Dynamic Prediction

Day 9

- Introduction to Survival Prediction
- Introduction to Dynamic Prediction
- Why? When? How?

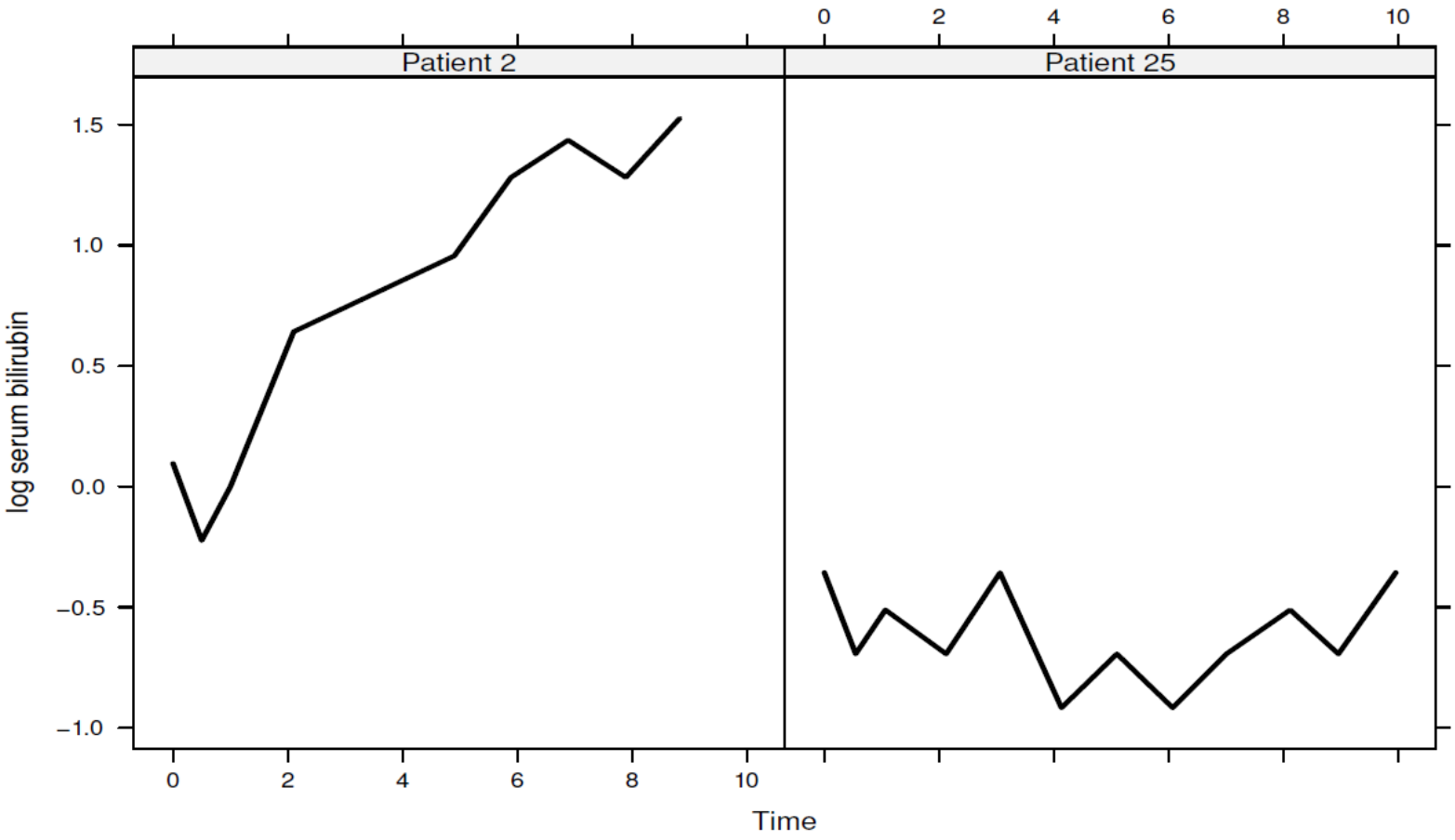
Dynamic Prediction

"What is the probability that the patient will be alive in 5 years, given what we know about them today?"

- Prediction is often conducted at some baseline time (treatment, diagnosis, etc.)
- During follow-up additional information may become available for a patient (biomarkers, intermediate event)
- **Dynamic prediction:** Incorporates changing patient information to predict survival probabilities at time points during a patient's follow-up
- Used by clinicians to make important personalized care decisions

Example: PBC Study

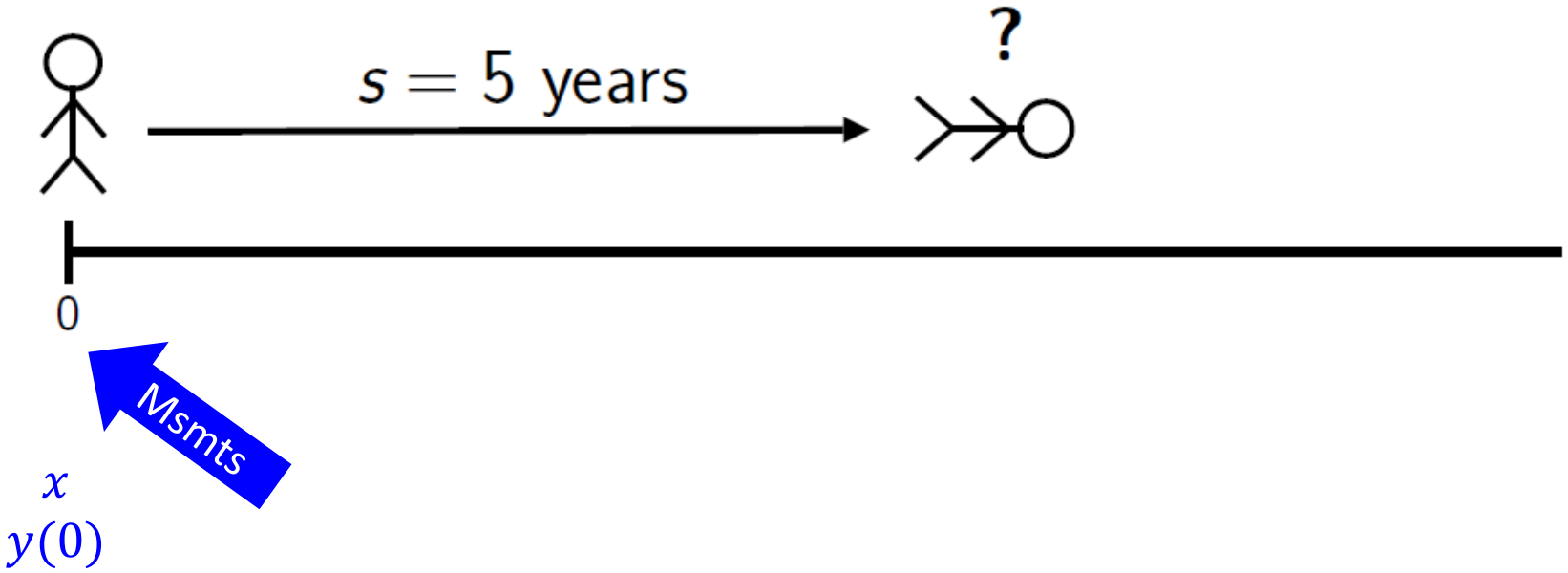
- **Goal:** Dynamic predictions of survival probabilities for Patients 2 and 25 from the PBC dataset



Will I be alive in 5 years?

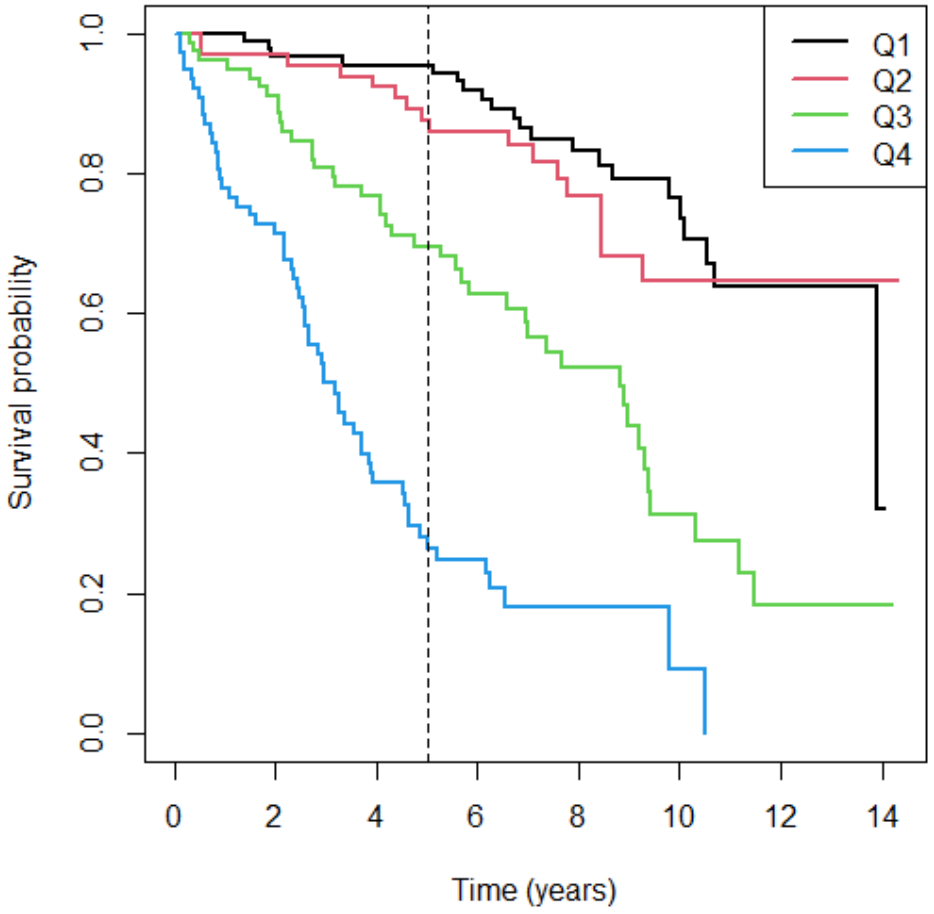
- Prediction window = 5 years
- Baseline covariates x (e.g., age, gender, treatment)
- Time-dependent covariates $y(t)$

$$S(5) = \Pr(T^* \geq 5 | x, y(0))$$



Traditional Survival Prediction: K-M

```
plot(survfit(Surv(years,status2)~ser.cut, data=pbc2.id), conf.int=FALSE,lwd=2, ylab="Survival probability", xl
ab="Time (years)", col=1:4)
legend("topright",c("Q1","Q2","Q3","Q4"),lty=1, lwd=2, col=1:4)
```



- Probability of surviving the next 5 years $\hat{S}(5)$
- Patient 2 (in Q2): 87.6% (95% CI: 79.9-96.0%)
- Patient 25 (in Q1): 95.6% (95% CI: 91.4-99.9%)
- May want to adjust for other baseline covariates using a survival model

Traditional Survival Prediction: Cox

```
mod.cox <- coxph(Surv(years, status2) ~ ser.cut + drug + age + sex, data=pbc2.id)
summary(mod.cox)
```

```
## Call:
## coxph(formula = Surv(years, status2) ~ ser.cut + drug + age +
##     sex, data = pbc2.id)
##
##   n= 312, number of events= 140
##
##                       coef exp(coef)  se(coef)      z Pr(>|z|)
## ser.cut(0.8,1.35]  0.240839  1.272317  0.328085  0.734    0.463
## ser.cut(1.35,3.4]  1.349634  3.856014  0.272772  4.948  7.5e-07 ***
## ser.cut(3.4,28]    2.450707 11.596540  0.269606  9.090  < 2e-16 ***
## drugD-penicil     -0.230931  0.793794  0.176599 -1.308    0.191
## age                0.045825  1.046891  0.008153  5.621  1.9e-08 ***
## sexfemale          0.127537  1.136027  0.235164  0.542    0.588
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## ser.cut(0.8,1.35]    1.2723    0.78597    0.6689     2.420
## ser.cut(1.35,3.4]    3.8560    0.25934    2.2592     6.581
## ser.cut(3.4,28]     11.5965    0.08623    6.8366    19.671
## drugD-penicil        0.7938    1.25977    0.5615     1.122
## age                  1.0469    0.95521    1.0303     1.064
## sexfemale            1.1360    0.88026    0.7165     1.801
##
## Concordance= 0.793  (se = 0.02 )
## Likelihood ratio test= 144  on 6 df,   p=<2e-16
## Wald test            = 145.5  on 6 df,   p=<2e-16
## Score (logrank) test = 177.6  on 6 df,   p=<2e-16
```

Traditional Survival Prediction: Cox

- Remember that the function that we are trying to predict is

$$\begin{aligned} S(t|x) &= \exp\{-\int_0^t h(s)ds\} \\ &= \exp\{-\int_0^t h_0(t)\exp(x'\beta)ds\} \\ &= \exp\{-H_0(t)\exp(x'\beta)\} \end{aligned}$$

- Thus, to get our predicted survival we want to compute

$$\hat{S}(t|x) = \exp\{-\hat{H}_0(t)\exp(x'\hat{\beta})\}$$

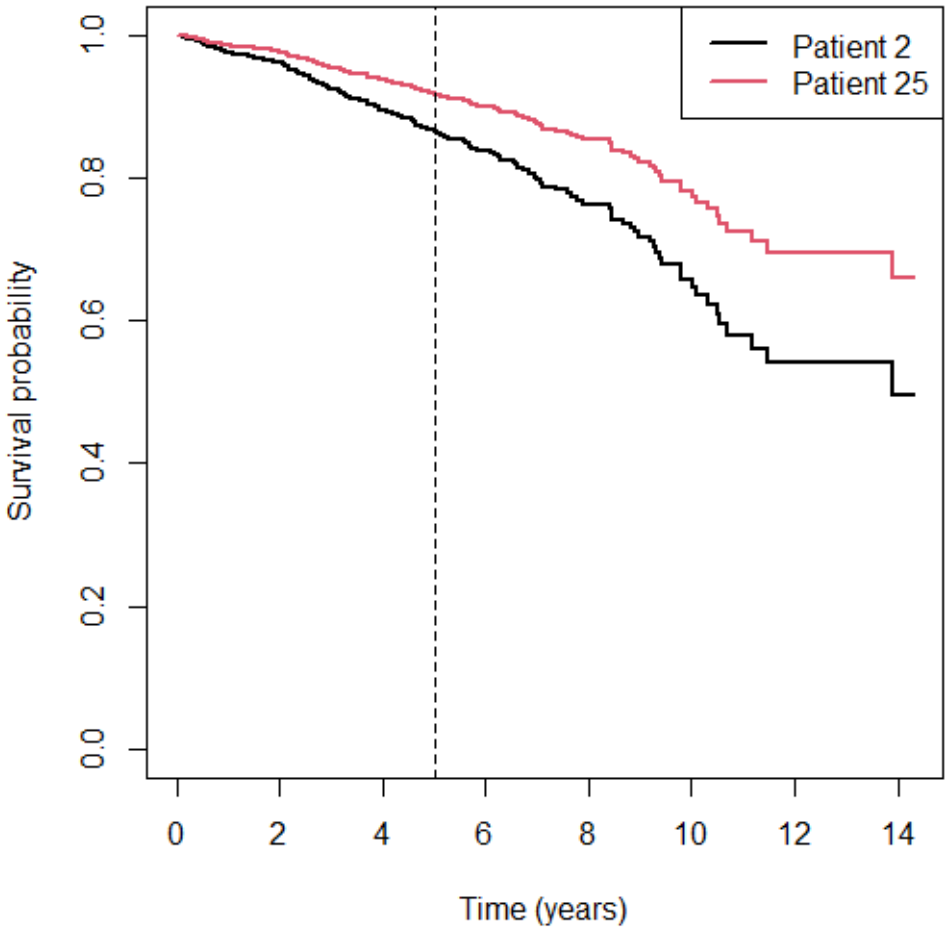
- Cox model does not estimate the baseline hazard, so we use a non-parametric estimator of the baseline hazard function

Traditional Survival Prediction: Cox

- We can estimate $\widehat{H}_0(t)$ using the Breslow estimator
- R does this using the "survfit" function applied to a "coxph" model

```
ND <- rbind(pbc2.id[pbc2.id$id==2,], pbc2.id[pbc2.id$id==25,])
survfit_ND <- survfit(mod.cox, newdata = ND)
plot(survfit_ND, col=1:2, lwd=2,
       ylab="Survival probability",
       xlab="Time (years)")
legend("topright", c("Patient 2", "Patient 25"), col=1:2, lwd=2, lty=1)
```

Traditional Survival Prediction: Cox

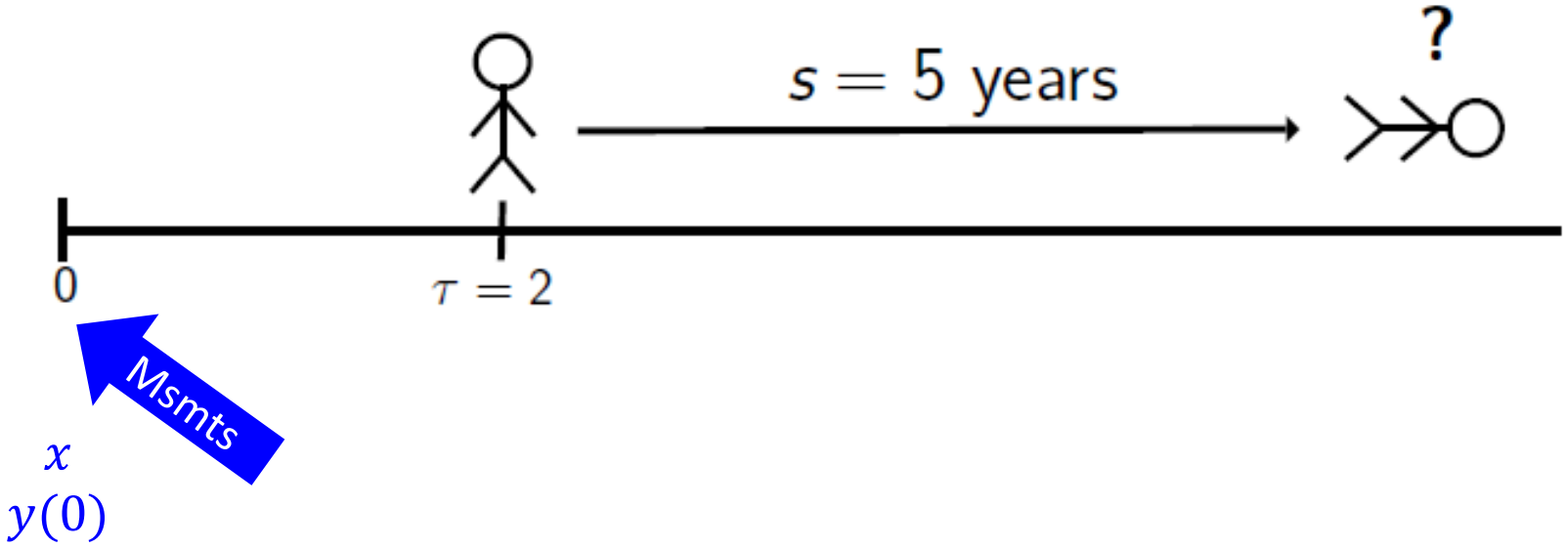


- Probability of surviving the next 5 years $\hat{S}(5)$
- Patient 2 (in Q2): 86.7% (95% CI: 80.3-93.5%)
- Patient 25 (in Q1): 91.9% (95% CI: 88.1-96.0%)
- Prediction uses their baseline marker value and baseline marker measurement

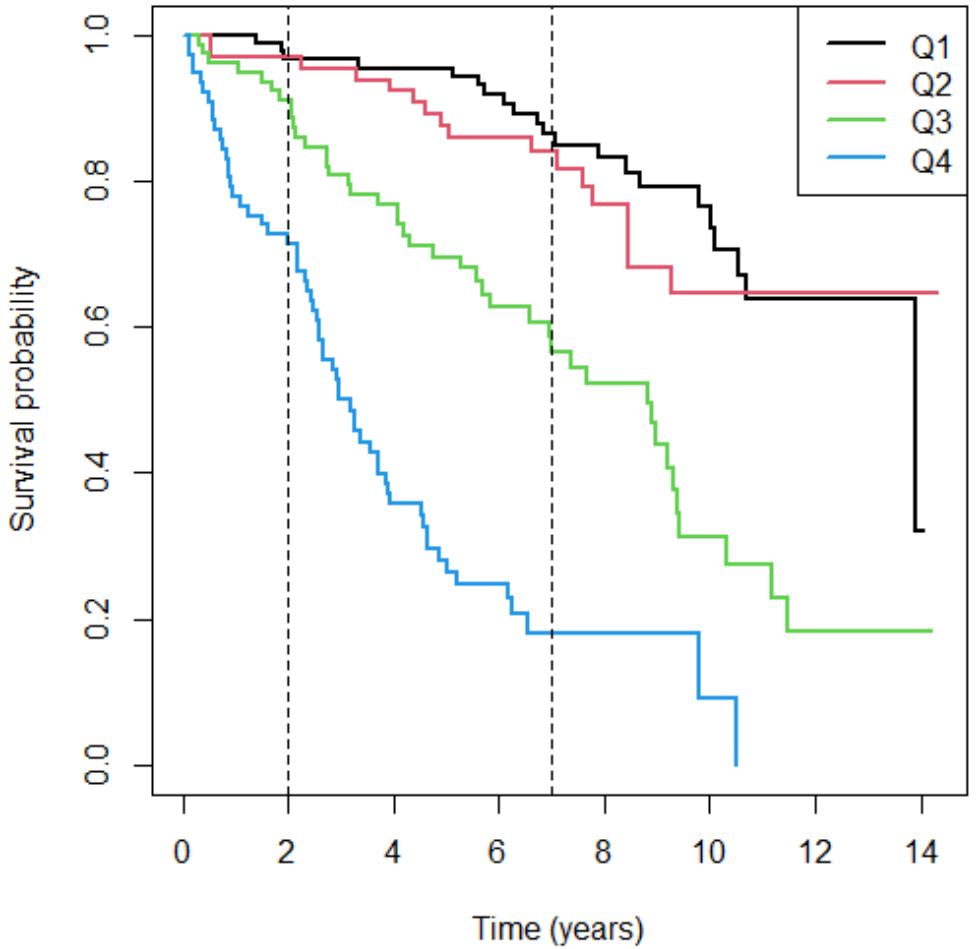
Will I be alive in 5 years?

- The patient is still alive in two years and comes back in
- Now want to make a prediction for the patient two years into their follow-up
- Need to compute their conditional survival probability

$$S(u|t) = \Pr(T^* \geq u | T^* > t, x, y(0)), \quad u > t$$



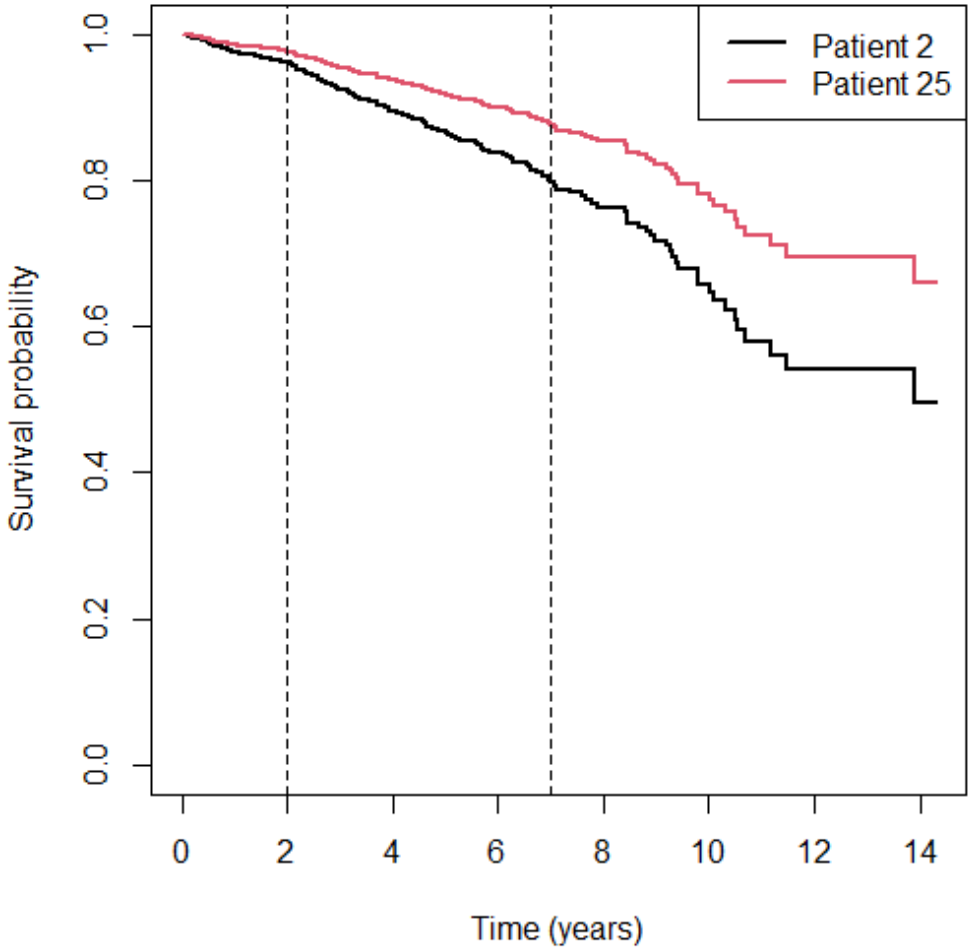
Conditional Survival: K-M



$$S(u|t) = S(u)/S(t)$$

- Given the patient survived 2 years, the probability of surviving the next 5 years $\hat{S}(7|2)$
- Patient 2 (in Q2): 86.6%
- Patient 25 (in Q1): 89.3%

Conditional Survival: Cox

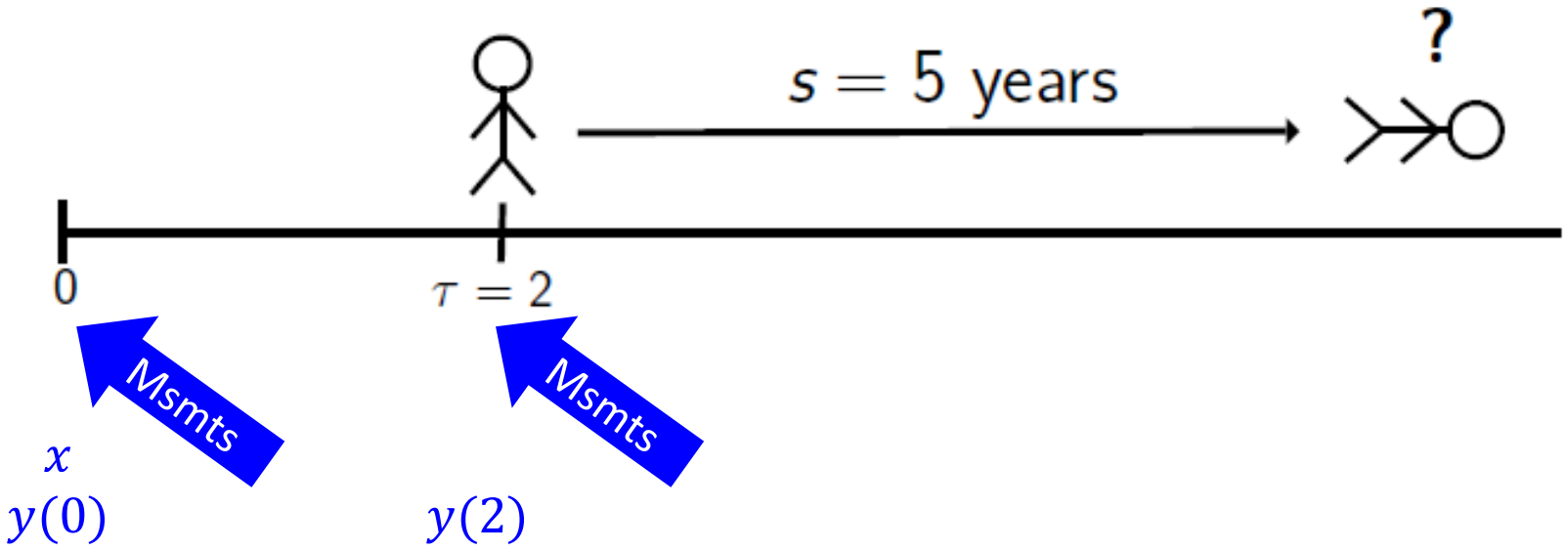


- Similarly from the Cox model
- Given the patient survived 2 years, the probability of surviving the next 5 years $\hat{S}(7|2)$
- Patient 2 (in Q2): 82.4%
- Patient 25 (in Q1): 89.4%

Will I be alive in 5 years?

- Still only using baseline information
- At the follow-up visit, we might learn more information about the patient (collect information on the longitudinal marker)

$$\Pr(T^* \geq u | T^* > t, x, y(t)), \quad u > t$$



Dynamic Prediction

- Interested in predicting survival probability for a patient that has provided a set of longitudinal measurements up to a specific time point t
- For a subject j we have available measurements up to time t

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

- Dynamic prediction of interest is

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

Survival Probabilities: Estimation

- We assume that the joint model has been fitted to the data
- Based on the fitted model we can estimate the conditional survival probabilities

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

- Where D_n is the sample on which the joint model was fitted

Survival Probabilities: Estimation

- $p_j(u|t)$ can be rewritten as

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

- M_j is the longitudinal history approximated by the mixed effects model (function of the random effects and parameters)
- A naïve estimator for $p_j(u|t)$ can be constructed by plugging-in the MLEs and the Empirical Bayes estimates

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

- Works well in practice, but standard errors are difficult to compute

Survival Probabilities: Estimation

- Instead, can use a Bayesian formulation of the problem

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

- The first part of the integrand we saw on the previous slide
- The second part, if the sample size is sufficiently large, we can approximate the posterior of the parameters by

$$\{\theta|\mathcal{D}_n\} \sim N(\hat{\theta}, \hat{var}(\hat{\theta}))$$

where $\hat{\theta}$ are the MLEs

Survival Probabilities: Estimation

- A Monte Carlo estimate of $p_j(u|t)$ can be obtained using the following simulation scheme

Step 1. Draw $\theta(l) \sim N(\hat{\theta}, \hat{var}(\hat{\theta}))$

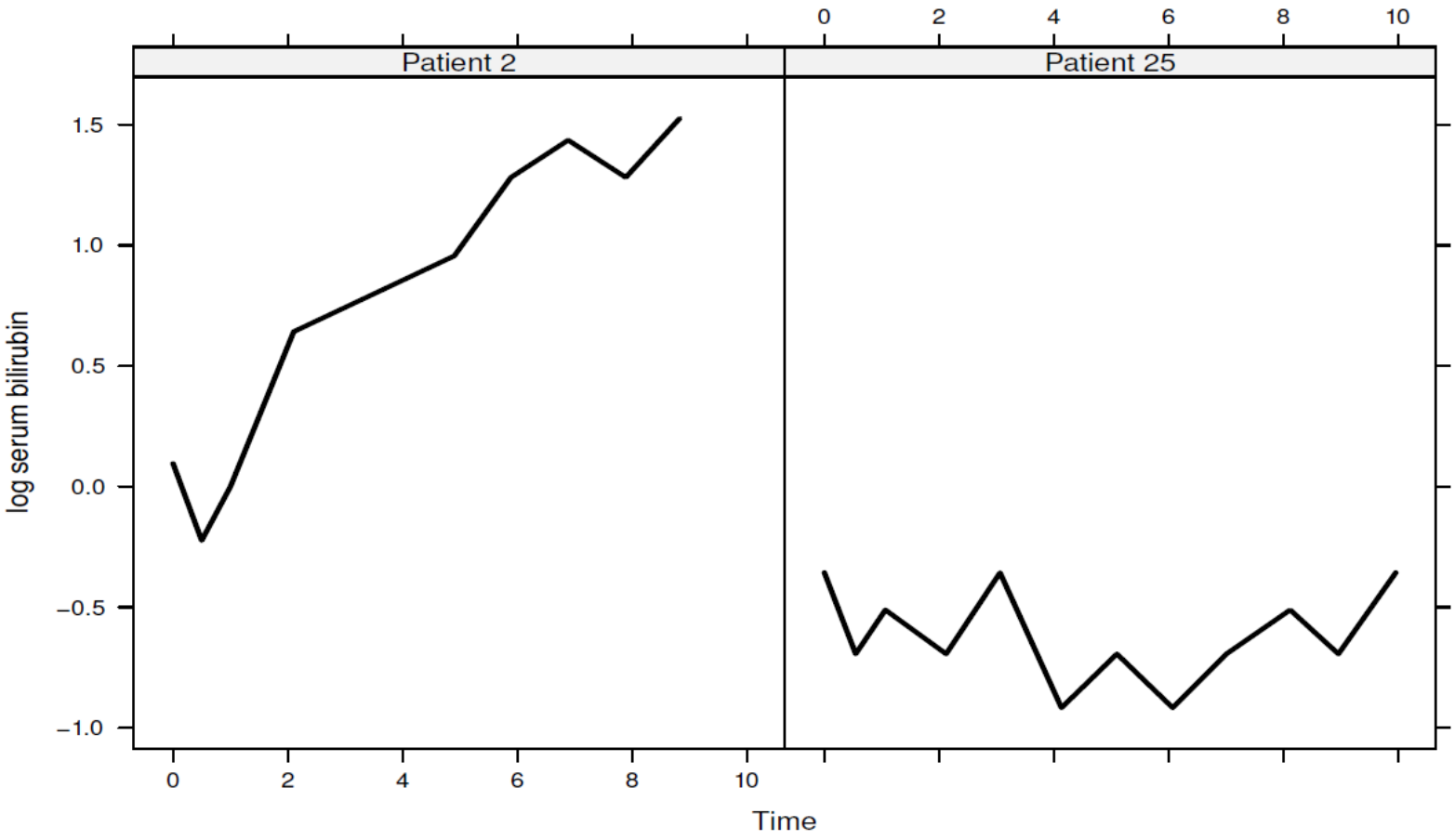
Step 2. Draw $b_j^{(l)} \sim \{b_j | T_j^* > t, \mathcal{Y}_j(t), \theta^{(l)}\}$

Step 3. Compute $p_j^{(l)}(u|t) = S_j\{u|\mathcal{M}_j(u, b_j^{(l)}, \theta^{(l)}; \theta^{(l)}\}/S_j\{t|\mathcal{M}_j(t, b_j^{(l)}, \theta^{(l)}; \theta^{(l)}\}$

- Repeat Steps 1-3, $l = 1, \ldots L$ times, where L is the number of Monte Carlo samples
- Can use the realizations $\{p_i^{(l)}(u|t), l = 1, \ldots, L\}$ to derive point estimates of $p_i(u|t)$
- E.g. $\hat{p}_i(u|t) = \text{median}\{p_i^{(l)}(u|t), l = 1, \ldots, L\}$

Example: PBC Study

- **Goal:** Dynamic predictions of survival probabilities for Patients 2 and 25 from the PBC dataset



Example: PBC Study

- **Goal:** Dynamic predictions of survival probabilities for Patients 2 and 25 from the PBC dataset
- We fit a joint model
- Longitudinal submodel
  - Fixed and random effects: cubic B-spline with boundary knots at 0 and 15 years
- Survival submodel
  - Treatment effect + current level of true serum bilirubin marker
  - Piecewise-constant baseline hazard with 7 intervals

```
lmeFit <- lme(log(serBilir) ~ bs(year, 4, Boundary.knots=c(0,15)), random = list(id = pdDiag(form = ~ bs(year,
4, Boundary.knots=c(0,15)))), data= pbc2)
survFit <- coxph(Surv(years, status2) ~ drug, data = pbc2.id, x = TRUE)
jointFit <- jointModel(lmeFit, survFit, timeVar="year", method = "piecewise-PH-aGH")
```

Example: PBC Study

- Based on the fitted joint model we estimate $p_j(u|t)$ for Patients 2 and 25
- We use 500 Monte Carlo samples, and we take the median of the Monte Carlo estimates as our estimate and the corresponding 95% pointwise confidence intervals

Example: PBC Study

```
pbc2[pbc2$id==2,c("id","year","serBilir","years","status2")]
```

```
##    id      year serBilir    years status2
## 3   2 0.0000000      1.1 14.15234       0
## 4   2 0.4983025      0.8 14.15234       0
## 5   2 0.9993429      1.0 14.15234       0
## 6   2 2.1027270      1.9 14.15234       0
## 7   2 4.9008871      2.6 14.15234       0
## 8   2 5.8892783      3.6 14.15234       0
## 9   2 6.8858833      4.2 14.15234       0
## 10  2 7.8907020      3.6 14.15234       0
## 11  2 8.8325485      4.6 14.15234       0
```

Example: PBC Study

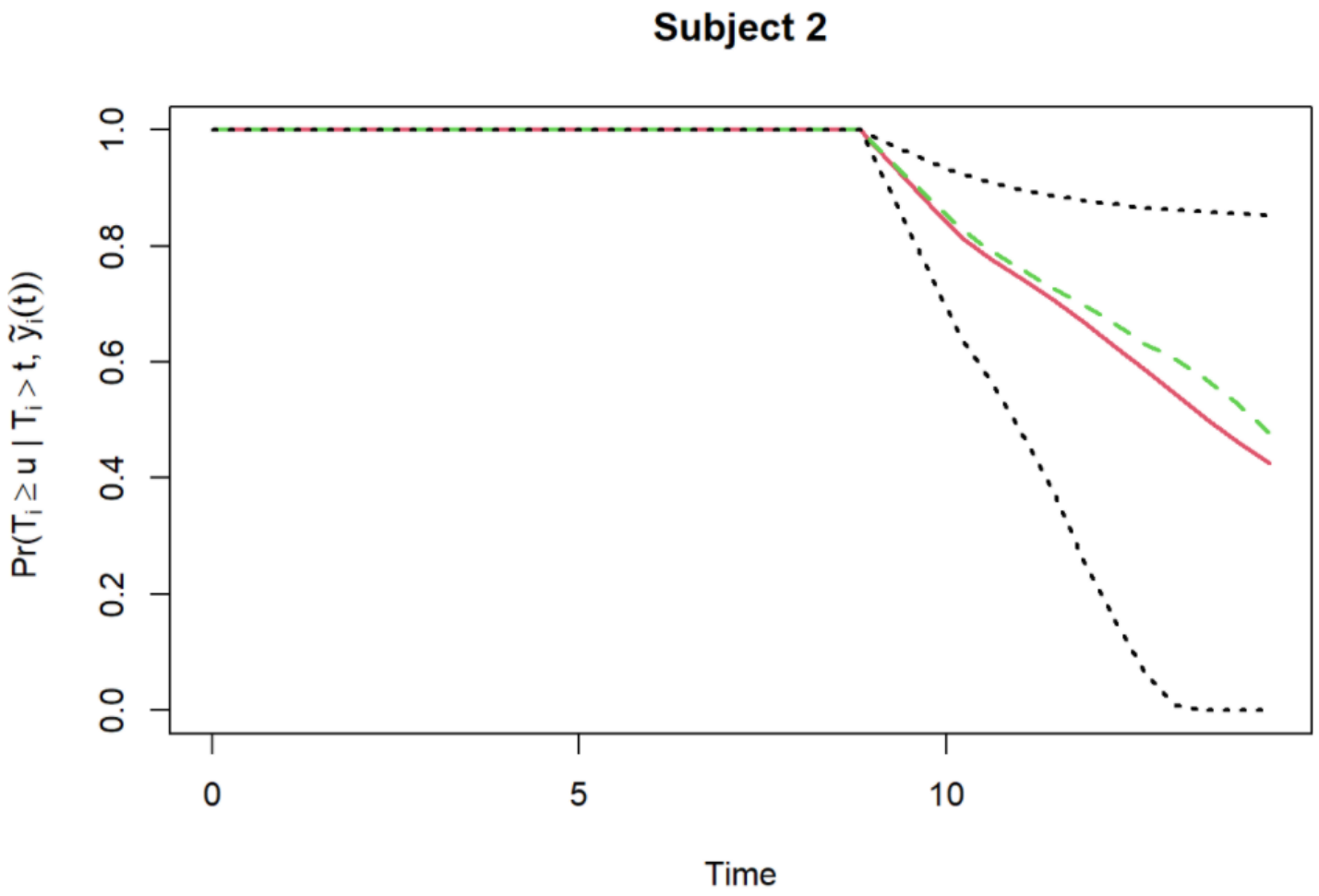
```
survProbs <- survfitJM(jointFit, newdata=pbc2[pbc2$id==2,], M=500)

survProbs
```

```
##
## Prediction of Conditional Probabilities of Event
##  based on 500 Monte Carlo samples
##
## $`2`
##      times   Mean Median  Lower  Upper
## 1   8.8325 1.0000 1.0000 1.0000 1.0000
## 1   8.9405 0.9851 0.9860 0.9721 0.9932
## 2   9.3609 0.9277 0.9327 0.8645 0.9682
## 3   9.7813 0.8711 0.8811 0.7521 0.9457
## 4  10.2017 0.8148 0.8290 0.6431 0.9250
## 5  10.6221 0.7764 0.7910 0.5639 0.9099
## 6  11.0425 0.7426 0.7583 0.4746 0.8962
## 7  11.4629 0.7067 0.7270 0.3761 0.8872
## 8  11.8833 0.6684 0.6985 0.2524 0.8785
## 9  12.3037 0.6278 0.6659 0.1552 0.8722
## 10 12.7241 0.5855 0.6297 0.0622 0.8662
## 11 13.1445 0.5427 0.6030 0.0082 0.8622
## 12 13.5649 0.5005 0.5685 0.0002 0.8585
## 13 13.9853 0.4613 0.5281 0.0000 0.8557
## 14 14.4057 0.4265 0.4770 0.0000 0.8537
```

Example: PBC Study

```
#dashed: median, solid=mean
plot(survProbs, lty = c(1:2,3,3), lwd=2, conf.int = TRUE)
```

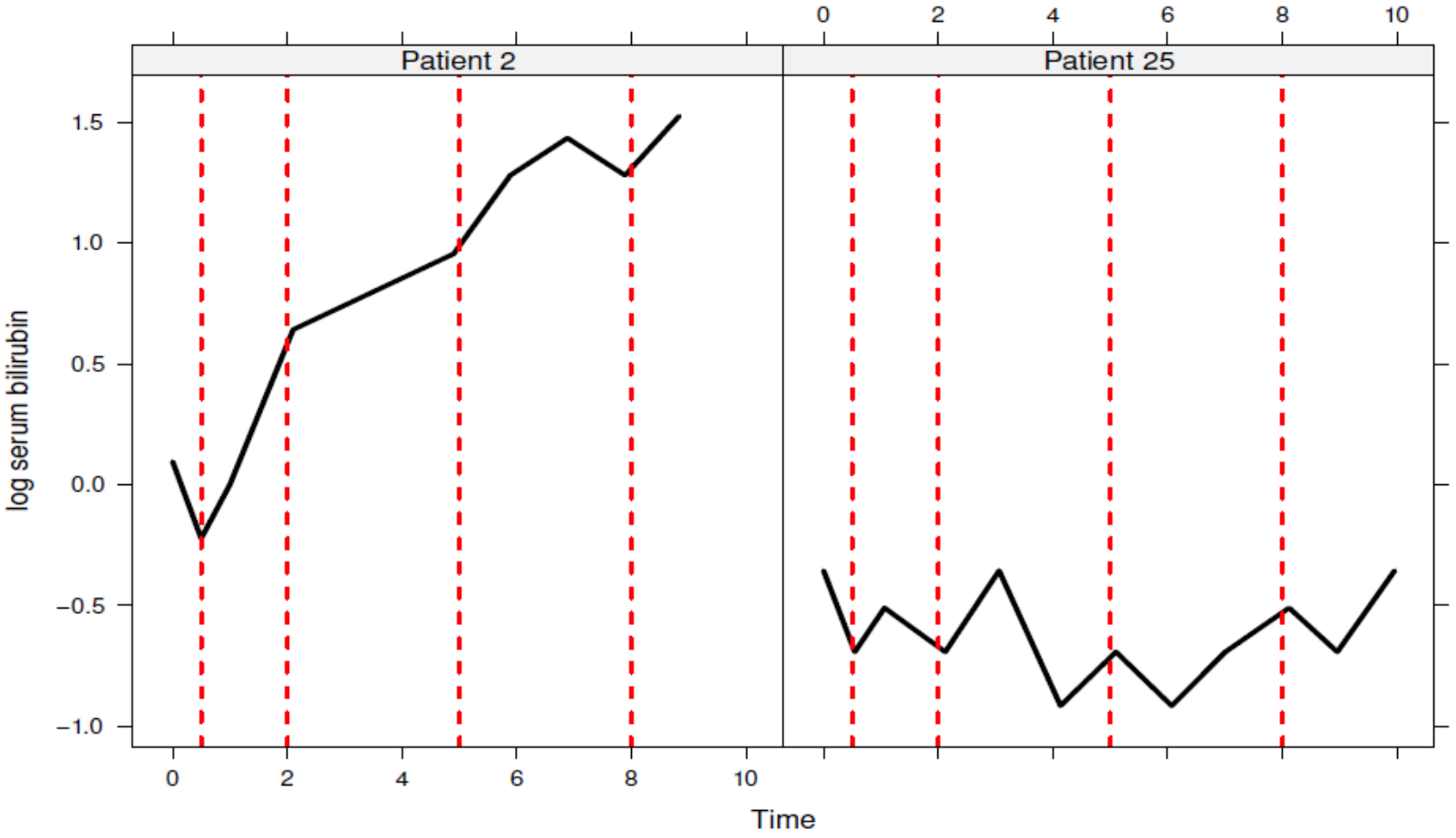


Example: PBC Study

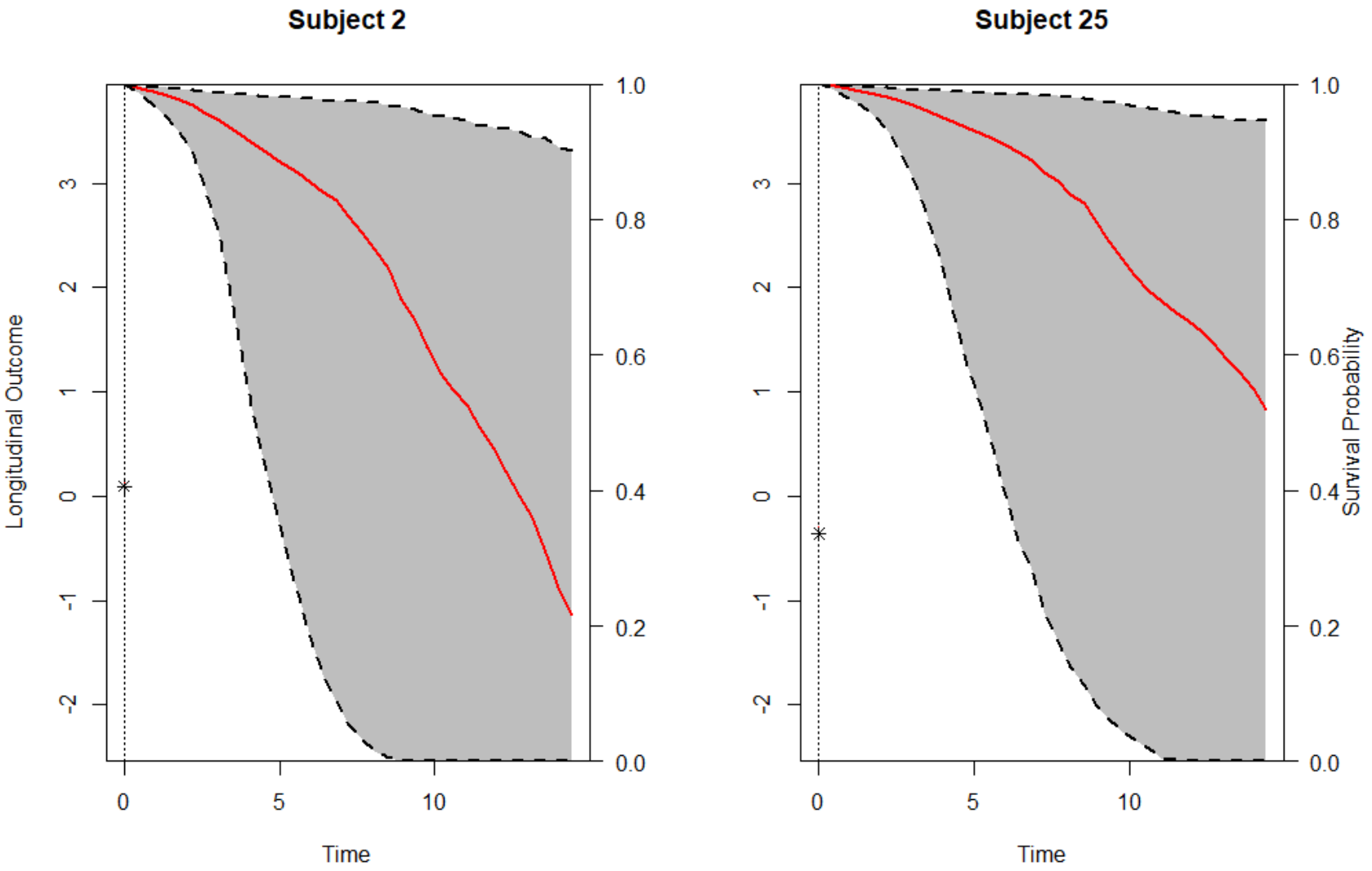
```
survfitJM(jointFit, newdata = pbc2[pbc2$id == 2, ],
          survTimes = c(14.5, 15), last.time = "years")
```

```
##
## Prediction of Conditional Probabilities of Event
##  based on 200 Monte Carlo samples
##
## $`2`
##     times   Mean Median  Lower  Upper
## 1 14.1523 1.0000 1.0000 1.0000 1.0000
## 1 14.5000 0.9019 0.9698 0.3378 0.9990
## 2 15.0000 0.8014 0.9229 0.0218 0.9979
```

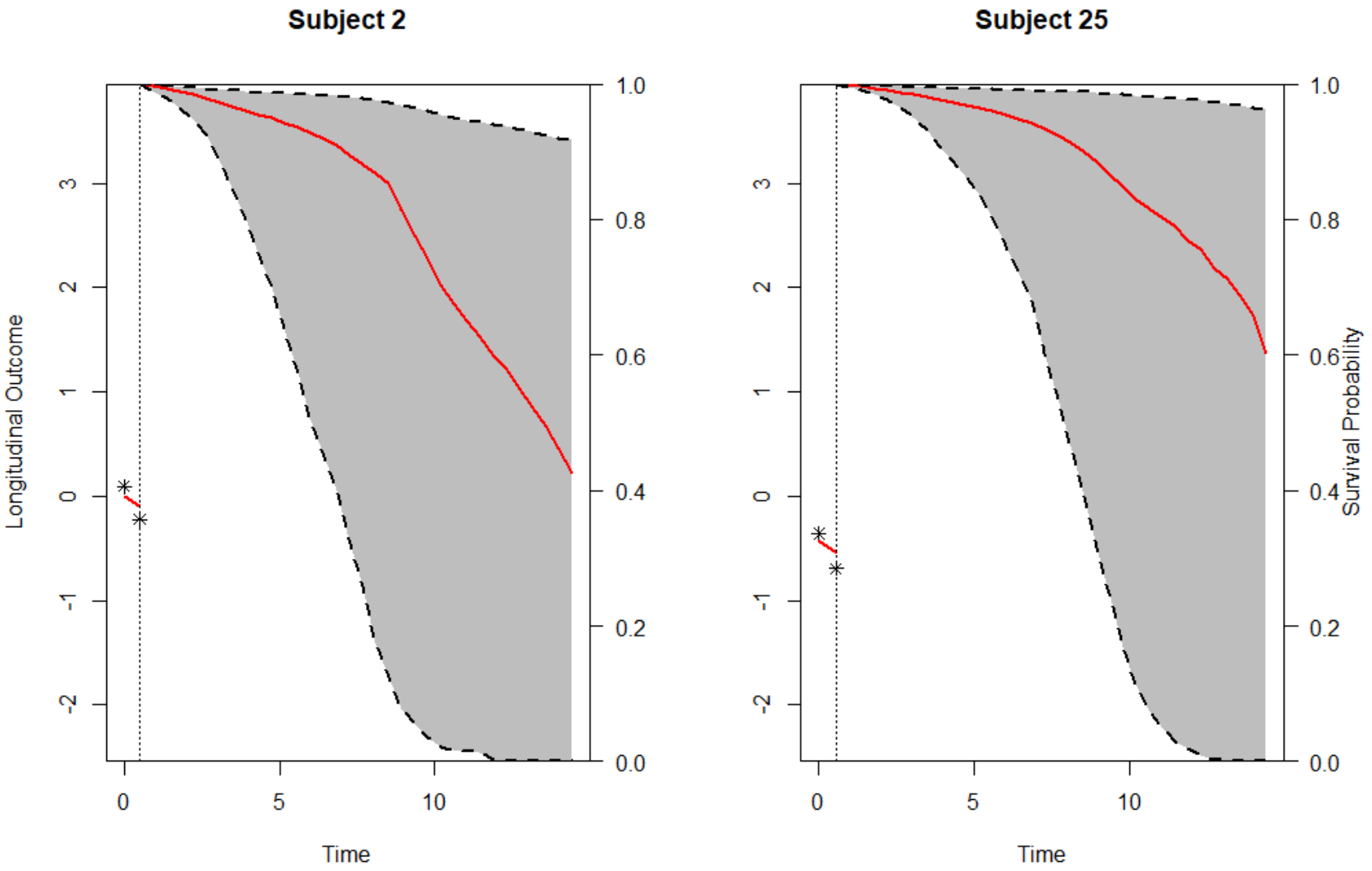
Example: PBC Study



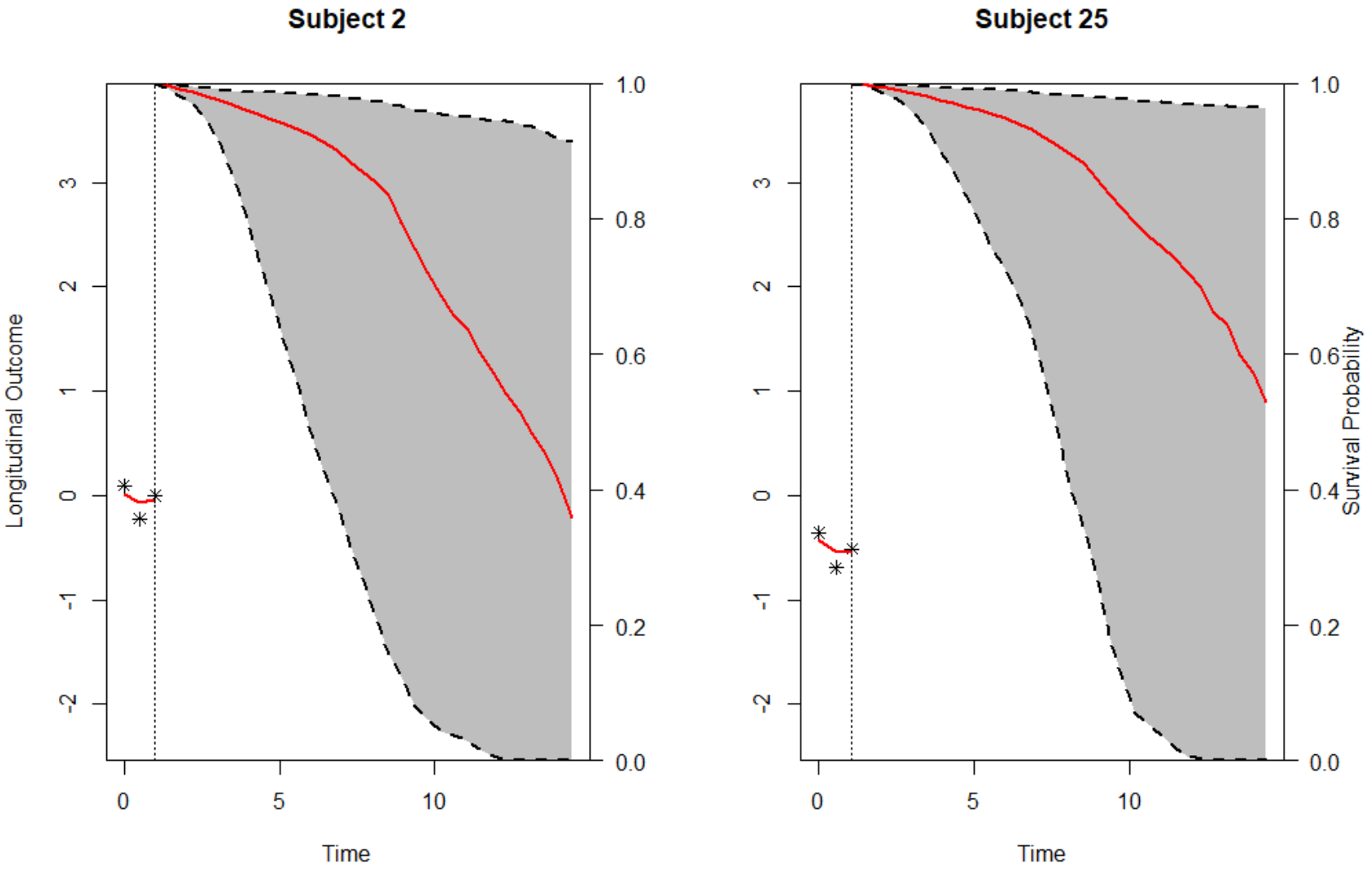
Example: PBC Study



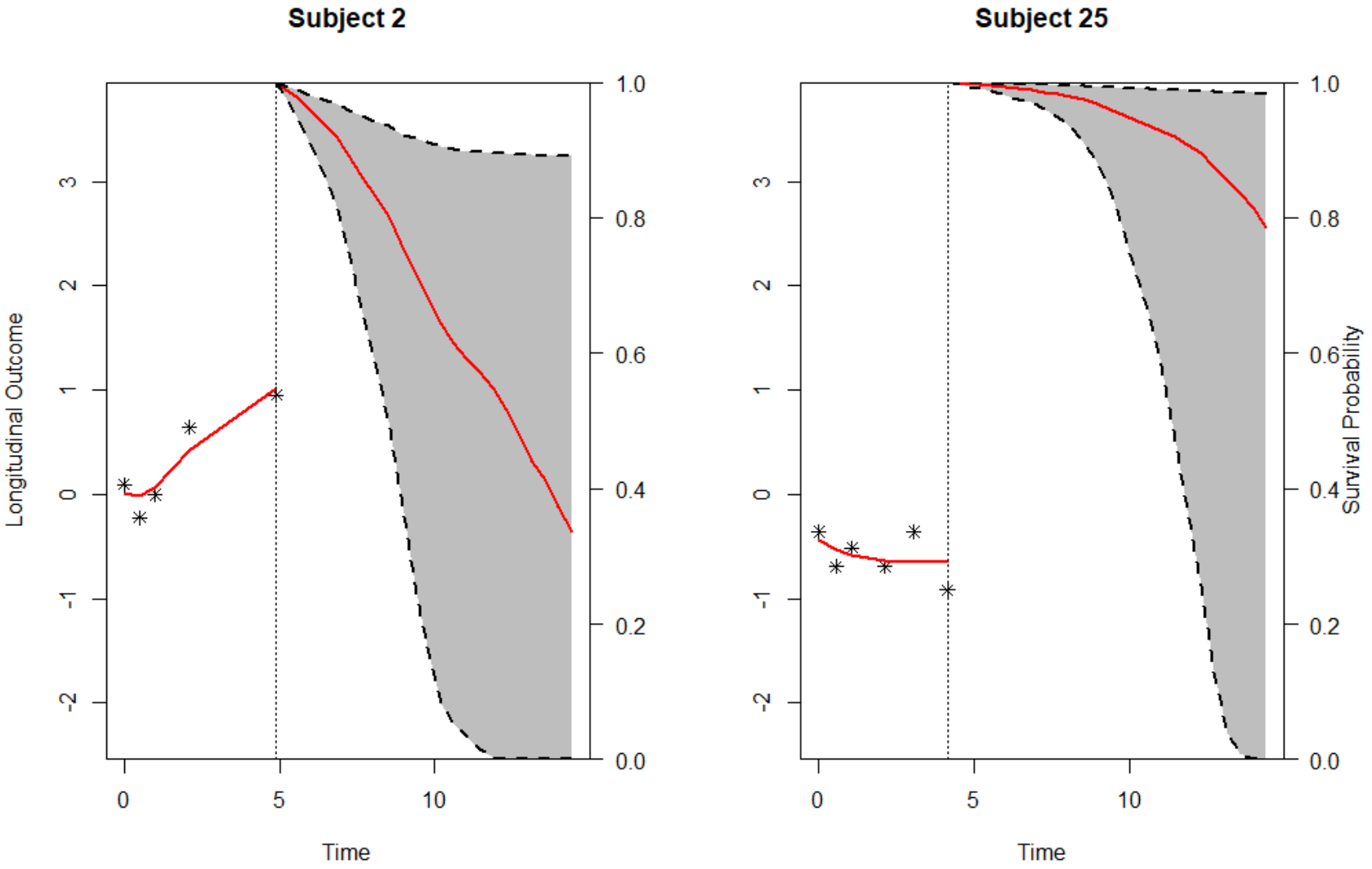
Example: PBC Study



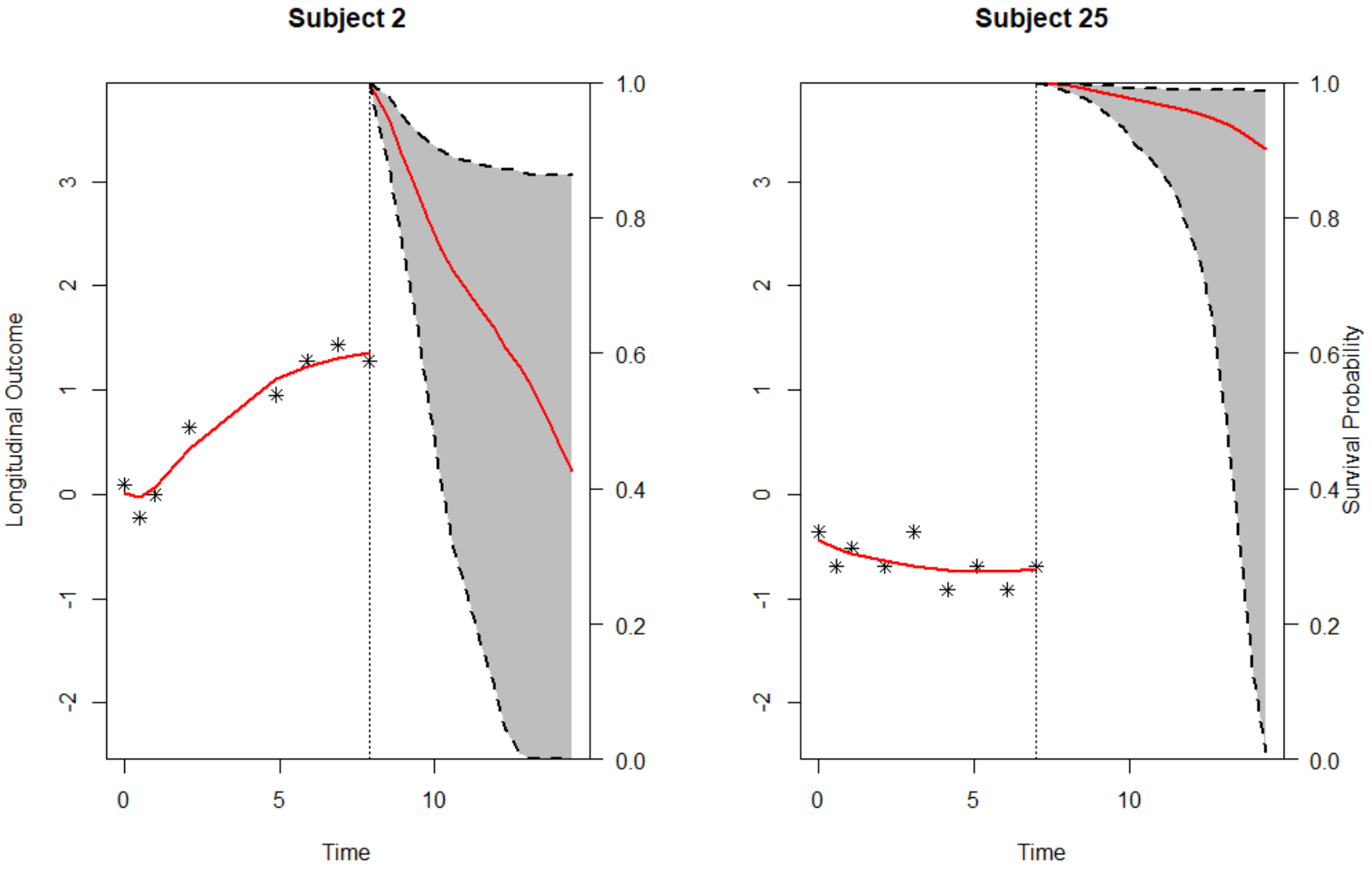
Example: PBC Study



Example: PBC Study



Example: PBC Study



Longitudinal Responses

- In some occasions, it may also be of interest to predict the longitudinal outcome
- We can proceed in the same manner as for the survival probabilities
- We have available marker measurements up to time point t

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

- We are interested in

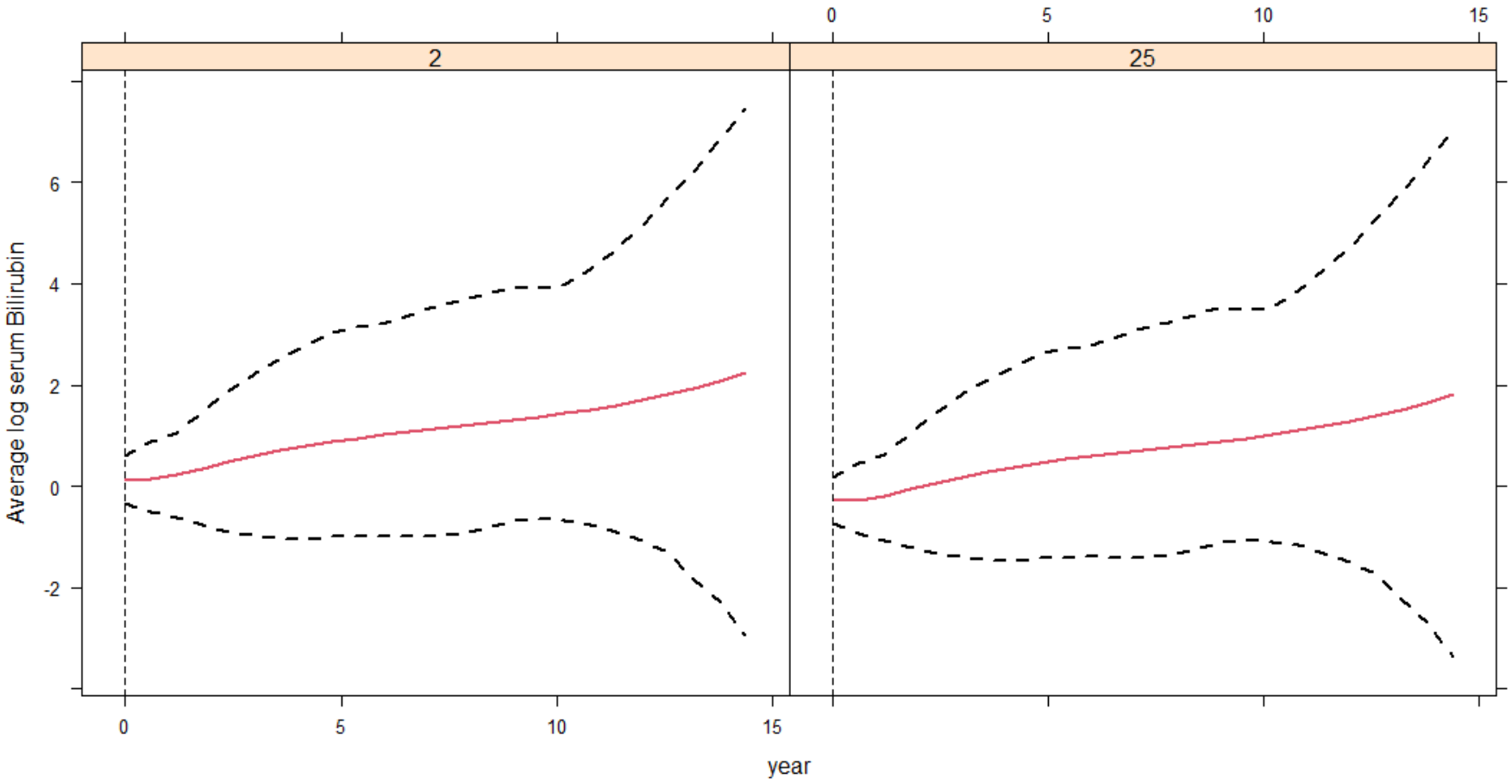
$$w_j(u|t) = E\{y_j(u)|T_j^* > t, \mathcal{Y}_j(t), \mathcal{D}_n\}, \quad u > t$$

Longitudinal Responses: Estimation

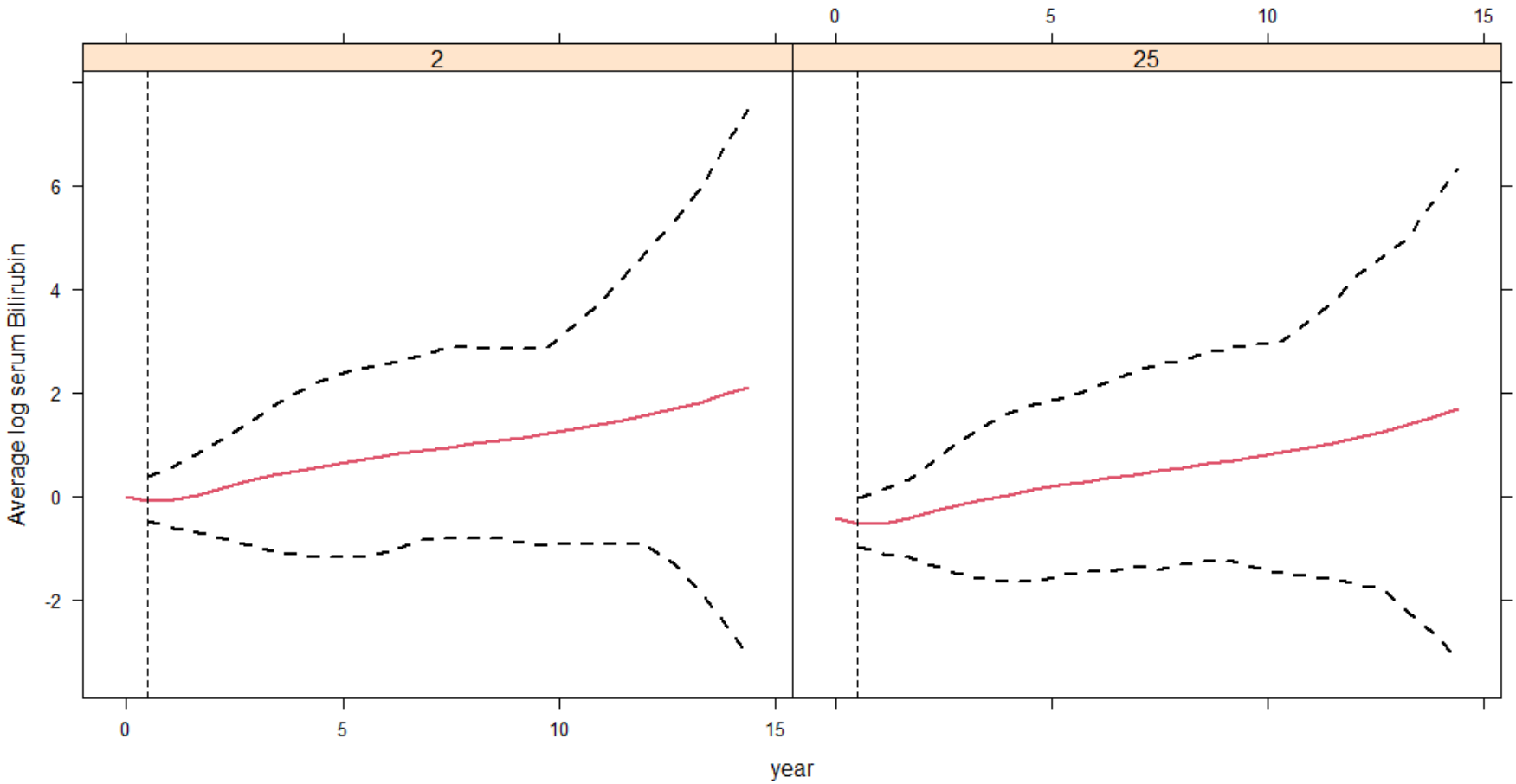
- We use a similar approach as for the survival probabilities
- A Monte Carlo estimate of $w_j(u|t)$ can be obtained using the following simulation scheme

$$\text{Step 1. Draw } \theta(l) \sim N(\hat{\theta}, v\hat{a}r(\hat{\theta}))$$
$$\text{Step 2. Draw } b_j^{(l)} \sim \{b_j | T_j^* > t, \mathcal{Y}_j(t), \theta^{(l)}\}$$
$$\text{Step 3. Compute } w_j^{(l)}(u|t) = x_j'(u)\beta^{(l)} + z_j'(u)b_j^{(l)}$$

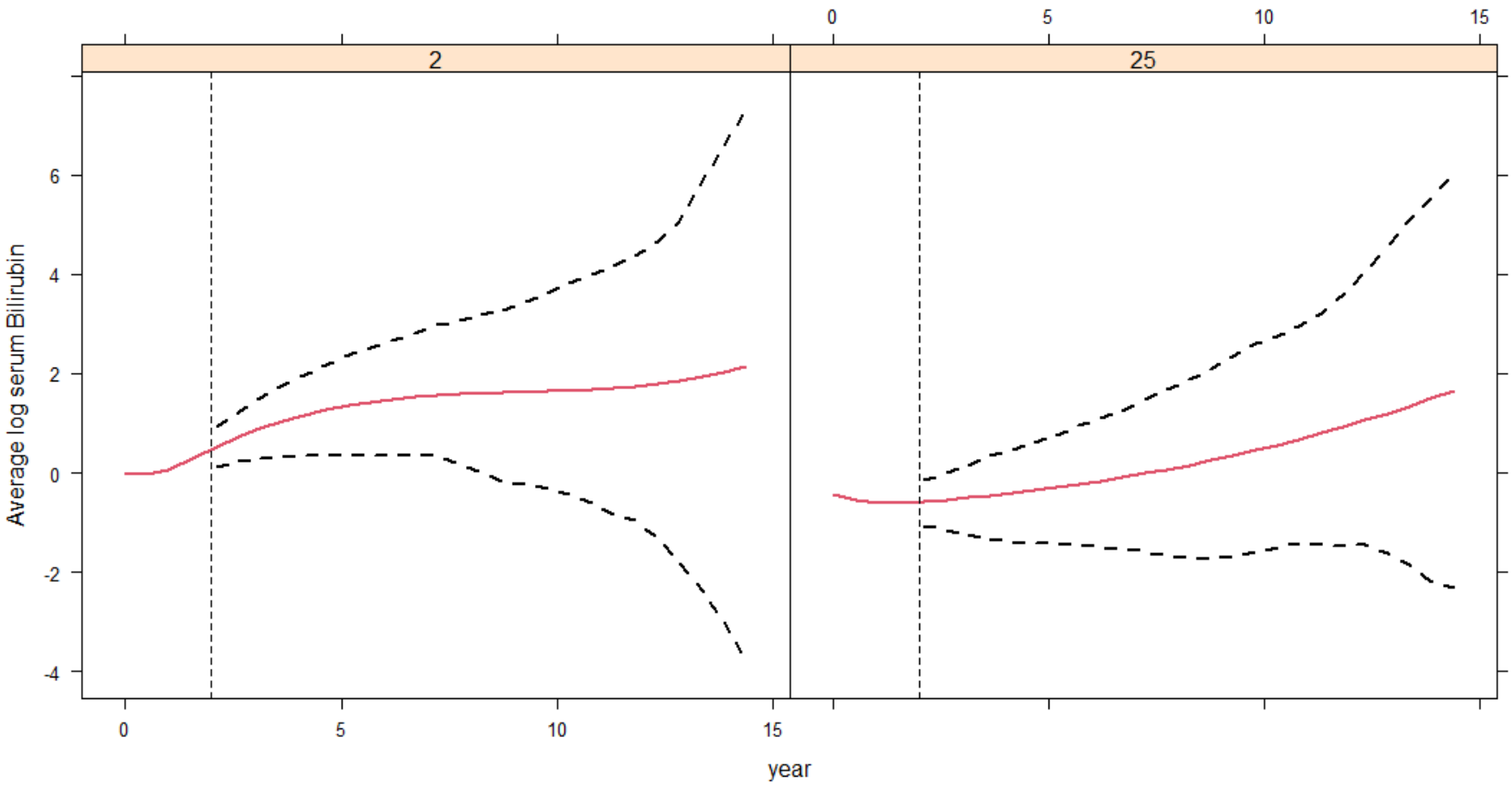
Example: PBC Study



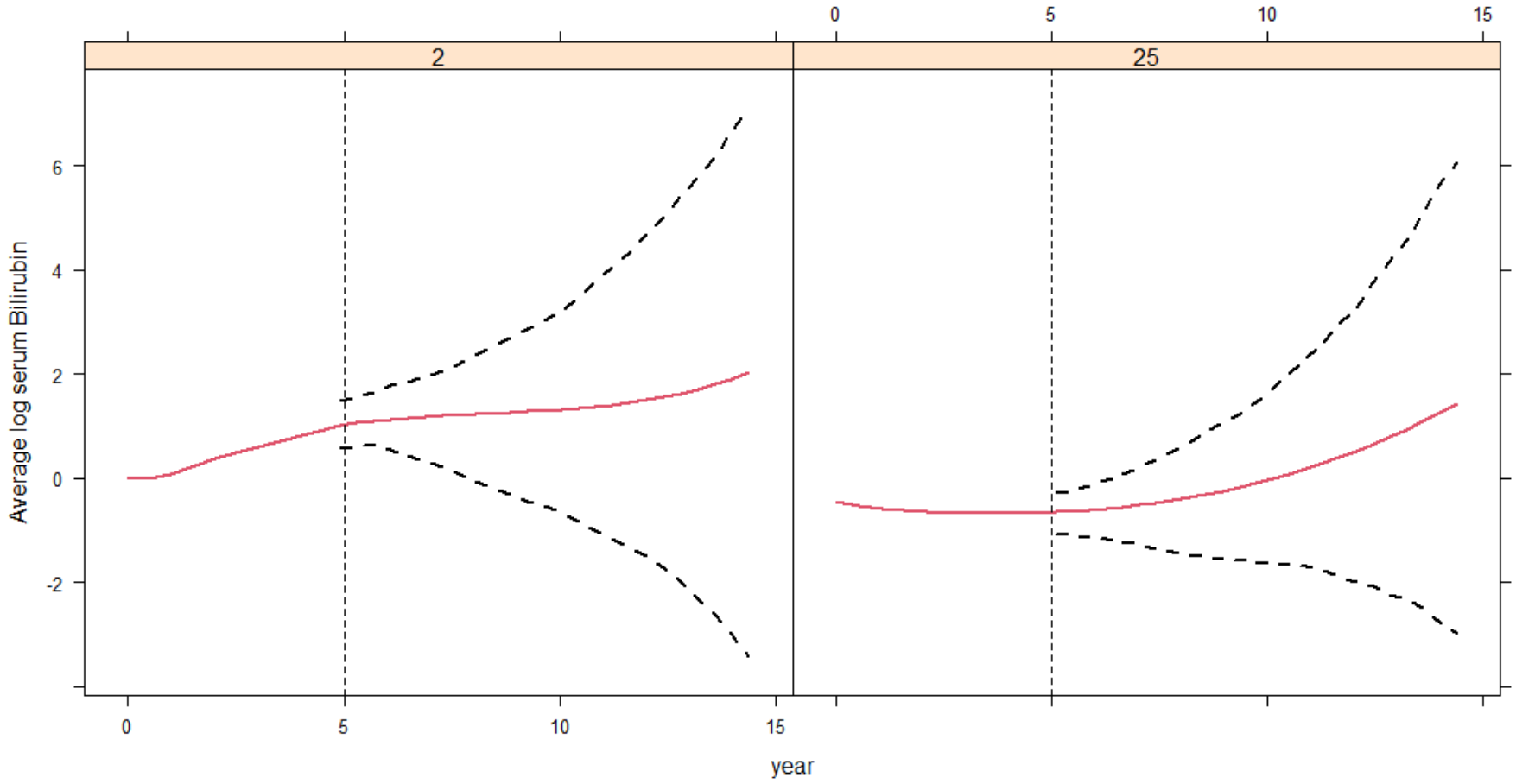
Example: PBC Study



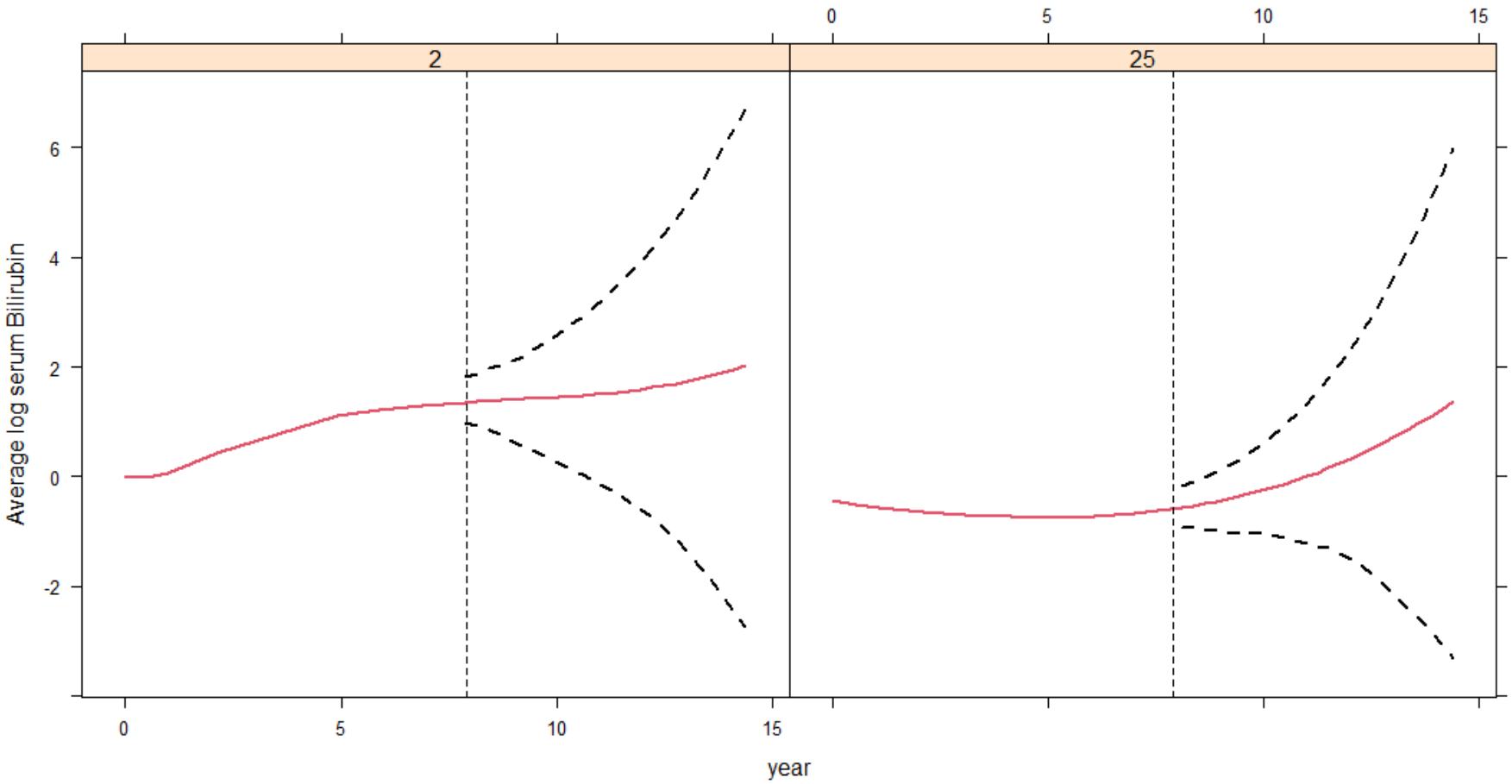
Example: PBC Study



Example: PBC Study



Example: PBC Study



Joint Modeling for Dynamic Prediction

- Valid prediction function that produces consistent predictions *(Jewell and Nielson, 1993)*
- Produces predictions at any time point during follow-up
- Handles irregular marker measurements
- Builds the subject-specific marker profile that extrapolates up to time t (doesn't need LOCF assumption)
- Integrates over future values of the marker
- Biologically seems like a logical approach
- Computationally intensive
- Other methods: Landmarking, Two-stage models, etc.

Performance Metrics

- Discrimination: C-index
- Calibration: Brier Score
- Time-dependent measures
  - Need to account for censoring
  - Blanche et al 2013
  - pec package
- Validation of discrimination and calibration measures can be achieved with standard re-sampling techniques
  - Cross-validation, Bootstrap
- Time-consuming to fit the joint model many times

Course Summary

- When do we need joint models for longitudinal and survival outcomes?
  - To handle endogenous time-varying covariates in a survival analysis context
  - To account for informative dropout in a longitudinal data analysis context
- How do joint models work?
  - A mixed model for the longitudinal outcome
  - A relative risk model for the event process
  - Explain interrelationships with shared random effects

Course Summary

- 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 (parameterization) between the longitudinal and survival process
  - Issues with convergence: Consider transforming covariates, modifying starting values, starting with simpler models and building up
- Extensions
  - Under the full conditional independence assumption we can easily extend the basic joint model
  - Multiple longitudinal outcomes, multiple failure times
  - Is a lot more computationally intensive

Course Summary

- 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 subject
  - Joint models are an excellent tool for personalized medicine
- So much more…
  - Different kinds of joint models
  - Predictive performance metrics