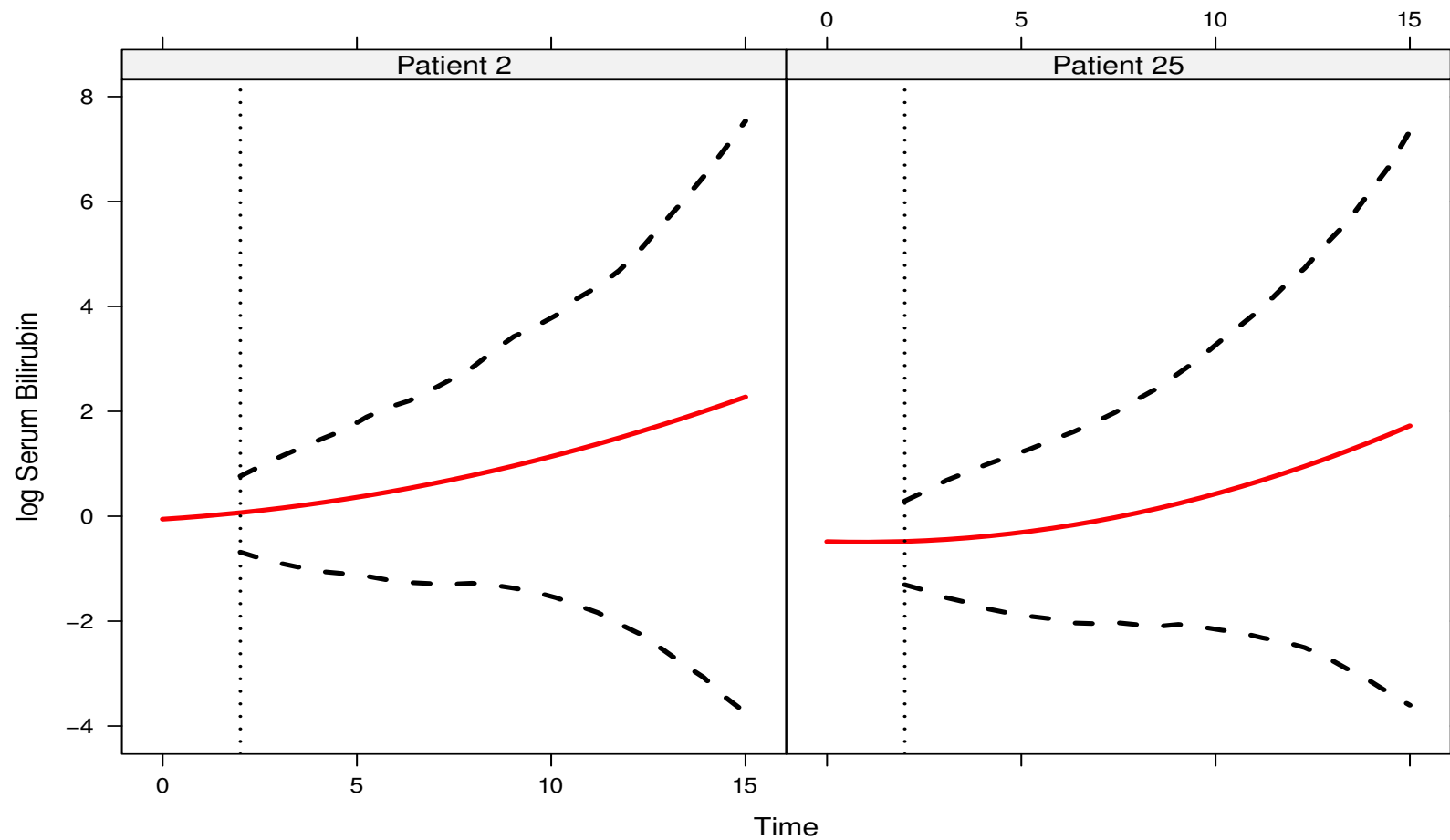
6.3 Longitudinal Responses: Estimation* (cont'd)



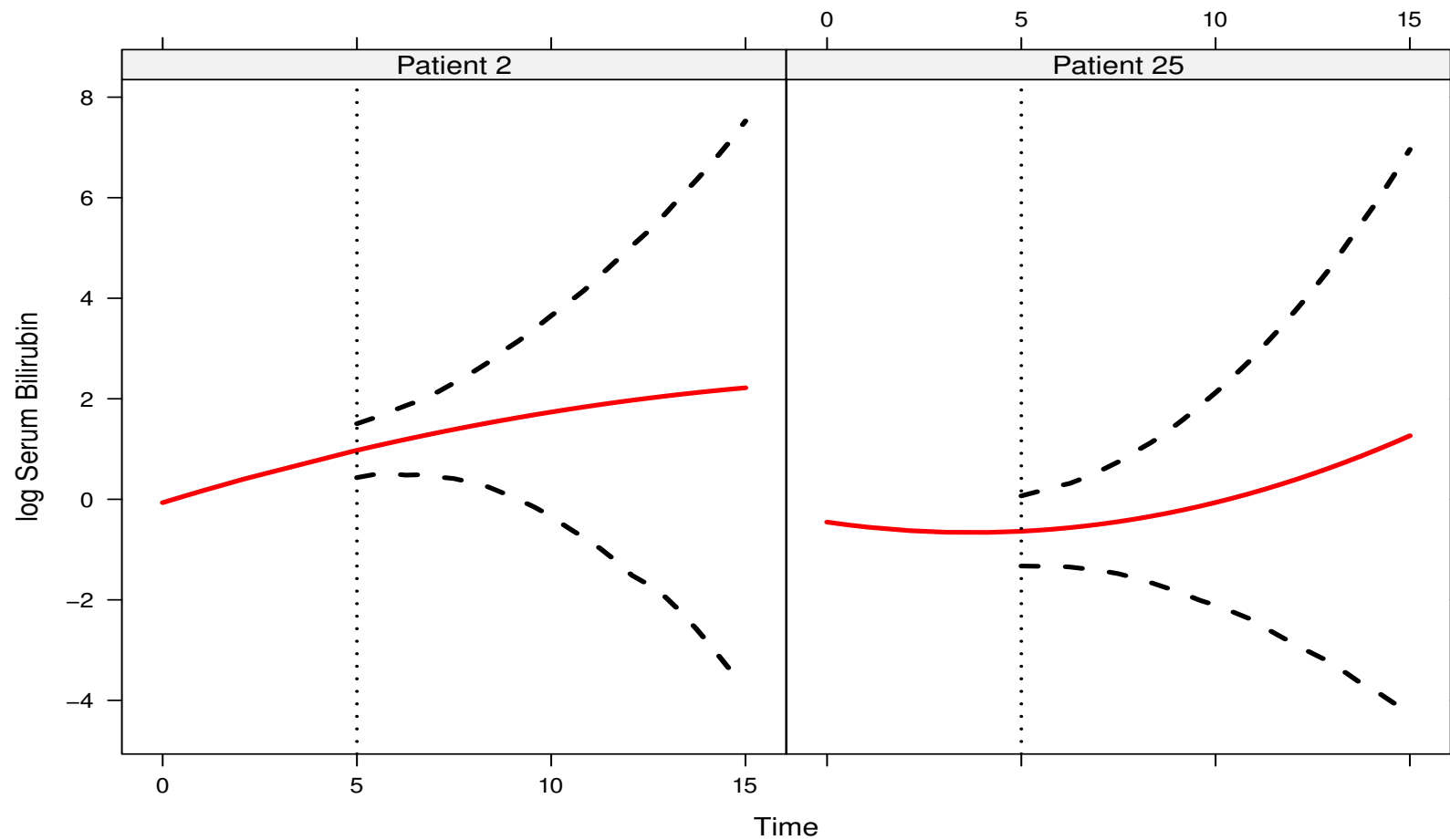
6.3 Longitudinal Responses: Estimation* (cont'd)



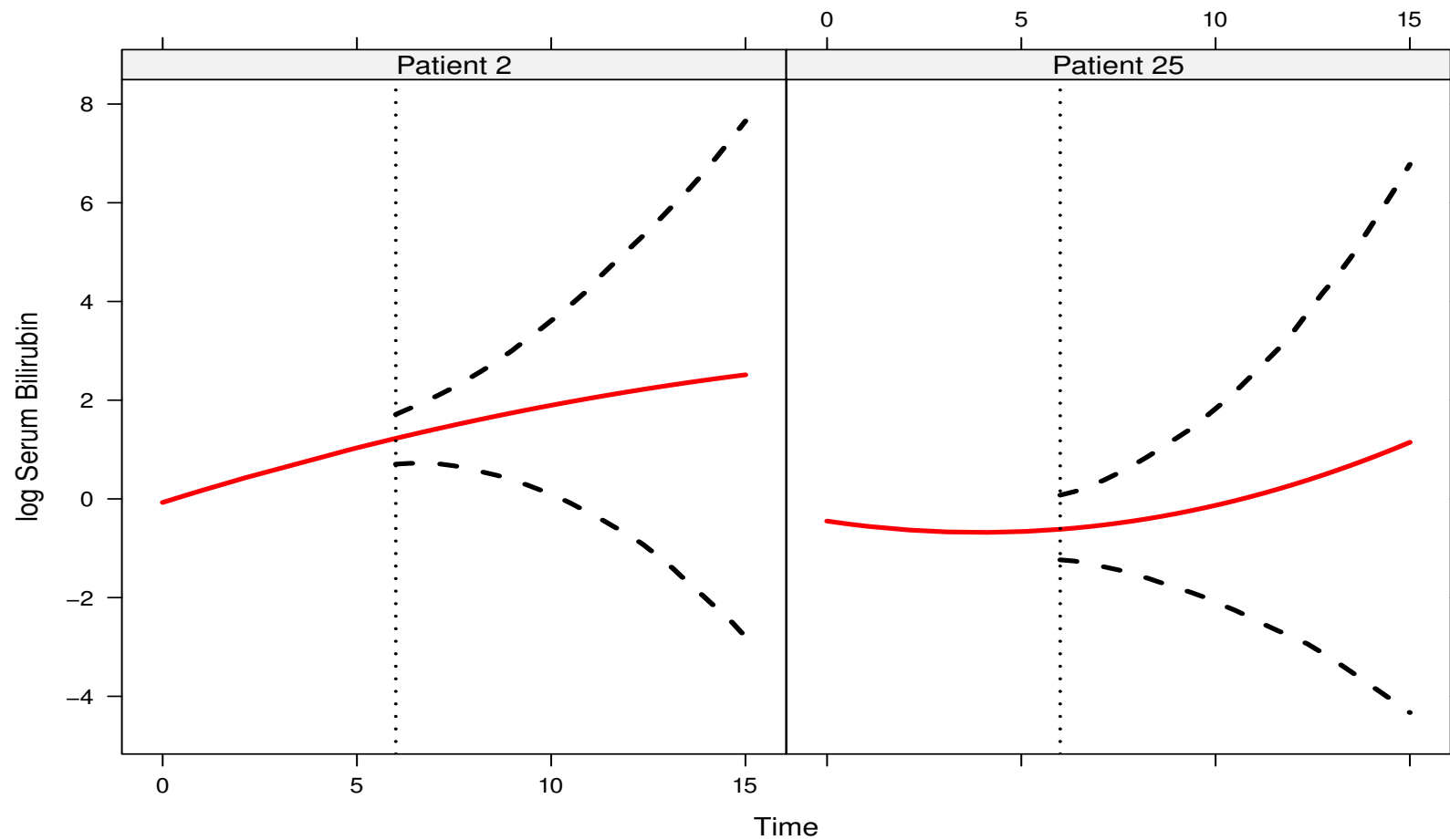
6.3 Longitudinal Responses: Estimation* (cont'd)



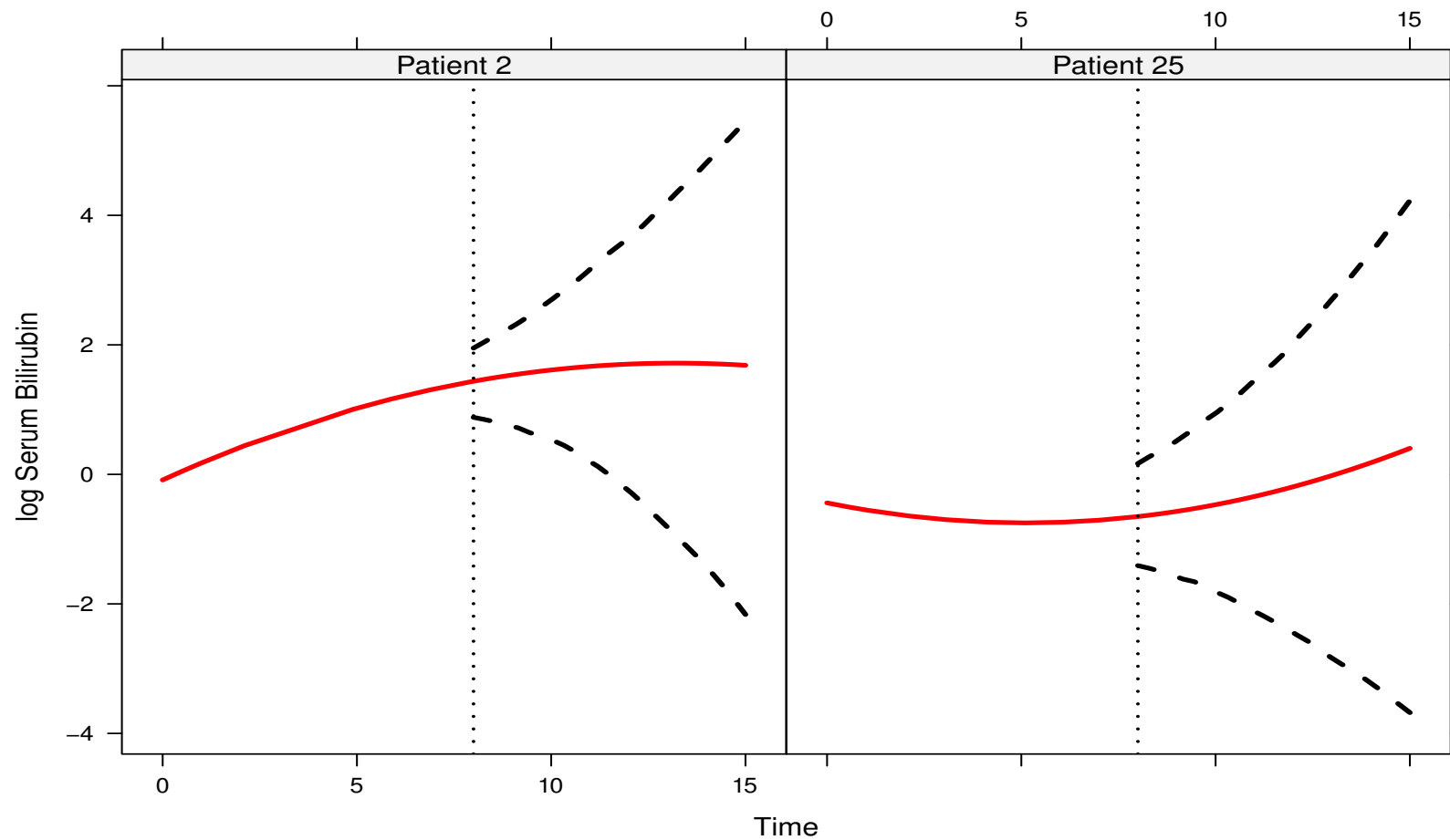
6.3 Longitudinal Responses: Estimation* (cont'd)



6.3 Longitudinal Responses: Estimation* (cont'd)



6.3 Longitudinal Responses: Estimation* (cont'd)



6.3 Longitudinal Responses: Estimation* (cont'd)

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

```
lfit <- predict(jointFit, newdata = pbc2[pbc2$id == "2", ],  
               type = "Subject", interval = "conf", returnData = TRUE)
```

```
lfit
```

```
xyplot(pred + low + upp ~ year, data = lfit, type = "l",  
        lty = c(1,2,2), col = c(2,1,1), lwd = 2)
```

6.3 Longitudinal Responses: Estimation* (cont'd)

R> Web interface using the **shiny** package

```
library("shiny")
```

```
JMbayes::runDynPred()
```

6.4 Importance of the Parameterization

- All previous predictions were based on the standard joint model

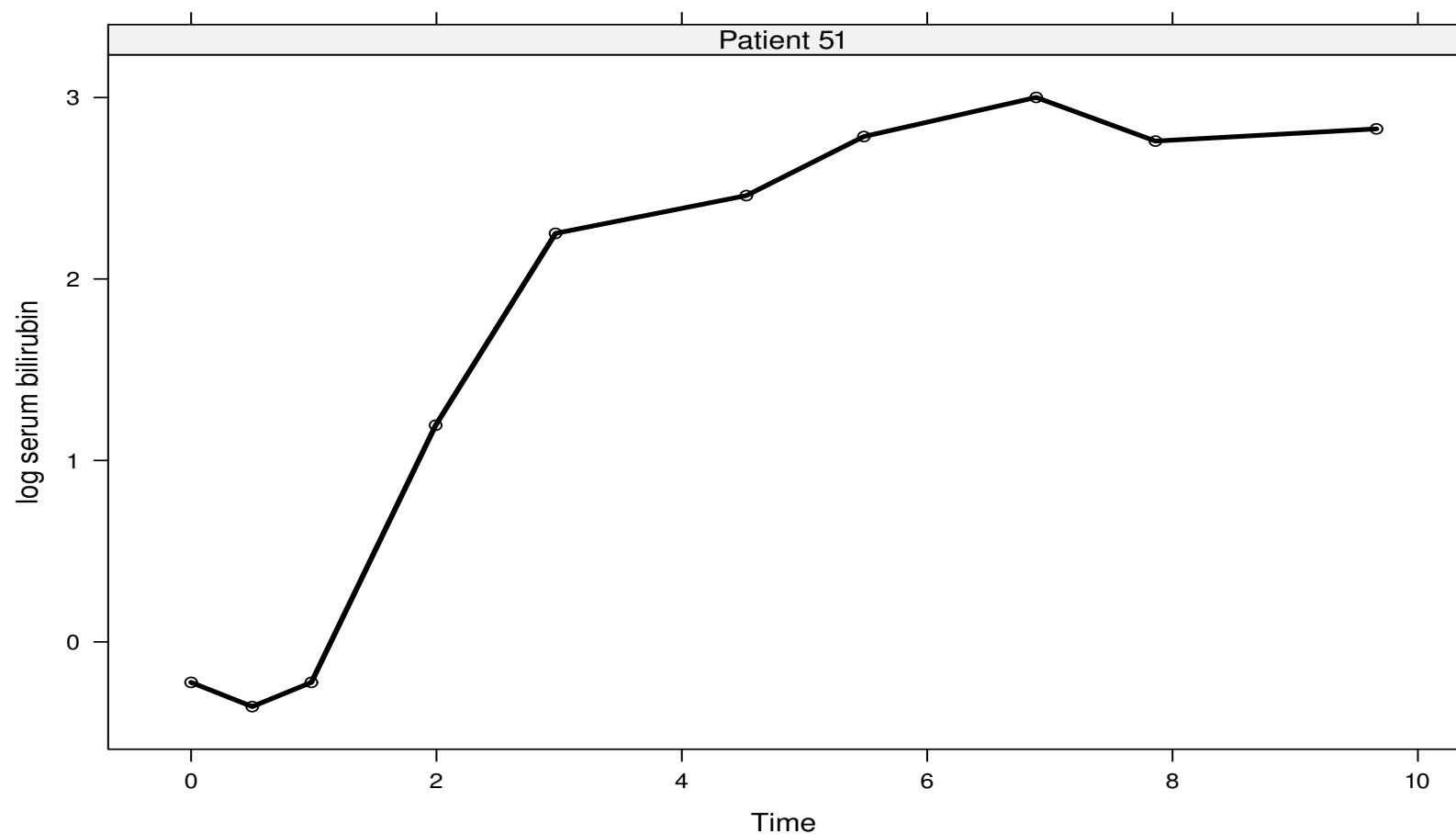
$$\left\{ \begin{array}{l} h_i(t \mid \mathcal{M}_i(t)) = h_0(t) \exp\{\gamma^\top w_i + \alpha m_i(t)\}, \\ y_i(t) = m_i(t) + \varepsilon_i(t) \\ y_i(t) = x_i^\top(t)\beta + z_i^\top(t)b_i + \varepsilon_i(t), \end{array} \right.$$

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

6.4 Importance of the Parameterization (cont'd)

- We have seen earlier that there are several alternative parameterizations (see Section 5.1)
- Relevant questions:
 - ▷ Does the assumed parameterization affect predictions?
 - ▷ Which parameterization is the most optimal?
- Example: We compare predictions for the longitudinal and survival outcomes under different parameterizations for Patient 51 from the PBC study

6.4 Importance of the Parameterization (cont'd)



6.4 Importance of the Parameterization (cont'd)

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

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

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

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

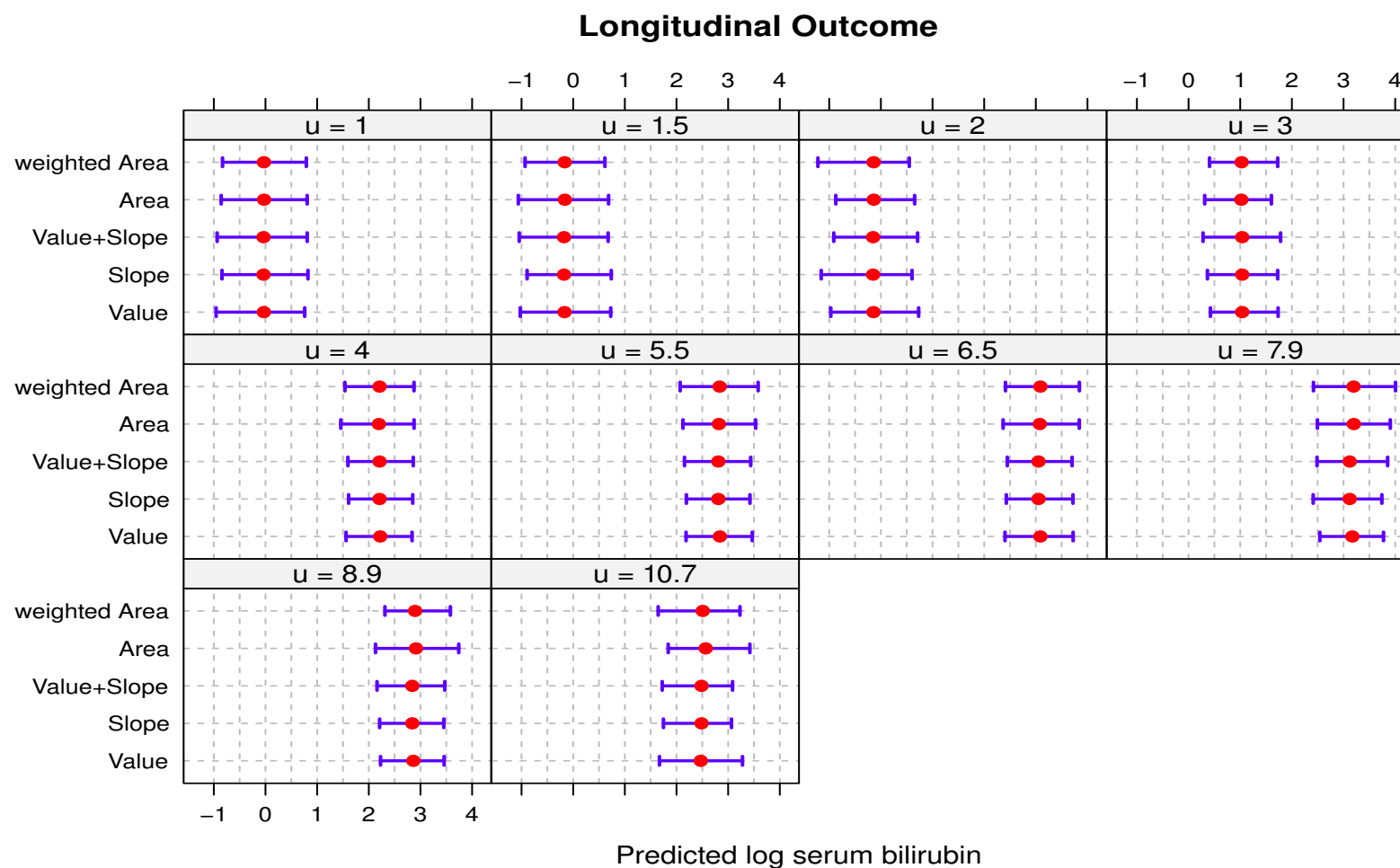
6.4 Importance of the Parameterization (cont'd)

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

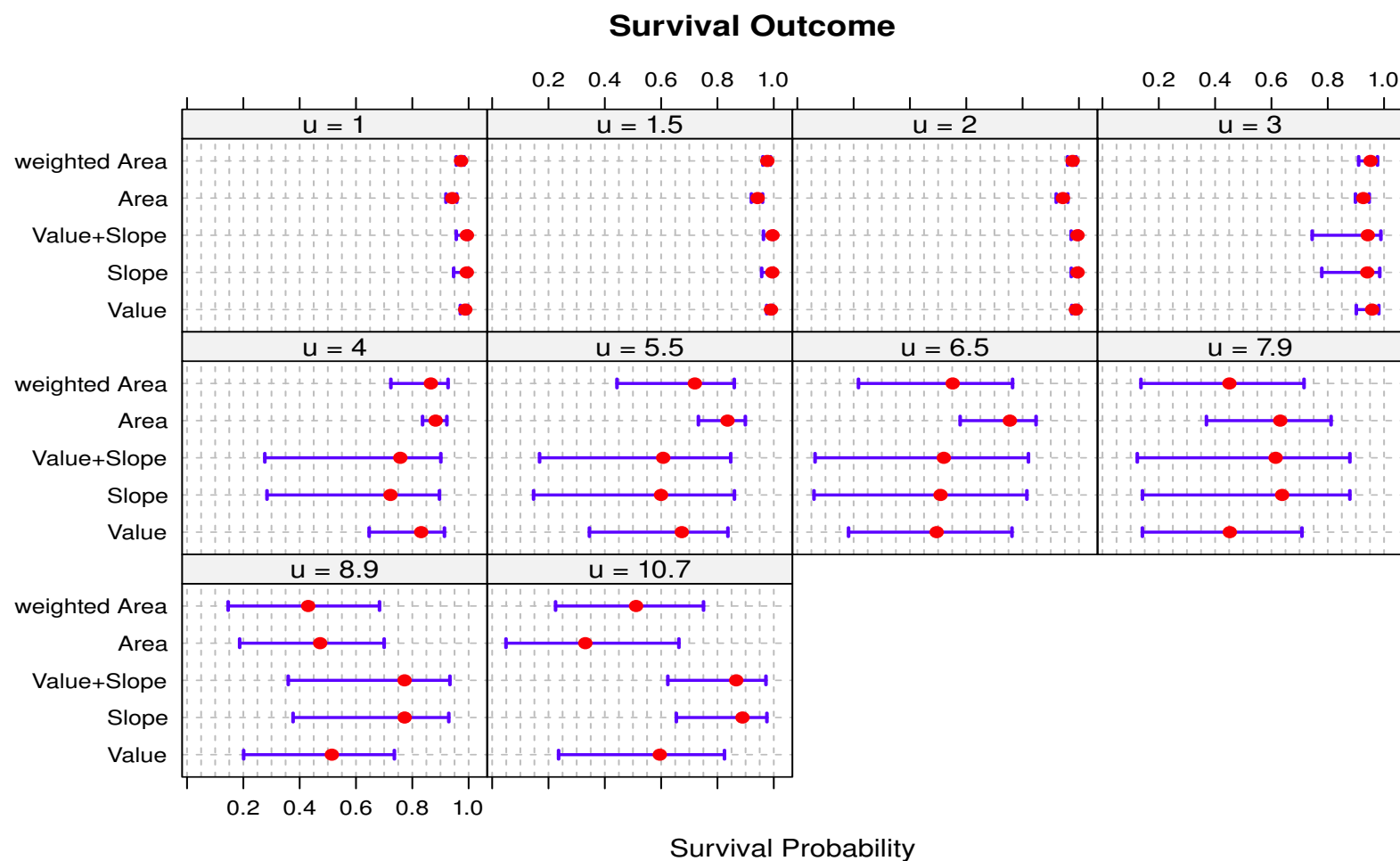
$$h_i(t) = h_0(t) \exp \left\{ \gamma \text{D-pnc}_i + \alpha_4 \int_0^t \phi(t-s) m_i(s) ds \right\},$$

where $\phi(\cdot)$ standard normal pdf

6.4 Importance of the Parameterization (cont'd)



6.4 Importance of the Parameterization (cont'd)



6.4 Importance of the Parameterization (cont'd)

- The chosen parameterization can influence the derived predictions
 - ▷ especially for the survival outcome
- My current work: How to optimally choose parameterization?
 - ▷ per subject (personalized medicine)
- Quite promising results from the Bayesian approach using Bayesian Model Averaging techniques
 - ▷ it can be done with package **JMbayes**,
 - ▷ it falls a bit outside the scope of this course, but
 - ▷ I can provide information if interested...

Chapter 7

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
 - ▷ use parametric but flexible models for the baseline hazard function
 - ▷ consider how to model the association structure between the two processes
⇒ Parameterization
- **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
- ▷ \Rightarrow joint models constitute an excellent tool for personalized medicine

- **What we did not cover**

- ▷ assessment of predictive performance
- ▷ diagnostics for joint models using residuals
- ▷ ...

The End!

7.2 Additional References

- Andrinopoulou, E.R., Eilers, P.H.C., Takkenberg, J.J.M. and Rizopoulos, D. (2017). Improved dynamic predictions from joint models of longitudinal and survival data with time-varying effects using P-splines. *Biometrics*, to appear.
- Andrinopoulou, E.R., Rizopoulos, D., Takkenberg, J.J.M. and Lesaffre, E. (2017). Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. *Statistical Methods in Medical Research*, to appear.
- Andrinopoulou, E.R., Rizopoulos, D., Takkenberg, J.J.M. and Lesaffre, E. (2014). Joint modeling of two longitudinal outcomes and competing risk data. *Statistics in Medicine* **33**, 3167–3178.
- 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.

7.2 Additional References (cont'd)

- 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.
- 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.
- Garcia-Hernandez, A. and Rizopoulos, D. (2017). %JM: A SAS macro to fit jointly generalized mixed models for longitudinal data and time-to-event responses. *Journal of Statistical Software*, to appear.
- 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.

7.2 Additional References (cont'd)

- 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.
- 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.
- Mauff, K., Steyerberg, E.W., Nijpels, G., van der Heijden, A. and Rizopoulos, D. (2017). Extension of the association structure in joint models to include weighted cumulative effects. *Statistics in Medicine*, to appear.
- 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.

7.2 Additional References (cont'd)

- 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. (2016). The R package JMBayes for fitting joint models for longitudinal and time-to-event data using MCMC. *Journal of Statistical Software* **72**(7), 1–45.
- 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.
- 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.

7.2 Additional References (cont'd)

- 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. and Lesaffre, E. (2014). Introduction to the special issue on joint modelling techniques. *Statistical Methods in Medical Research* **23**, 3–10.
- Rizopoulos, D., Molenberghs, G. and Lesaffre, E. (2017). Dynamic Predictions with Time-Dependent Covariates in Survival Analysis using Joint Modeling and Landmarking. *Biometrical Journal*, to appear.
- Rizopoulos, D., Taylor, J.M.G., van Rosmalen, J., Steyerberg, E.W. and Takkenberg, J.J.M. (2016). Personalized screening intervals for biomarkers using joint models for longitudinal and survival data. *Biostatistics* **17**, 149–164.
- 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.
- 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.

7.2 Additional References (cont'd)

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

7.2 Additional References (cont'd)

- 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.
- 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.

7.2 Additional References (cont'd)

- 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.
- 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.

7.2 Additional References (cont'd)

- 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.

Practicals

Practical 1: A Simple Joint Model

- We will fit a simple joint model to the PBC dataset
- Start R and load package **JM**, using `library(JM)`
- The longitudinal (long format) and survival information for the PBC patients can be found in data frames `pb2` and `pb2.id`. The variables that we will need are:

Practical 1: A Simple Joint Model (cont'd)

▷ pbc2

- * `id`: patient id number
- * `serBilir`: serum bilirubin
- * `year`: follow-up times in years

▷ pbc2.id

- * `years`: observed event times in years
- * `status`: 'alive', 'transplanted', 'dead'
- * `drug`: treatment indicator

Practical 1: A Simple Joint Model (cont'd)

- **T1:** Fit the linear mixed effects model for log serum bilirubin using function `lme()`, assuming simple linear evolutions in time for each subject, i.e., a simple random-intercepts and random-slopes structure and different average evolutions per treatment group (see pp. 31–35)

$$y_i(t) = \beta_0 + \beta_1 t + \beta_2 \{\text{D-penic}_i \times t\} + b_{i0} + b_{i1} t + \varepsilon_i(t)$$

- **T2:** Create the indicator for the composite event (i.e., ‘alive’ = 0, ‘transplanted’ or ‘dead’ = 1) using the code

```
pb2c.id$status2 <- as.numeric(pbc2.id$status != "alive")
```

Practical 1: A Simple Joint Model (cont'd)

- **T3:** Fit the Cox PH model using `coxph()` that includes only treatment as baseline covariate, remember to set `x = TRUE` (see pp. 55–56)
- We want to fit the joint model

$$\left\{ \begin{array}{l} y_i(t) = m_i(t) + \varepsilon_i(t) \\ \quad = \beta_0 + \beta_1 t + \beta_2 \{\text{D-penic}_i \times t\} + b_{i0} + b_{i1} t + \varepsilon_i(t), \quad \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2), \\ h_i(t) = h_0(t) \exp\{\gamma \text{D-penic}_i + \alpha m_i(t)\}, \end{array} \right.$$

Practical 1: A Simple Joint Model (cont'd)

- **T4:** Fit this joint model based on the fitted linear mixed and Cox models using function `jointModel()` (see pp. 93–95)
 - ▷ with piecewise-constant baseline hazard & the (pseudo) adaptive GH rule
- **T5:** Use the `summary()` method to obtain a detailed output of the fitted joint model – interpret the results
- **T6:** Produce 95% confidence intervals for the parameters in the longitudinal submodel, and for the hazard ratios in the survival submodel using function `confint()` (the `parm` argument of `confint()` can take as values `"all"` (default), `"Longitudinal"` and `"Event"`)

Practical 1: A Simple Joint Model (cont'd)

- This model assumes that the strength of the association between the level of serum bilirubin and the risk of the composite event is the same in the two treatment groups
- To relax this additivity assumption we will add the interaction effect between serum bilirubin and treatment

$$\left\{ \begin{array}{l} y_i(t) = m_i(t) + \varepsilon_i(t) \\ \quad = \beta_0 + \beta_1 t + \beta_2 \{\text{D-penic}_i \times t\} + b_{i0} + b_{i1} t + \varepsilon_i(t), \quad \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2), \\ \\ h_i(t) = h_0(t) \exp[\gamma \text{D-penic}_i + \alpha_1 m_i(t) + \alpha_2 \{\text{D-penic}_i \times m_i(t)\}], \end{array} \right.$$

Practical 1: A Simple Joint Model (cont'd)

- To fit this model with package **JM** we need to define the `interFact` argument of `jointModel()`. This should be a named `list` with two elements:
 - ▷ `value`: a formula with the factors for which we wish to calculate the interaction terms
 - ▷ `data`: the data frame used to fit the Cox model
- T7: Define this list and fit the corresponding joint model. Use the `summary()` method to obtain a detailed output and interpret the results

Practical 1: A Simple Joint Model (cont'd)

- Based on the fitted joint model we can test for three treatment effects, namely
 - ▷ in the longitudinal process:

$$H_0 : \beta_2 = 0$$

- ▷ in the survival process:

$$H_0 : \gamma = \alpha_2 = 0$$

- ▷ in the joint process:

$$H_0 : \beta_2 = \gamma = \alpha_2 = 0$$

Practical 1: A Simple Joint Model (cont'd)

- We would like test these hypotheses using likelihood ratio tests
- T8: Fit the three joint models under the corresponding H_0 , and use function `anova()` to perform the LRTs (this function accepts as a first argument the joint model under the null, and as second the joint model under the alternative)

Practical 2: 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 **JM**, using `library(JM)`
- 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:

Practical 2: 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

Practical 2: 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

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

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

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

$h_0(t)$ taken piecewise-constant

Practical 2: Dynamic Predictions (cont'd)

- **T1:** Fit the linear mixed model using `lme()`, the Cox model using `coxph()`, and the corresponding joint model using `jointModel()`
- 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, ]
```

Practical 2: Dynamic Predictions (cont'd)

- **T3:** Using the first measurement of Patient 155, and the fitted joint model calculate his conditional survival probabilities using function `survfitJM()` and plot it using the `plot` method (see p. 158)
- **T4:** Repeat the same procedure by including each time the next measurement of Patient 155 and see how his survival probabilities evolve dynamically in time as extra prothrombin measurements are recorded
 - ▷ check arguments `conf.int` and `fill.area` of the `plot()` method for including the 95% confidence intervals

Practical 2: Dynamic Predictions (cont'd)

- **T5:** Similarly, produce predictions for future longitudinal responses of Patient 155 using the `predict()` method for fitted joint models (see p. 165)
 - ▷ first using only the first measurement,
 - ▷ and following update the predictions after each new longitudinal measurement has been recorded