Dynamic Predictions in a Joint Model of Two Longitudinal Outcomes and Competing Risk Data. An Application in Heart Valve Data.

Eleni-Rosalina Andrinopoulou
Department of Biostatistics and CardioThoracic surgery, Erasmus Medical Center

e.andrinopoulou@erasmusmc.nl

BAYES

May 21-23th, 2013

Heart Valve Data

- 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
  - ▷ Two longitudinal responses: Aortic Gradient (continuous) and Aortic Regurgitation (ordinal)
  - ▷ Time-to-event response: time to death/reoperation (competing risks setting)

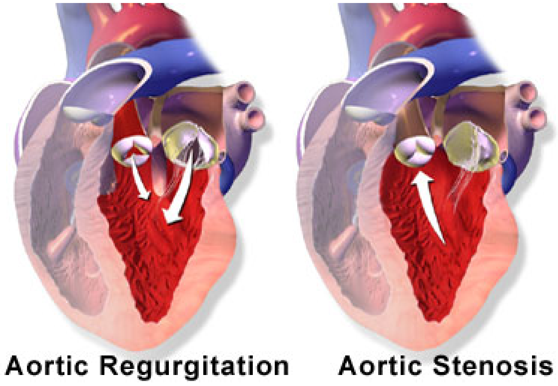
Heart Valve Data - (cont'd)

- Aortic gradient and aortic regurgitation are measurements of the valve function

  $\Downarrow$

  expecting biologically interrelationship.

- Both events (reoperation and death) are highly associated with the disease condition of the patient and correlated to each other.



Statistical Models - Heart Valve data

Joint model to the Heart Valve data:

- Linear mixed-effects model for the longitudinal continuous outcome.
- Mixed-effects continuation ratio model for the longitudinal ordinal outcome.
- Cause-specific hazard model for the competing risk failure time data.

Statistical Models - Heart Valve data (cont'd)

Let Y_{1i} and Y_{2i} represent the repeated measurements of AG and AR for the i-th patient, $i = 1, \ldots, n$.

- Linear mixed-effects model:

$$Y_{1i} = X_{1i}\beta_1 + Z_{1i}b_{1i} + \epsilon_i, \qquad \epsilon_i \sim (0, V)$$

  - ▷ $X_{1i}\beta_1$ denotes the fixed part
  - ▷ $Z_{1i}b_{1i}$ denotes the random part

Statistical Models - Heart Valve data (cont'd)

- Mixed-effects continuation ratio model:

$$\pi_j = \mathsf{P}(Y_{2i} = j \mid Y_{2i} \leq j) = \frac{\exp\left(\theta_j + X_{2i}\beta_2 + Z_{2i}b_{2i}\right)}{1 + \exp\left(\theta_j + X_{2i}\beta_2 + Z_{2i}b_{2i}\right)},$$

where j represents the category of the ordinal response

- ▷ $X_{2i}\beta_2$ denotes the fixed part
- ▷ $Z_{2i}b_{2i}$ denotes the random part

Statistical Models - Heart Valve data (cont'd)

- To account for the correlation between the two longitudinal outcomes we assume multivariate random effects

$$\begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, D = \begin{pmatrix} \Sigma_{b1} & \Sigma_{b12} \\ \Sigma_{b12} & \Sigma_{b2} \end{pmatrix}\right)$$

Statistical Models - Heart Valve data (cont'd)

Let T_i denote the observed failure time for the i-th patient and $\delta_i = 0, 1, 2$ the event indicator.

- Proportional hazard model:

$$\begin{cases} h_i^{d}(t) = h_0^{d}(t)\exp\{w_i^\top \gamma_d + (\tilde{\beta}_1 + b_{1i})^\top \alpha_{1d} + (\tilde{\beta}_2 + b_{2i})^\top \alpha_{2d}\}, \\ \\ h_i^{r}(t) = h_0^{r}(t)\exp\{w_i^\top \gamma_r + (\tilde{\beta}_1 + b_{1i})^\top \alpha_{1r} + (\tilde{\beta}_2 + b_{2i})^\top \alpha_{2r}\}, \end{cases}$$

where $\alpha_{1d}^\top$, $\alpha_{2d}^\top$, $\alpha_{1r}^\top$,$\alpha_{2r}^\top$ are the coefficient that link the longitudinal and survival part.

Specifically, $\alpha_1^\top$ are the coefficients of the continuous longitudinal outcome and $\alpha_2^\top$ are the coefficients of the ordinal longitudinal outcomes.

Statistical Models - Heart Valve data (cont'd)

- **Assumption**: full conditional independence, given the random effects

$$\mathsf{P}(T_i, \delta_i, Y_{1i}, Y_{2i}^* \mid b_{1i}, b_{2i}; \theta) = \mathsf{P}(T_i, \delta_i \mid b_{1i}, b_{2i}; \theta)\mathsf{P}(Y_{1i} \mid b_{1i}, b_{2i}; \theta)\mathsf{P}(Y_{2i}^* \mid b_{1i}, b_{2i}; \theta)$$

$$\mathsf{P}(Y_{1i} \mid b_{1i}; \theta) = \prod_j \mathsf{P}\{Y_{1i}(t_{ij}) \mid b_{1i}; \theta\}$$

$$\mathsf{P}(Y_{2i}^* \mid b_{1i}; \theta) = \prod_j \mathsf{P}\{Y_{2i}^*(t_{ij}) \mid b_{2i}; \theta\}$$

The random effects b_{1i} and b_{2i} underlie both the longitudinal and survival processes.

Statistical Models - Heart Valve data (cont'd)

Models for the Heart Valve data

- continuous longitudinal submodel
  - ▷ Fixed part: intercept, B-splines for time, age and gender
  - ▷ Random part: B-splines for time
- ordinal longitudinal submodel
  - ▷ Fixed part: intercept, time, age and gender
  - ▷ Random part: random intercept
- survival submodel
  - ▷ age
  - ▷ piecewise constant baseline hazard

Bayesian estimation - Heart Valve data

- Bayesian formulation for the proposed joint model

- Posterior inferences using a Markov chain Monte Carlo (MCMC) algorithm

  ▷ The posterior distribution is

$$\mathsf{P}(\theta \mid Y_{1i}, Y_{2i}^*, T_i, \delta_i) \propto \mathsf{P}(T_i, \delta_i \mid b_{1i}, b_{2i}, \theta_s)\mathsf{P}(Y_{1i} \mid b_{1i}, \theta_{Y_1})\mathsf{P}(Y_{2i}^* \mid b_{2i}, \theta_{Y_2^*}) \mathsf{P}(b_{1i} \mid \theta_{Y_1})\mathsf{P}(b_{2i} \mid \theta_{Y_2^*})\mathsf{P}(\theta_{Y_1})\mathsf{P}(\theta_{Y_2^*})\mathsf{P}(\theta_s)$$

Bayesian estimation - Heart Valve data (cont'd)

- The likelihood for the survival part is

$$
\begin{aligned}
&P(T_i, \delta_i \mid b_{1i}, b_{2i}; \theta_s) \\
&\; = \; [h_{0k}(T_i) \exp\{w_i^\top \gamma_k + (\tilde{\beta}_1 + b_{1i})^\top \alpha_{1k} + (\tilde{\beta}_2 + b_{2i})^\top \alpha_{2k}\}]^{I(\delta_i = k)} \\
&\; \times \; \exp\Big\{ - \sum_{q=1}^{m} [h_{0d}(T_i) \exp\{w_i^\top \gamma_d + (\tilde{\beta}_1 + b_{1i})^\top \alpha_{1d} + (\tilde{\beta}_2 + b_{2i})^\top \alpha_{2d}\} \\
&\; + \; h_{0r}(T_i) \exp\{w_i^\top \gamma_r + (\tilde{\beta}_1 + b_{1i})^\top \alpha_{1r} + (\tilde{\beta}_2 + b_{2i})^\top \alpha_{2r}\}] T_{iq} \Big\}
\end{aligned}
$$

Bayesian estimation - Heart Valve data (cont'd)

- priors
  - ▷ β_1, β_2, b_1, b_2, α_1, α_2 ⇒ normal prior distributions
  - ▷ D^{-1} ⇒ wishart prior distribution
  - ▷ τ, $h_0^d(t)$, $h_0^r(t)$ ⇒ gamma prior distributions

- 110000 iterations with 20000 burn in and 20 thinning

Dynamic predictions - Heart Valve data

Increased interest towards dynamic predictions

Focus on the prediction of

- cumulative incidence function
- longitudinal outcomes

Useful from the clinical point of view

Dynamic predictions - Heart Valve data (cont'd)

- New patient l that has provided us

$$\tilde{Y}_{1l}(t) = \{Y_{1l}(s), 0 \leq s < t\} \quad \text{and} \quad \tilde{Y}_{2l}(t) = \{Y_{2l}(s), 0 \leq s < t\}$$

- The interest lies on

$$\pi_{lk}(u, t; \theta) = \mathsf{P}(T^*_{lk} < u \mid \cup_{k=1}^{K} T^*_{lk} > t, \tilde{Y}_{1l}(t), \tilde{Y}_{2l}(t), D_n; \theta),$$

where $u > t$ and D_n denotes the sample on which the joint model was fitted

Dynamic predictions - Heart Valve data (cont'd)

- Conditional probabilities

$$\pi_{lk}(u, t; \theta) =$$

$$\int P(T^*_{lk} < u \mid \cup_{k=1}^{K} T^*_{lk} > t, \tilde{Y}_{1l}(t), \tilde{Y}_{2l}(t); \theta) \times p(b_l \mid \cup_{k=1}^{K} T^*_{lk} > t, \tilde{Y}_{1l}(t), \tilde{Y}_{2l}(t); \theta) db_l$$

Dynamic predictions - Heart Valve data (cont'd)

- Conditional Independence Assumption

$$\pi_{lk}(u, t; \theta) =$$

$$\int P(T^*_{lk} < u \mid \cup_{k=1}^K T^*_{lk} > t; \theta) \times p(b_l \mid \cup_{k=1}^K T^*_{lk} > t, \tilde{Y}_{1l}(t), \tilde{Y}_{2l}(t); \theta) db_l$$

Dynamic predictions - Heart Valve data (cont'd)

- Conditional Independence Assumption

$$\pi_{lk}(u, t; \theta) =$$

$$\int \frac{CIF(u, t, b_l; \theta)}{S(t, b_l; \theta)} \times p(b_l \mid \cup_{k=1}^{K} T_{lk}^* > t, \tilde{Y}_{1l}(t), \tilde{Y}_{2l}(t); \theta) db_l$$

Dynamic predictions - Heart Valve data (cont'd)

- Conditional Independence Assumption

$$\pi_{lk}(u, t; \theta) =$$

$$\int \frac{CIF(u, t, b_l; \theta)}{S(t, b_l; \theta)} \times p(b_l \mid \cup_{k=1}^{K} T_{lk}^* > t, \tilde{Y}_{1l}(t), \tilde{Y}_{2l}(t); \theta) db_l$$

- Estimation of these probabilities is based on a Monte Carlo scheme

$$\int \pi_{lk}(u, t; \theta) p(\theta \mid D_n) d\theta$$

Dynamic predictions - Heart Valve data (cont'd)

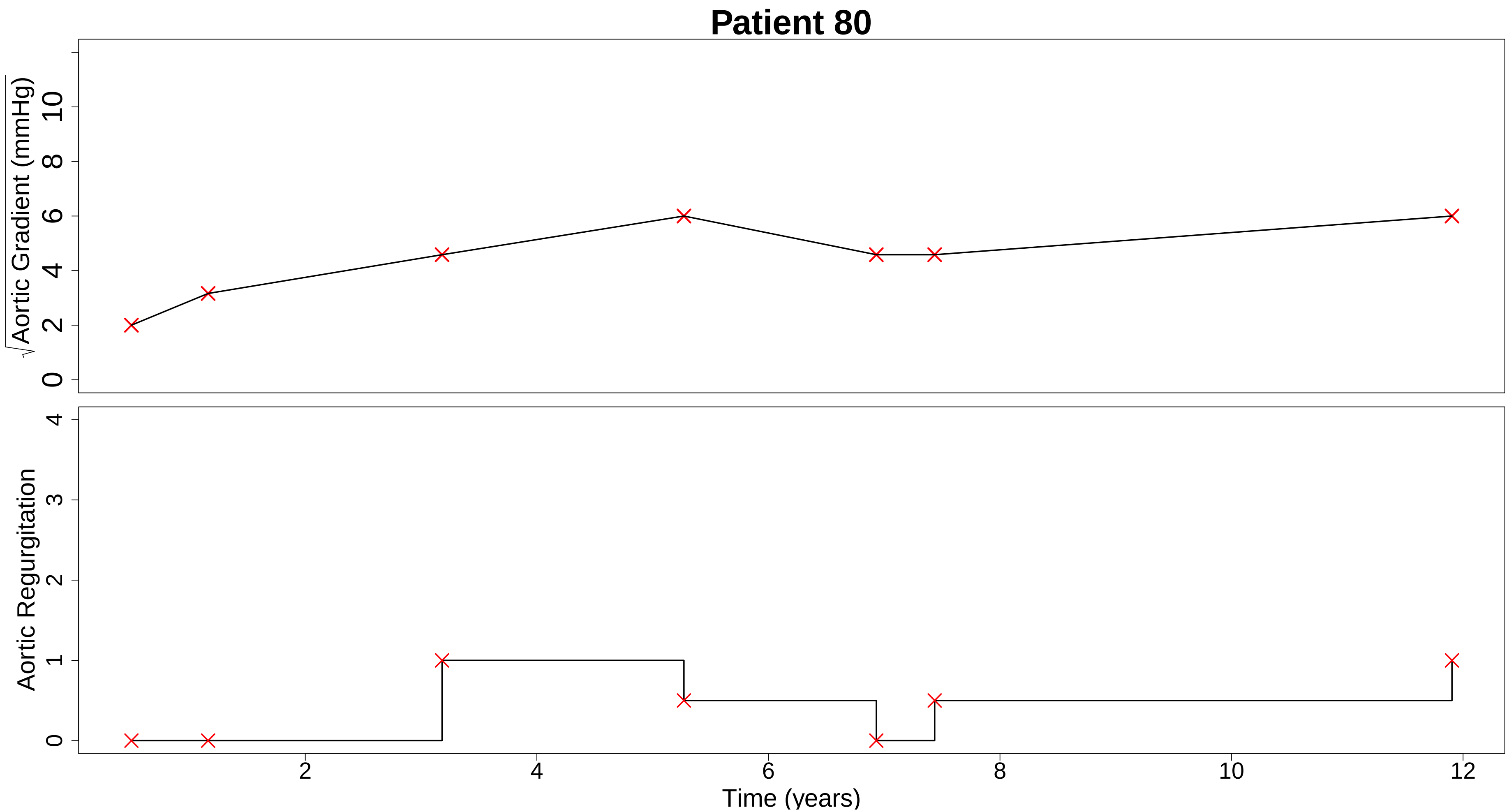
- We estimate the conditional probabilities using a MCMC algorithm. Specifically,

  - ▷ Randomly select θ^* from the MCMC of the joint model
  - ▷ Draw $b_l^* \sim \{b_l | \cup_{k=1}^{K} T_l^* > t, \tilde{Y}_{1l}(t), \tilde{Y}_{2l}(t), \theta^*\}$
  - ▷ Compute $\pi_{lk}(u, t, b_l^*; \theta^*) = CIF(u, t, b_l^*; \theta^*) / S(t, b_l^*; \theta^*)$

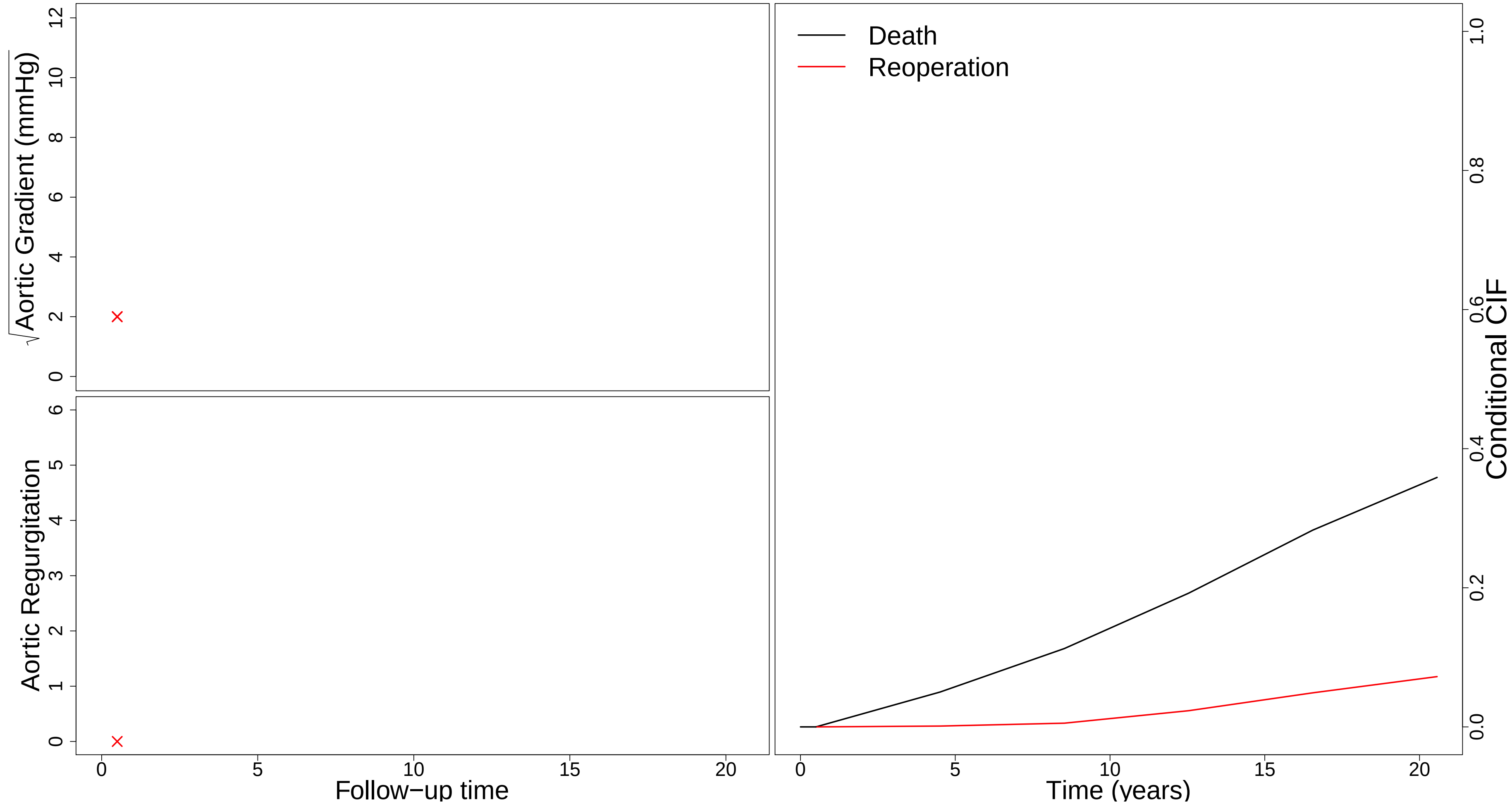
- Repeat for a patient N times

- Then, estimation of the conditional probabilities: mean/median$\{\pi_{lk}(u, t, b_l^*; \theta^*)\}$

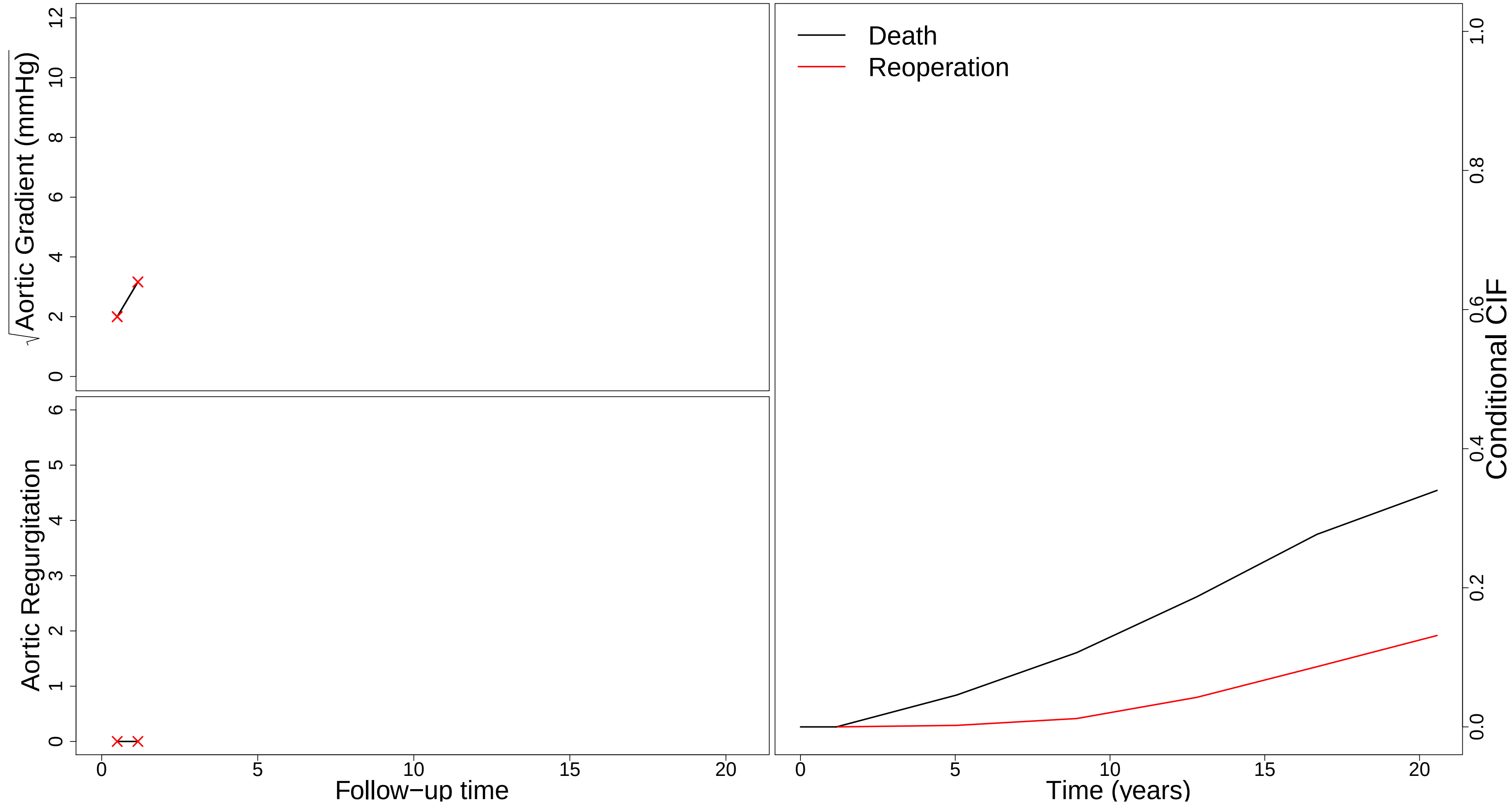
Dynamic predictions - Heart Valve data (cont'd)



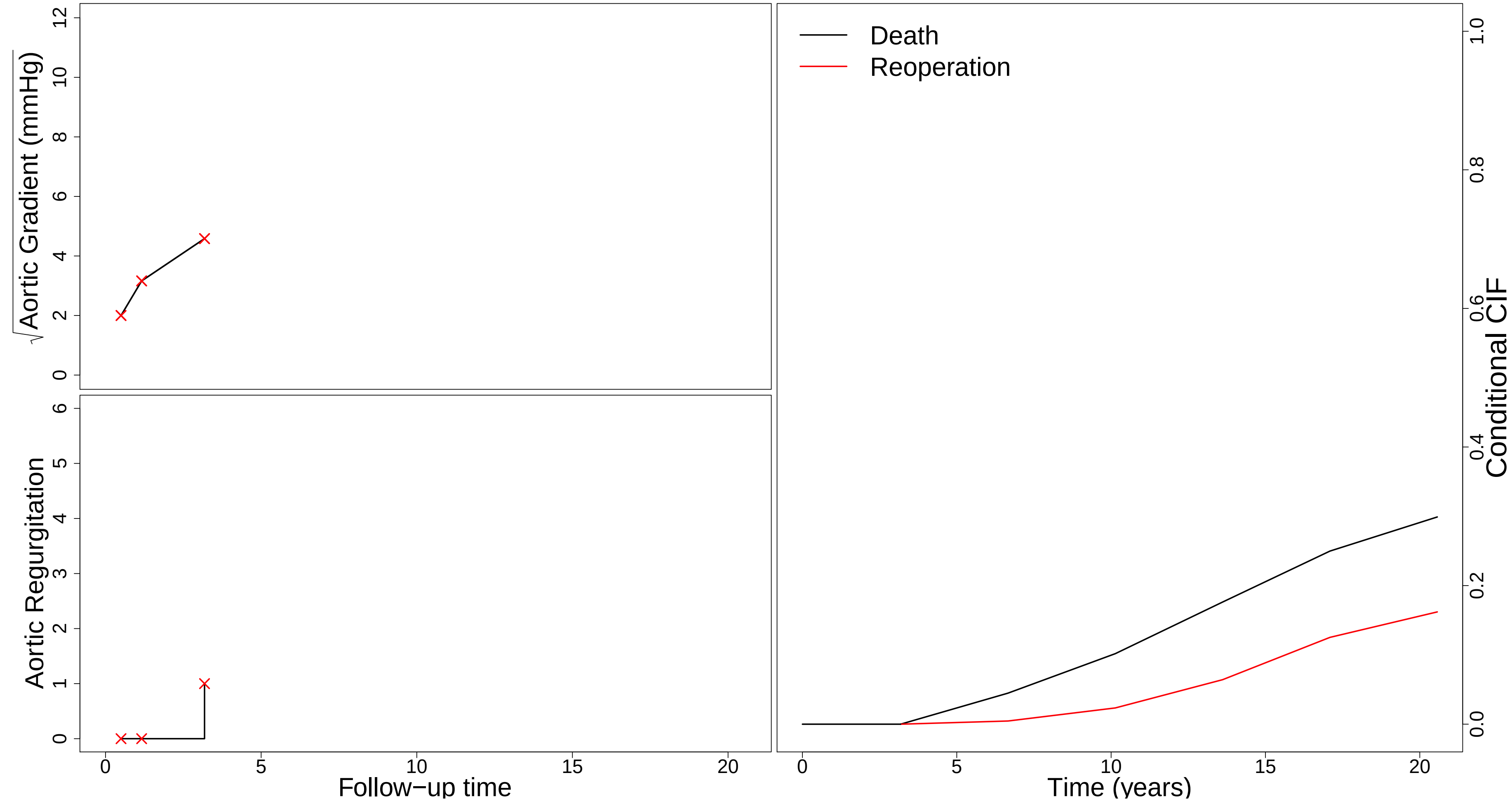
Dynamic predictions - Heart Valve data (cont'd)



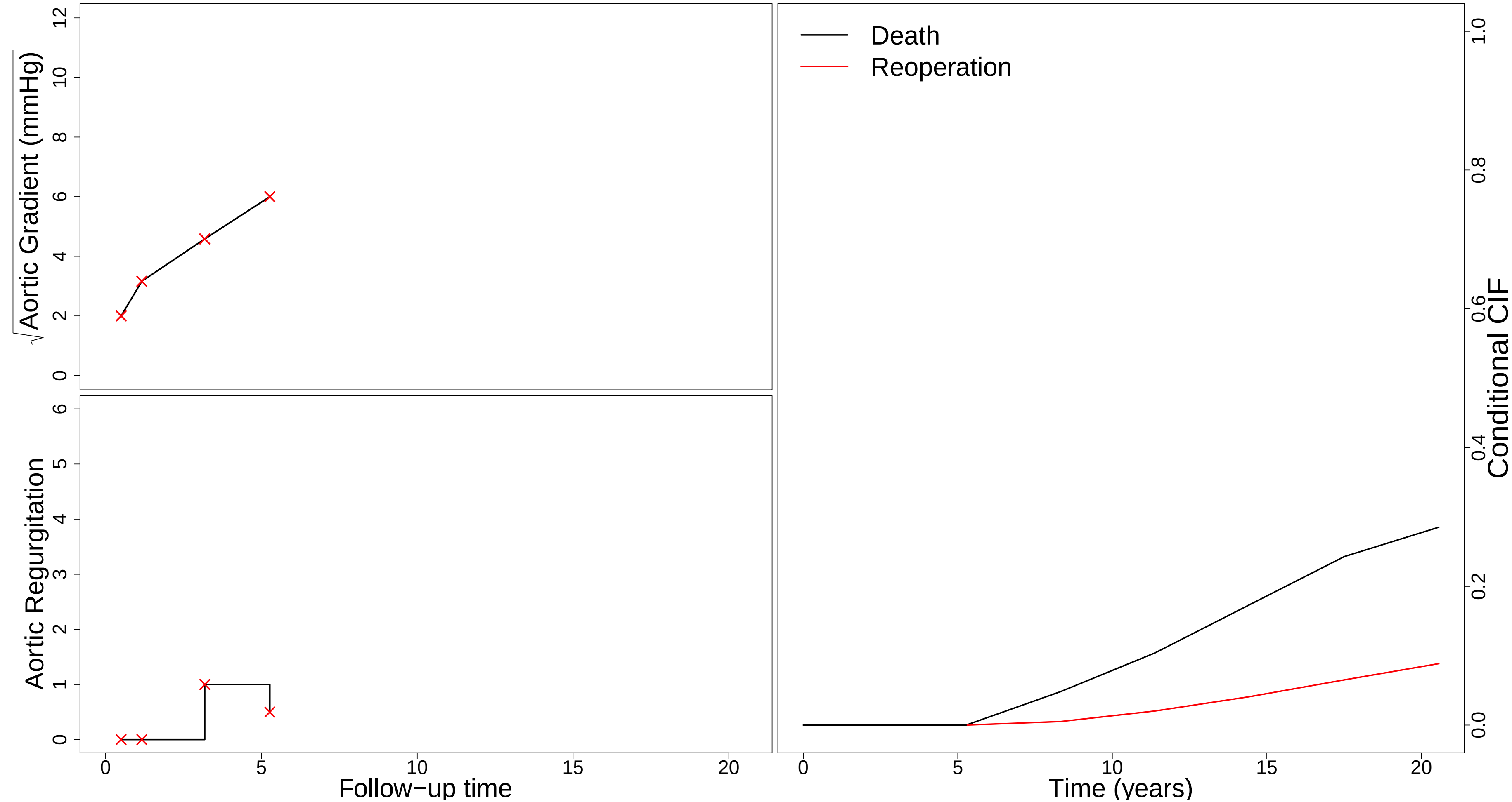
Dynamic predictions - Heart Valve data (cont'd)



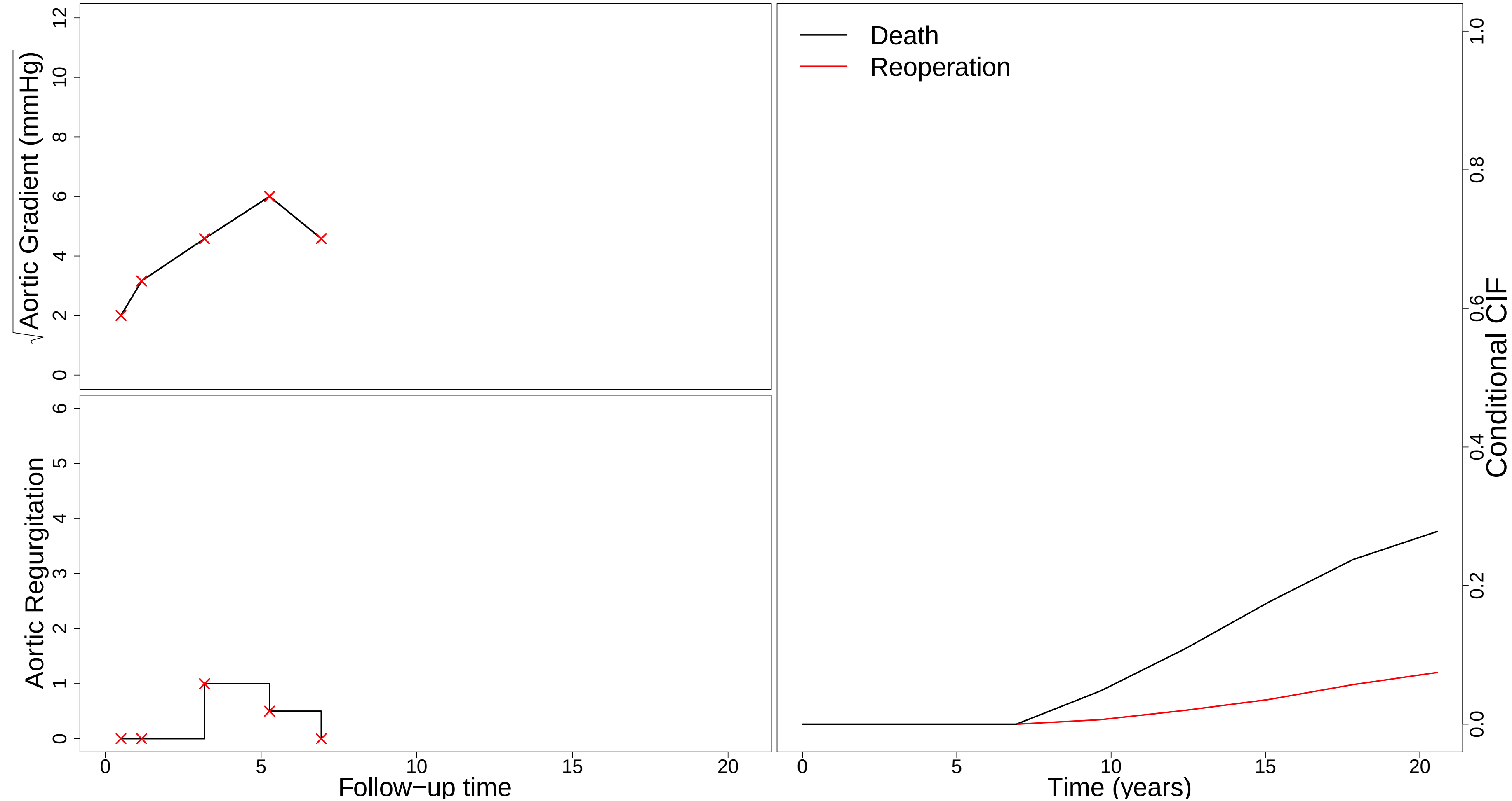
Dynamic predictions - Heart Valve data (cont'd)



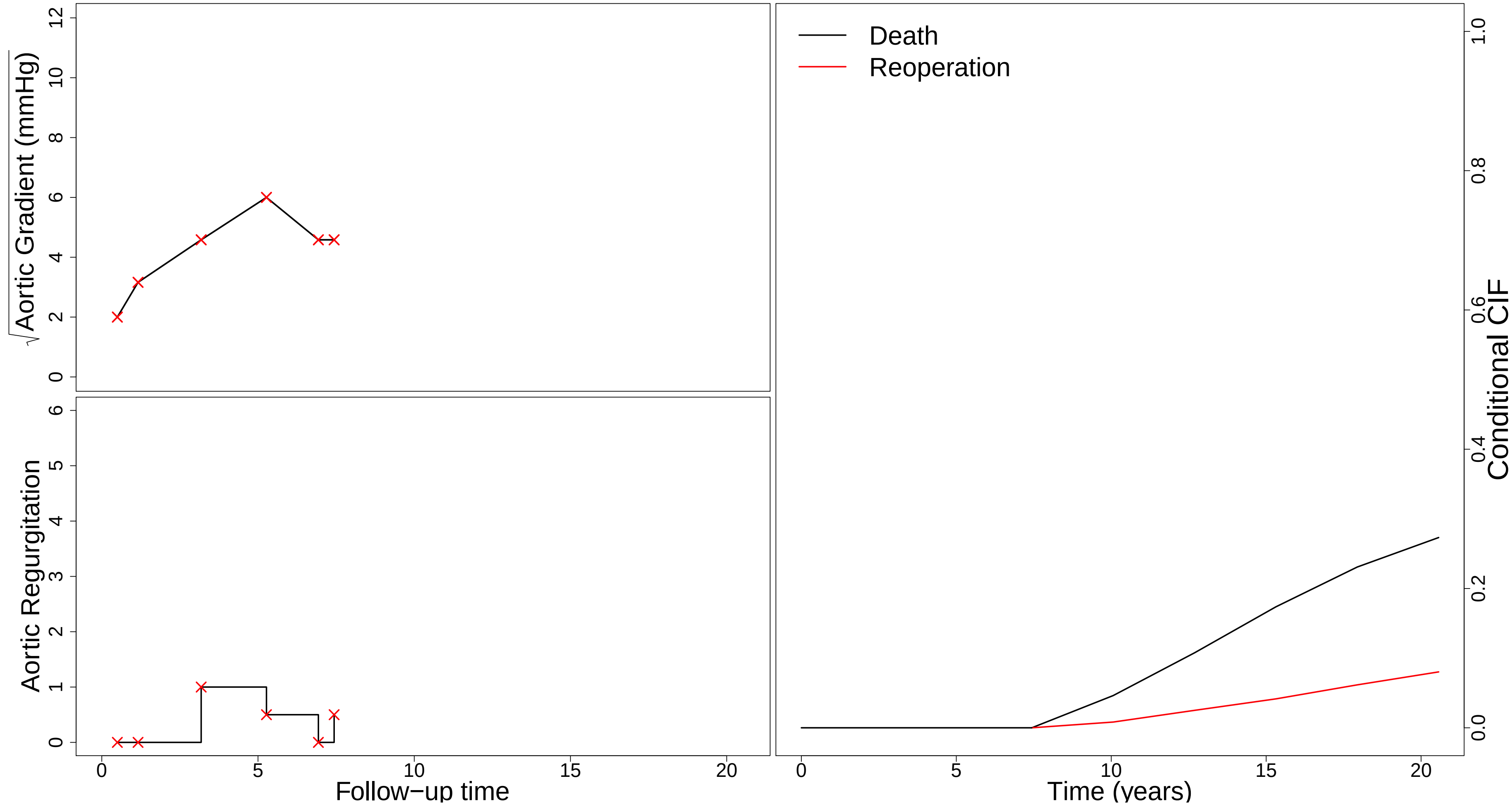
Dynamic predictions - Heart Valve data (cont'd)



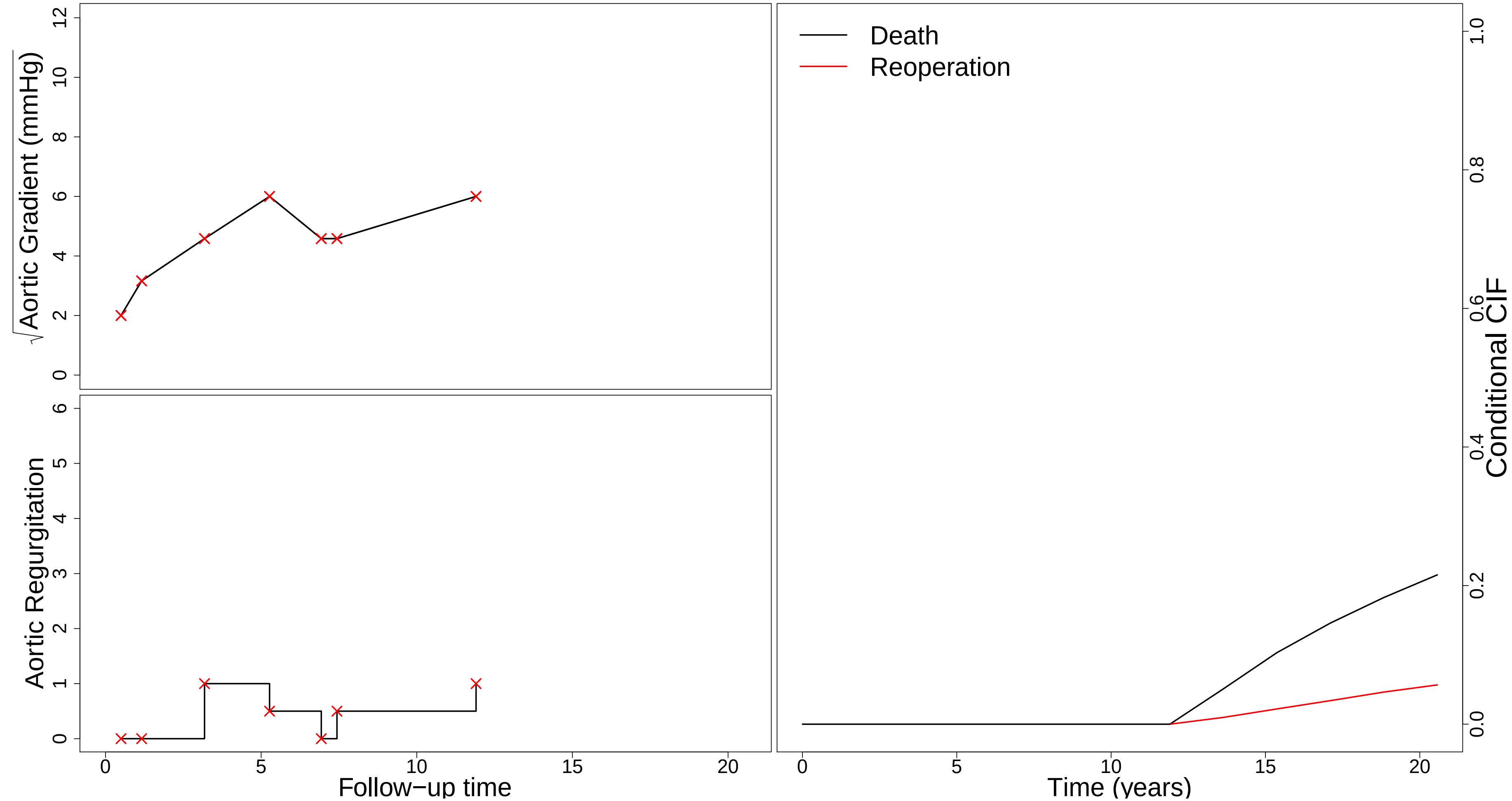
Dynamic predictions - Heart Valve data (cont'd)



Dynamic predictions - Heart Valve data (cont'd)



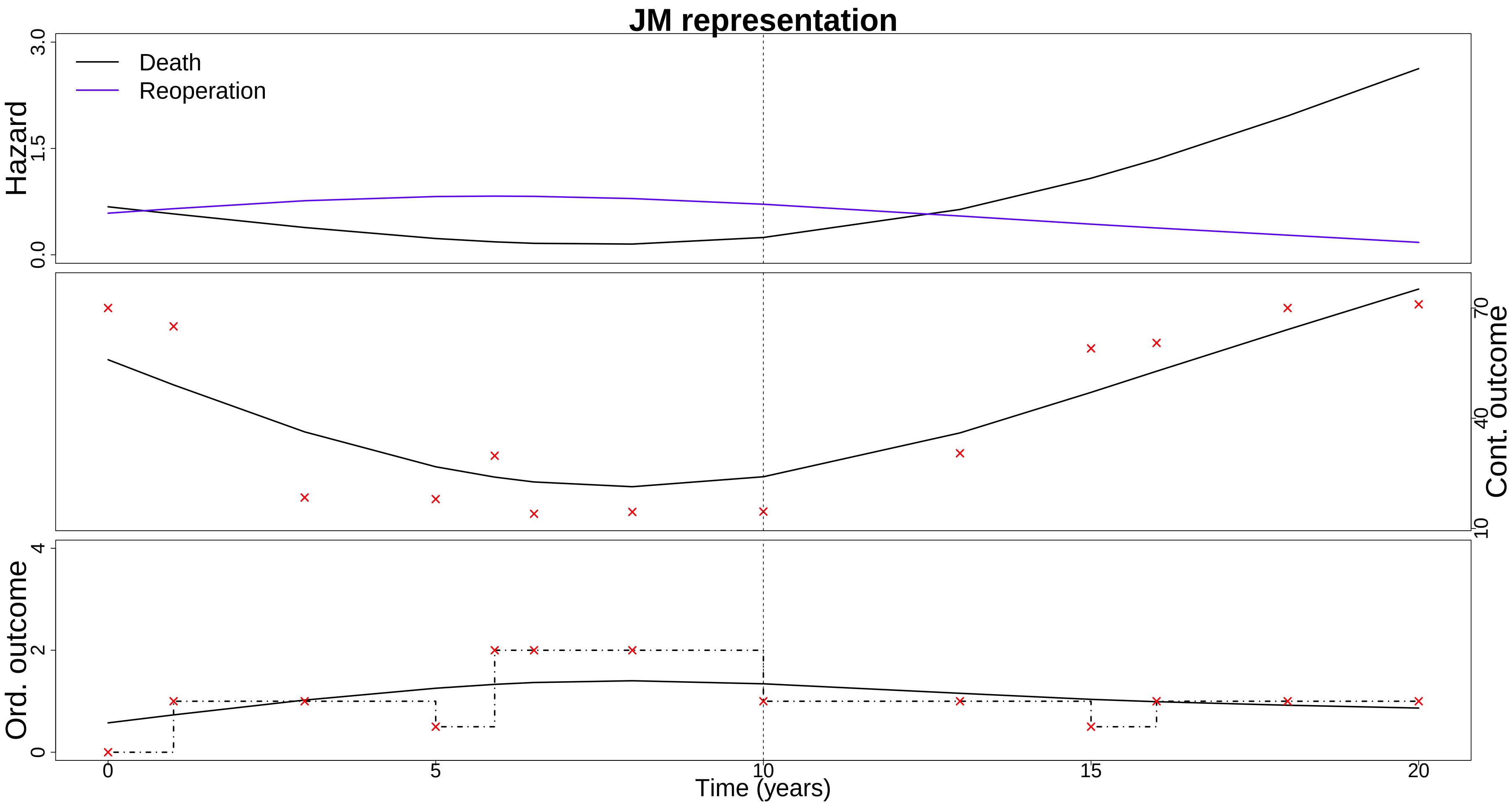
Dynamic predictions - Heart Valve data (cont'd)



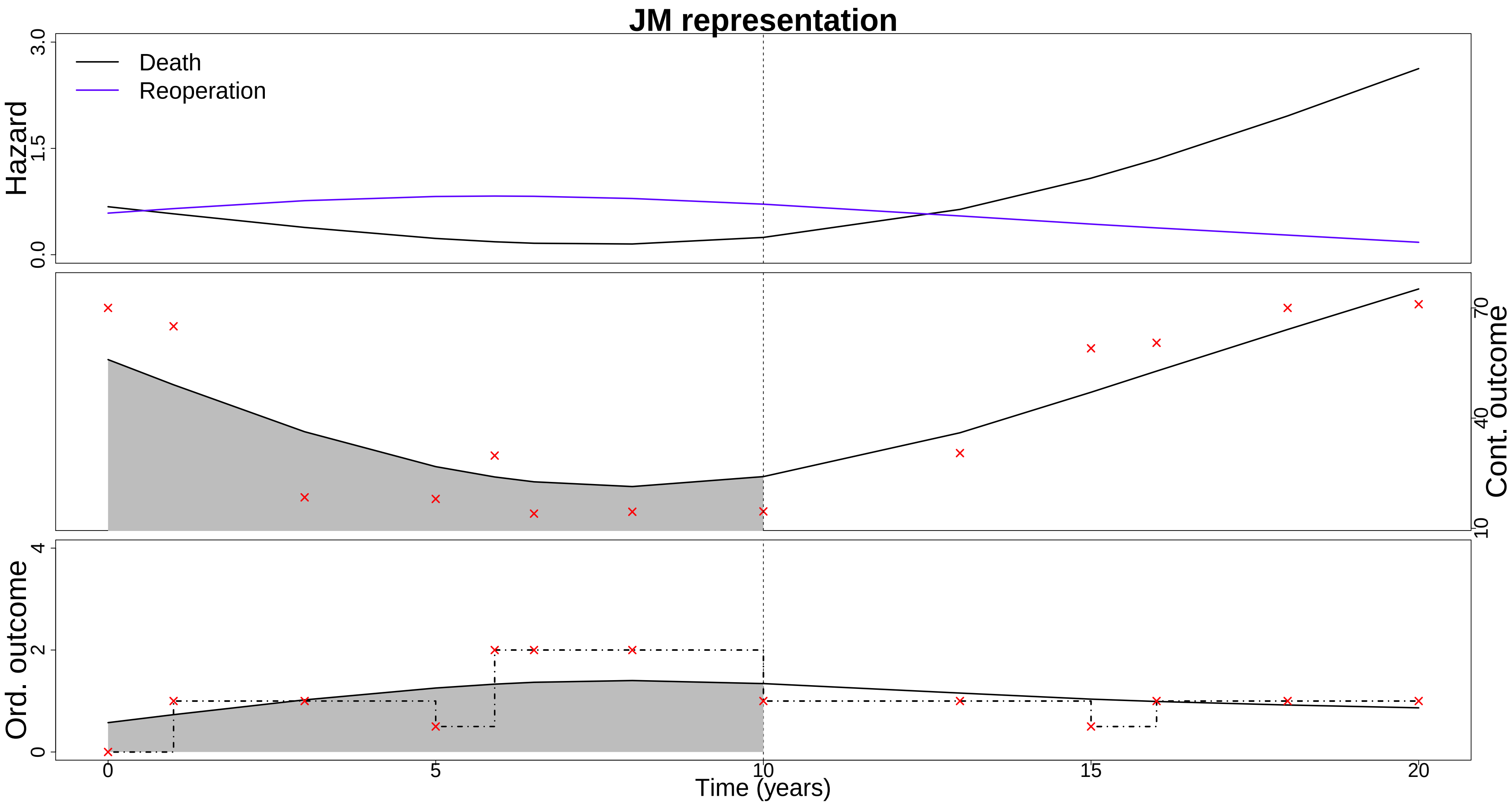
Working on...

- Time-dependent parameterizations
- Bayesian model averaging

Working on... (cont'd)



Working on... (cont'd)



Working on... (cont'd)

- Different parameterizations - time dependent

$$h_{ik}(t) = h_{0k}(t) \exp\{w_i^\top \gamma_k + \alpha_{1k} f_{1i}(t) + \alpha_{2k} f_{2i}(t)\},$$

$$h_{ik}(t) = h_{0k}(t) \exp\{w_i^\top \gamma_k + \alpha_{1k} \int_0^t f_{1i}(s) ds + \alpha_{2k} \int_0^t f_{2i}(s) ds\},$$

where $f_{1i}(t)$ and $f_{2i}(t)$ denote the true and unobserved value of the longitudinal outcomes at time t.

Working on... (cont'd)

- Combinations!

$$h_{ik}(t) = h_{0k}(t)\exp\{w_i^\top \gamma_k + \alpha_{1k} f_{1i}(t) + \alpha_{2k}^\top(\tilde\beta_2 + b_{2i})\},$$

$$h_{ik}(t) = h_{0k}(t)\exp\{w_i^\top \gamma_k + \alpha_{1k}^\top(\tilde\beta_1 + b_{1i}) + \alpha_{2k} f_{2i}(t)\},$$

$$h_{ik}(t) = h_{0k}(t)\exp\{w_i^\top \gamma_k + \alpha_{1k} f_{1i}(t) + \alpha_{2k}\int_0^t f_{2i}(s)ds\}.$$

...

Working on... (cont'd)

Is there one single model appropriate for the data?

Solution: Bayesian Model Averaging

Combine models with different

- association structures
- covariates

...

Thank you!