Fitting Joint Models in R using Packages JM and JMbayes

Dimitris Rizopoulos
Department of Biostatistics, Erasmus Medical Center, the Netherlands

d.rizopoulos@erasmusmc.nl

Joint Statistical Meetings

August 11th, 2015

1.1 Introduction

- 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

1.2 Illustrative Case Study

- Aortic Valve study: Patients who received a human tissue valve in the aortic position
  - ▷ data collected by Erasmus MC (from 1987 to 2008);
    77 received sub-coronary implantation; 209 received root replacement

- Outcomes of interest:
  - ▷ death and re-operation → composite event
  - ▷ aortic gradient

1.3 Research Questions

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

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

1.3 Research Questions (cont'd)

- Focus on multiple outcomes
  - ▷ Complex effect estimation: how strong is the association between the longitudinal **outcome** and the hazard rate of death?
  - ▷ Handling implicit outcomes: focus on the longitudinal outcome but with **dropout**

1.3 Research Questions (cont'd)

In the Aortic Valve dataset:

- Research Question:
  - ▷ Can we utilize available aortic gradient measurements to predict survival/re-operation

1.4 Goals

- 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, ...

1.4 Goals (cont'd)

- **Goals** of this talk:
  - ▷ Introduce joint models
    - * definition
    - * association structures
    - * dynamic predictions
  - ▷ Illustrate software capabilities in R

2.1 Joint Modeling Framework

- To answer our questions of interest we need to postulate a model that relates
  - ▷ the aortic gradient with
  - ▷ the time to death or re-operation

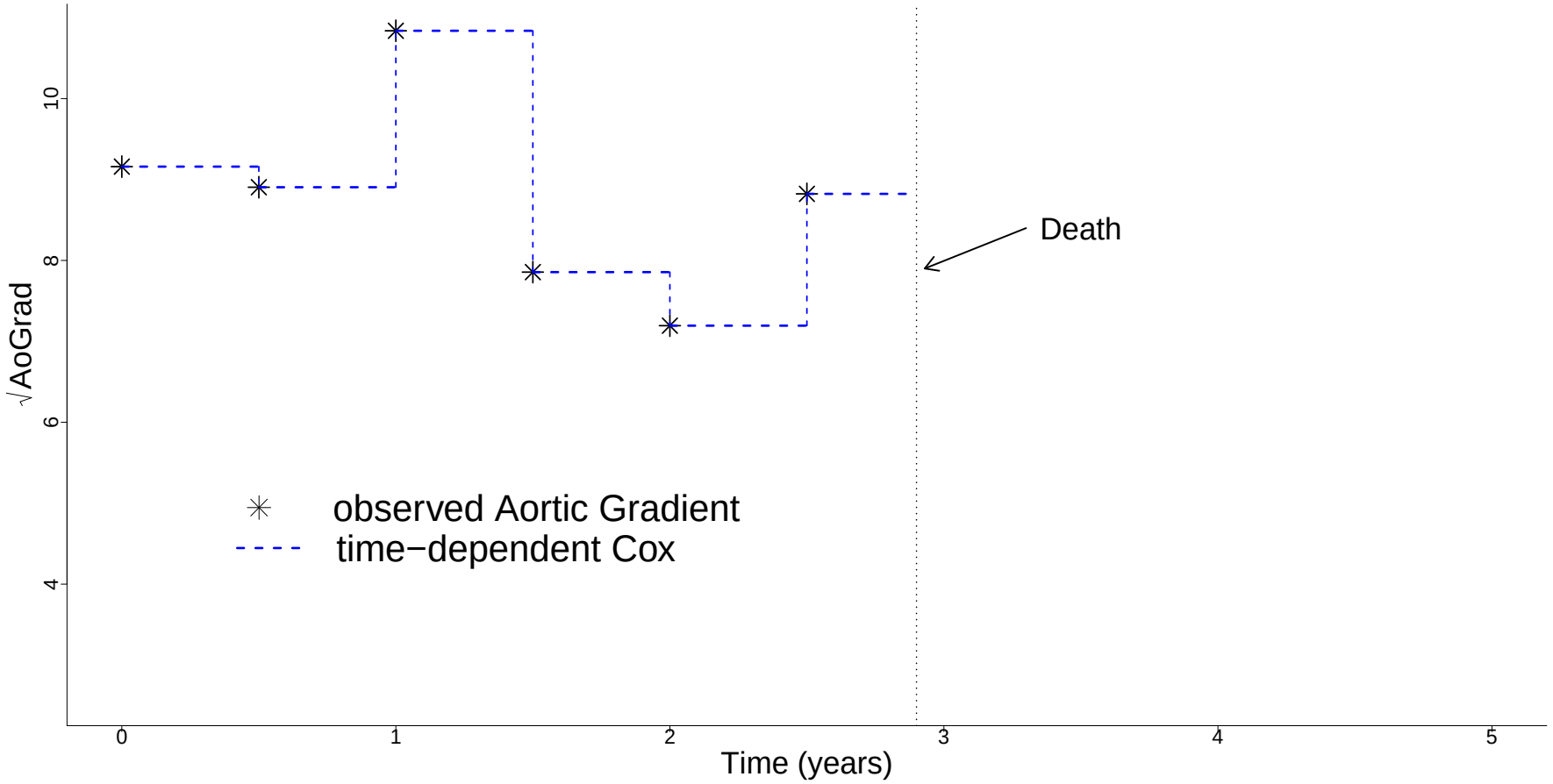
- Problem: Aortic gradient is an **endogenous** time-dependent covariate (Kalbfleisch and Prentice, 2002, Section 6.3)
  - ▷ Measurements on the same patient are correlated
  - ▷ Endogenous (aka internal): the future path of the covariate up to any time $t > s$ IS affected by the occurrence of an event at time point s

2.1 Joint Modeling Framework (cont'd)

- What is special about endogenous time-dependent covariates
  - ▷ measured with error
  - ▷ the complete history is not available
  - ▷ existence directly related to failure status

- What if we use the Cox model?
  - ▷ the association size can be severely underestimated
  - ▷ true potential of the marker will be masked

2.1 Joint Modeling Framework (cont'd)



2.1 Joint Modeling Framework (cont'd)

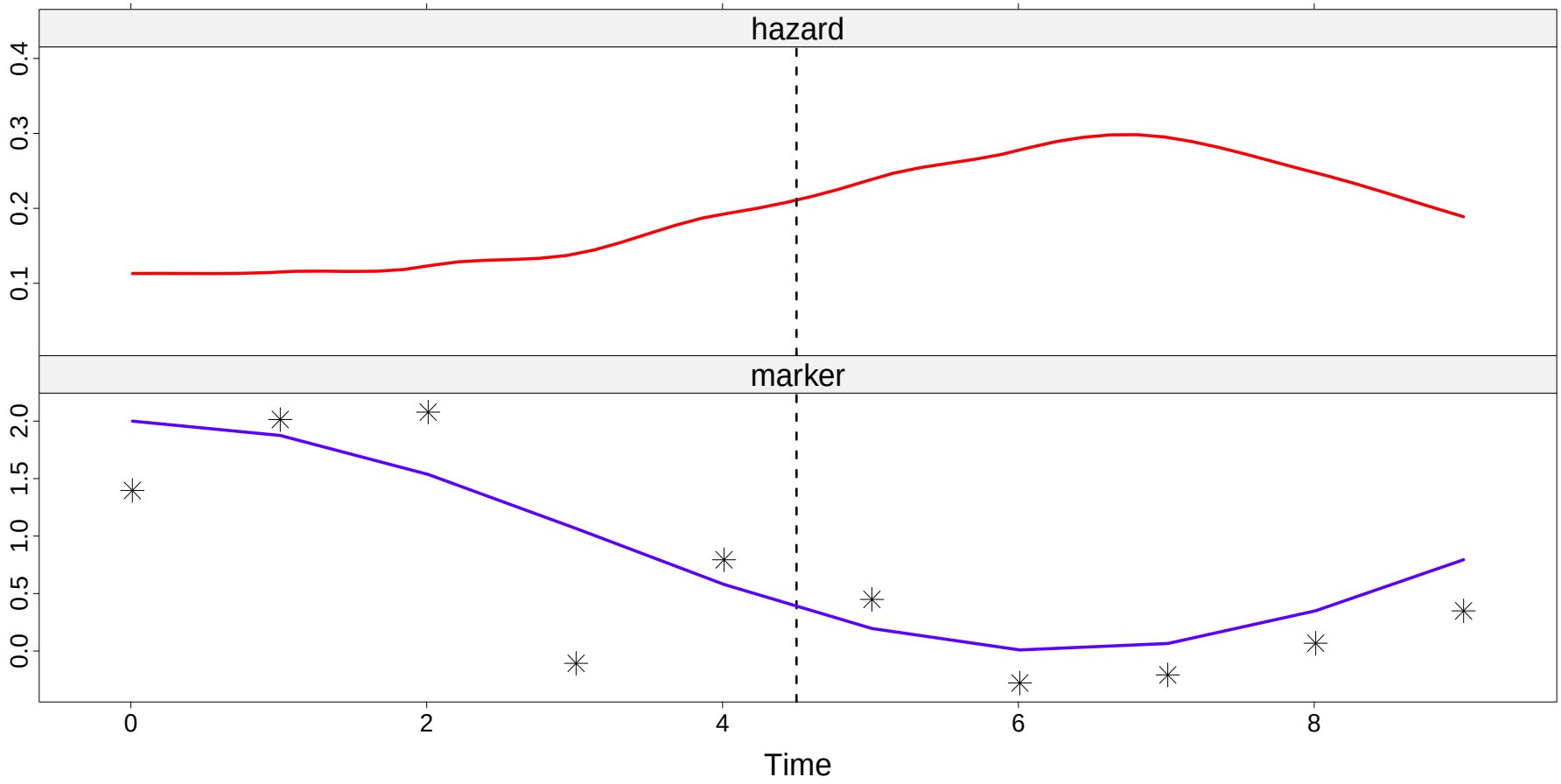
- To account for the special features of these covariates a new class of models has been developed

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 is **not** assumed constant between visits

2.1 Joint Modeling Framework (cont'd)



2.1 Joint Modeling Framework (cont'd)

- We define a standard joint model

  ▷ Survival Part: Relative risk model

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

where

* $m_i(t)$ = the *true* & *unobserved* value of aortic gradient at time t
* α quantifies the effect of aortic gradient on the risk for death/re-operation
* w_i baseline covariates

2.1 Joint Modeling Framework (cont'd)

▷ Longitudinal Part: Reconstruct $\mathcal{M}_i(t) = \{m_i(s), 0 \leq s < t\}$ using $y_i(t)$ and a mixed effects model (we focus on continuous markers)

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

where

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

3.1 Joint Models in 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)
```

3.1 Joint Models in 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

3.1 Joint Models in 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)
```

3.1 Joint Models in 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
  - ▷ ...

3.1 Joint Models in 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)

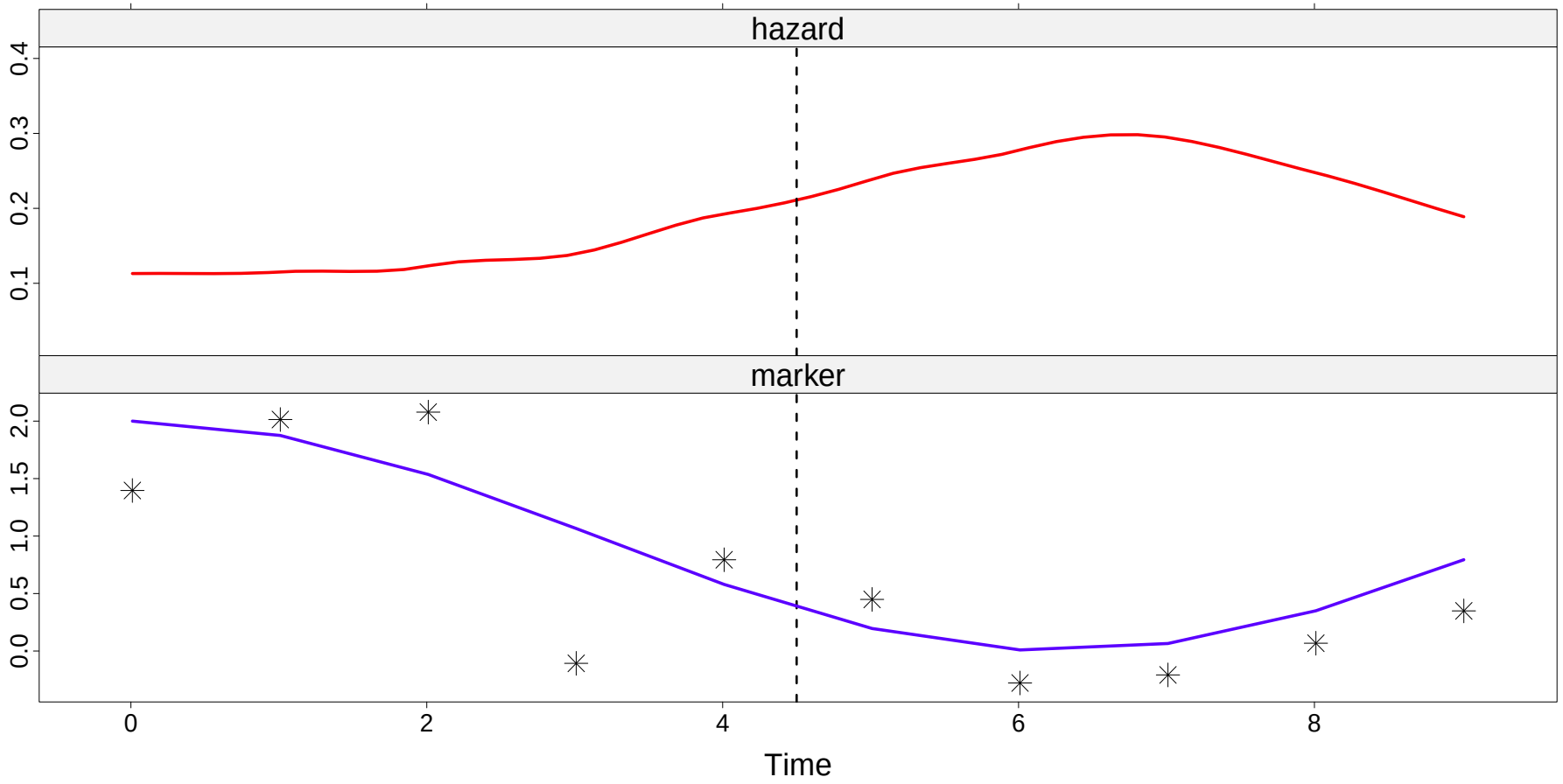
4.1 Association Structures

- The standard assumption is

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

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

4.1 Association structures (cont'd)



4.1 Association Structures (cont'd)

- The standard assumption is

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

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

Is this the only option? Is this the most optimal for prediction?

4.2 Time-dependent Slopes

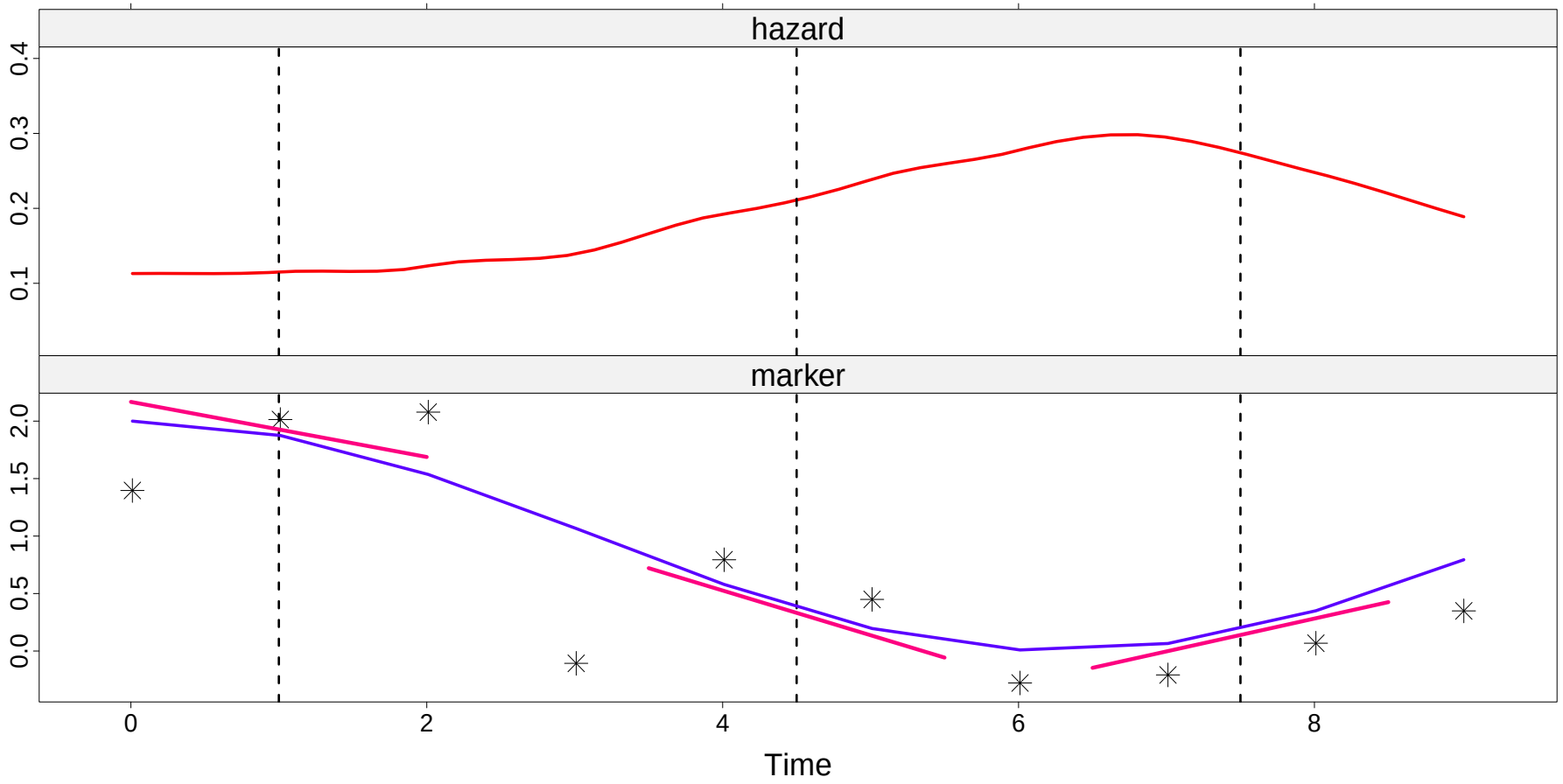
- The hazard for an event at t is associated with both the current value and the slope of the trajectory at t (Ye et al., 2008, Biometrics):

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

where

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

4.2 Time-dependent Slopes (cont'd)



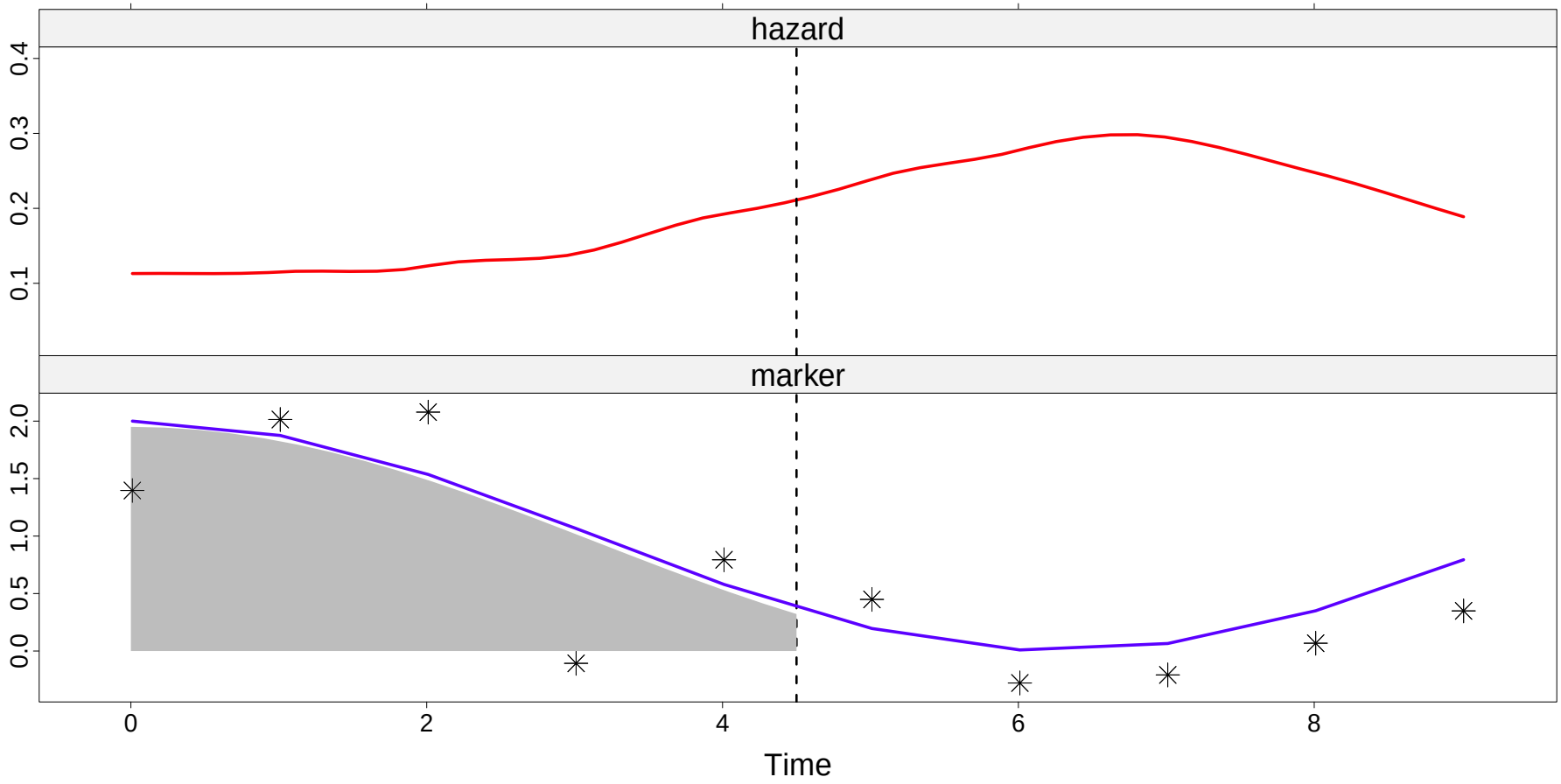
4.3 Cumulative Effects

- The hazard for an event at t is associated with 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 $\mathcal{M}_i(t)$

4.3 Cumulative Effects (cont'd)



4.5 Association Structures in JM & JMbayes

- Both package give options to define the aforementioned association structures
  - ▷ in **JM** via arguments `parameterization` & `derivForm`
  - ▷ in **JMbayes** via arguments `param` & `extraForm`

- **JMbayes** also gives the option for general transformation functions, e.g.,

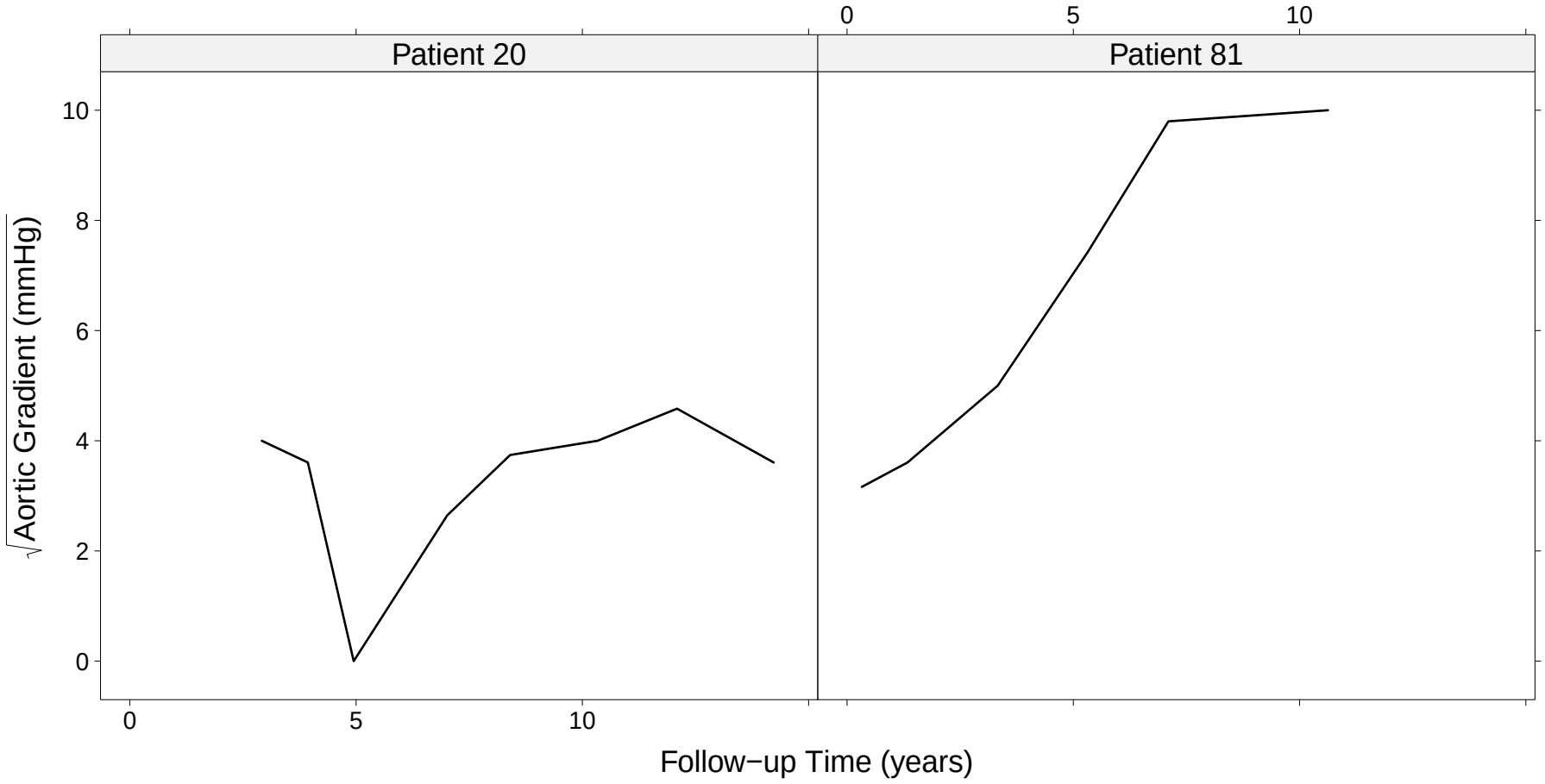
$$h_i(t \mid \mathcal{M}_i(t)) = h_0(t) \exp\{\gamma^\top w_i + \alpha_1 m_i(t) + \alpha_2 m_i(t) \times Treat_i + \alpha_3 m_i'(t) + \alpha_3 (m_i'(t))^2\},$$

5.1 Predictions – Definitions

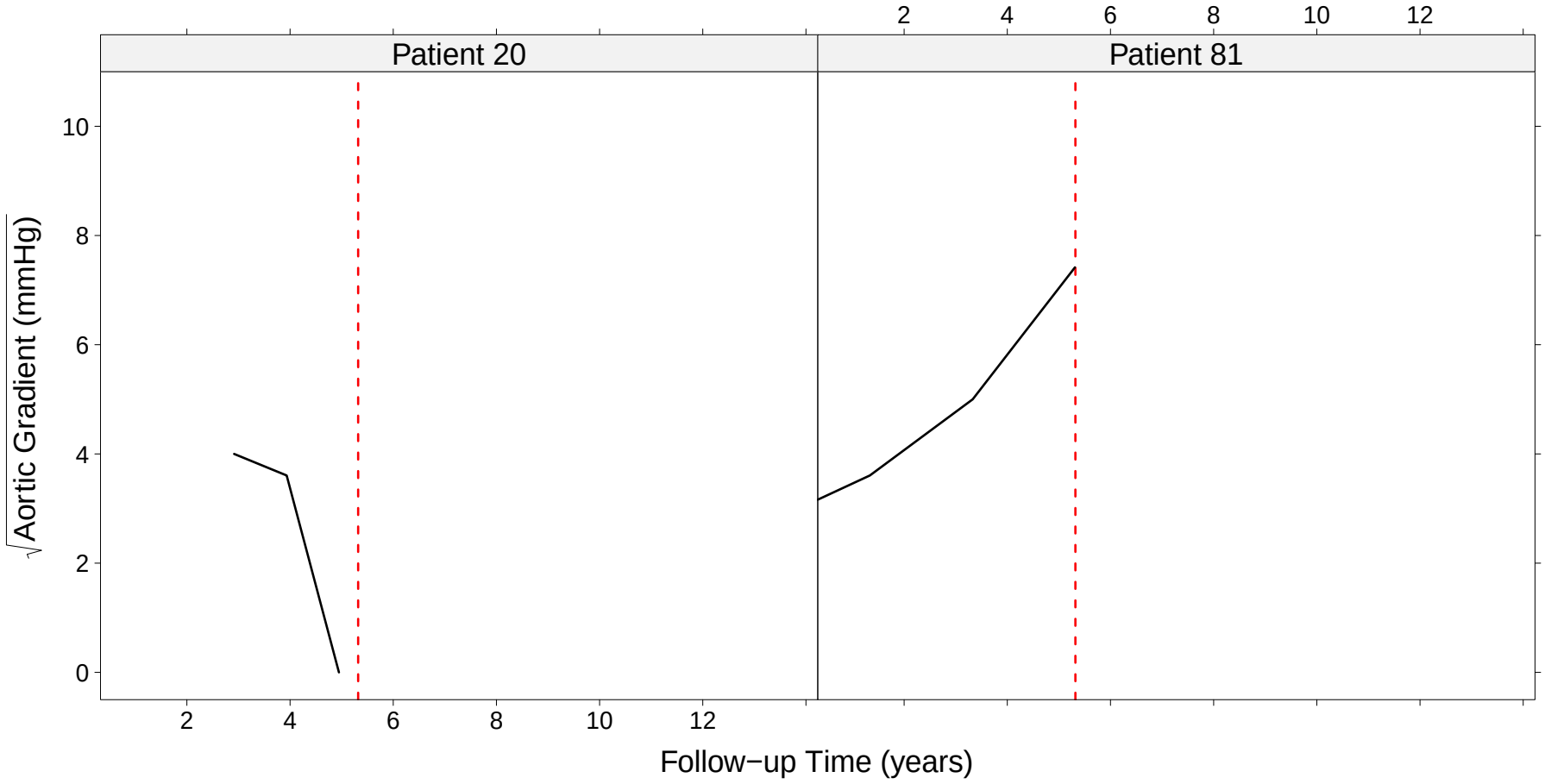
- We are interested in predicting survival probabilities for a new patient j that has provided a set of aortic gradient measurements up to a specific time point t

- Example: We consider Patients 20 and 81 from the Aortic Valve dataset

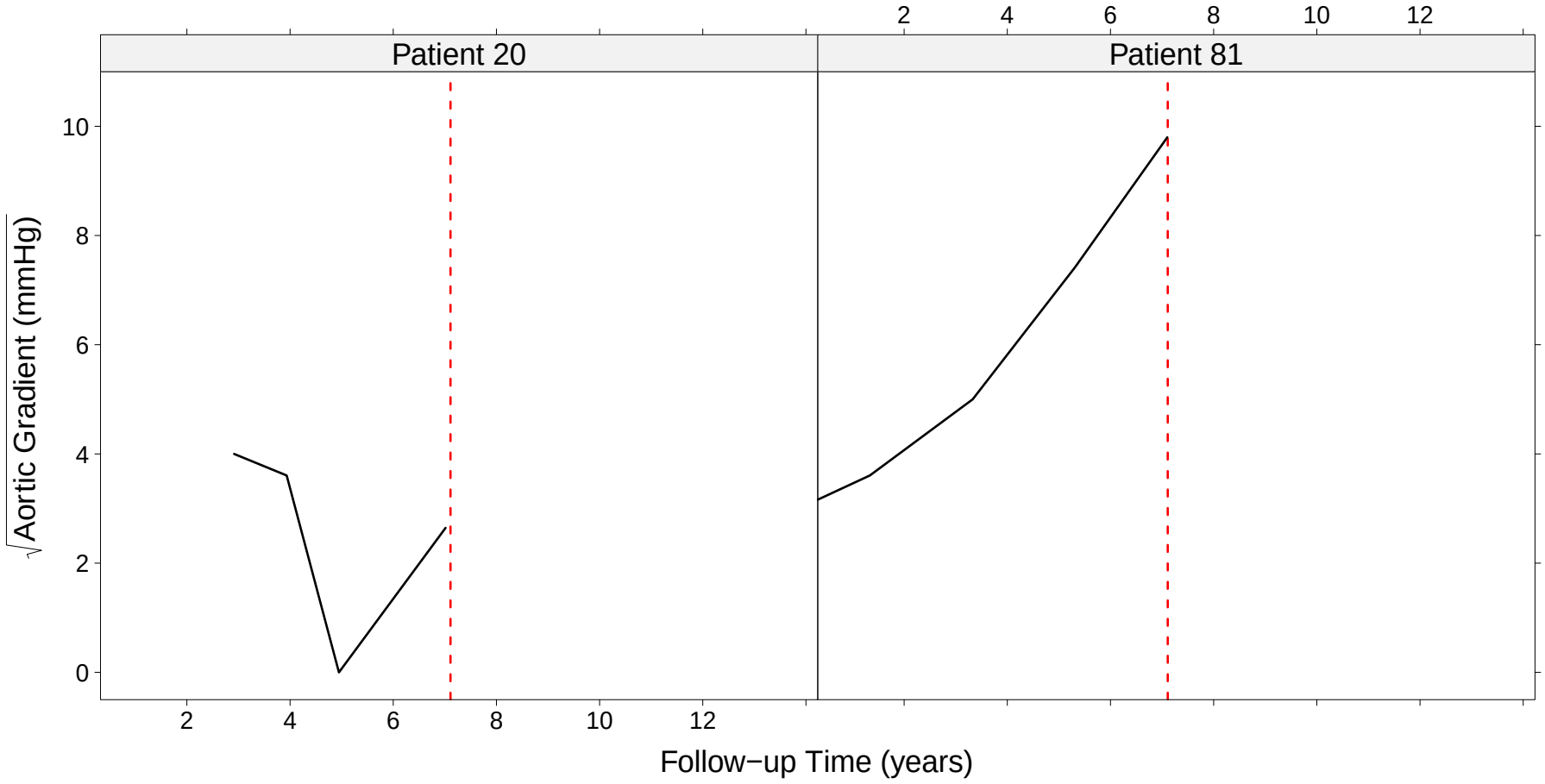
5.1 Predictions – Definitions (cont'd)



5.1 Predictions – Definitions (cont'd)



5.1 Predictions – Definitions (cont'd)

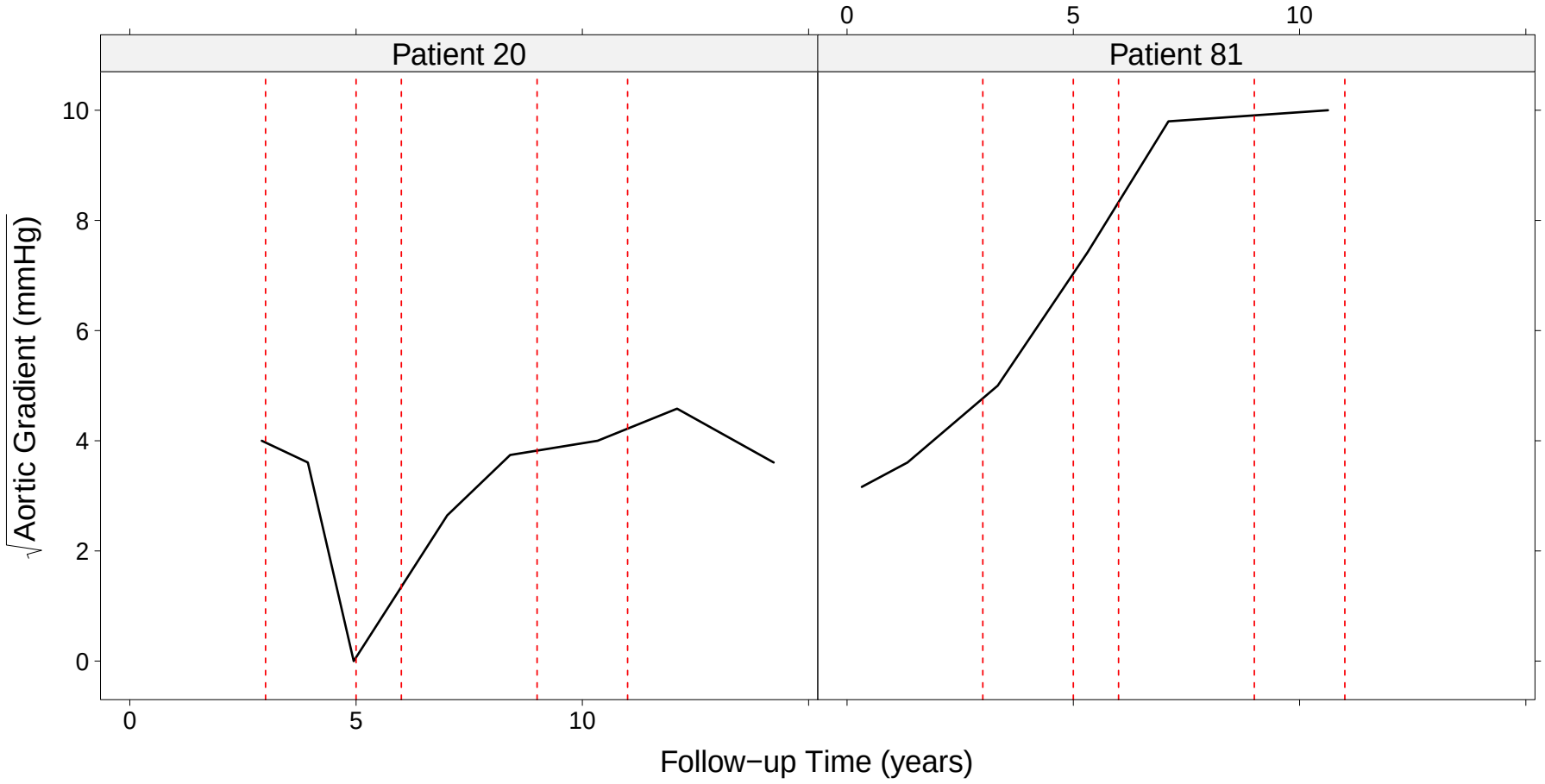


5.1 Predictions – Definitions (cont'd)

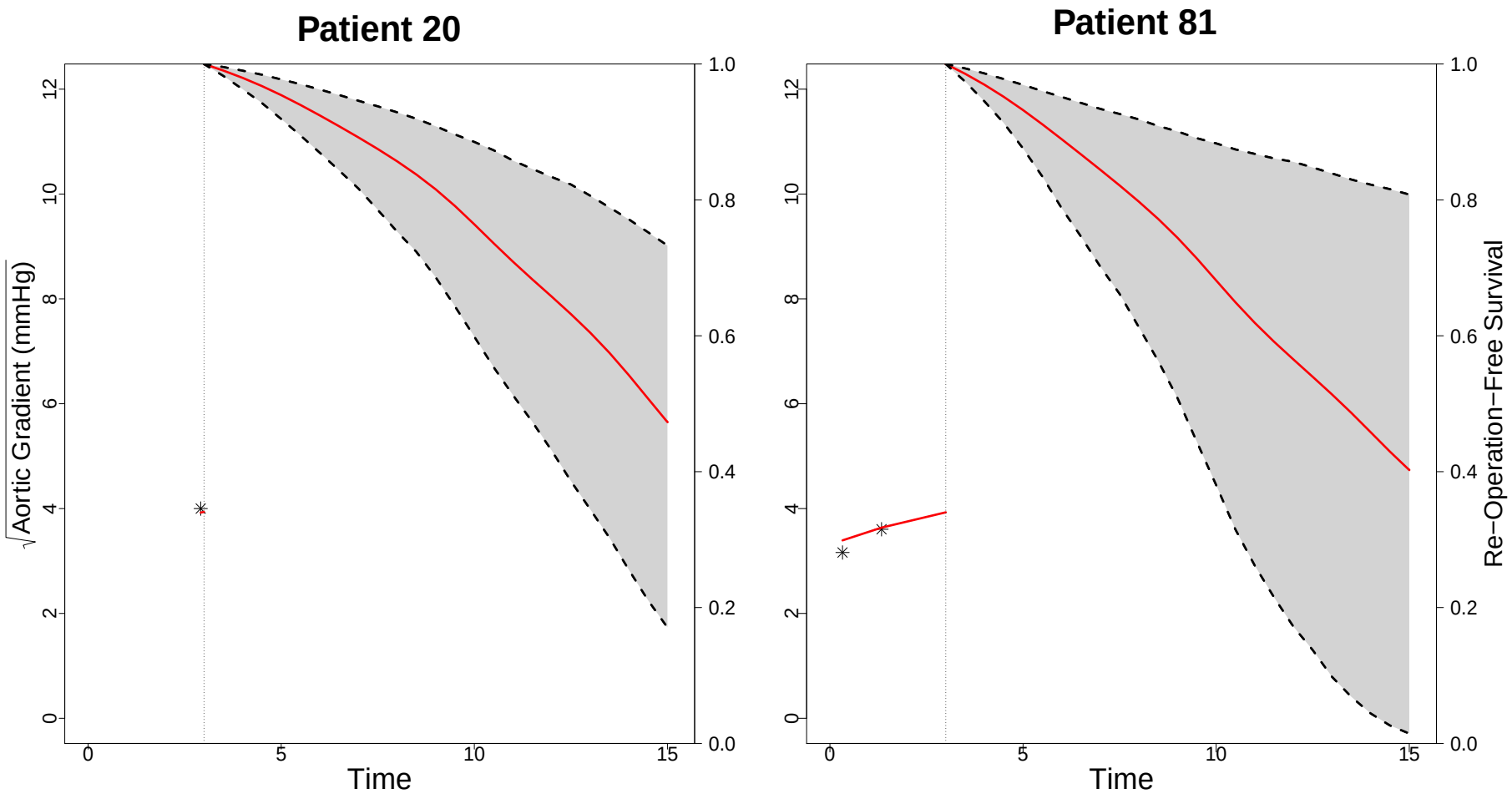
- What do we know for these patients?
  - ▷ a series of aortic gradient measurements
  - ▷ patient are event-free up to the last measurement

- **Dynamic Prediction** survival probabilities are dynamically updated as additional longitudinal information is recorded

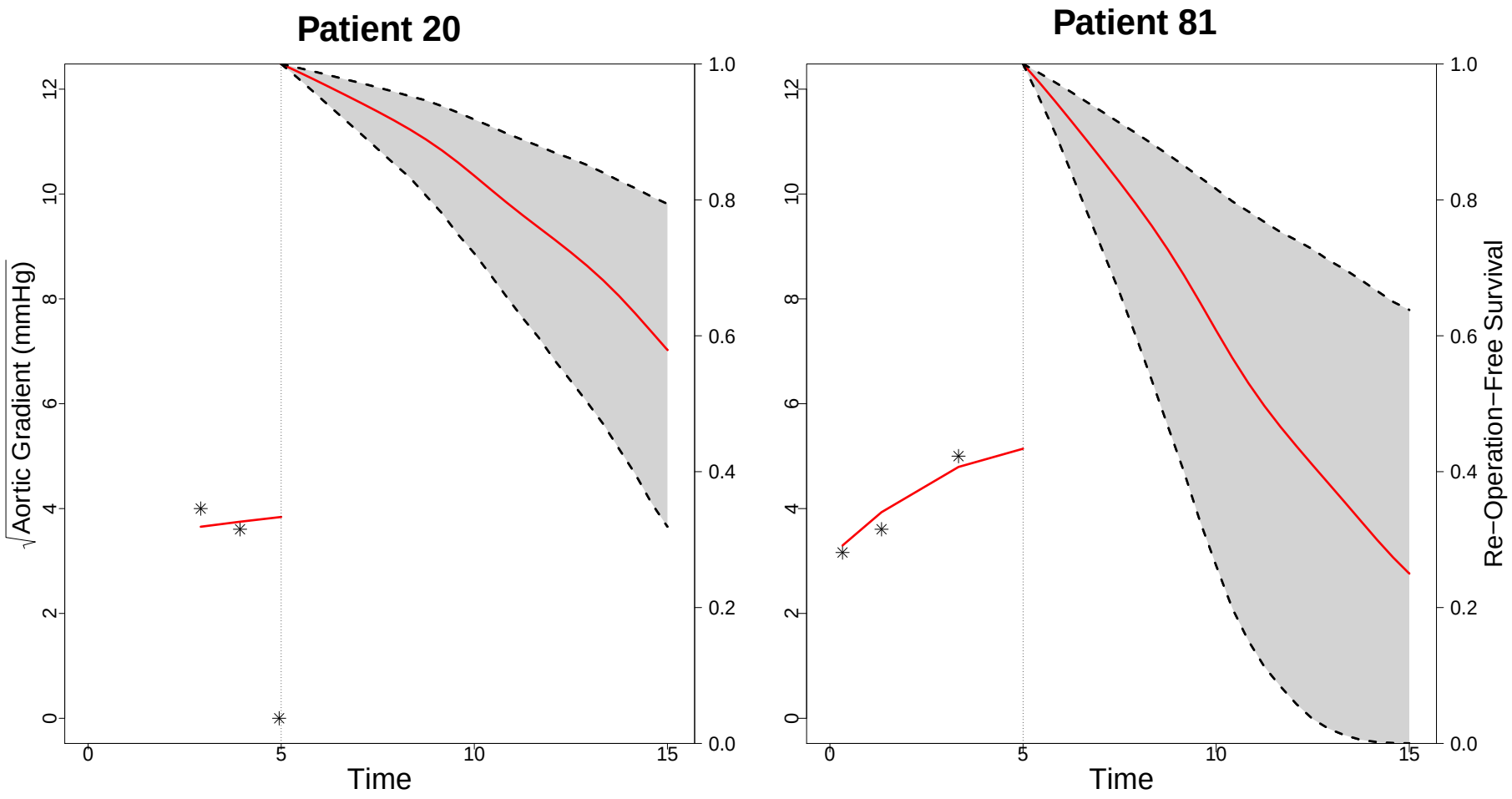
5.3 Dyn. Predictions – Illustration



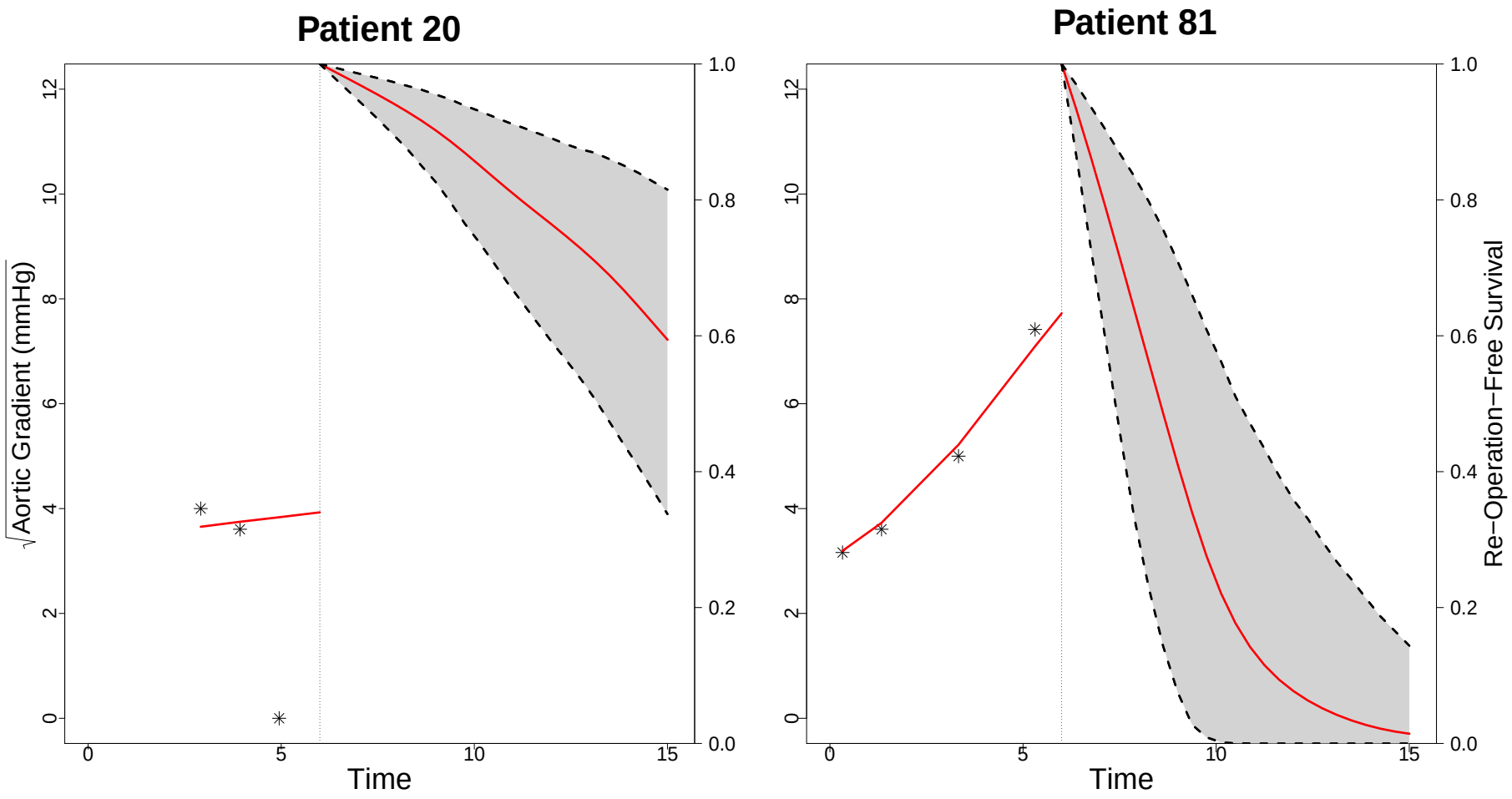
5.3 Dyn. Predictions – Illustration (cont'd)



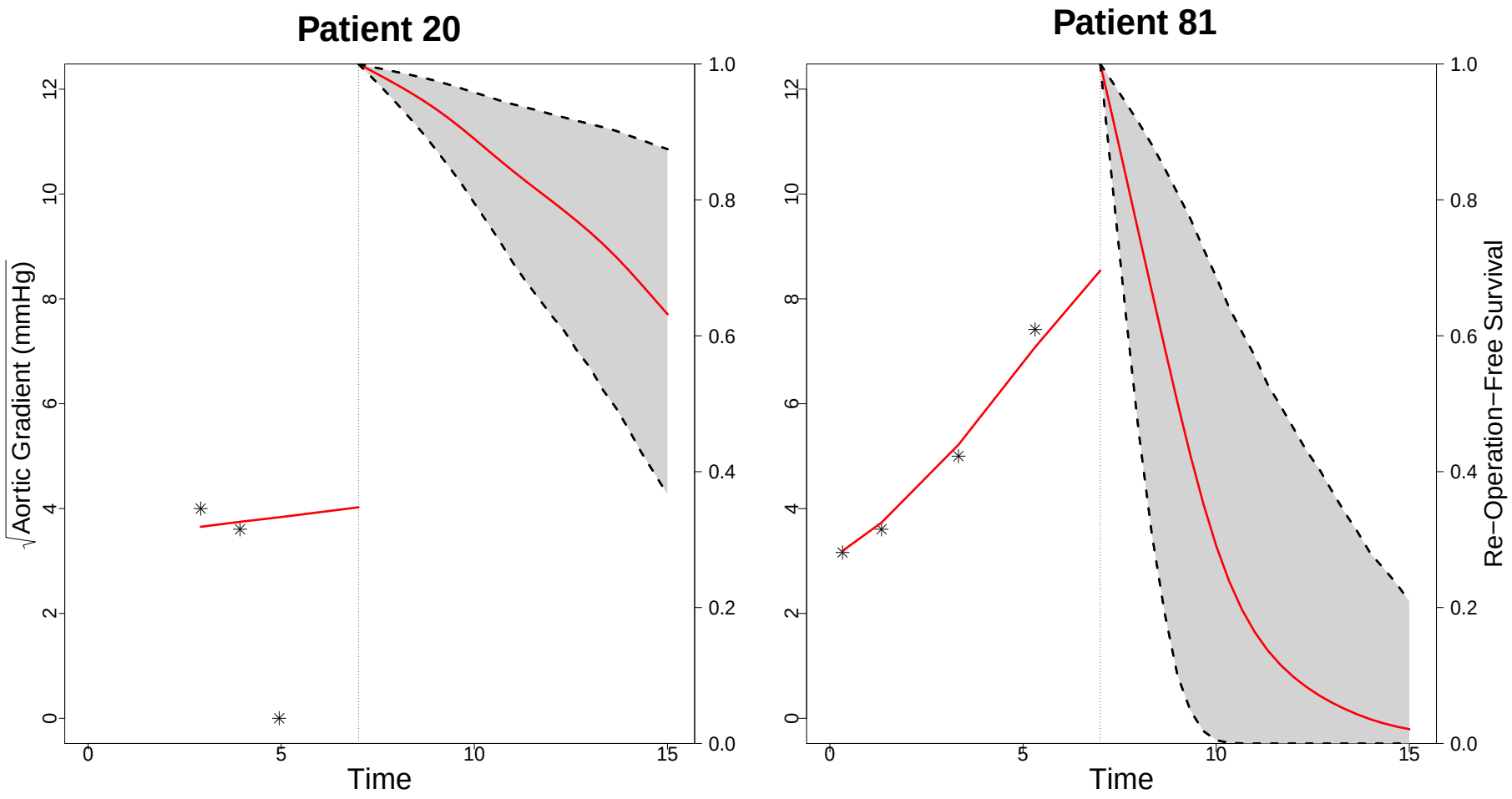
5.3 Dyn. Predictions – Illustration (cont'd)



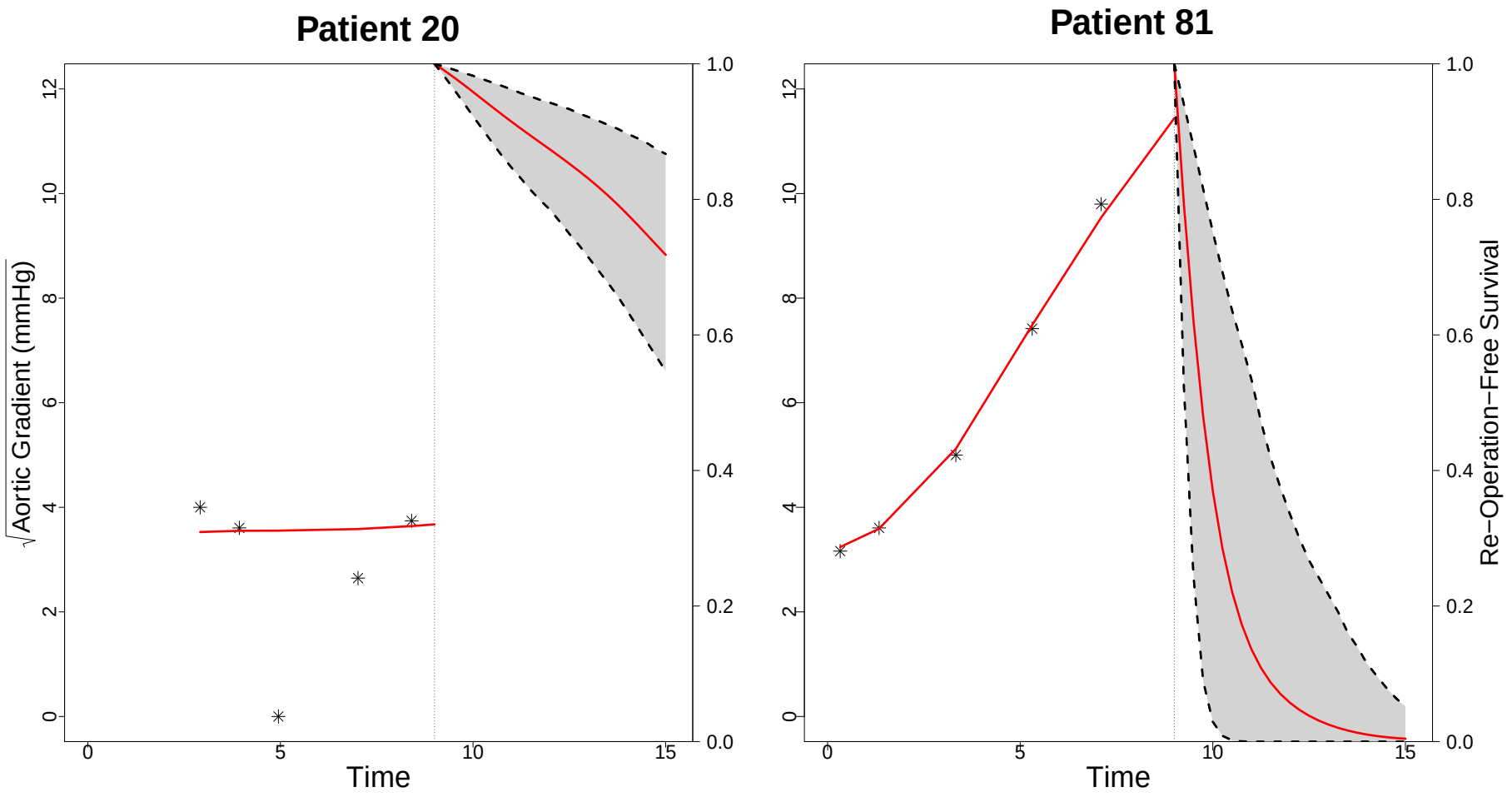
5.3 Dyn. Predictions – Illustration (cont'd)



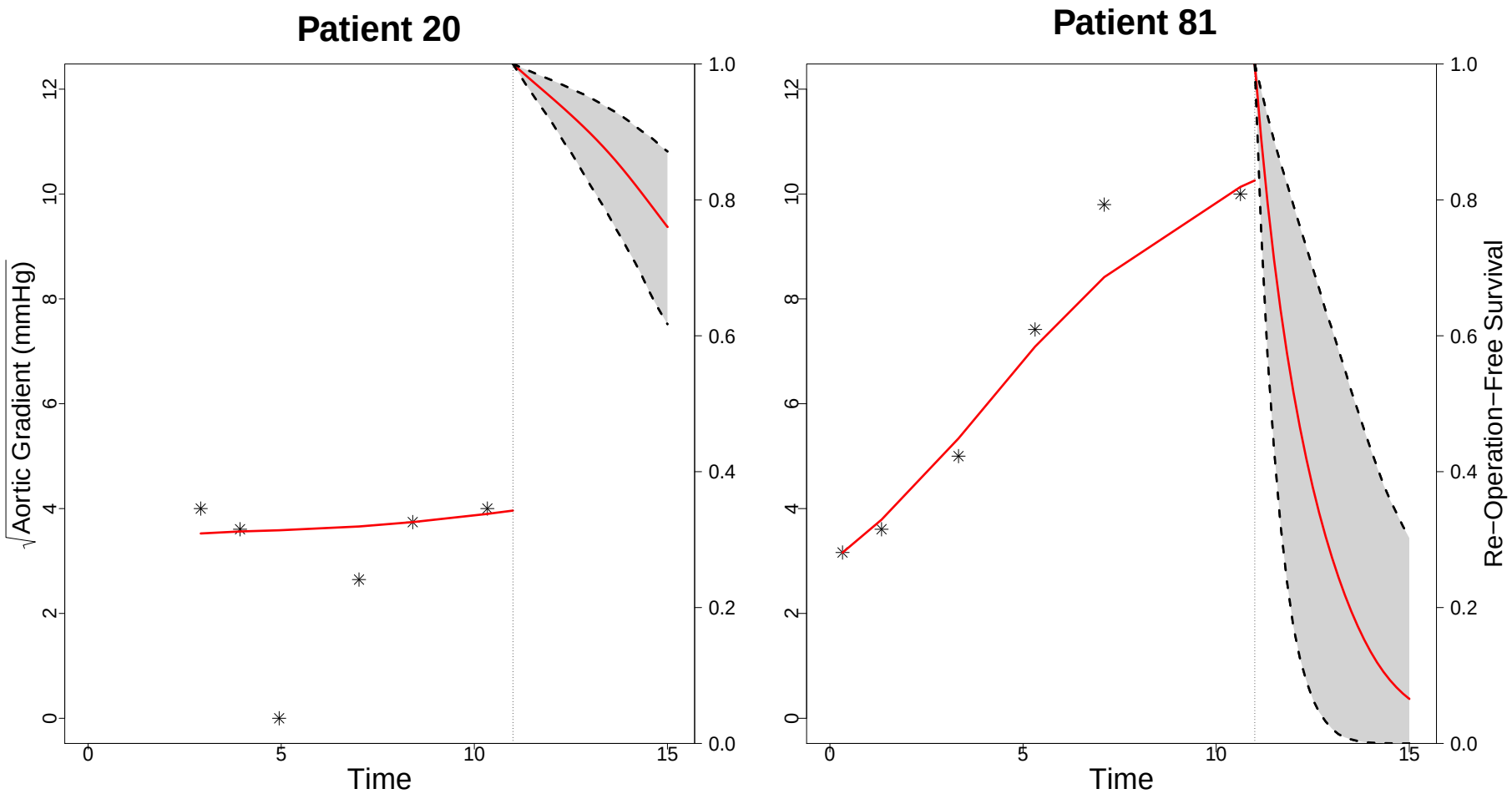
5.3 Prediction Survival – Illustration (cont'd)



5.3 Dyn. Predictions – Illustration (cont'd)



5.3 Dyn. Predictions – Illustration (cont'd)



5.2 Predictions – JM & JMbayes

- 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)

# shiny app in JMbayes
JMbayes::runDynPred()
```

6. Resources

- **JM**
  - ▷ JSS paper (http://www.jstatsoft.org/v35/i09/)
  - ▷ book for joint models (http://jmr.r-forge.r-project.org/)

- **JMbayes**
  - ▷ JSS paper (http://arxiv.org/abs/1404.7625; to appear)

Thank you for your attention!