Assessing Risk Indicators in Clinical Practice with Joint Models of Longitudinal and Time-to-Event Data

Eleni-Rosalina Andrinopoulou
Department of Biostatistics, Erasmus Medical Center

e.andrinopoulou@erasmusmc.nl

Center for Clinical & Translational Science & Training, University of Cincinnati

September 21th

Outline

▷ Introduction to Joint Models

▷ Recent Applications in Joint Models

▷ Individualized Predictions

Introduction to Joint Models

Introduction to Joint Models

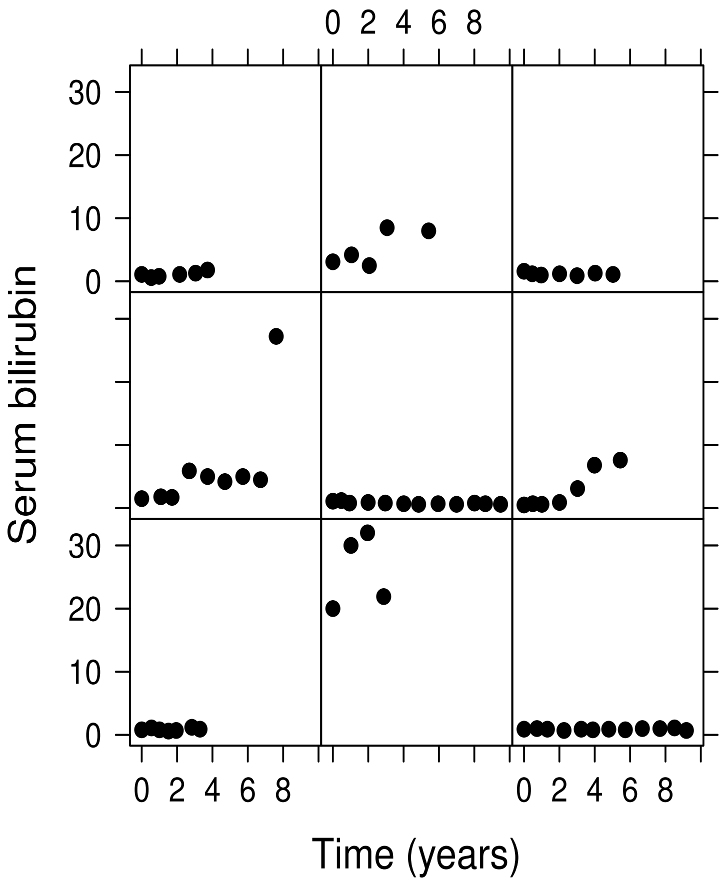
- Often in clinical studies multiple outcomes are collected

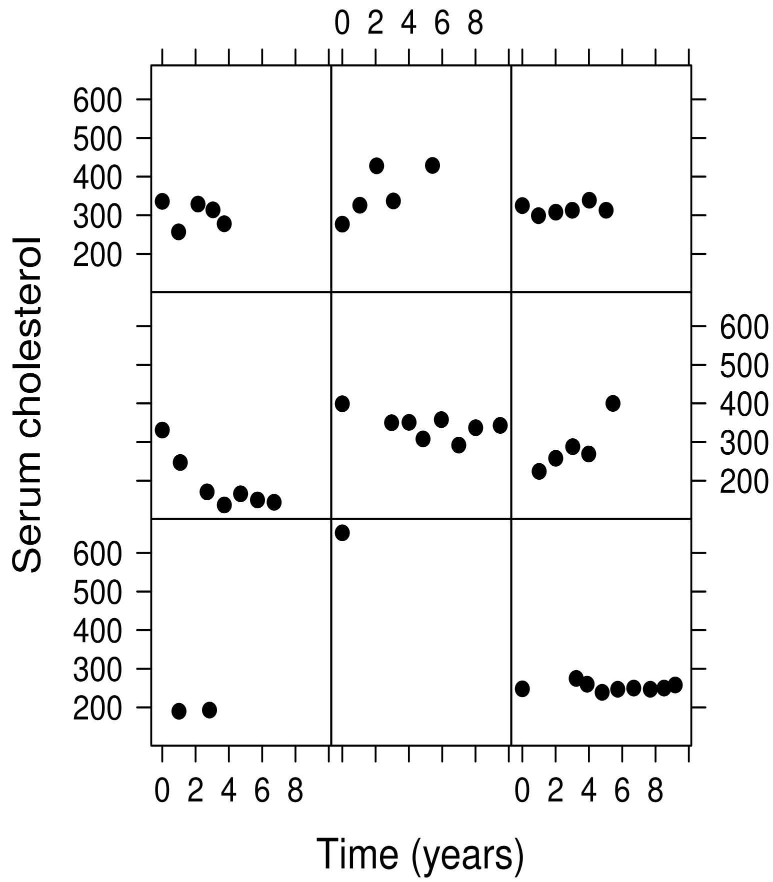
- Type of data
  - ▷ Longitudinal responses
  - ▷ Time-to-event data

Motivation - Data set 1

- 312 patients with primary **biliary cirrhosis**, a rare autoimmune liver disease, at Mayo Clinic
  - ▷ Patients were 50 years and older, 88% females and 50% D-penicil
  - ▷ Median number of visits per individual is 6
  - ▷ Longitudinal responses: **serum bilirubin** and **serum cholesterol** in mg/dl
  - ▷ Time-to-event response: time-to-**death** (45%)

Motivation - Data set 1 (con't)

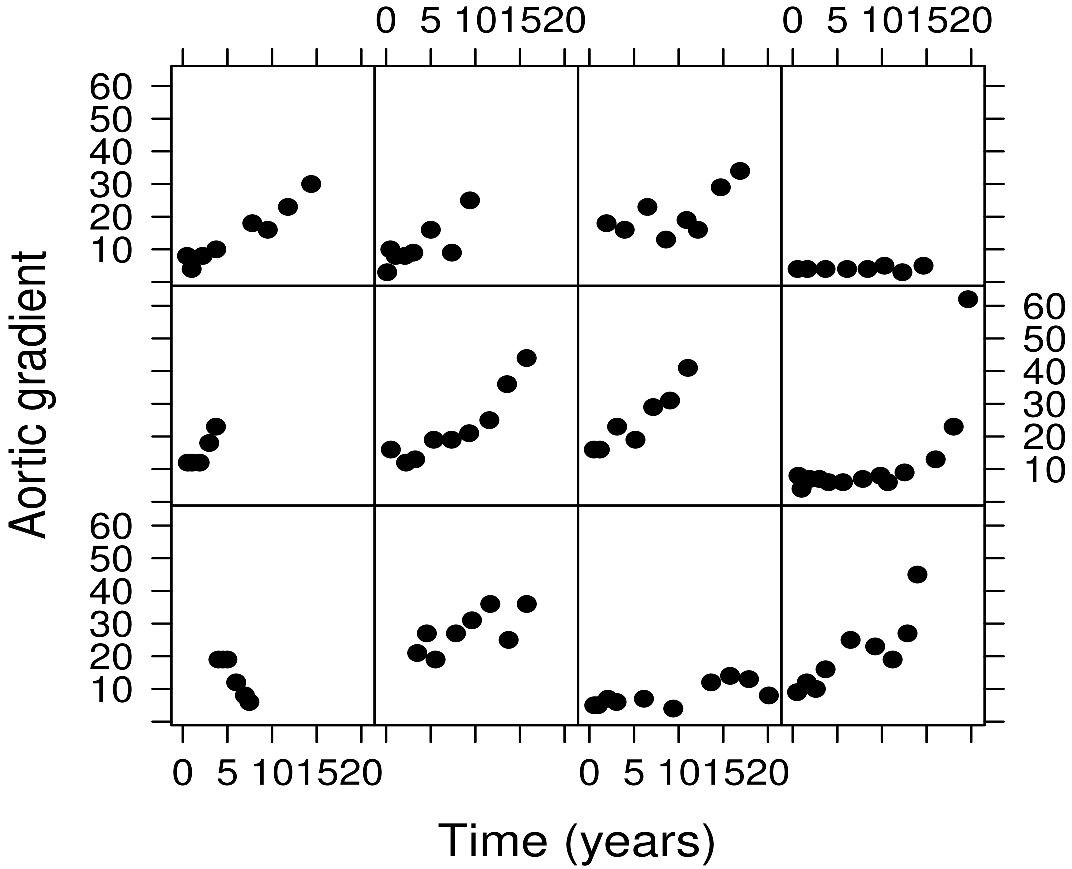




Motivation - Data set 2

- 286 patients who received **human tissue valve in aortic position** in Erasmus University Medical Center (Department of Cardio-Thoracic Surgery)
  - ▷ Patients were 16 years and older
  - ▷ Echo examinations scheduled at 6 months and 1 year postoperatively and biennially thereafter
  - ▷ Longitudinal response: **aortic gradient**
  - ▷ Time-to-event response: time-to-**death/reoperation** (54%)

Motivation - Data set 2 (con't)

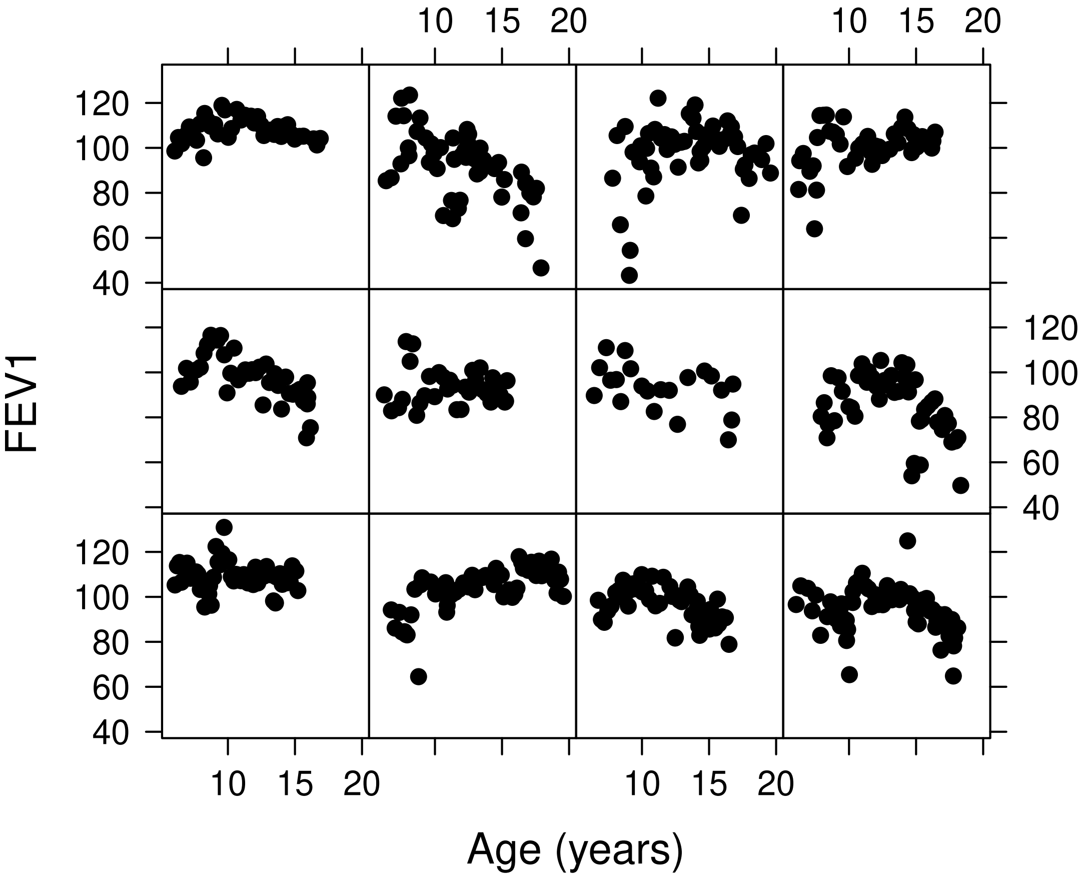


Motivation - Data set 3

- 1016 patients with **cystic fibrosis (CF)** in the US Cystic Fibrosis Foundation Patient Registry

  - ▷ Data recorded during childhood and adolescence (ages: 6-21 years) between January 1, 2003, and December 31, 2015
  - ▷ Median follow-up visit equal to 6 (range 1-93)
  - ▷ Longitudinal response: **FEV**$_1$
  - ▷ Time-to-event response: time-to-**first exacerbation** (70%)

Motivation - Data set 3 (con't)



Research Questions

- How can we utilize all available longitudinal measurements?

- How can we better predict the survival outcomes?

How can use all available information?

Introduction to Joint Models (cont'd)

- Special features should be taken into account

Longitudinal data

▷ Correlation between measurements obtained from the same patients

▷ Biological variation of the outcome

▷ Unbalanced datasets

Survival data

▷ Censored data (partial information for the event times)

Jointly

▷ Association between all outcomes

Introduction to Joint Models (cont'd)

- Frequently used analysis

  ▷ **Separate** analysis per outcome

  Mixed-effects models for the longitudinal outcomes

  Cox models for the time-to-event outcomes

  ▷ Naive **joint** analysis

  Cox model using the last observation

  Cox model using the mean or the slope of the repeated covariate

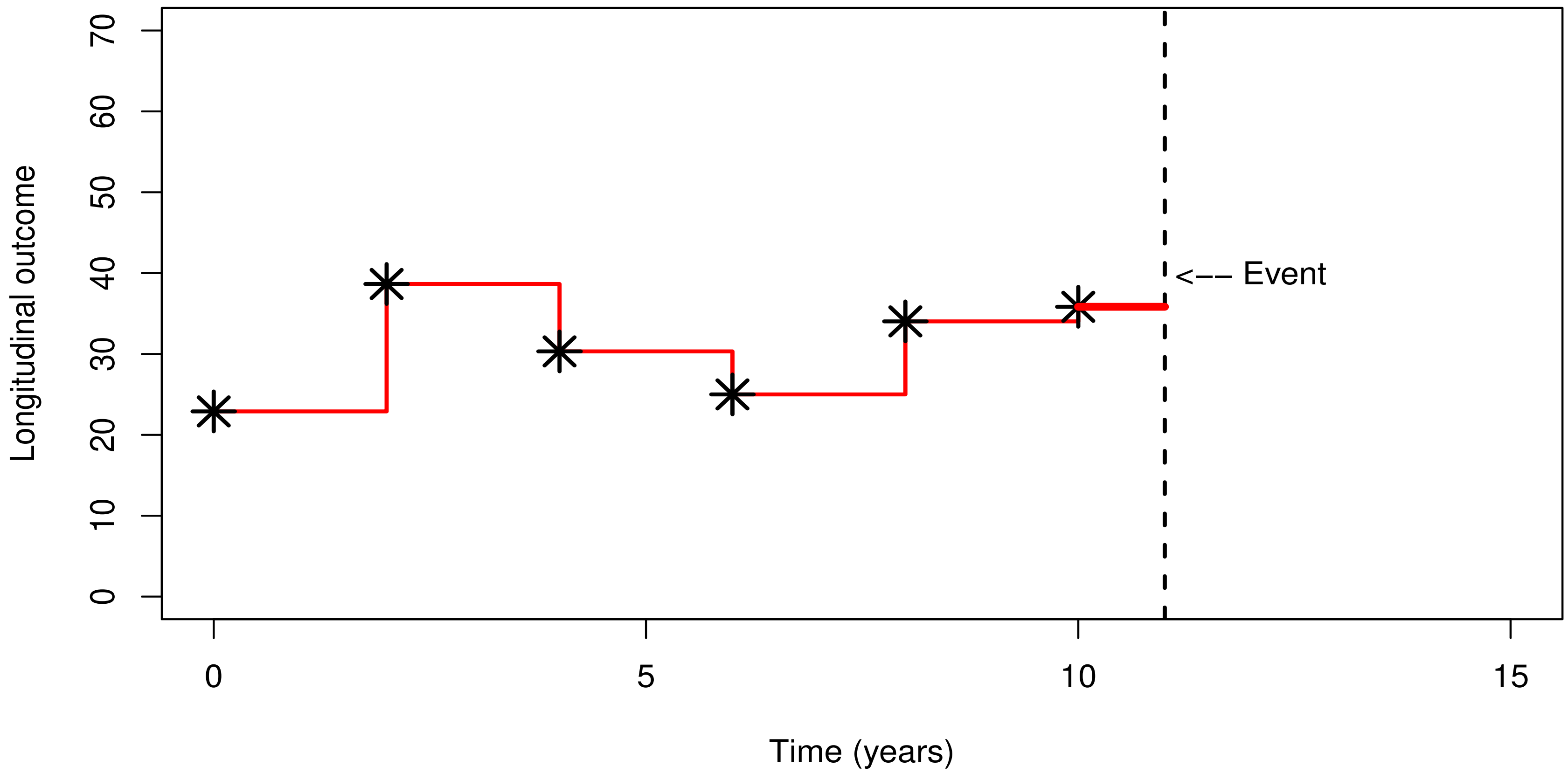
  Time-dependent Cox model

Introduction to Joint Models (cont'd)

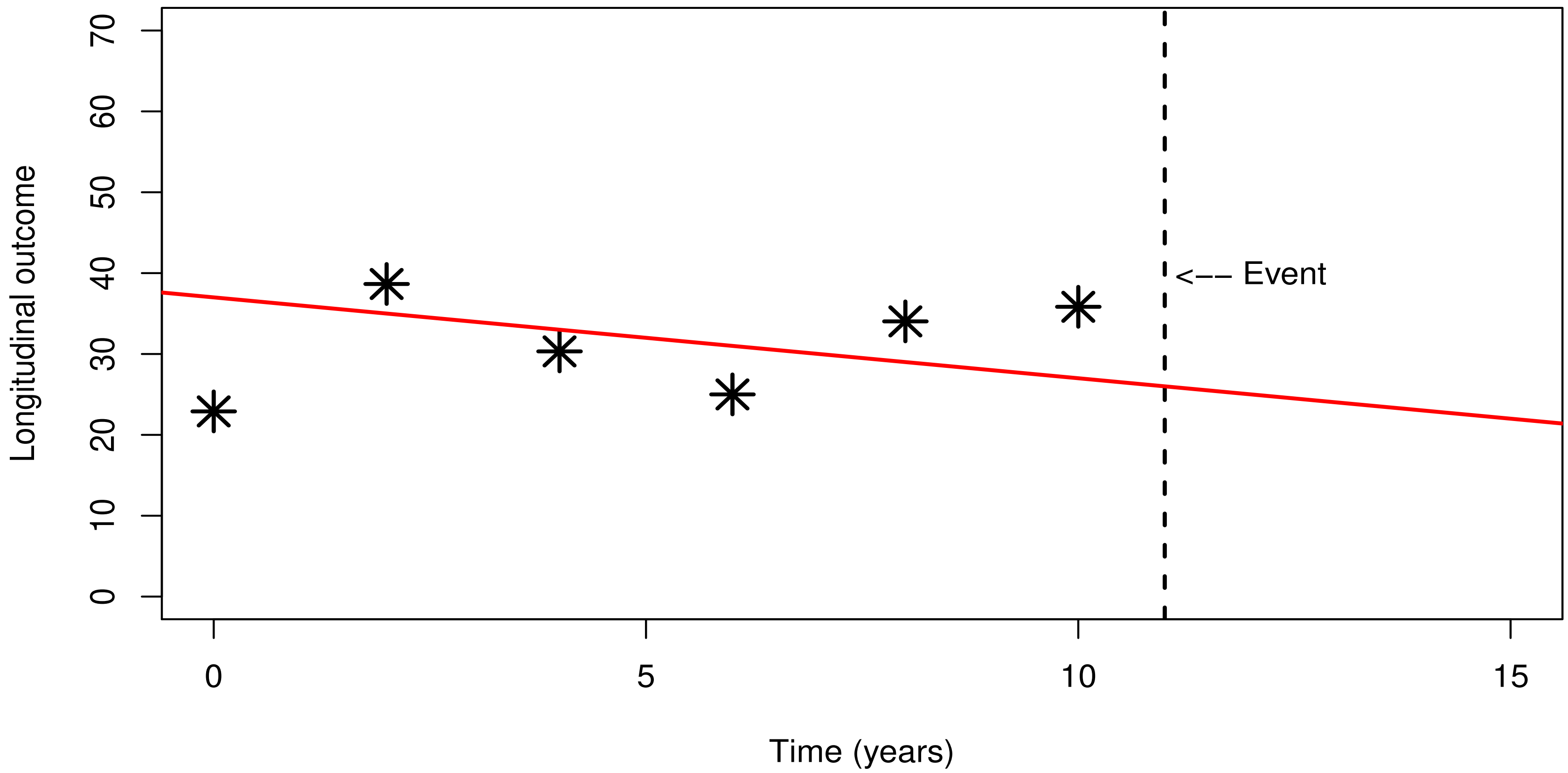
- Time-dependent Cox models are suitable only for **exogenous** covariates, not got **endogenous**

  - ▷ A time-varying covariate is **exogenous** if its value at any time point t is not affected by an event occurring at an earlier time point $s < t$ (period of the year, enviromental variables)
  - ▷ On the other hand all covariates measured on the patient (e.g., biomarkers) are **endogenous**

Introduction to Joint Models (cont'd)



Introduction to Joint Models (cont'd)



Introduction to Joint Models (cont'd)

- Let y_i represent the repeated measurements of an outcome for the i-th patient, $i = 1, \ldots, n$

$\boxed{\text{Mixed-effects model:}}$

$$y_i(t) = x_i^\top(t)\beta + z_i^\top(t)b_i + \epsilon_i(t) = \eta_i(t) + \epsilon_i(t), \quad b_i \sim N(0, D) \text{ and } \epsilon_i(t) \sim N(0, \Sigma_i)$$

where

- ▷ $x_i^\top(t)\beta$ denotes the fixed part
- ▷ $z_i^\top(t)b_i$ denotes the random part

Introduction to Joint Models (cont'd)

- Let T_i denote the observed failure time for the i-th patient and $\delta = 0, 1$ the event indicator

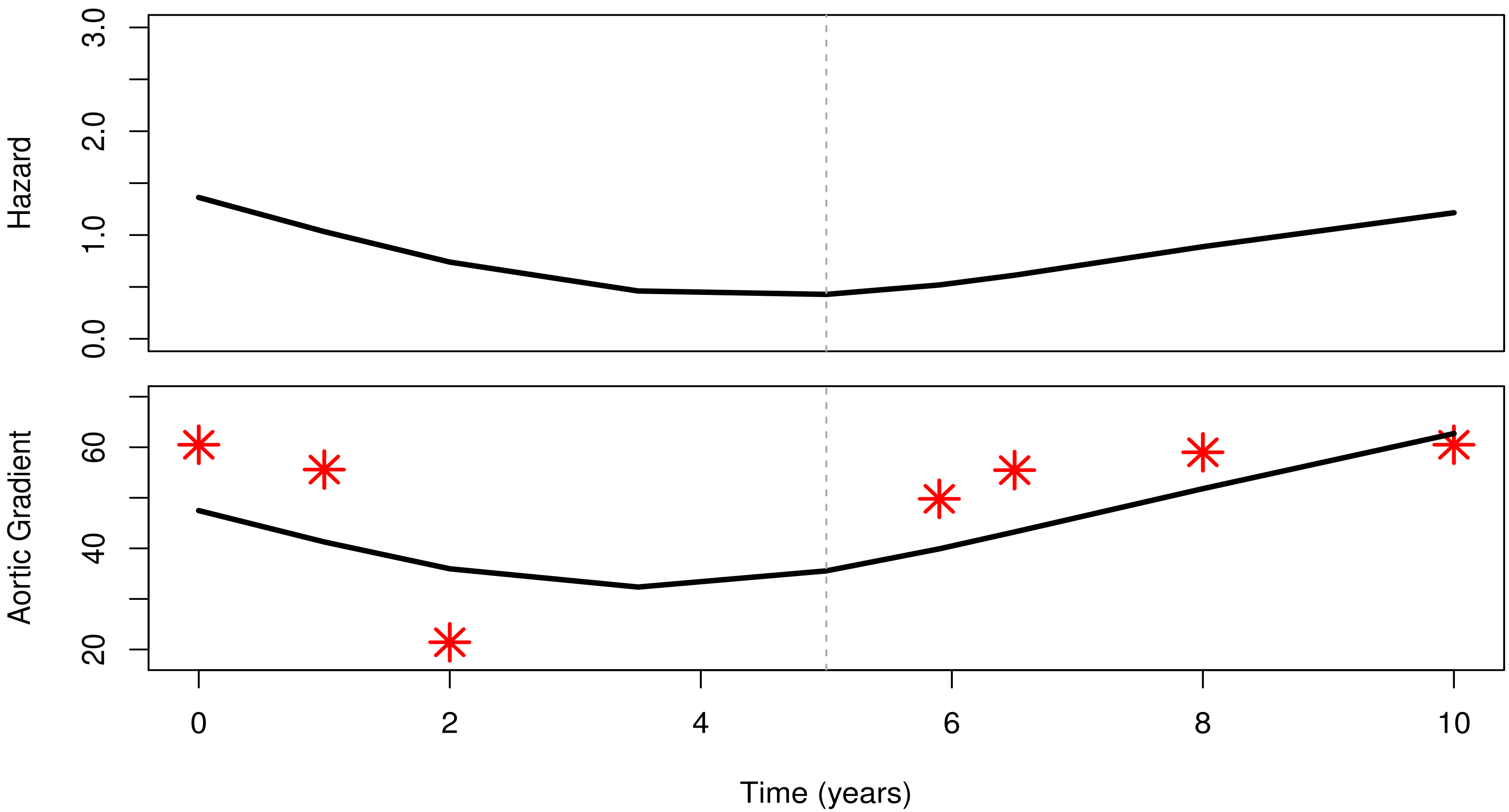
  $\boxed{\text{Cox model:}}$

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

  where

  - ▷ $\gamma^\top \omega_i$ denotes the baseline covariates with their coefficients
  - ▷ $\eta_i(t)$ denotes the value of the time-dependent covariate at time t
  - ▷ α measures the association between the longitudinal outcome at time t and the hazard for an event at the same time point

Introduction to Joint Models (cont'd)



Recent Applications in Joint Models

Shrinkage Approach

Motivated by the biliary cirrhosis data:

- Longitudinal responses:
  - ▷ **serum bilirubin**
  - ▷ **serum cholesterol**
- Time-to-event response:
  - ▷ time-to-**death**

Andrinopoulou, E. R. and Rizopoulos, D. (2016). Bayesian shrinkage approach for a joint model of longitudinal and survival outcomes assuming different association structures. Statistics in Medicine, 35(26), 4813-4823.

Shrinkage Approach (cont'd)

- In the standard joint model we assume that the underying value of the longitudinal biomarker is associated with the survival outcome at a time point t

Is that option always correct?

Shrinkage Approach (cont'd)

- Inappropriate modelling 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. patient who smoked had higher probability of survival)
  - ▷ most of those who had died were smokers but many stopped smoking at the last follow-up before they died

Shrinkage Approach (cont'd)

We need to carefully consider which longitudinal outcomes and which functional forms we will include

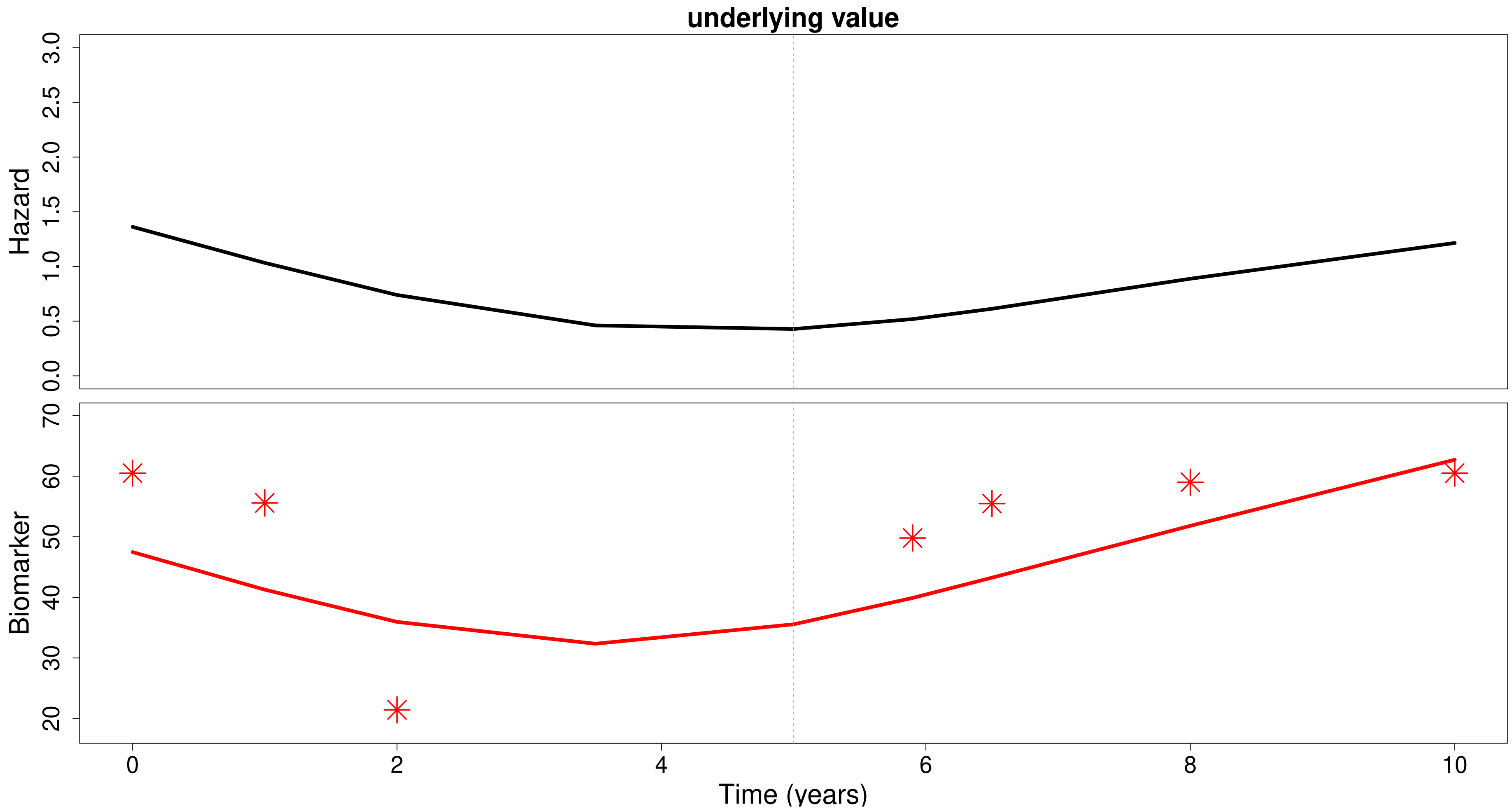
- Let's investigate that ...

Shrinkage Approach (cont'd)

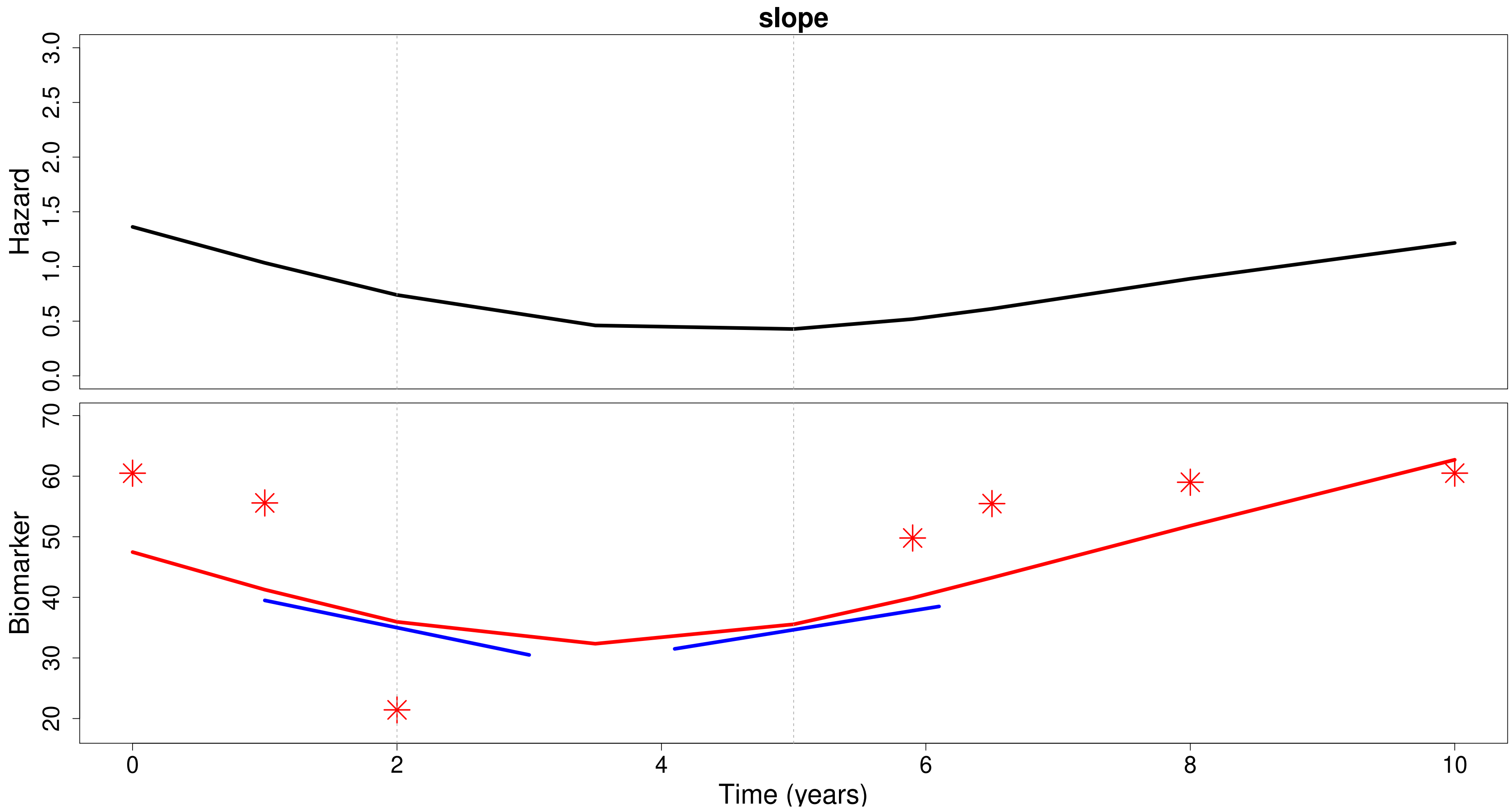
$\boxed{\text{Different parameterizations}}$ Let us assume $k = 1, \ldots, K$ longitudinal outcomes

$$\begin{aligned}
M_1 : h_i(t) &= h_0(t) \exp\{\gamma^\top w_i + \sum_{k=1}^{K} \alpha_{k1} \eta_{ik}(t)\}, \\
M_2 : h_i(t) &= h_0(t) \exp\{\gamma^\top w_i + \sum_{k=1}^{K} \alpha_{k2} \eta'_{ik}(t)\}, \\
M_3 : h_i(t) &= h_0(t) \exp\{\gamma^\top w_i + \sum_{k=1}^{K} \alpha_{k3} \int_0^t \eta_{ik}(s) ds\} \\
&\vdots
\end{aligned}$$

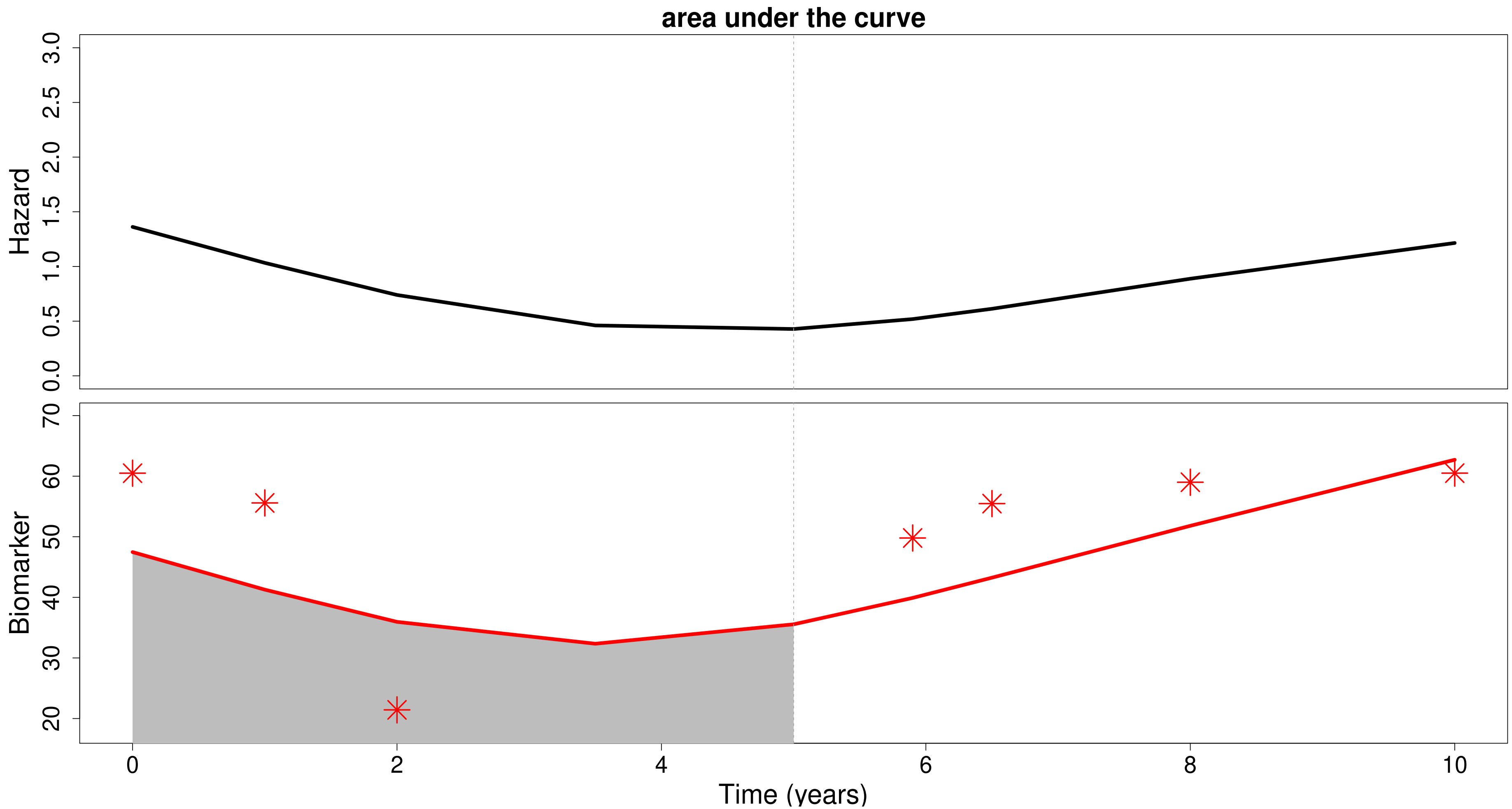
Shrinkage Approach (cont'd)



Shrinkage Approach (cont'd)



Shrinkage Approach (cont'd)



Shrinkage Approach (cont'd)

Extension of the standard JM

$$h_i(t) \;=\; h_0(t) \exp\left[\gamma^\top w_i + \sum_{k=1}^{K}\sum_{j=1}^{J} f_j\{\eta_{ik}(t), \alpha_{kj}\}\right],$$

where

- ▷ $i = 1, \ldots, n$ represents the patient,
- ▷ $k = 1, \ldots, K$ represents the longitudinal outcome
- ▷ $j = 1, \ldots, J$ represents the parameterization

Shrinkage Approach (cont'd)

For every longitudinal outcome which features are more predictive for survival?

$\Downarrow$

High dimensional model

$\Downarrow$

Variable selection problem

Shrinkage Approach (cont'd)

For every longitudinal outcome which features are more predictive for survival?

$\Downarrow$

High dimensional model

$\Downarrow$

Variable selection problem

$\Downarrow$

Penalties

Shrinkage Approach (cont'd)

- We employed a Bayesian approach and used Markov chain Monte Carlo (MCMC) methods to estimate the parameters of the proposed joint model

  ▷ Shrinkage priors for the association parametes

  ▷ Priors that give a high probability of being near 0

  Bayesian lasso

  Bayesian ridge

  Horseshoe

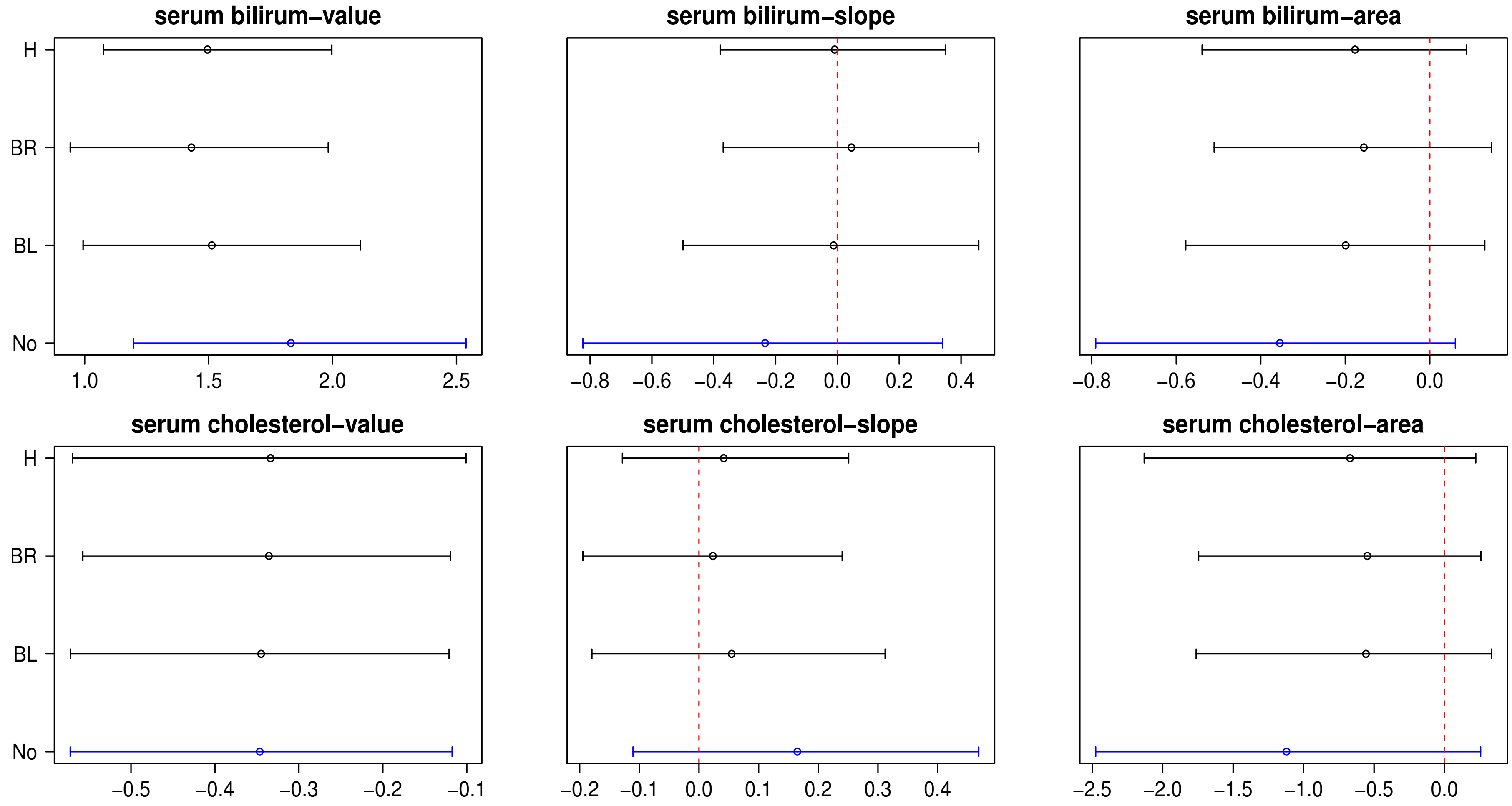
Shrinkage Approach (cont'd)

Analysis:

- **Longitudinal submodels** (Serum bilirubin and Serum cholesterol):
  - ▷ Fixed part: splines for time and gender
  - ▷ Random part: splines for time
- **Survival submodel** (Time-to-death):
  - ▷ underlying value, slope and the area under the curve for serum bilirubin and serum cholesterol
  - ▷ age and gender

Shrinkage Approach (cont'd)

Results:



Time-Varying Effects

Motivated by the heart data:

- Longitudinal response:
  - ▷ **aortic gradient**
- Time-to-event response:
  - ▷ time-to-**death/reoperation**

Andrinopoulou, E. R., Eilers, P. H., Takkenberg, J. J. and Rizopoulos, D. (2017). Improved dynamic predictions from joint models of longitudinal and survival data with time-varying effects using P-splines. Biometrics, doi: 10.1111/biom.12814.

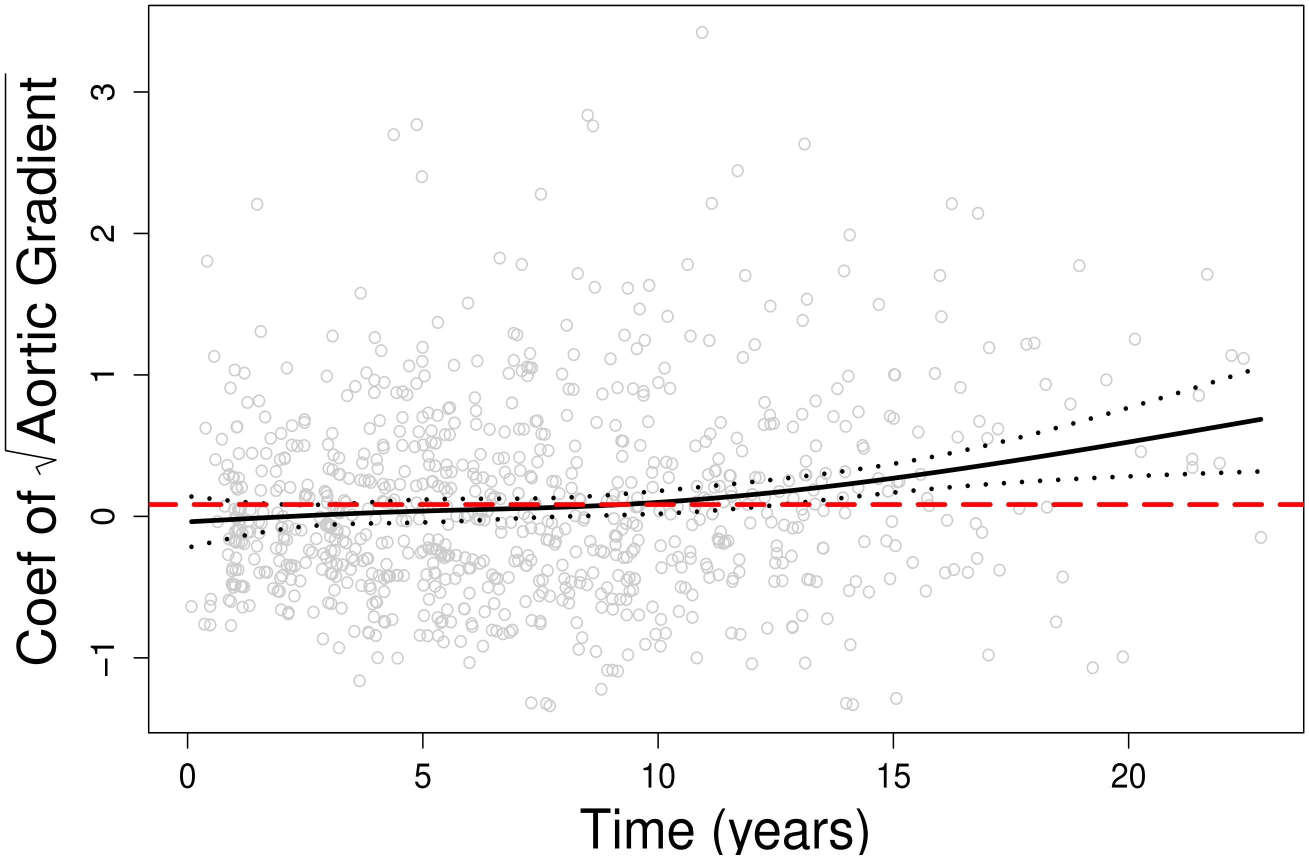
Time-Varying Effects (cont'd)

- Standard joint models assume a constant regression coefficient for the effect of the covariates.

  ▷ when treatment is initiated, the strength of the association between the longitudinal and survival outcomes may also change

Time-Varying Effects (cont'd)

Proportional hazard assumption - Time dependent Cox model



Time-Varying Effects (cont'd)

A time-varying coefficient joint model

Time-Varying Effects (cont'd)

Specifically,

$$h_i(t) \;=\; h_0(t)\exp[\gamma^\top w_i + f_j\{\lambda_j(t), \eta_i(t)\}],$$

where

- w_i is a vector of baseline covariates with a corresponding vector of regression coefficients γ
- $f_j\{\lambda_j(t_i), \eta_i(t)\}$ is the form of association (j $=$ 1, . . . , J) between the longitudinal and the survival outcomes $\rightarrow$ **underlying value, slope or area under the curve**

Time-Varying Effects (cont'd)

- We consider estimation of the function $\lambda_j(t)$ using the regression P-spline method, where

$$\lambda_j(t) = \sum_{\ell=1}^{L} \alpha_{j\ell} B_\ell(t, \nu),$$

where

  - $\alpha_{j\ell}$ is a set of parameters that capture the strength of association between the longitudinal and survival outcomes
  - $B_\ell(t_i, \nu)$ denotes the q-th basis function of a B-spline with knots $\nu_1, \ldots, \nu_Q$

- The idea behind the P-spline method is to assume a high number of knots and penalize the coefficients to tackle the problem of the large number or parameters

Time-Varying Effects (cont'd)

- We employ a Bayesian approach and use Markov chain Monte Carlo (MCMC) methods to estimate the parameters of the proposed joint model

- The penalty from the frequentist penalized likelihood translates into a prior distribution

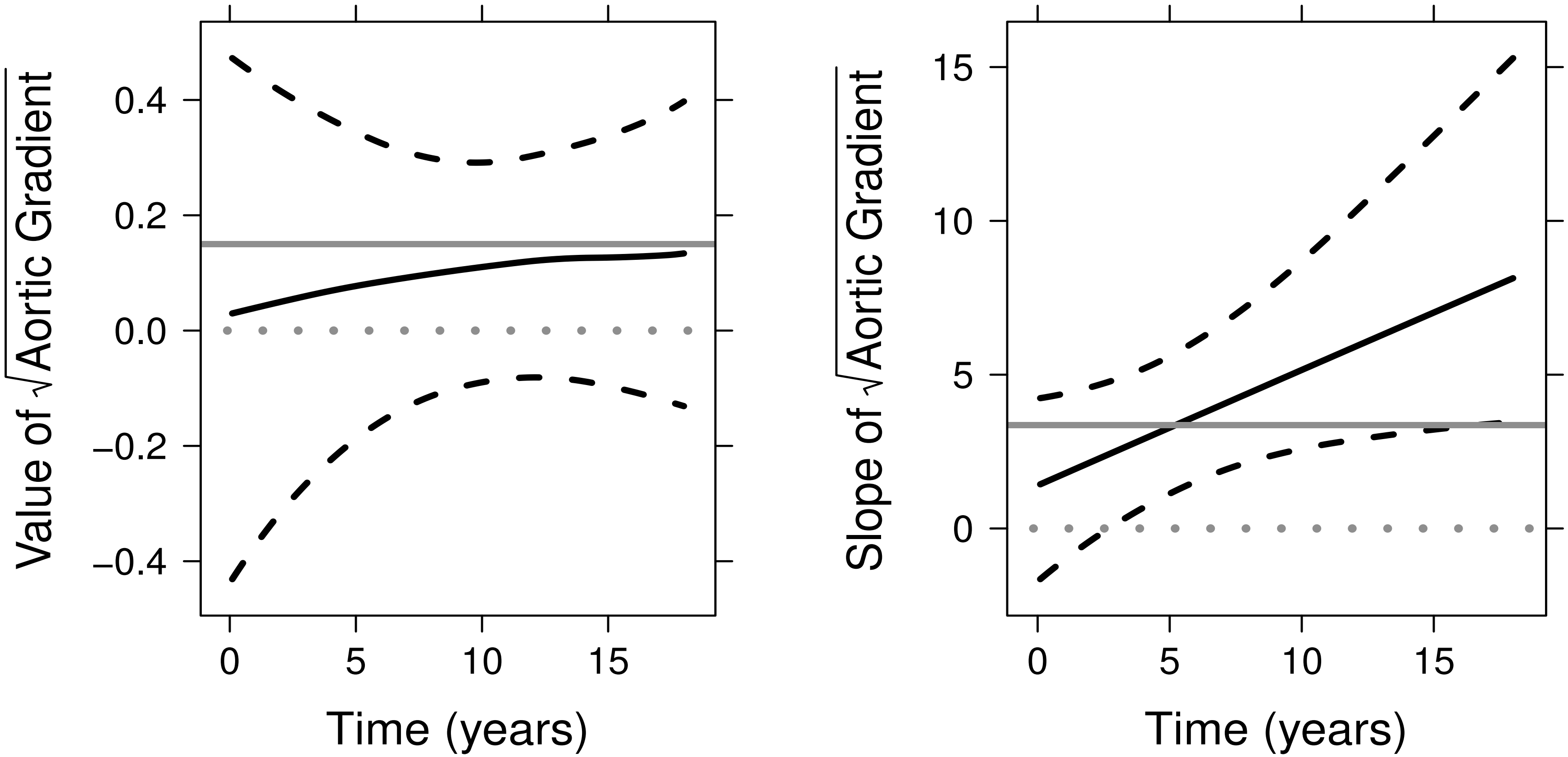
Time-Varying Effects (cont'd)

Analysis:

- **Longitudinal submodel** (Aortic gradient):
  - ▷ Fixed part: splines for time and gender
  - ▷ Random part: splines for time
- **Survival submodel** (Time-to-death/reoperation):
  - ▷ value and slope of aortic gradient
  - ▷ gender

Time-Varying Effects (cont'd)

Results:



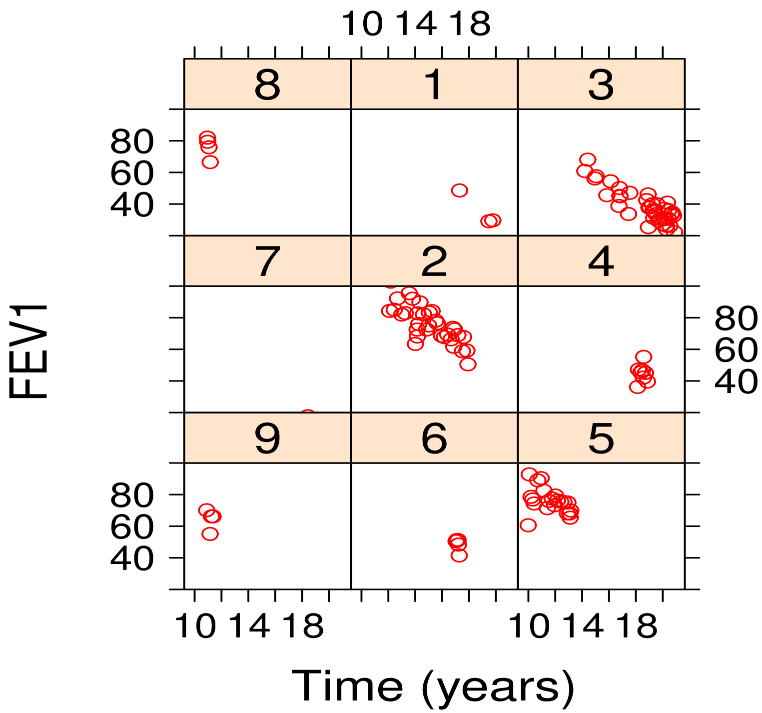
Latent Class Joint Model

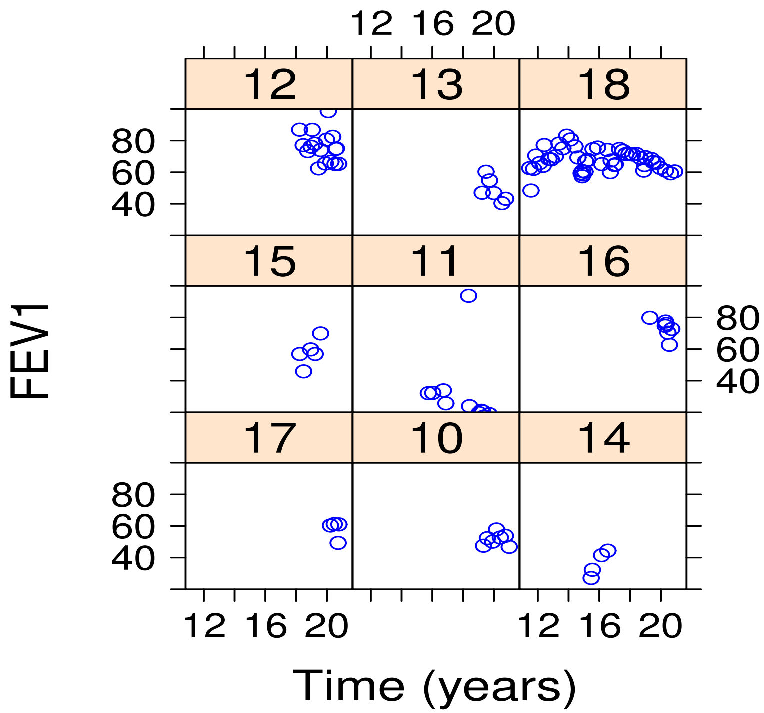
Motivated by the CF data:

- Longitudinal response:
  - ▷ **FEV$_1$**
- Time-to-event response:
  - ▷ time-to-**first exacerbation**

Andrinopoulou, E.R., Nasserinejad, K., Szczesniak, R. and Rizopoulos, D. (2018). Integrating Latent Classes in the Bayesian Shared Parameter Joint Model of Longitudinal and Survival Outcomes. arXiv preprint arXiv:1802.10015.

- Standard joint models assume homoscedasticity
  - ▷ Latent sup-populations ⇒ **FEV$_1$**





Latent Class Joint Model (cont'd)

- Special features
  - ▷ Time-to-**first exacerbation** is associated with **FEV**$_1$

$$\Downarrow$$

Joint modeling of longitudinal and survival data

- ▷ shared parameter joint models
- ▷ joint latent class models

Latent Class Joint Model (cont'd)

- Special features
  - ▷ Time-to-**first exacerbation** is associated with **FEV$_1$**

$$\Downarrow$$

Joint modeling of longitudinal and survival data

- ▷ shared parameter joint models
  - **+ quantify the strength of the association**
  - **– does not allow for latent sub-populations**
- ▷ joint latent class models

Latent Class Joint Model (cont'd)

- Special features
  - ▷ Time-to-**first exacerbation** is associated with **FEV**$_1$

$$\Downarrow$$

Joint modeling of longitudinal and survival data

- ▷ shared parameter joint models
- ▷ joint latent class models
  - \+ **postulates the existence of sub-populations**
  - − **does not quantify the strength of the association**

Latent Class Joint Model (cont'd)

Combine

$\swarrow \qquad \searrow$

$\boxed{\textbf{shared parameter joint model}}$ $\boxed{\textbf{joint latent class model}}$

Latent Class Joint Model (cont'd)

- **Longitudinal data**
  - ▷ Let y_i represent the repeated measurements of an outcome for the i-th patient, $i = 1, \ldots, n$

Mixed-effects submodel:

$y_{ic}(t) = x_i^\top(t)\beta_c + z_i^\top(t)b_{ic} + \epsilon_i(t) = \eta_{ic}(t) + \epsilon_i(t), \quad b_{ic} \sim N(0, D_c)$ and $\epsilon_i(t) \sim N(0, \Sigma_i)$

where

- ▷ $x_i^\top(t)\beta_c$ denotes the fixed part
- ▷ $z_i^\top(t)b_{ic}$ denotes the random part
- ▷ c indicates the class $c = 1, \ldots, C$

Latent Class Joint Model (cont'd)

- **Survival data**
  - ▷ Let T_i denote the observed failure time for the i-th patient and $\delta = 0, 1$ the event indicator

    Survival submodel:

    $$h_{ic}(t) = h_{0c}(t) \exp[\gamma_c^\top \omega_i + \alpha_c \eta_{ic}(t)],$$

    where
    - ▷ ω_i denotes the baseline covariates
    - ▷ $\eta_{ic}(t)$ denotes the longitudinal outcome
    - ▷ α_c association parameter

Latent Class Joint Model (cont'd)

- **Problem**:
  - ▷ **What is the optimal number of classes**
- **Solution**: *Commonly used*
  - ▷ **BIC**: Frequentist
  - ▷ **DIC**: Bayesian

Latent Class Joint Model (cont'd)

- **Problem**:
  - ▷ **What is the optimal number of classes**
- **Solution**: *Commonly used*
  - ▷ **BIC**: Frequentist
  - ▷ $\boxed{\textbf{DIC}\text{: Bayesian}}$ $\rightarrow$ **time-consuming**

Latent Class Joint Model (cont'd)

- **Problem**:
  - ▷ **What is the optimal number of classes**
- **Solution**: *Rousseau and Mengersen (RM), Journal of Royal Statistical Society, 2011*
  - ▷ **Step 1:** fit the model with a high number of classes

Latent Class Joint Model (cont'd)

- **Problem**:
  - ▷ **What is the optimal number of classes**
- **Solution**: *Rousseau and Mengersen (RM), Journal of Royal Statistical Society, 2011*
  - ▷ **Step 2:** obtain the number of non-empty classes at each iteration as:

$$c_{opt} = C - \sum_{c=1}^{C} I\left(\frac{N_c}{N} \leq \psi\right),$$

where N_c is the number of observations in class c, N is the total number of observations and ψ is a predefined number

Latent Class Joint Model (cont'd)

- **Problem**:
  - ▷ **What is the optimal number of classes**
- **Solution**: *Rousseau and Mengersen (RM), Journal of Royal Statistical Society, 2011*
  - ▷ **Step 3:** calculate the posterior mode of the non-empty classes

Latent Class Joint Model (cont'd)

- **Problem**:
  - ▷ **What is the optimal number of classes**
- **Solution**: *Rousseau and Mengersen (RM), Journal of Royal Statistical Society, 2011*
  - ▷ **Step 4:** fit the model with the optimal number of classes

Latent Class Joint Model (cont'd)

- We employed a Bayesian approach and used Markov chain Monte Carlo (MCMC) methods to estimate the parameters of the proposed joint model.
  For the class we assume,

$$\pi_{ic} \sim Dirichlet(A_c)$$

  - ▷ $A_c = A_1, \ldots, A_C$
  - ▷ $A_c < d/2$
  - ▷ d number of class-specific parameters

Latent Class Joint Model (cont'd)

Analysis:

- **Longitudinal submodels** (FEV$_1$):
  - ▷ Fixed part: splines for age and gender, (*other clinical characteristics*)
  - ▷ Random part: splines for age
- **Survival submodel** (Time-to-first exacerbation):
  - ▷ value of FEV_1
  - ▷ gender

Latent Class Joint Model (cont'd)

- Optimal number of classes

  RM approach: 6 classes

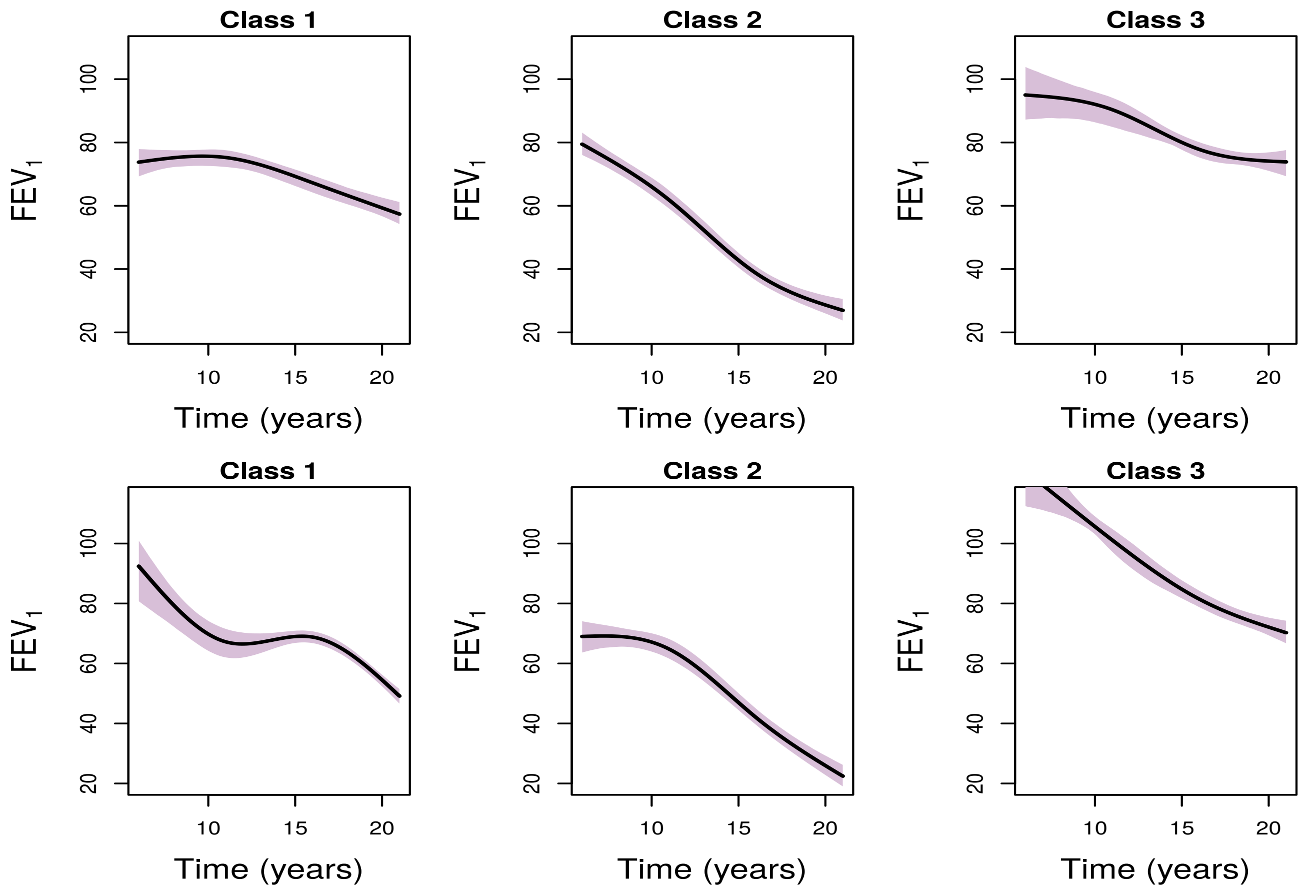
  ▷ cut off (ψ) 8% $\rightarrow$ **3 classes**

  ▷ cut off (ψ) 10% $\rightarrow$ **3 classes**

  ▷ cut off (ψ) 15% $\rightarrow$ **3 classes**

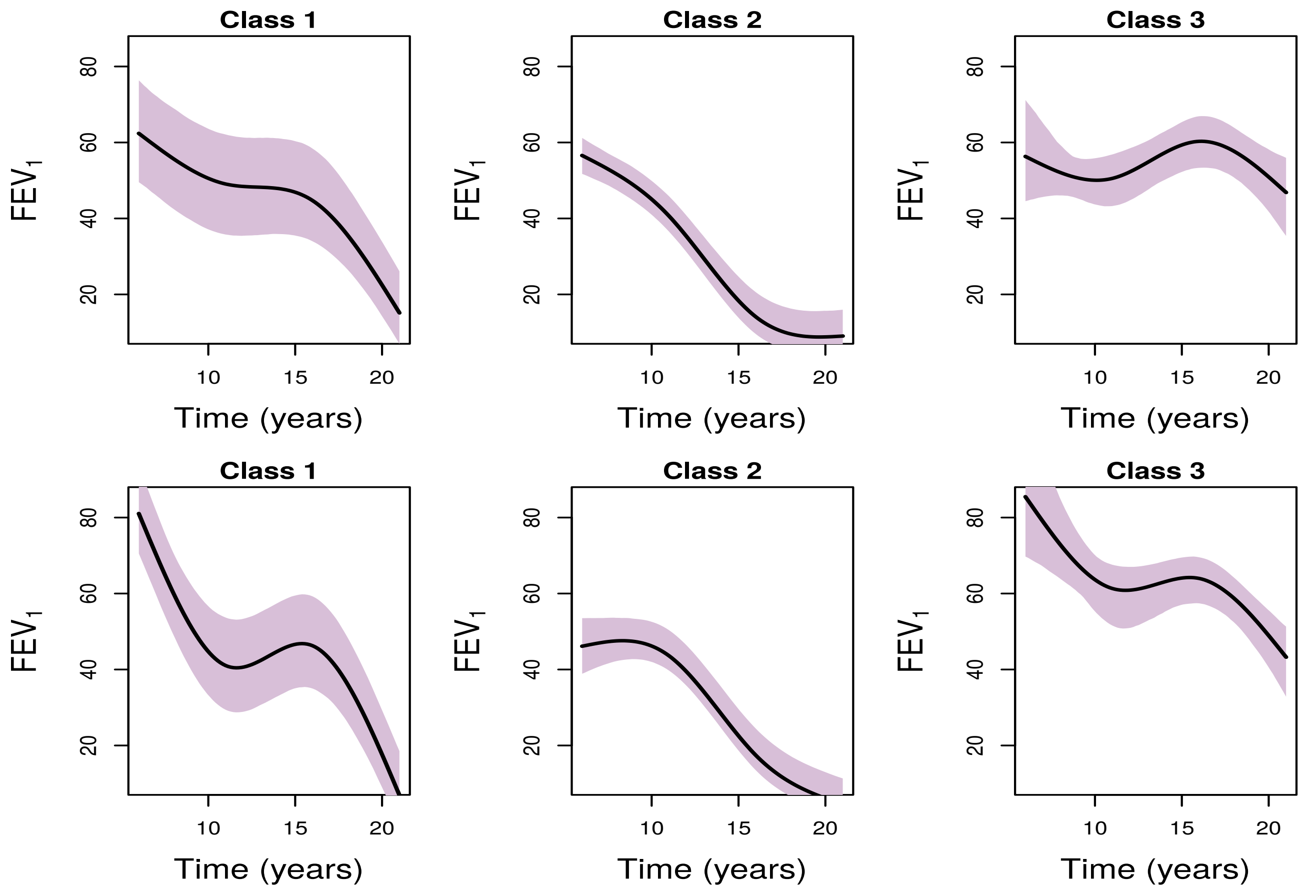
Latent Class Joint Model (cont'd)

Results: Less sick patients (females/males)



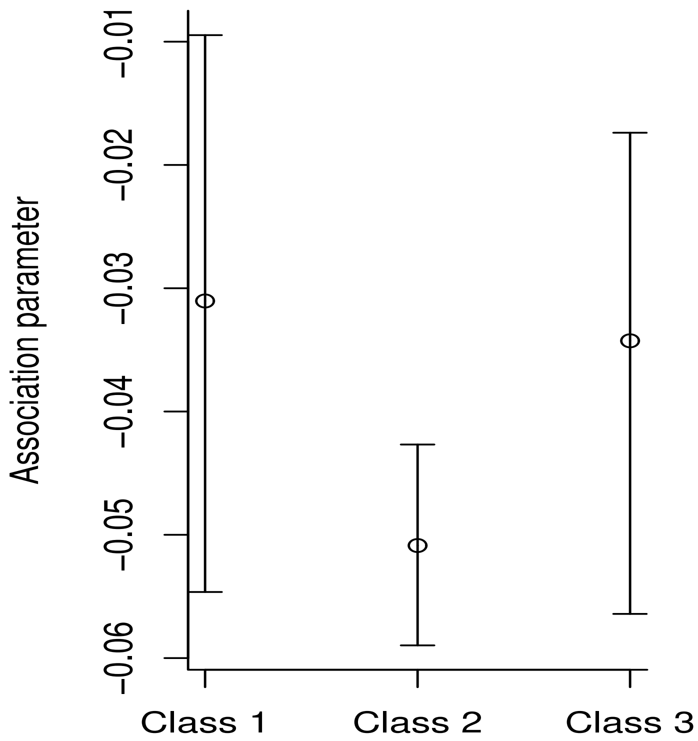
Latent Class Joint Model (cont'd)

Results: More sick patients (females/males)



Latent Class Joint Model (cont'd)

Results:



Individualized Predictions

Clinical Decision Making

- Guide clinical decision making $\rightarrow$ use **complete** biomarker information

  - **Dynamic nature:**

    different progression per patient

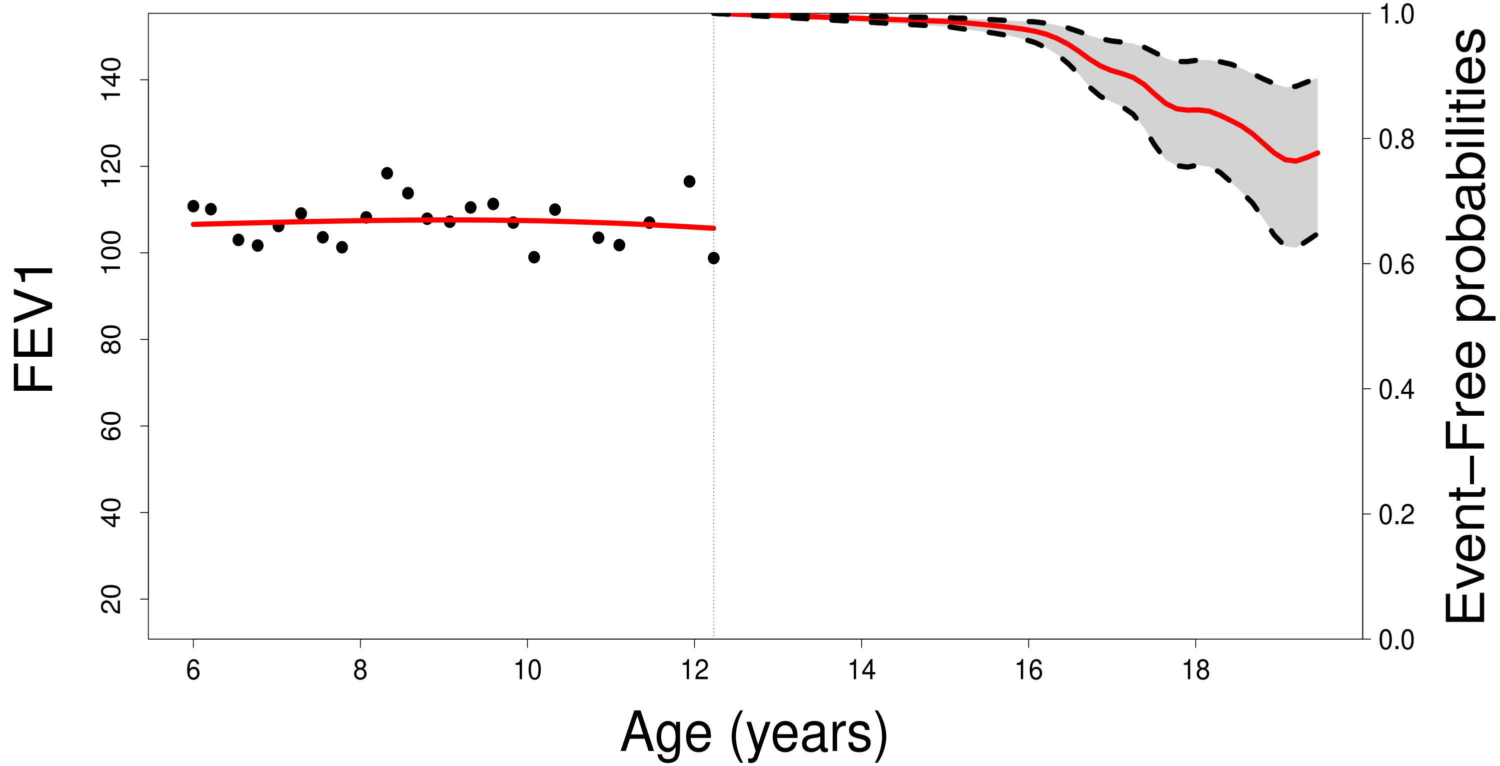
    progression changes over time within the same patient

$\boxed{\textbf{Repeated measurements provide better predictions}}$

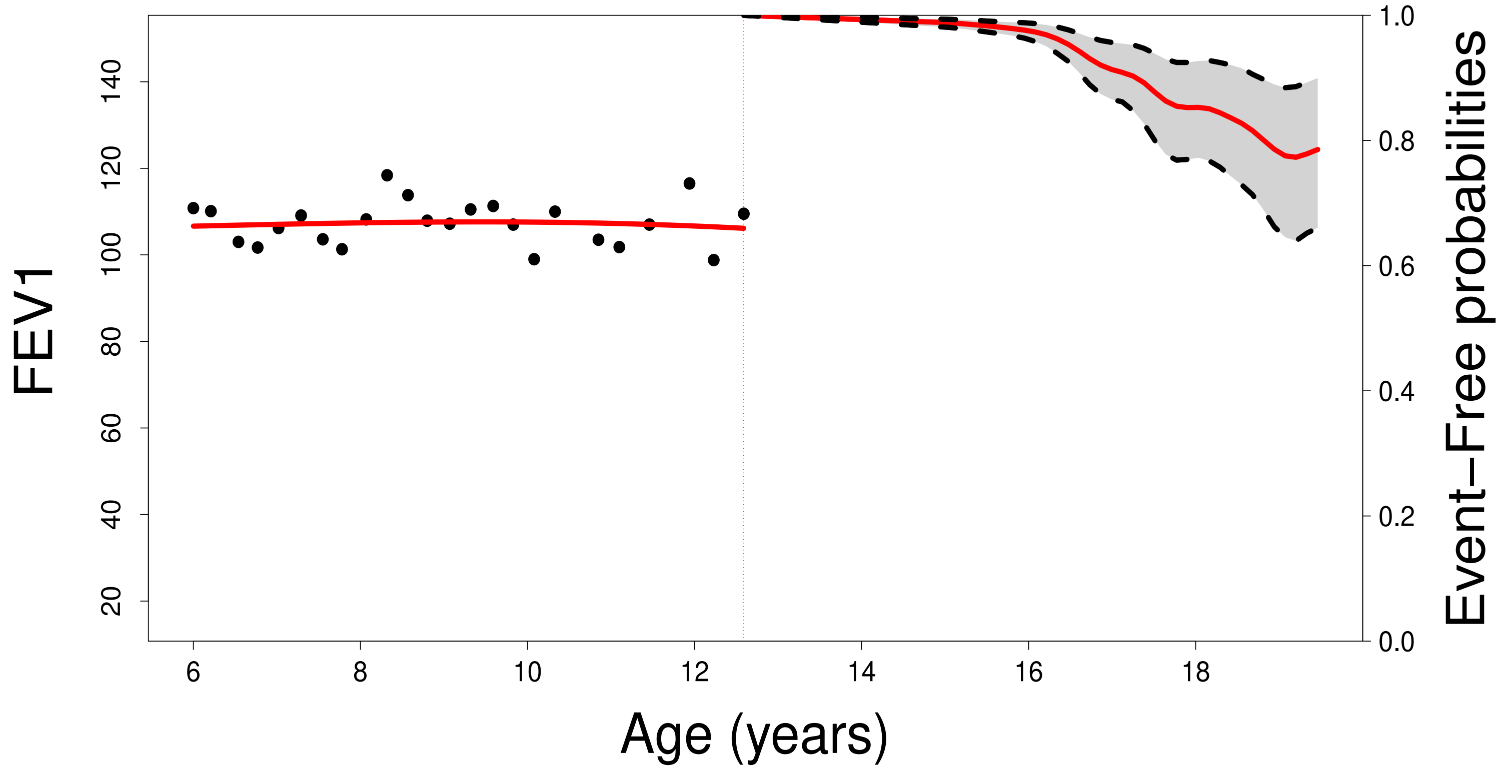
Clinical Decision Making (cont'd)

Subject-specific risk predictions based on the joint model $\rightarrow$ useful tool

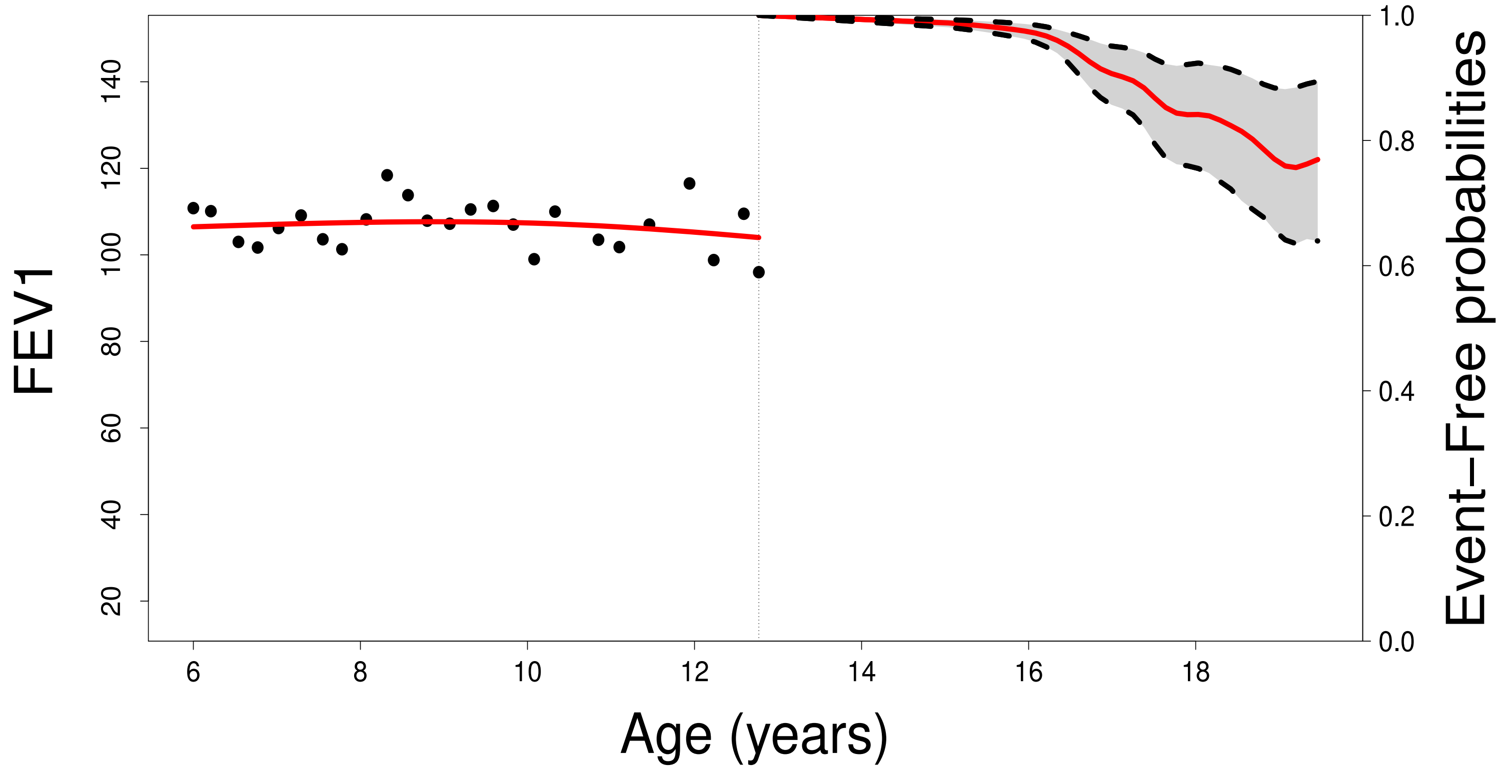
Dynamic Predictions



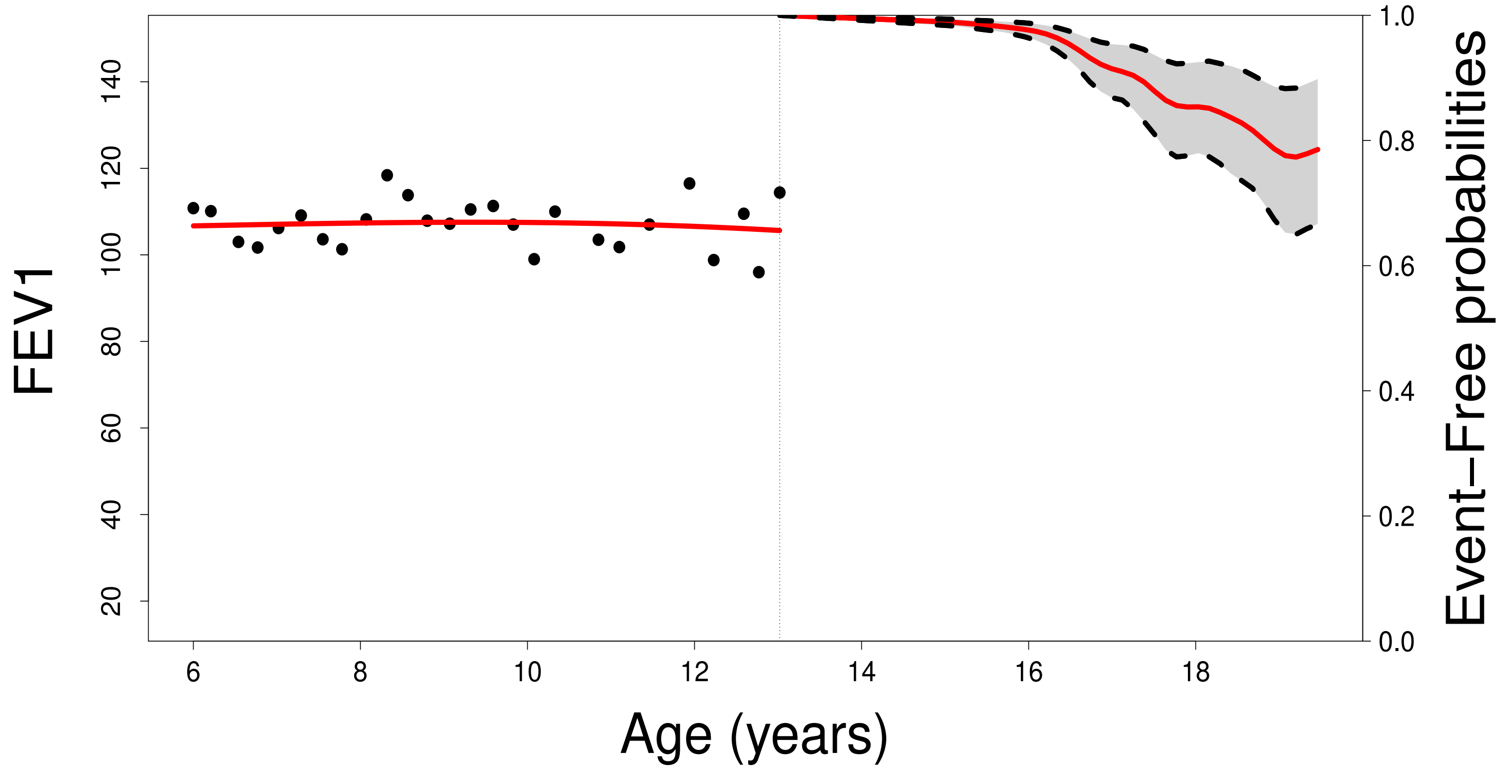
Dynamic Predictions



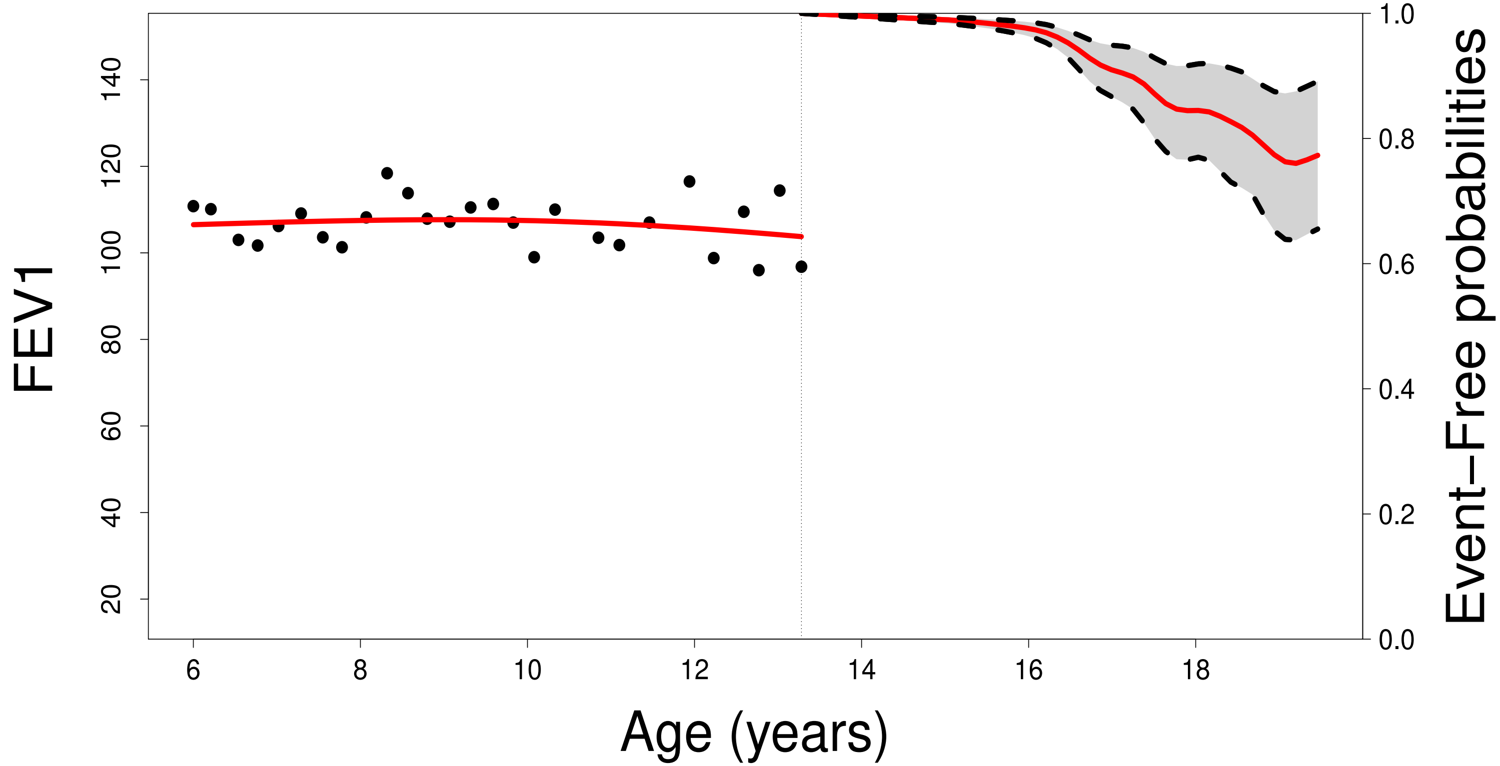
Dynamic Predictions



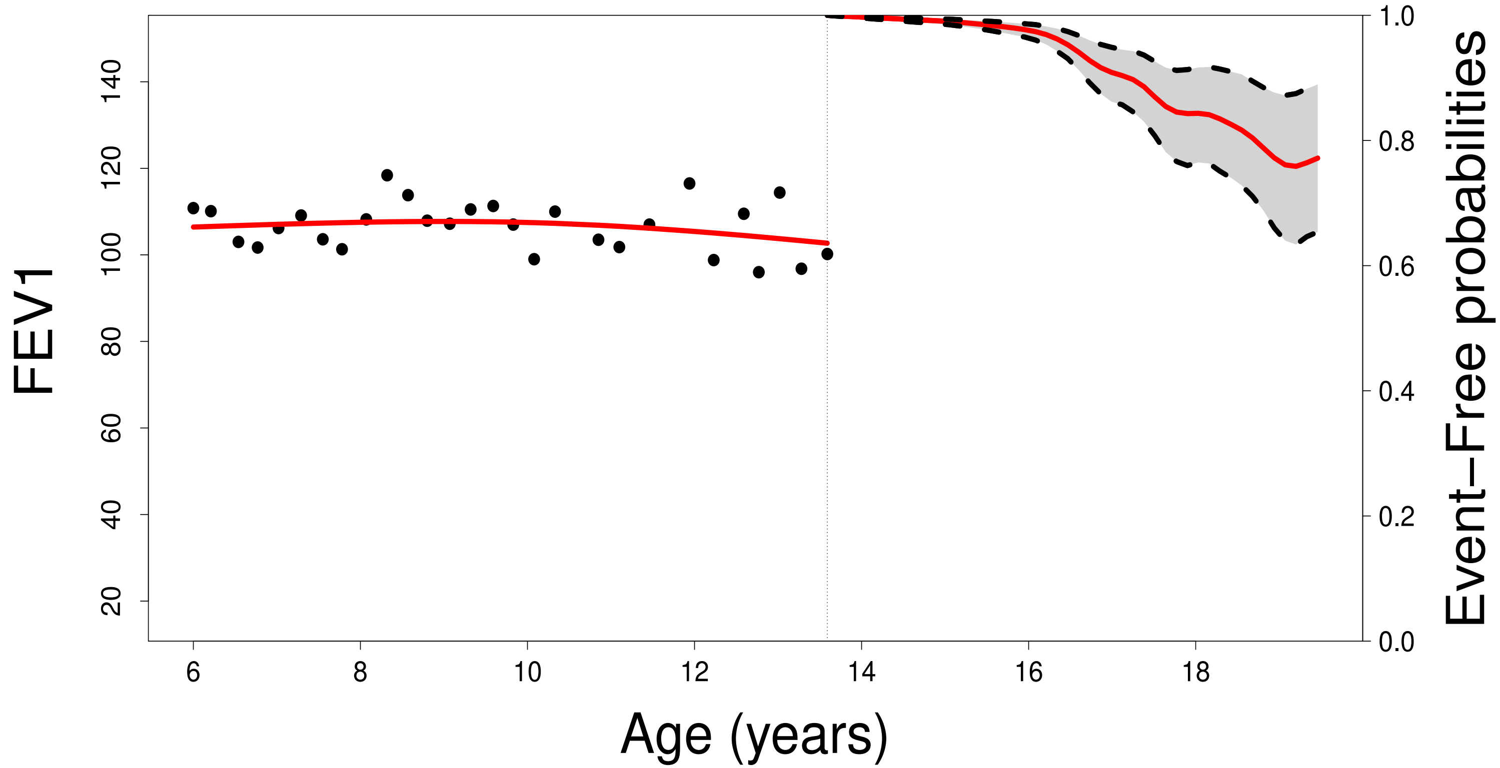
Dynamic Predictions



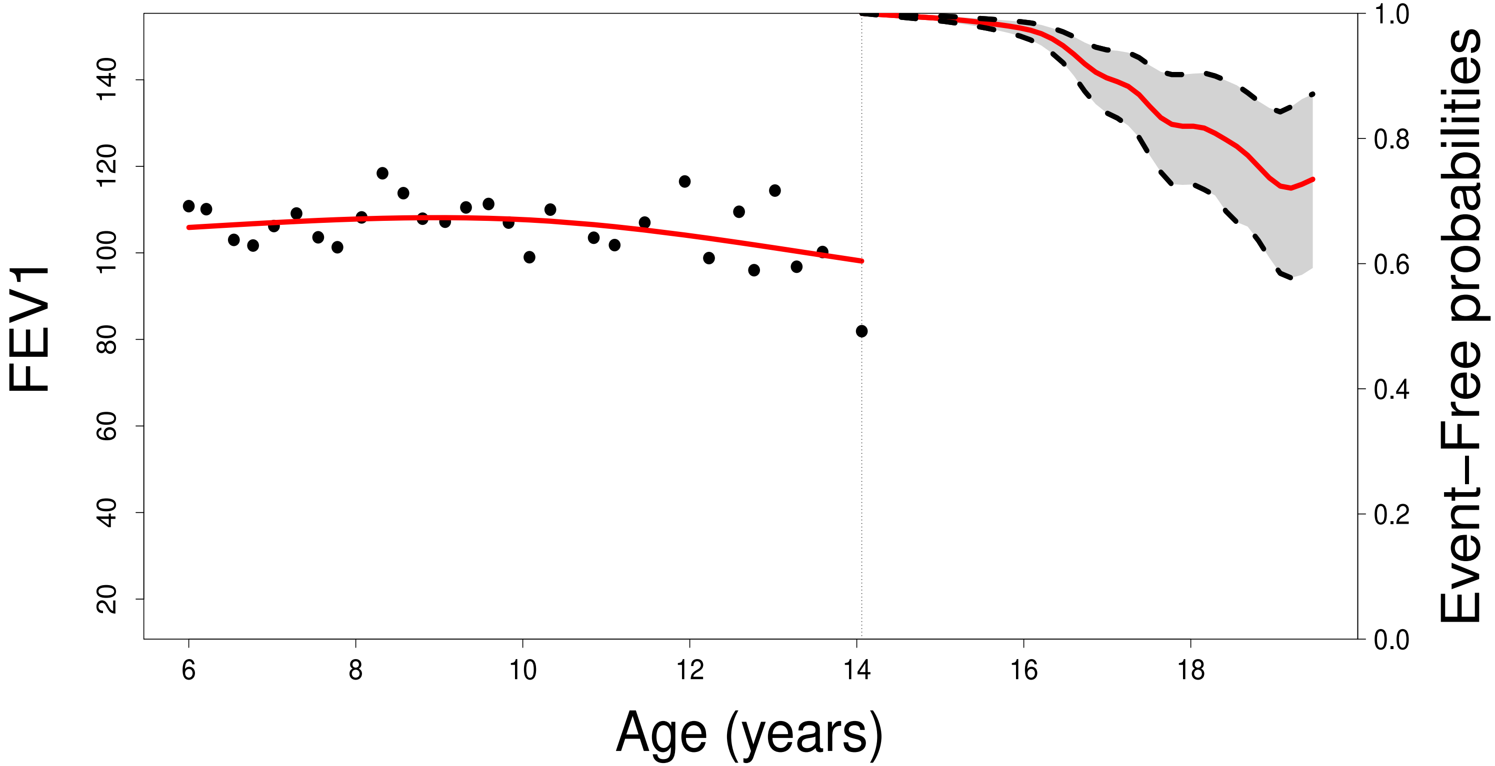
Dynamic Predictions



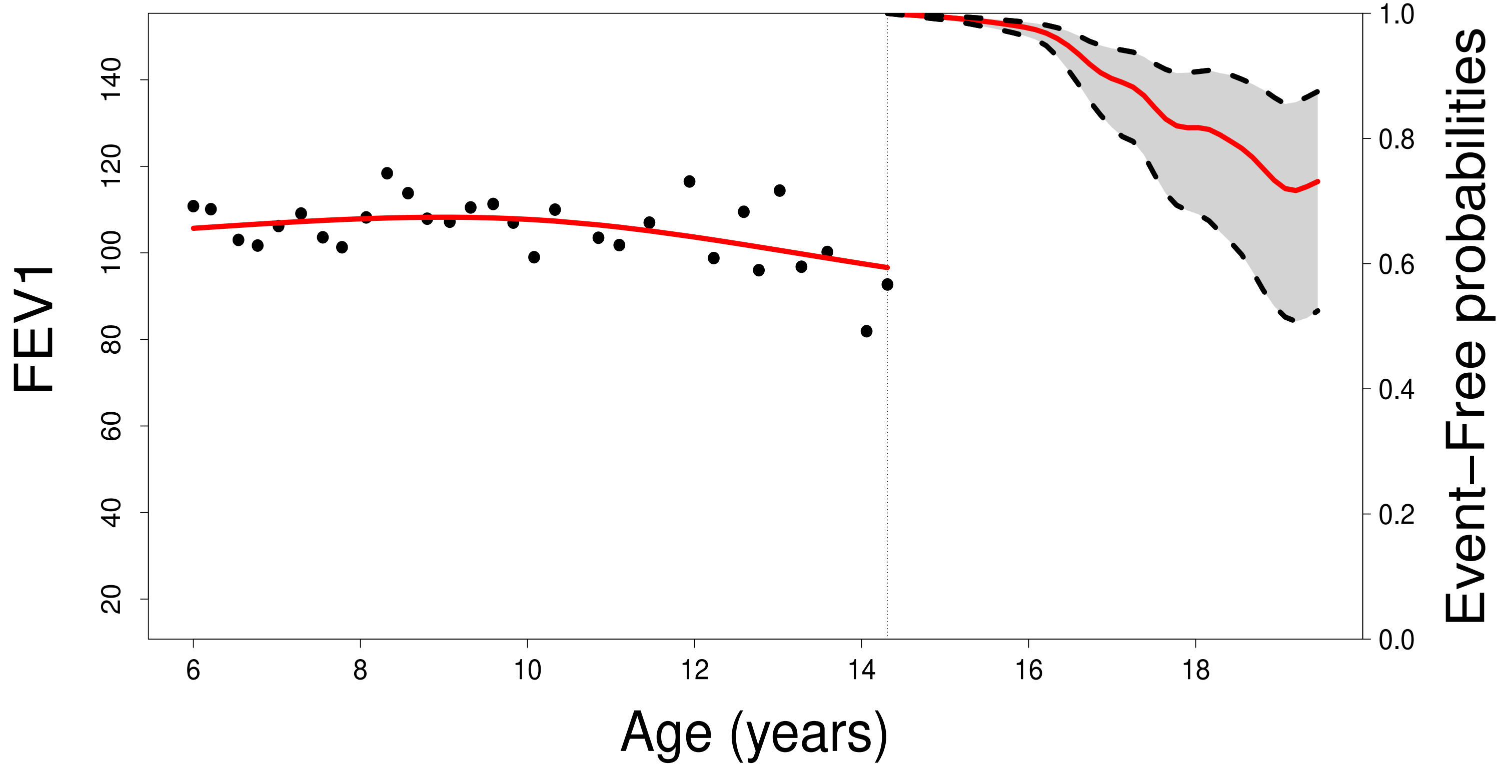
Dynamic Predictions



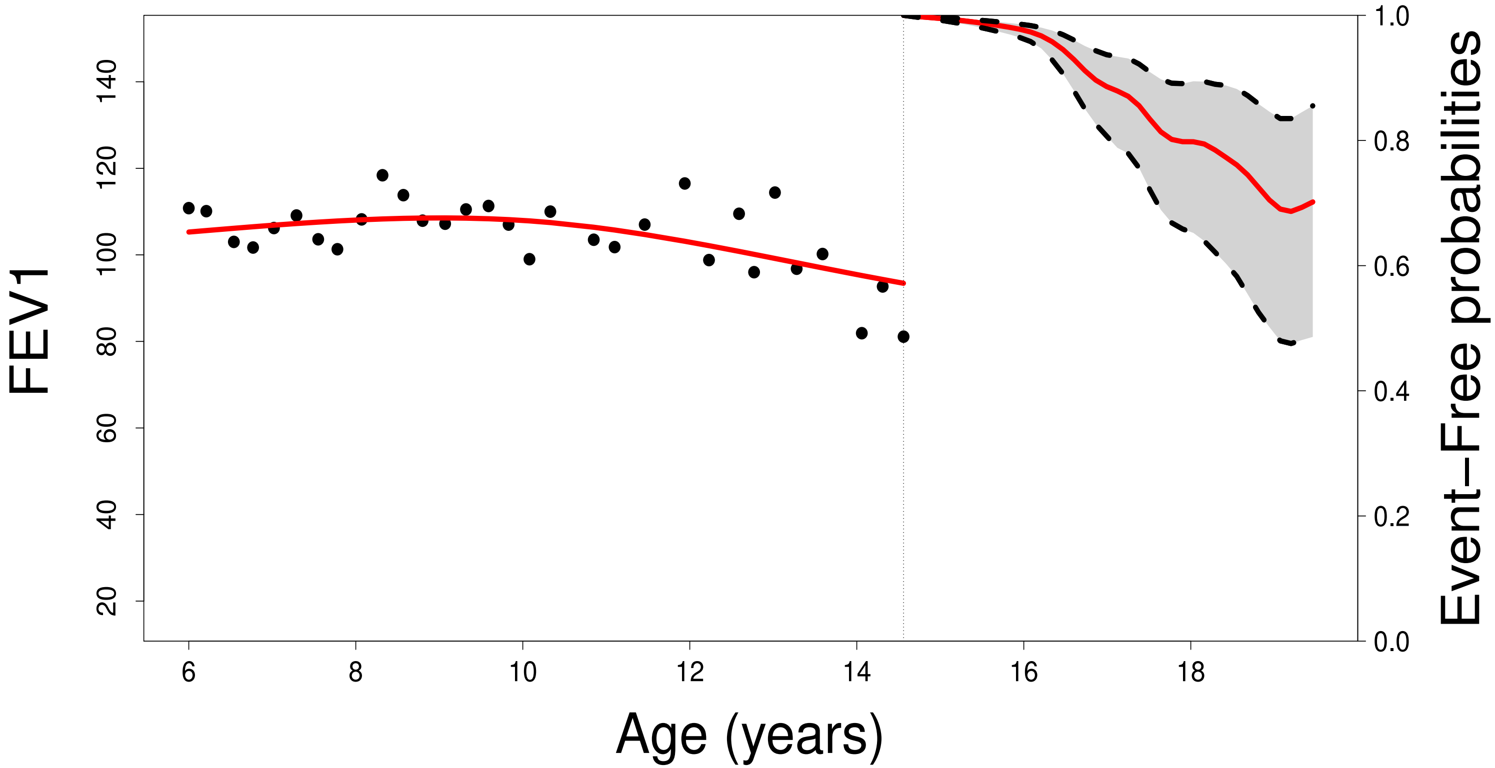
Dynamic Predictions



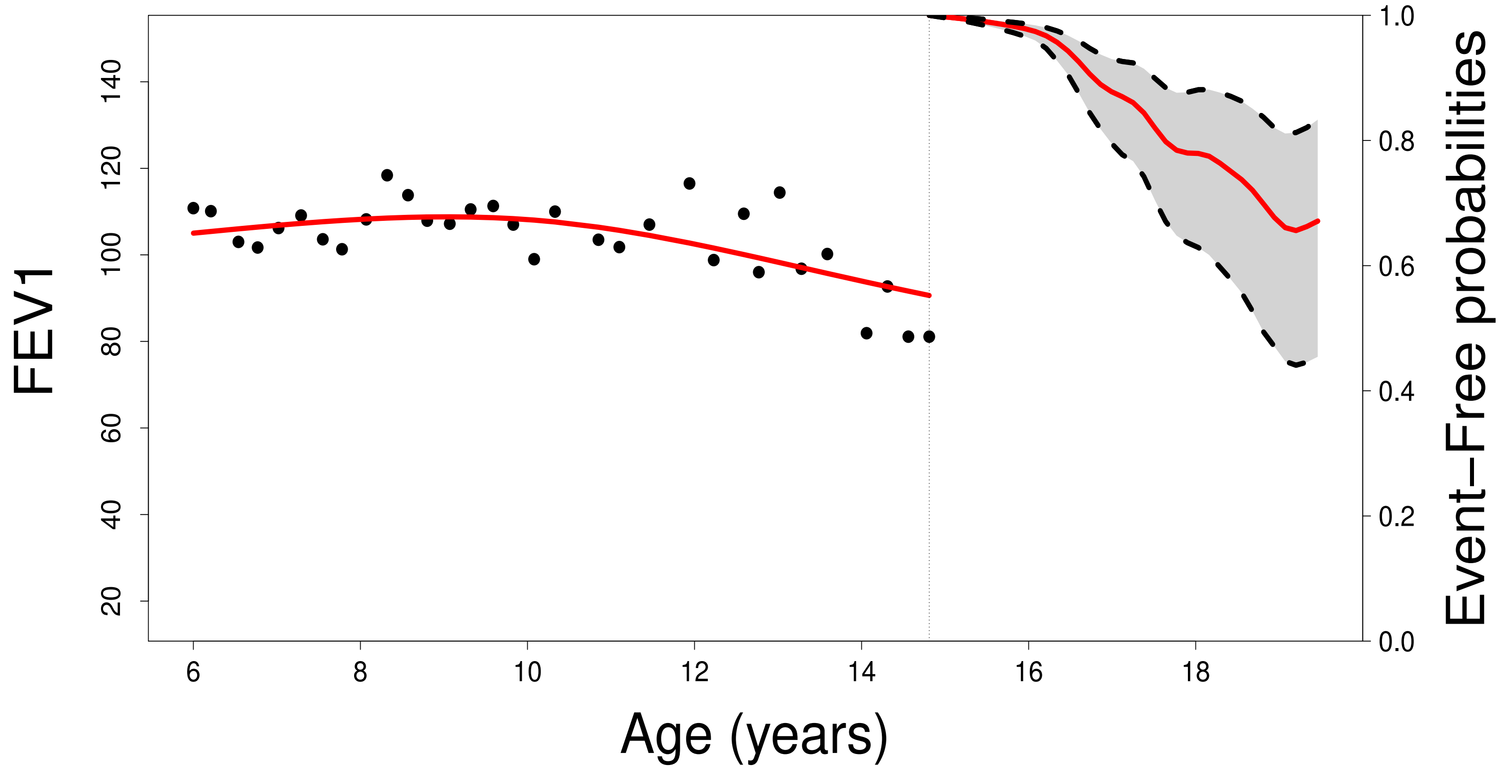
Dynamic Predictions



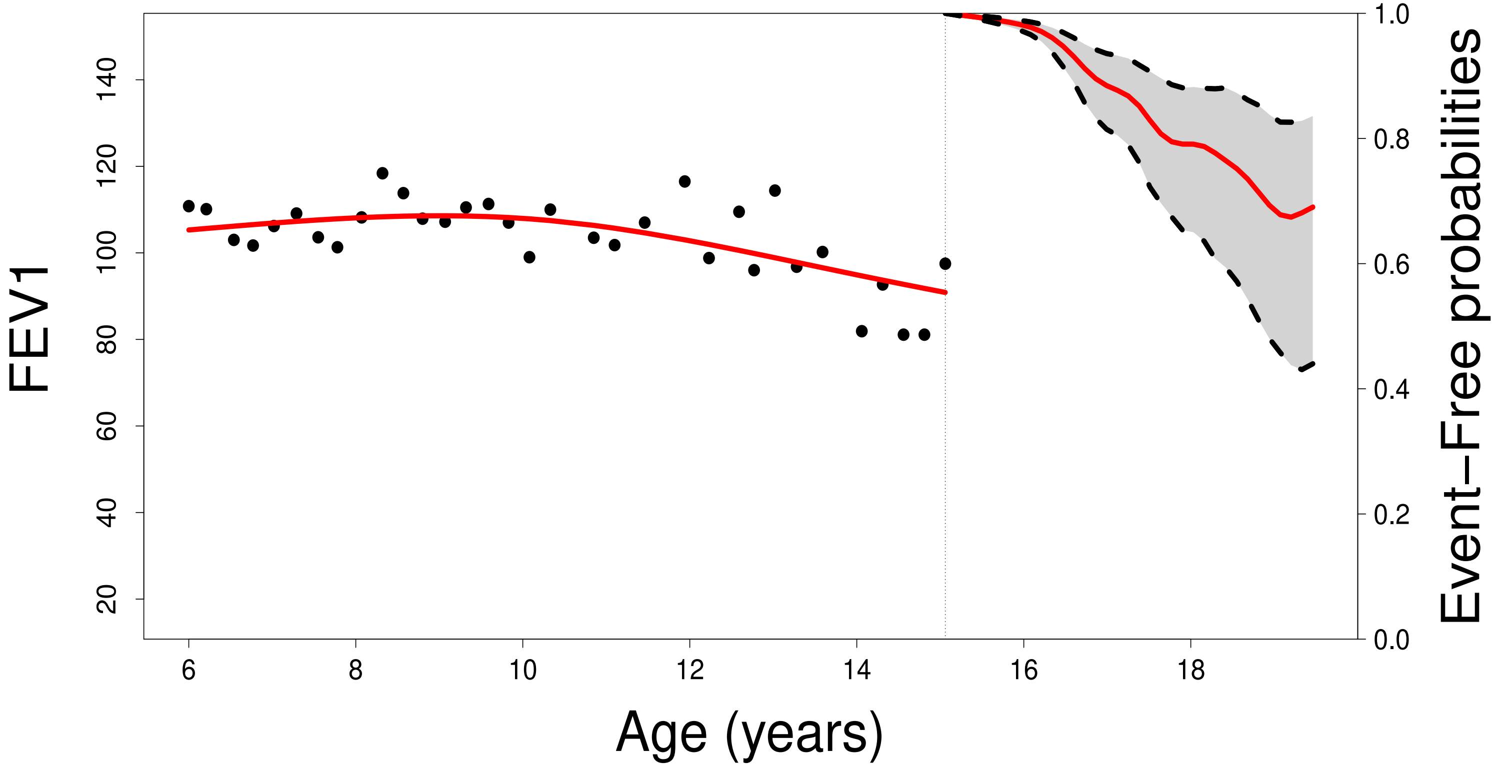
Dynamic Predictions



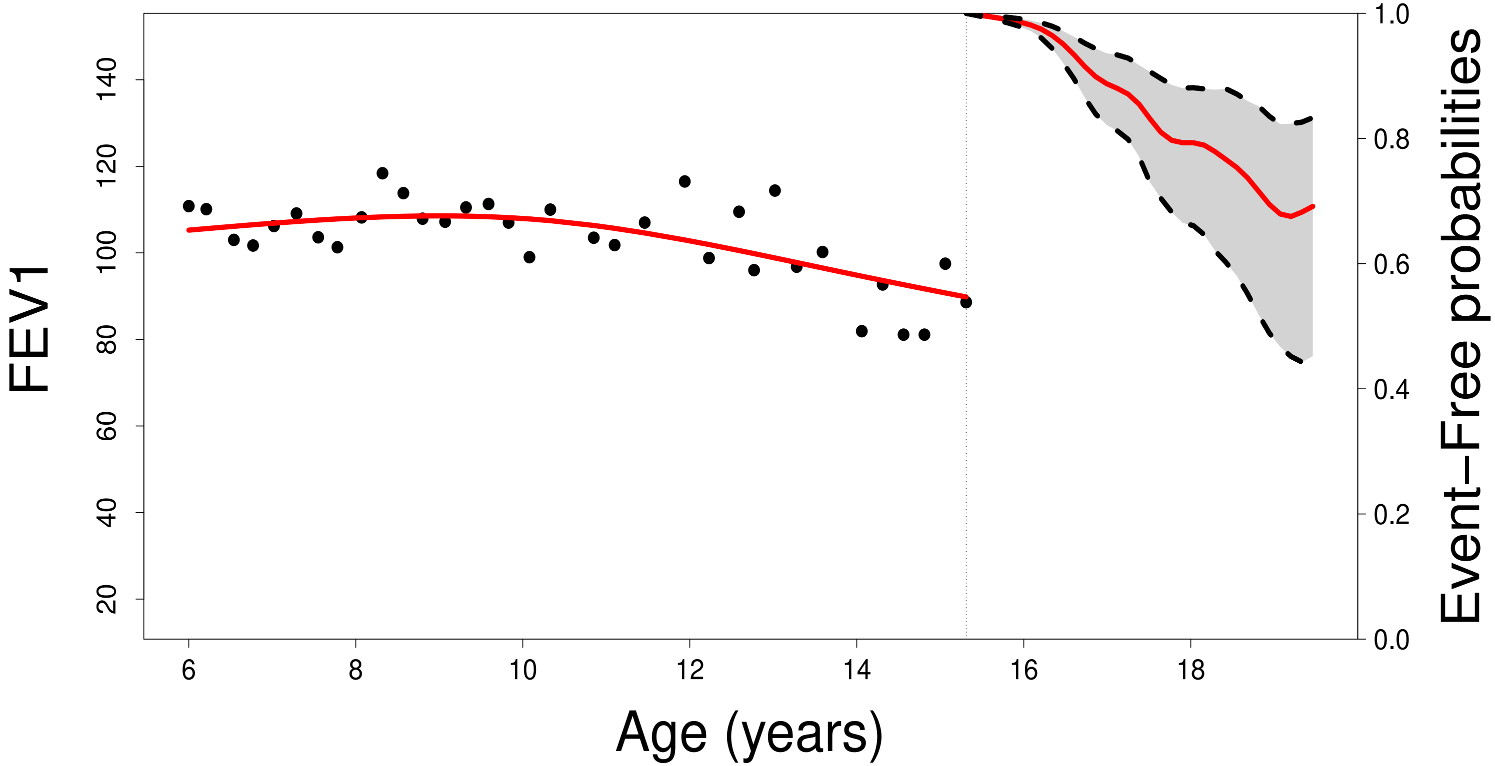
Dynamic Predictions



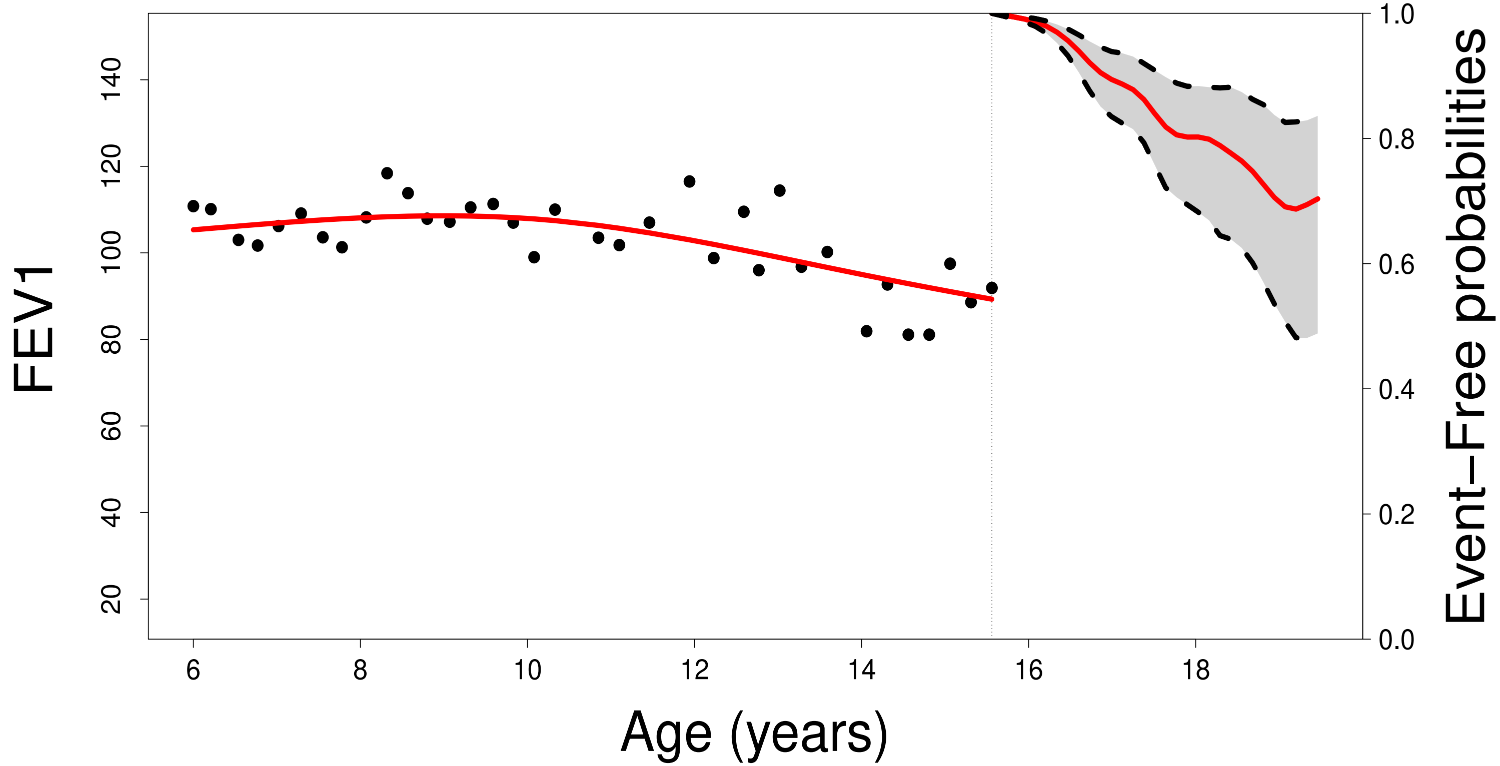
Dynamic Predictions



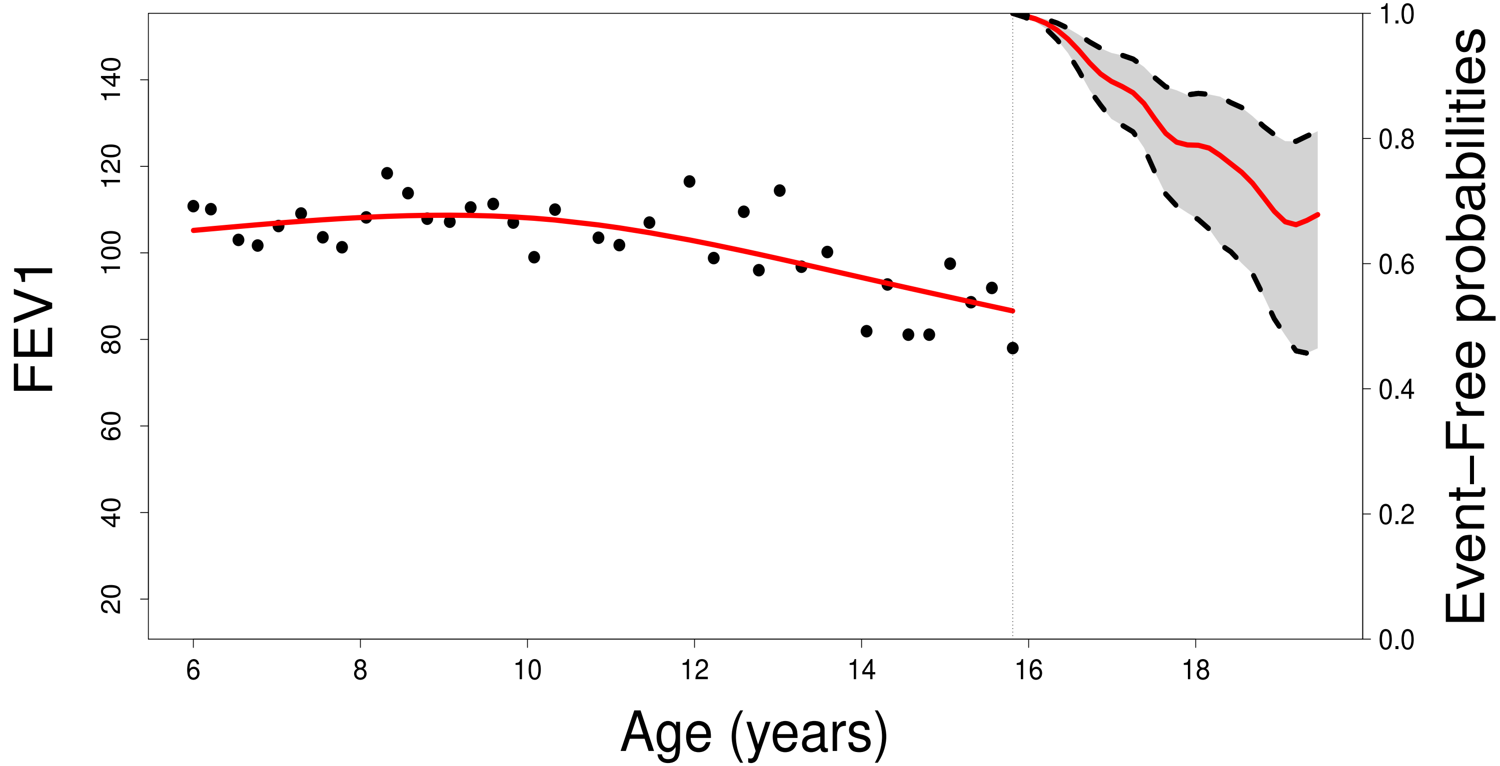
Dynamic Predictions



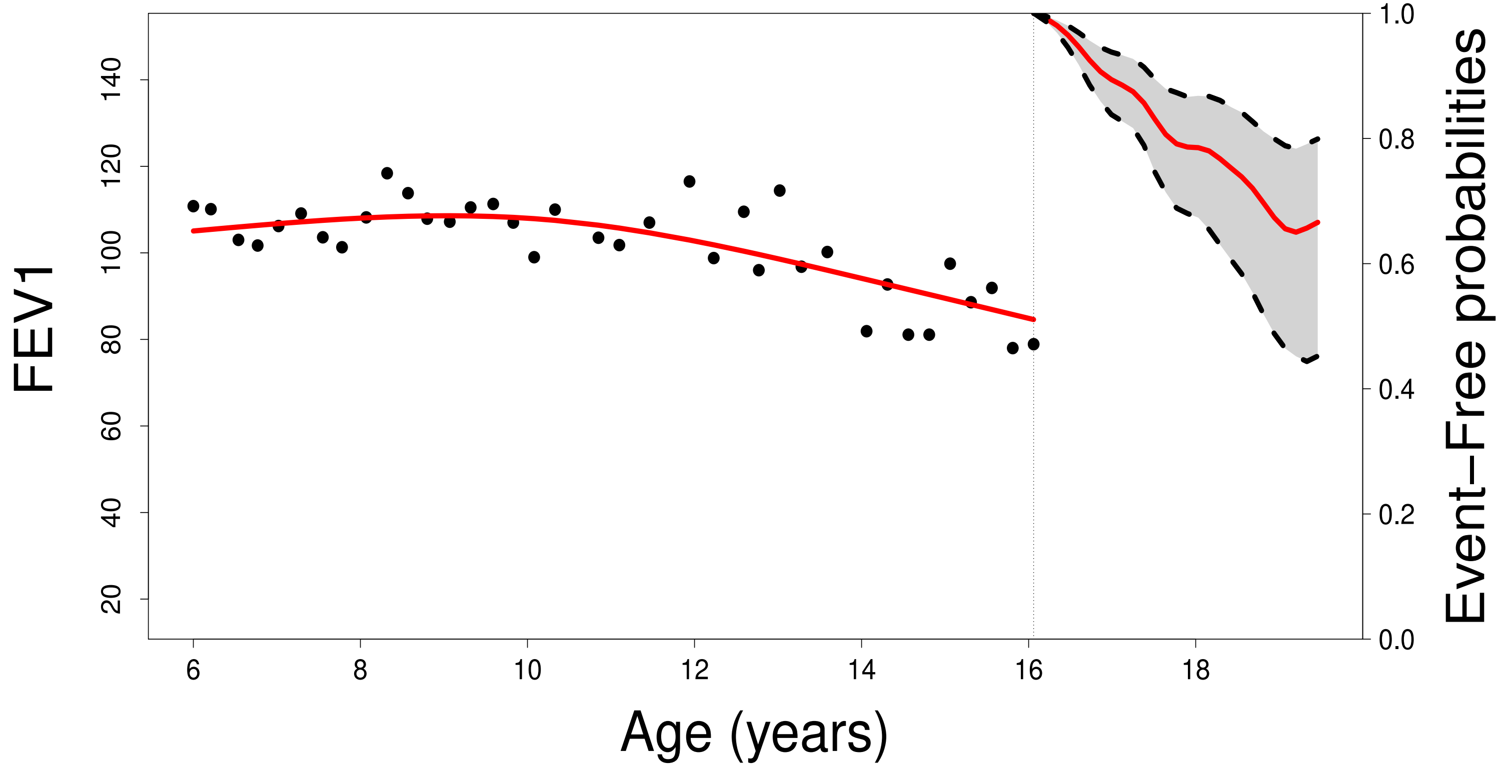
Dynamic Predictions



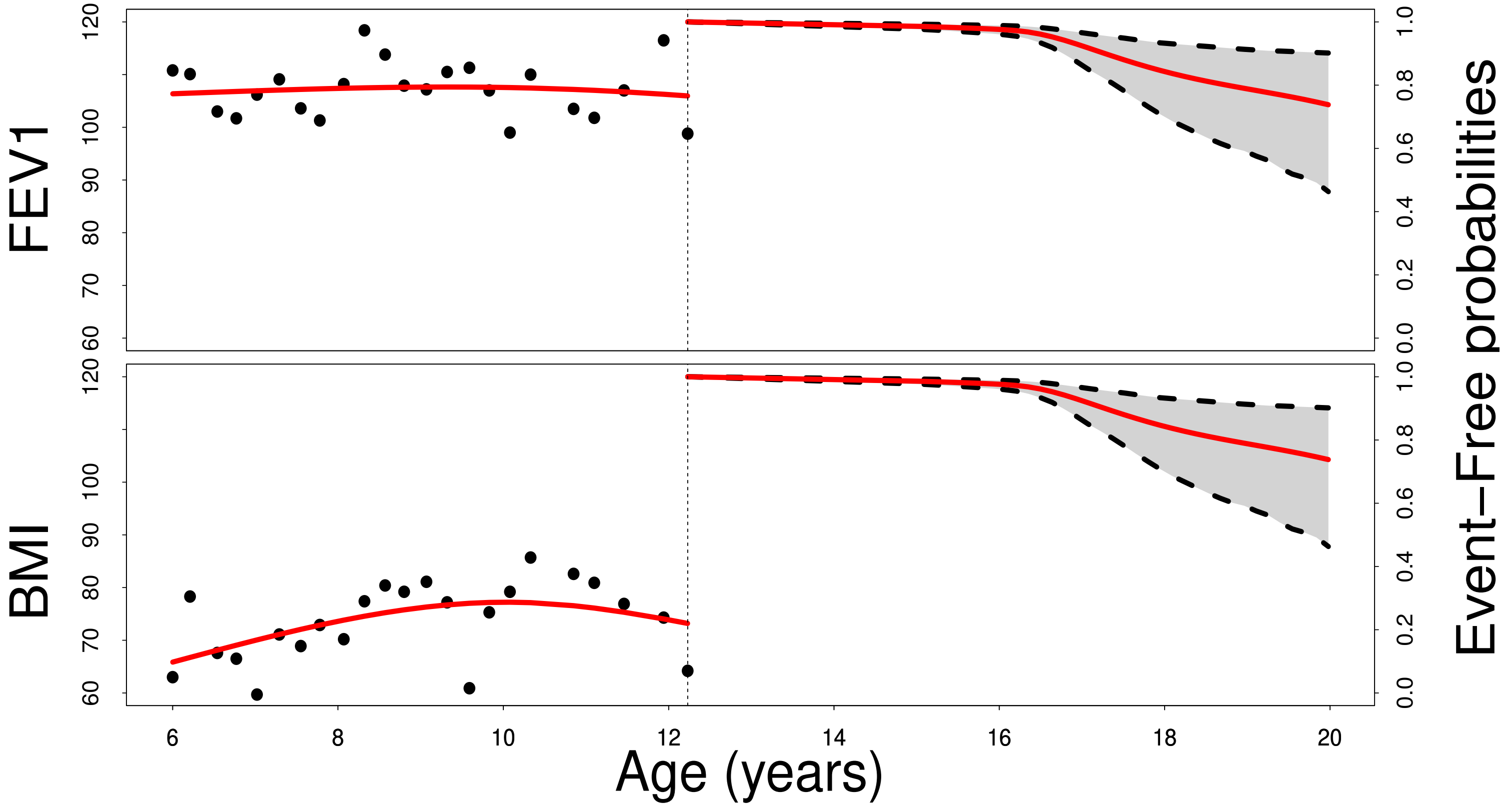
Dynamic Predictions



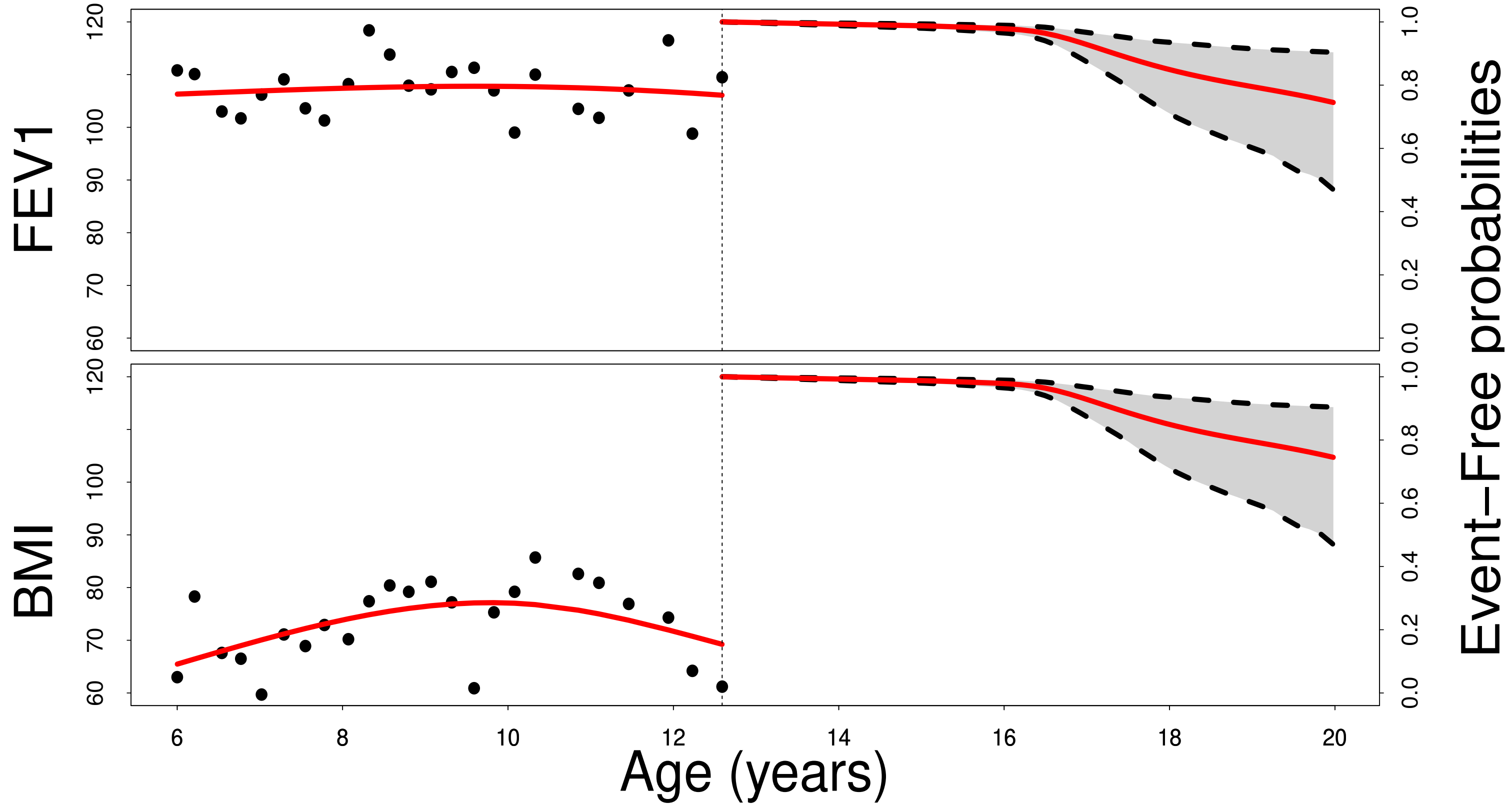
Dynamic Predictions



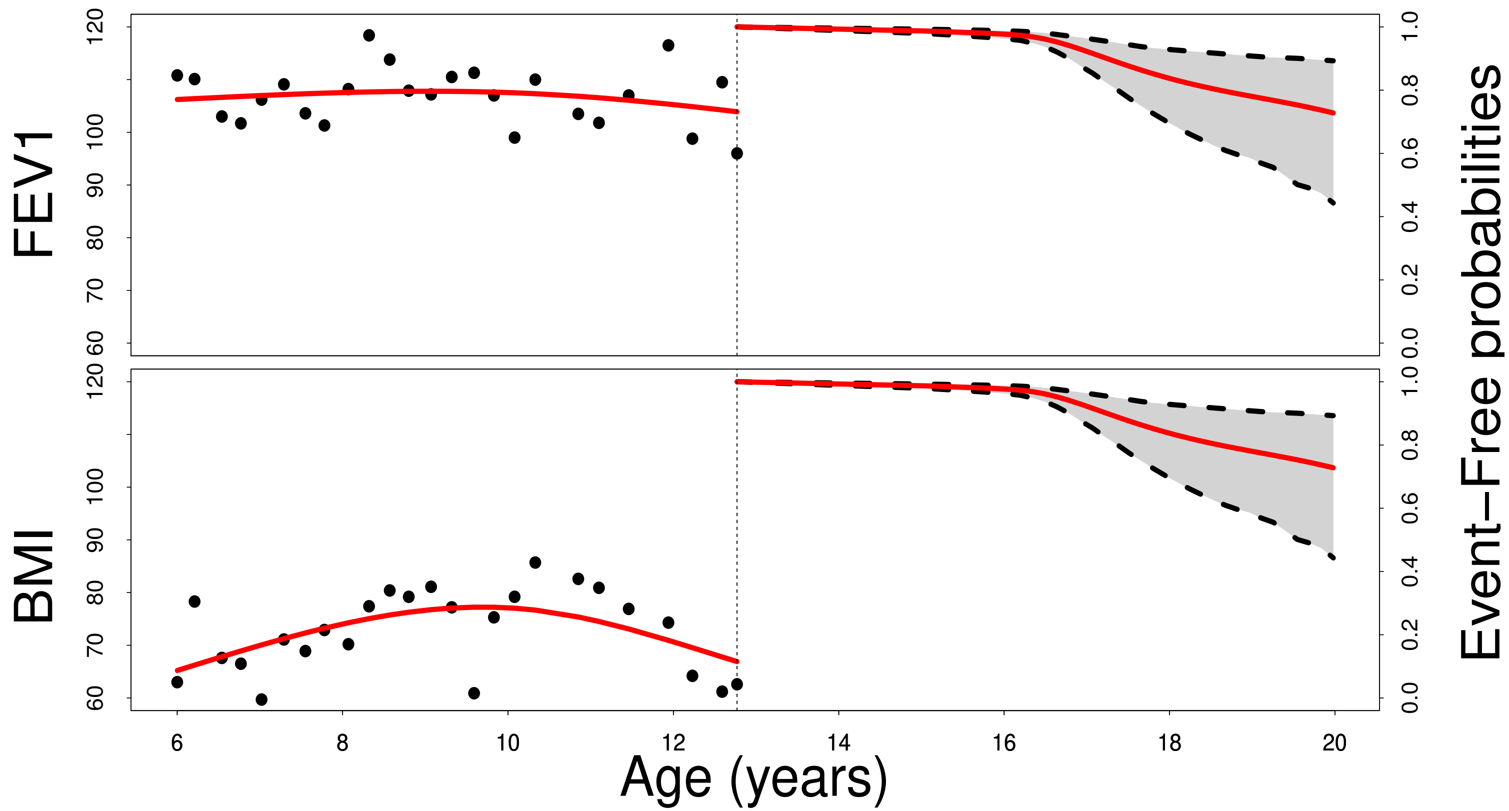
Dynamic Predictions (cont'd)



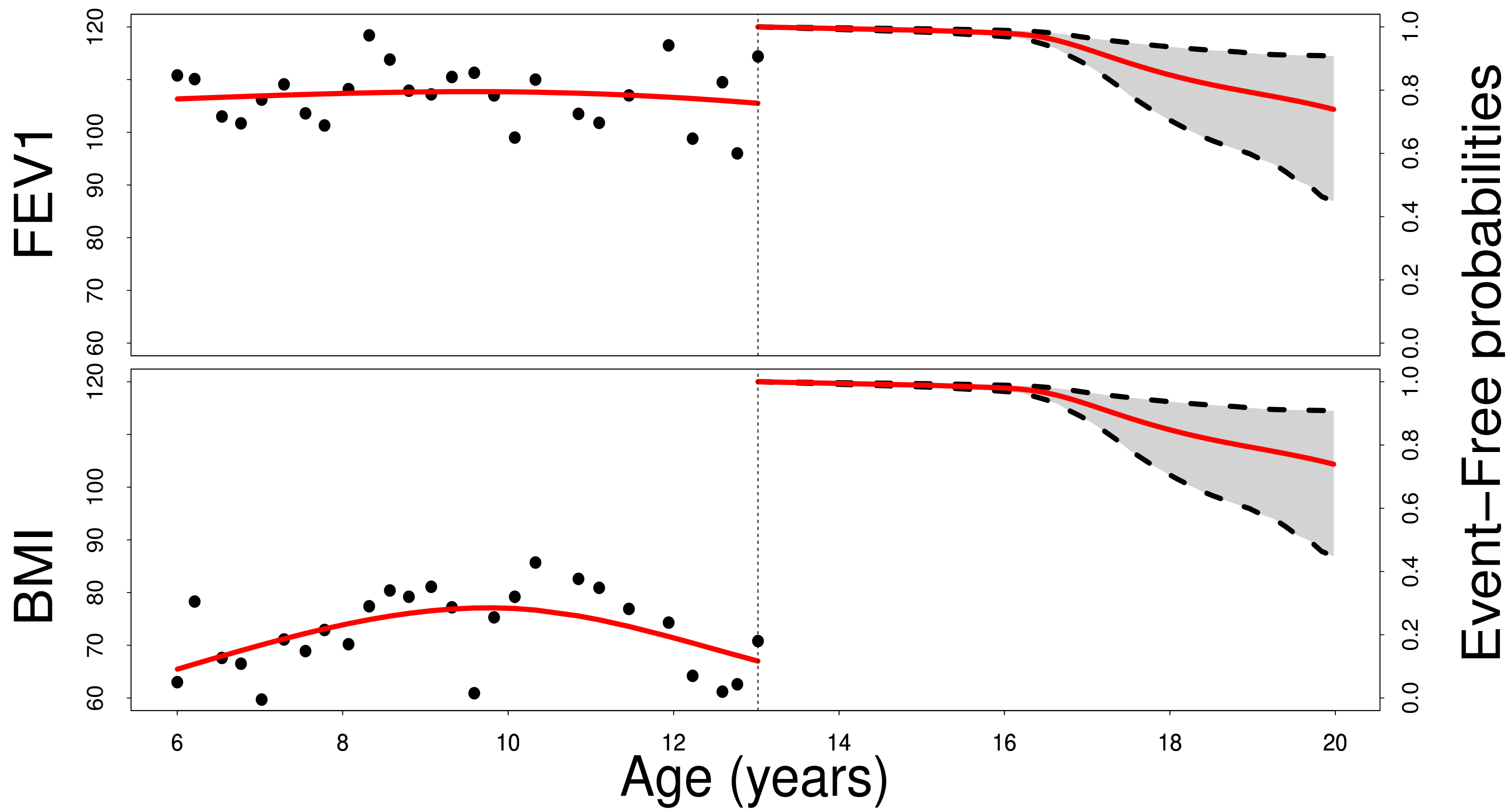
Dynamic Predictions (cont'd)



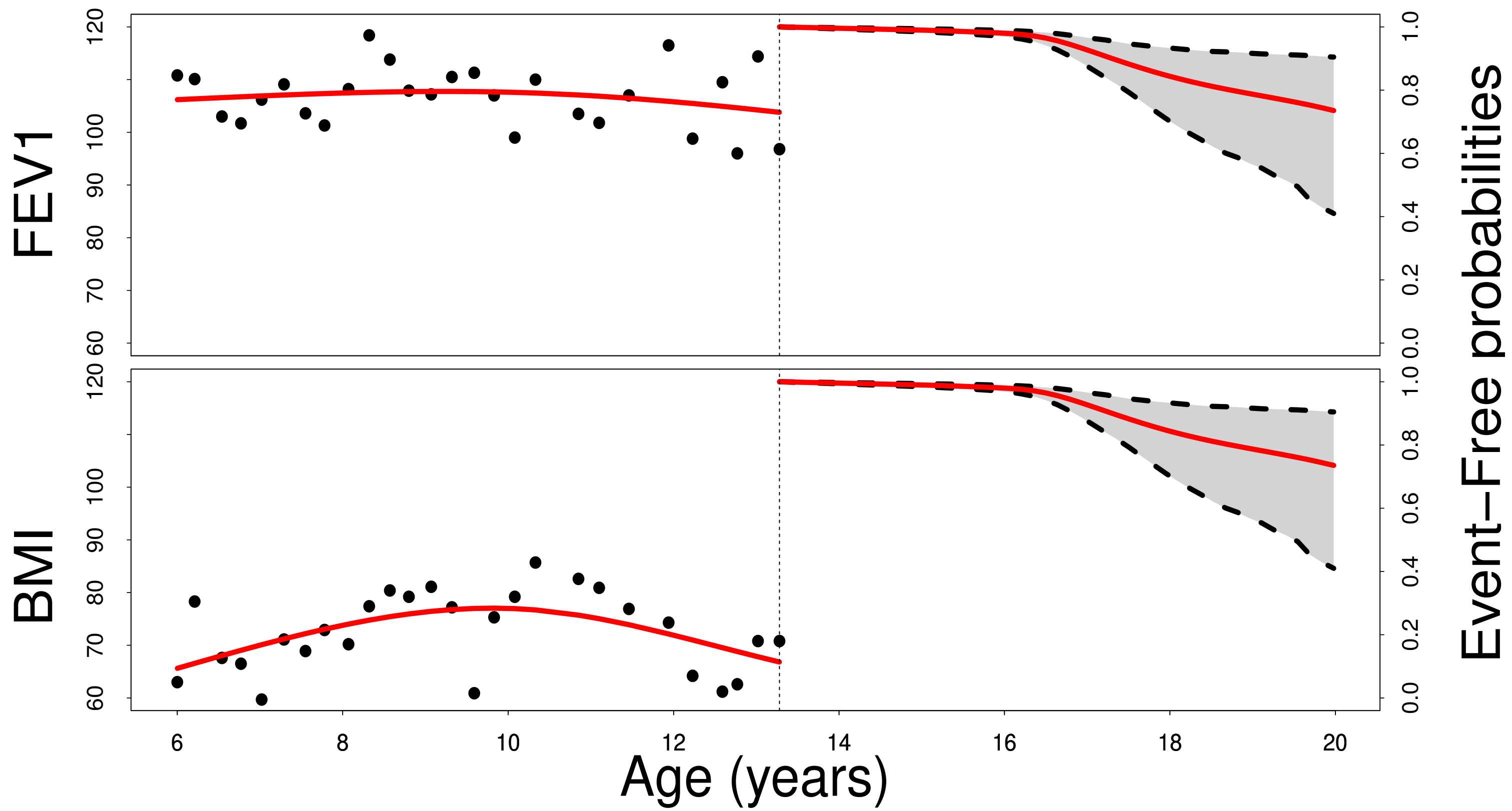
Dynamic Predictions (cont'd)



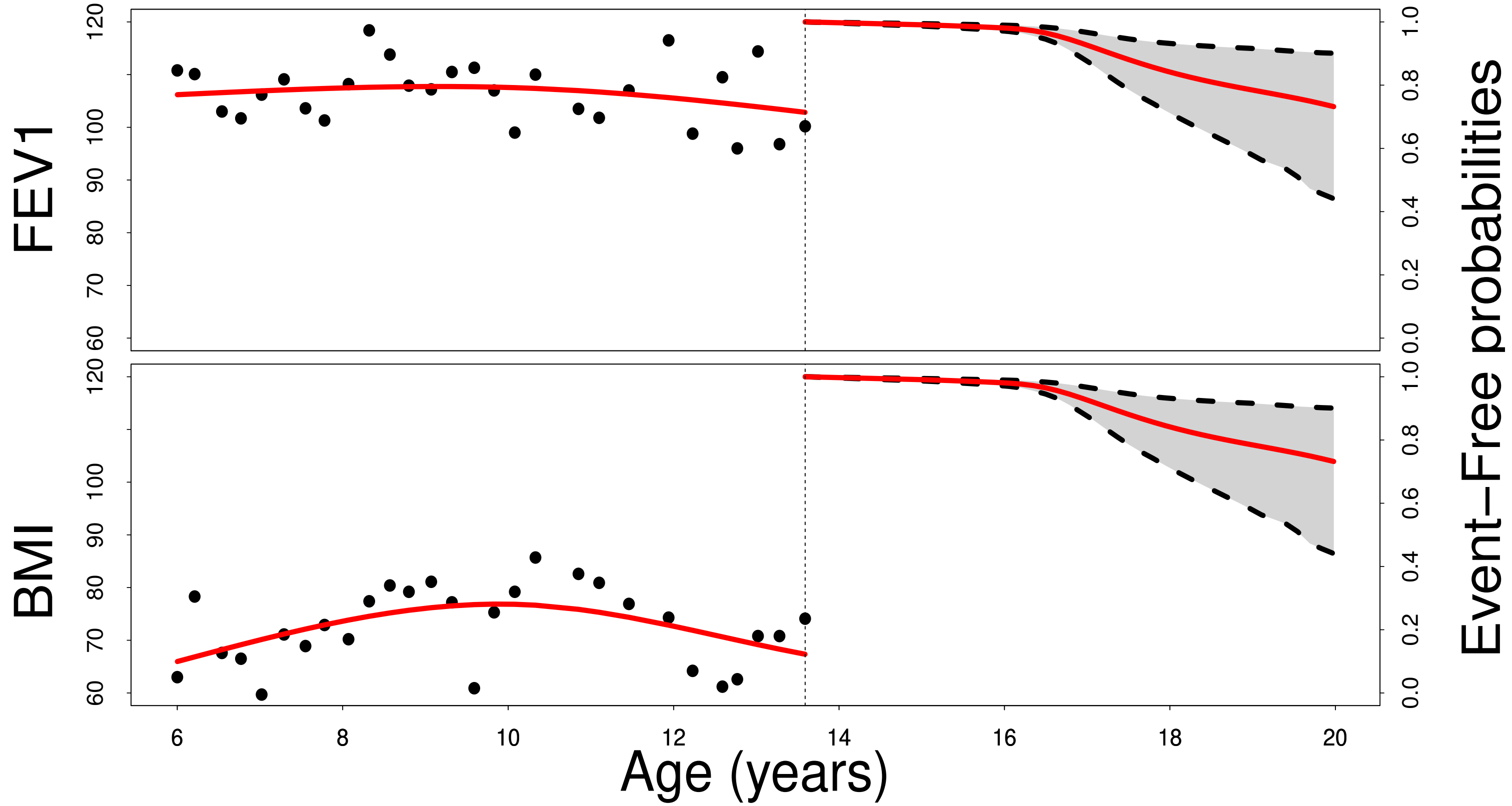
Dynamic Predictions (cont'd)



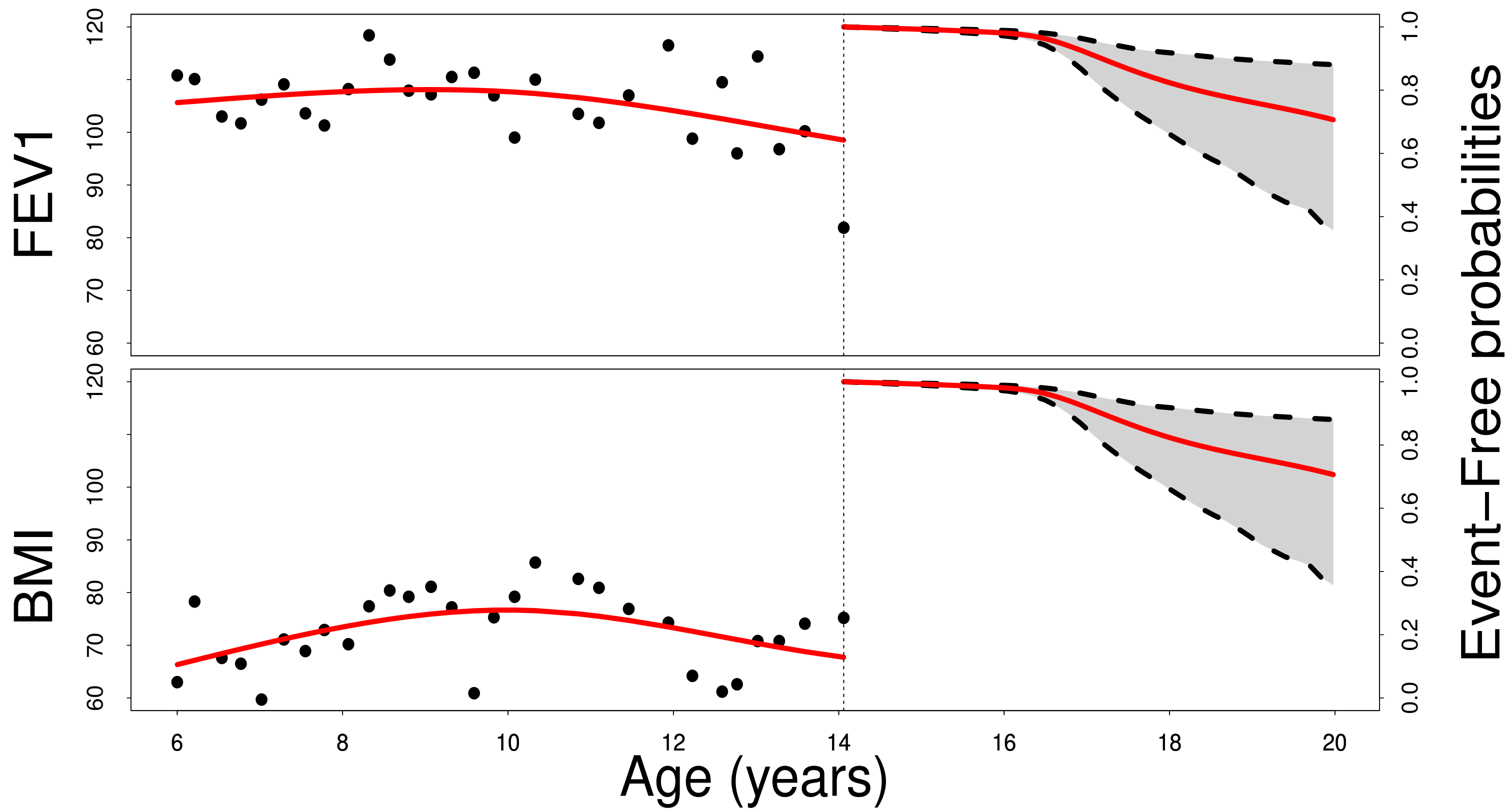
Dynamic Predictions (cont'd)



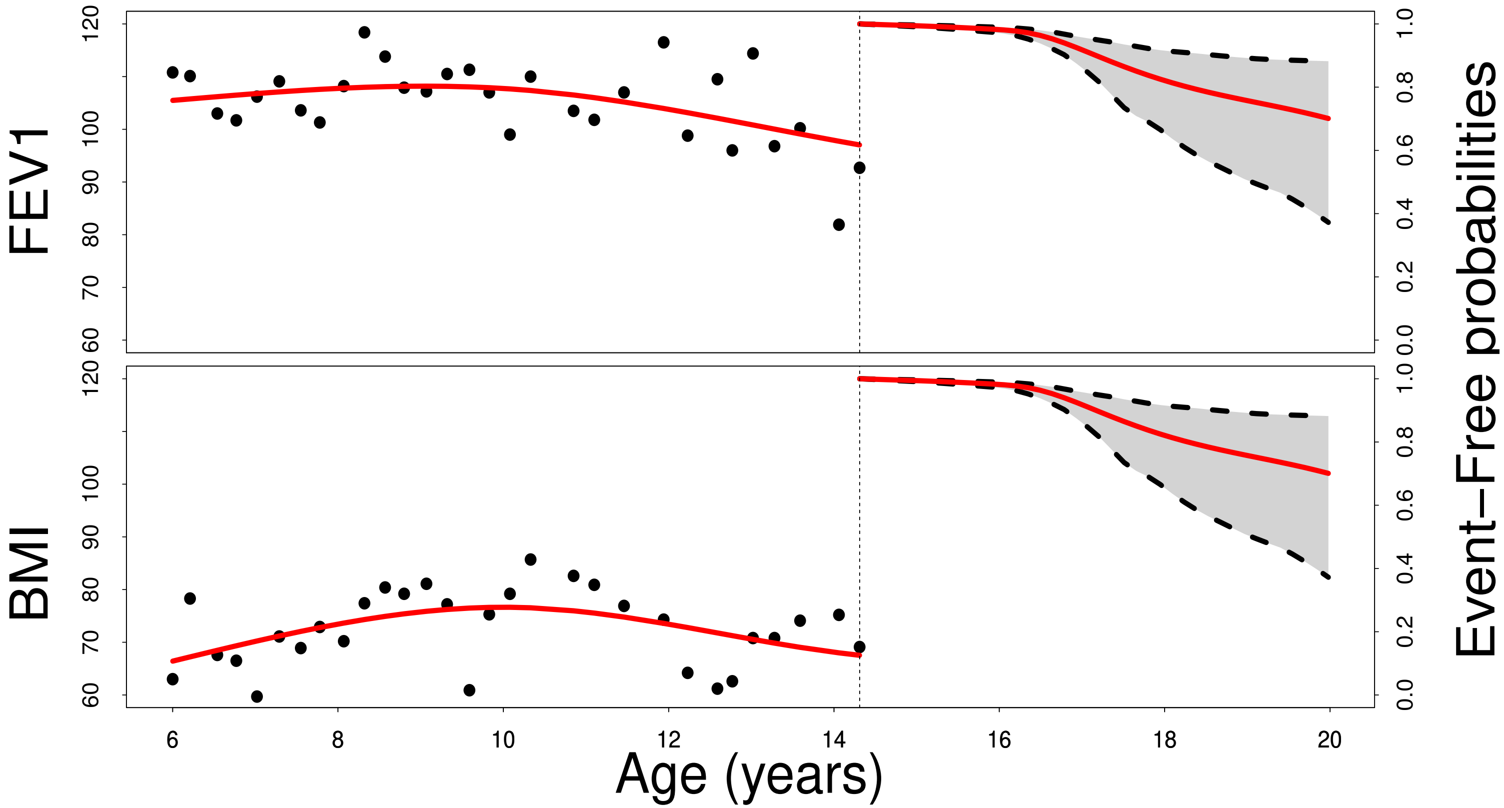
Dynamic Predictions (cont'd)



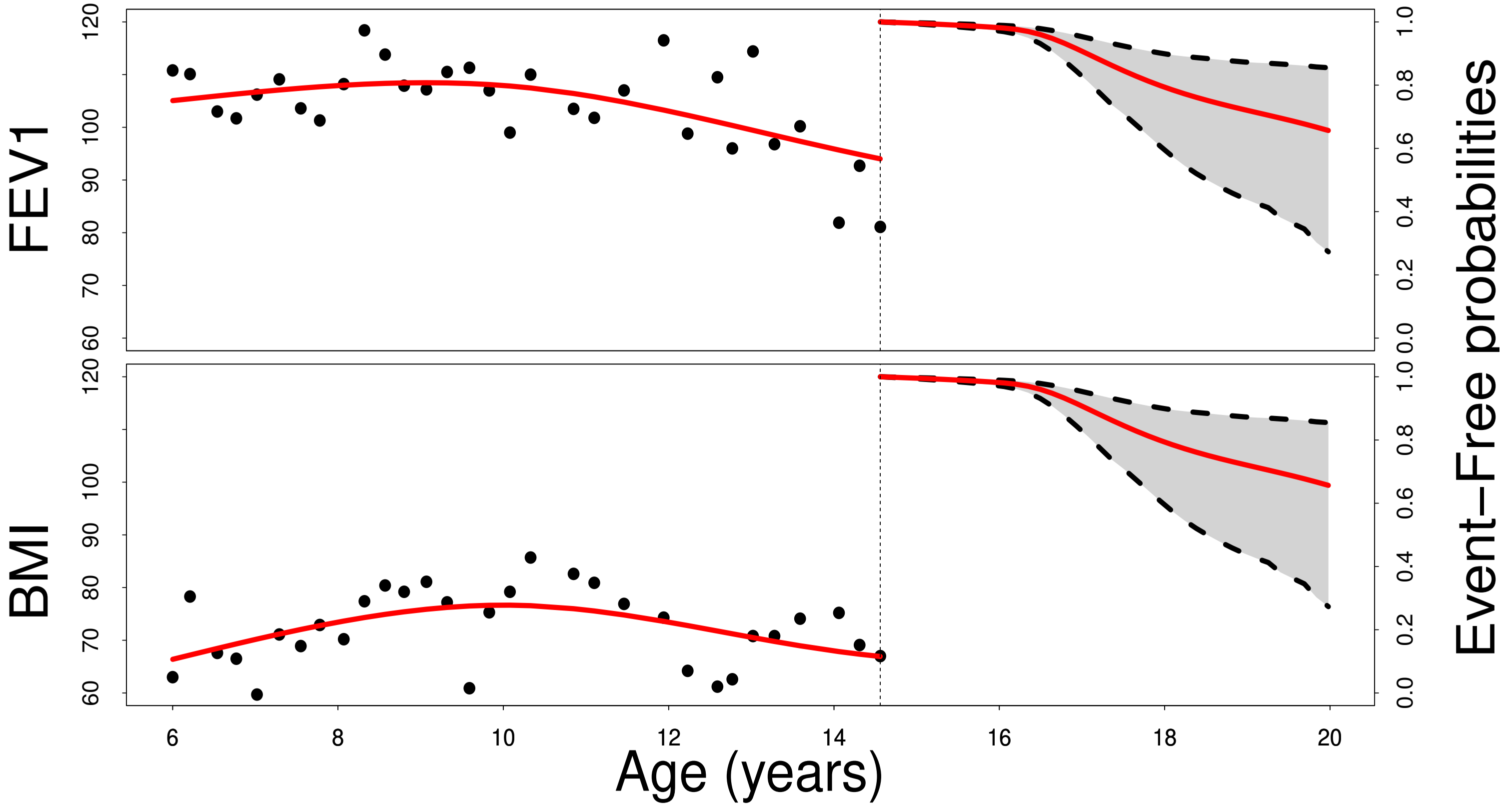
Dynamic Predictions (cont'd)



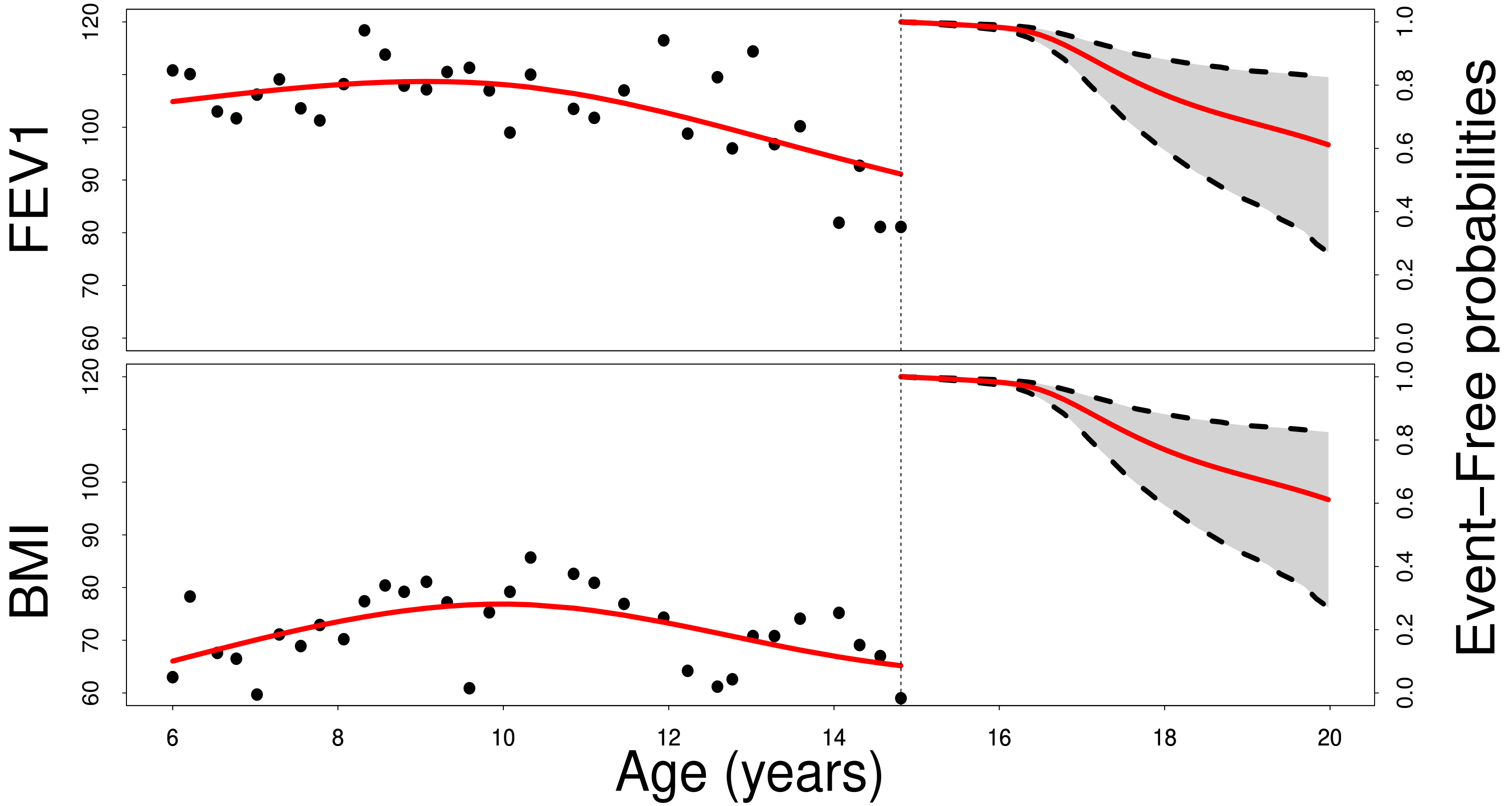
Dynamic Predictions (cont'd)



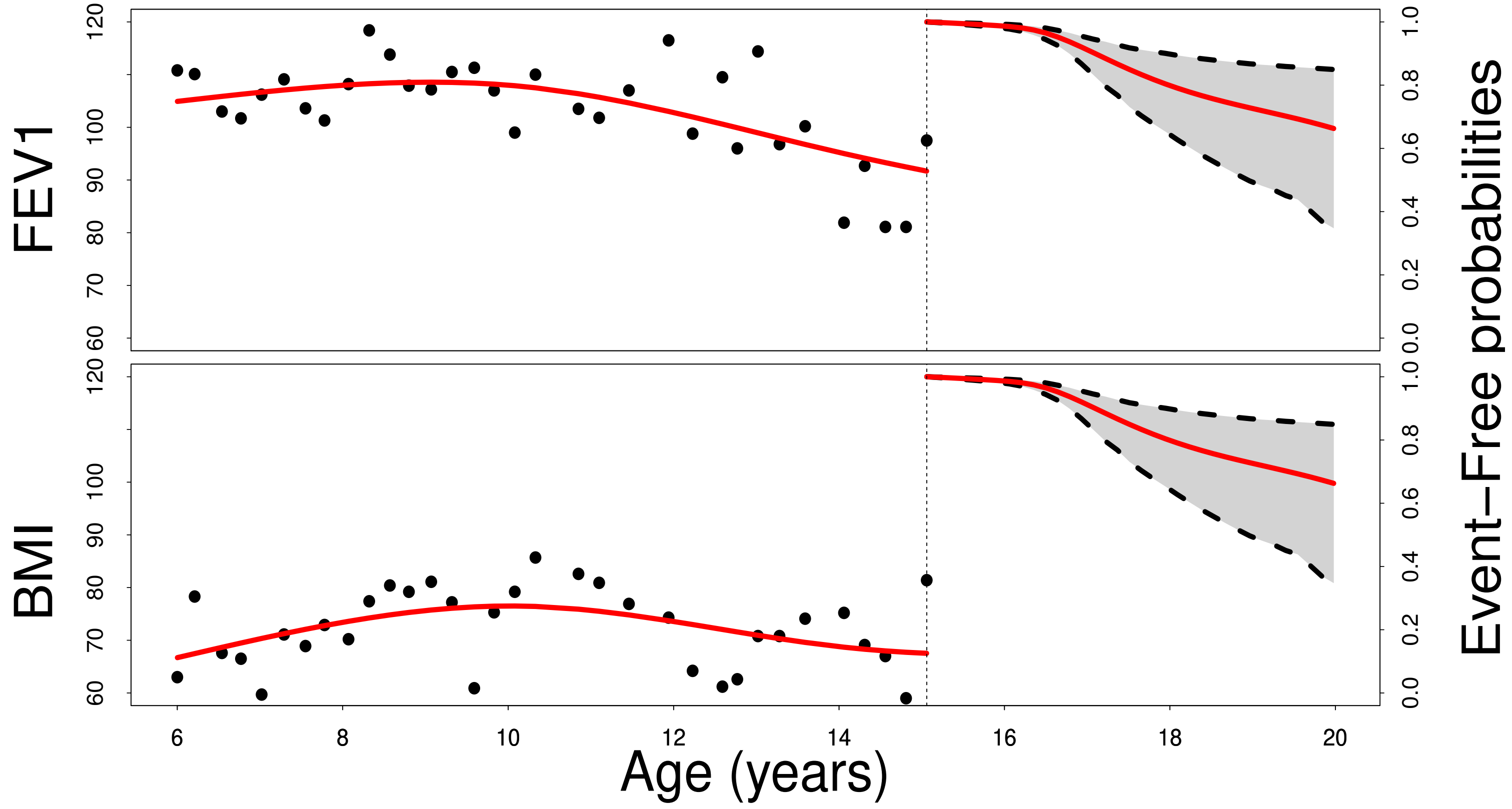
Dynamic Predictions (cont'd)



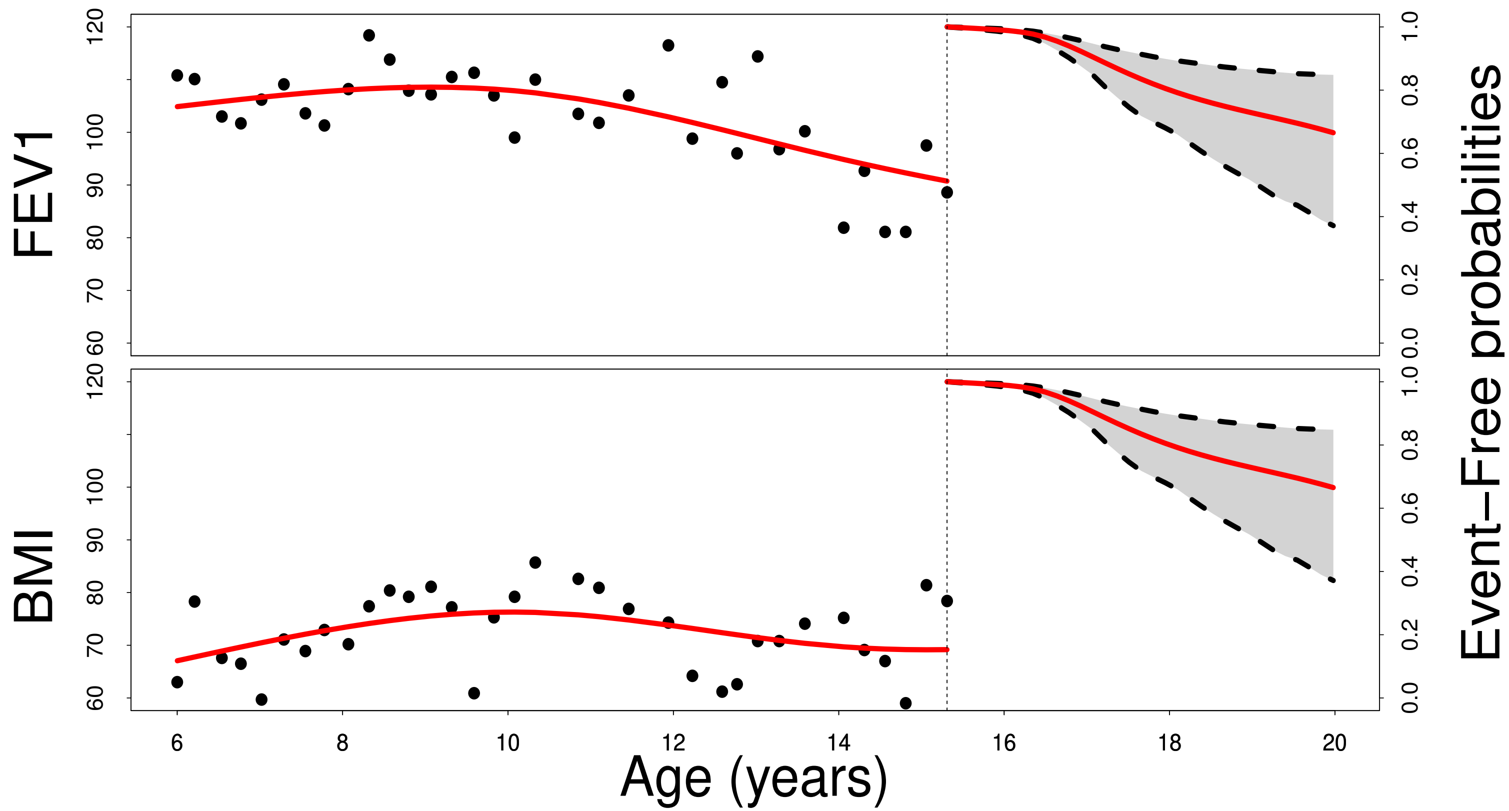
Dynamic Predictions (cont'd)



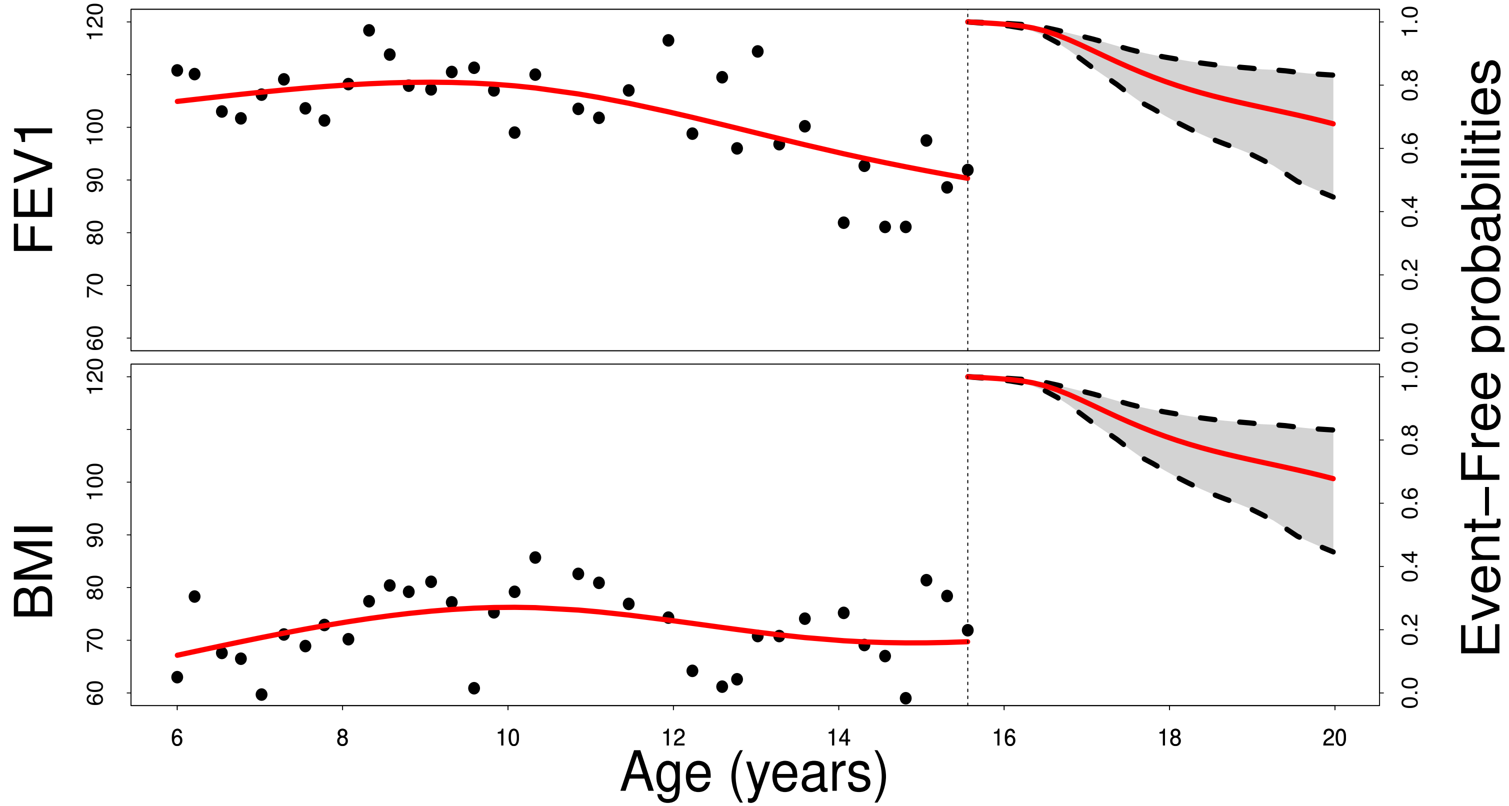
Dynamic Predictions (cont'd)



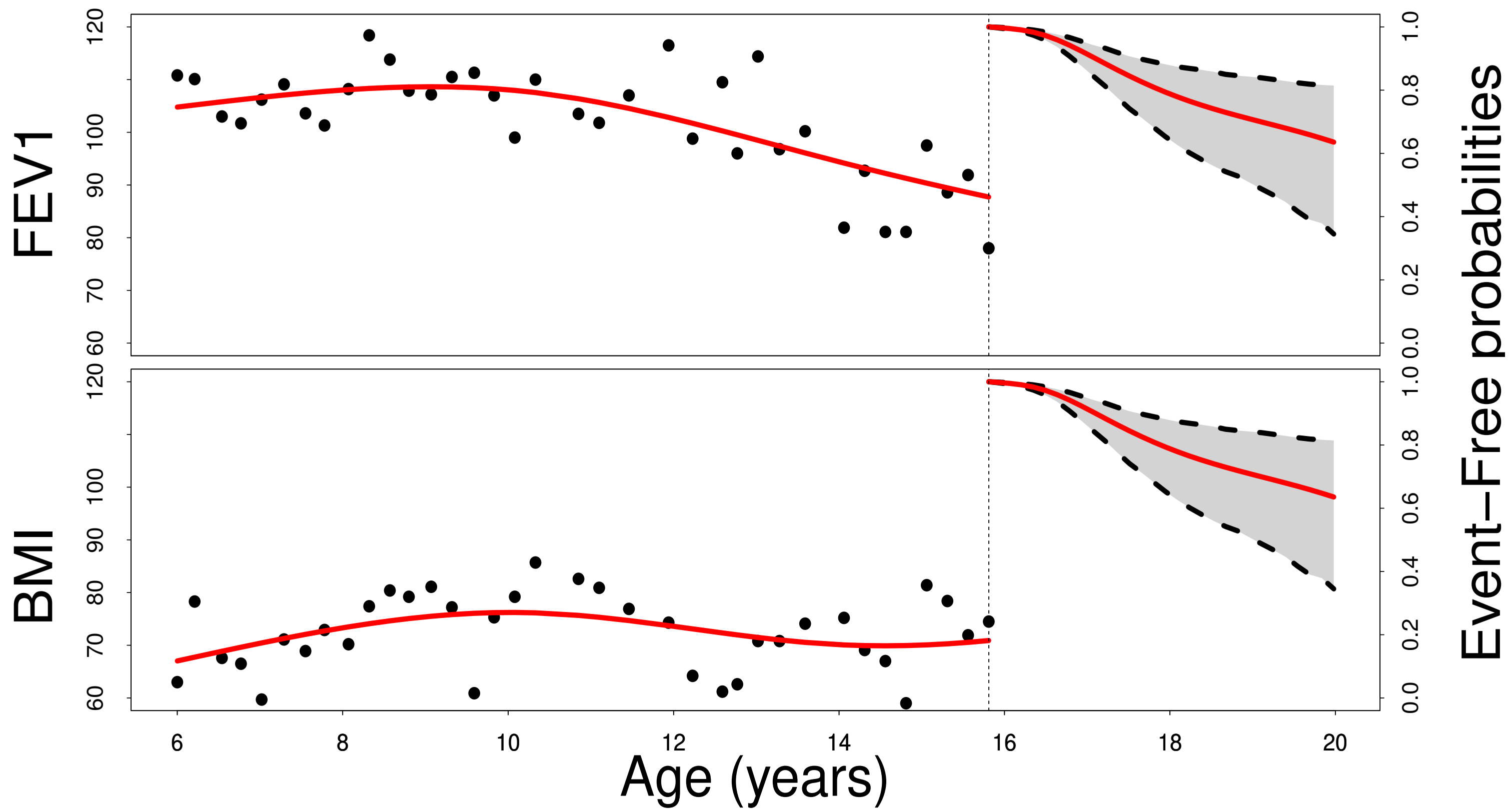
Dynamic Predictions (cont'd)



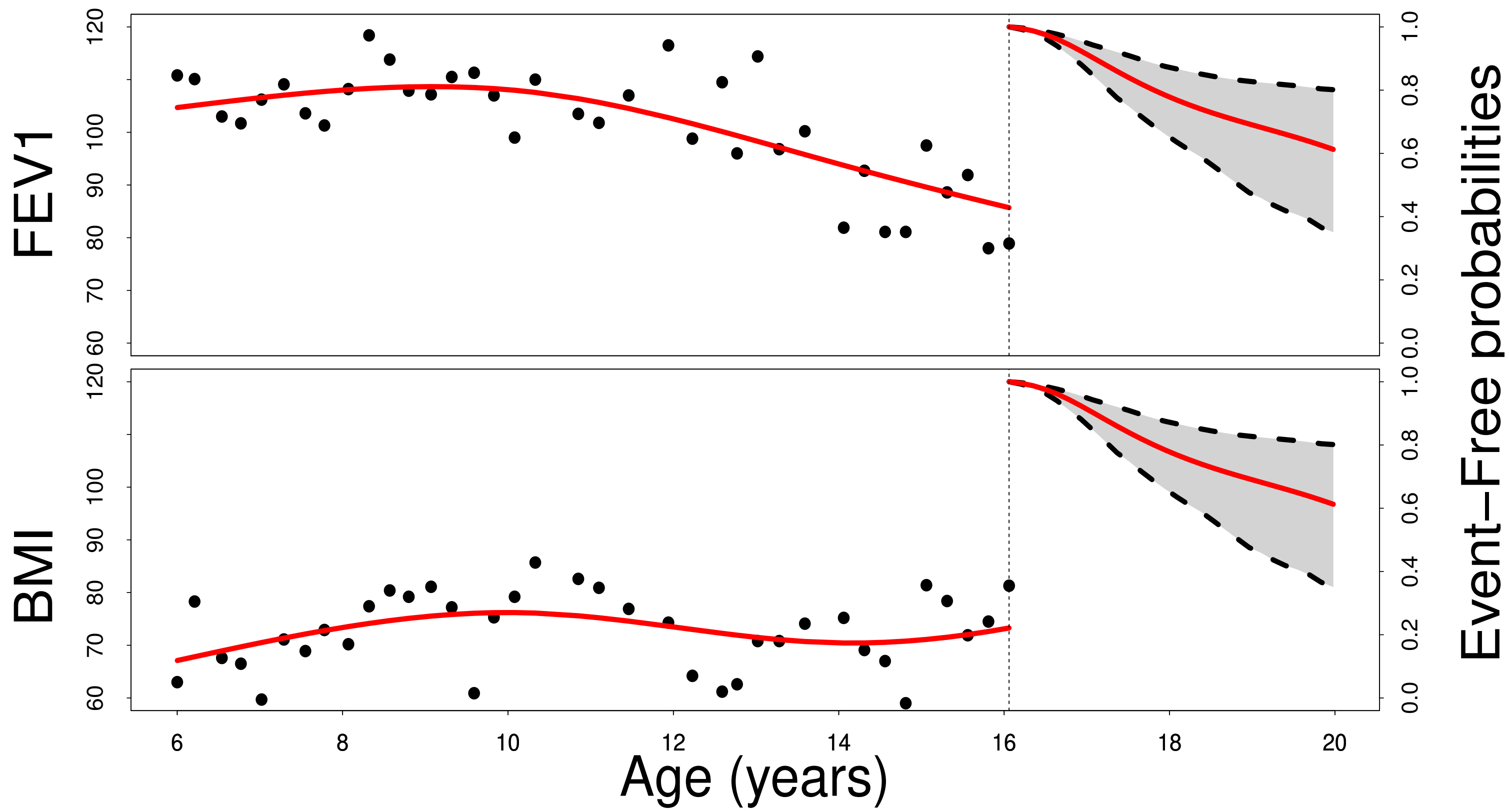
Dynamic Predictions (cont'd)



Dynamic Predictions (cont'd)



Dynamic Predictions (cont'd)



Dynamic Predictions (cont'd)

ANDRINOPOULOU, E. R., RIZOPOULOS, D., TAKKENBERG, J. J. AND LESAFFRE, E. (2015). COMBINED DYNAMIC PREDICTIONS USING JOINT MODELS OF TWO LONGITUDINAL OUTCOMES AND COMPETING RISK DATA. STATISTICAL METHODS IN MEDICAL RESEARCH, 6(4):1787-1801.

ANDRINOPOULOU, E. R., RIZOPOULOS, D., GELEIJNSE, M. L., LESAFFRE, E., BOGERS, A. J. AND TAKKENBERG, J. 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, 15(1), 1.

Discussion

- Joint models - popular framework
- Software
  - ▷ `JM`, `JMbayes`: **R**
  - ▷ `joineR`, `joineRML`: **R**
  - ▷ `stjm`: **Stata**
  - ▷ `JMFit`: **SAS**

Thank you!

Any questions?