[Part 2] – Model Development for the Prediction of Survival Times using Longitudinal Measurements

- Aasthaa Bansal PhD
- Pharmaceutical Outcomes Research & Policy Program
- University of Washington

Outline

- Background, Motivating Example
- Joint Models
- Partly Conditional Models
- Comparison

Background

(1) Longitudinal measurements: Mixed effects models, GEE

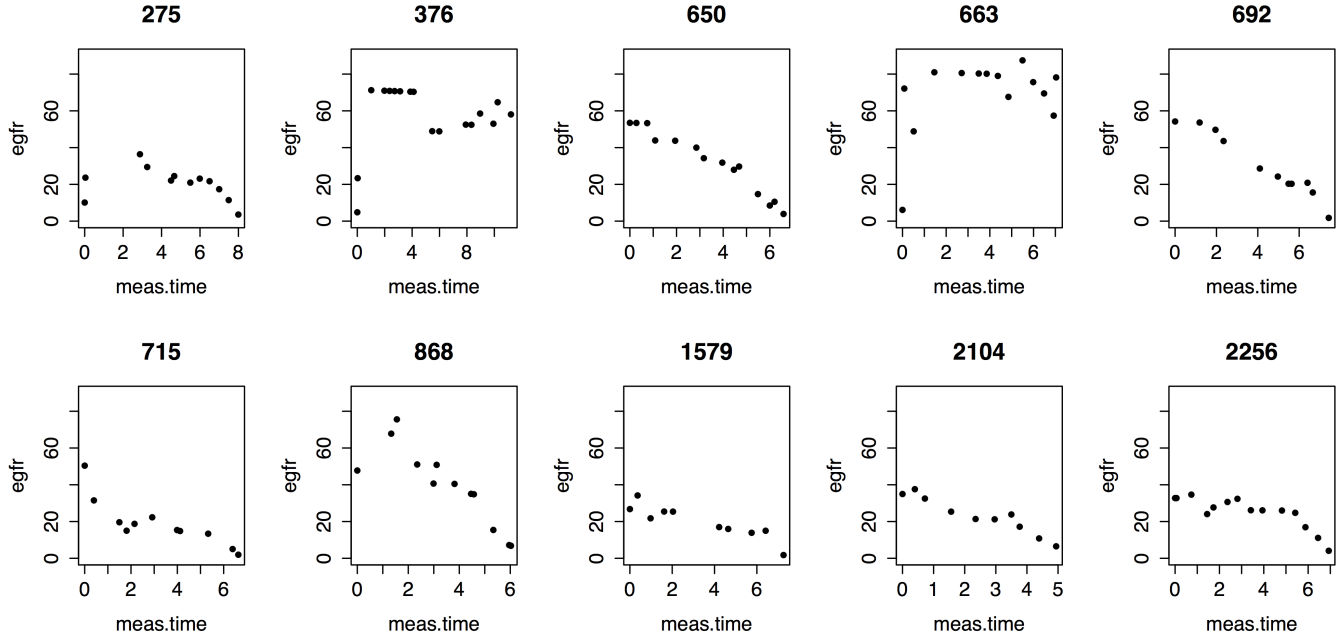
(2) Event data (survival outcomes): Cox model

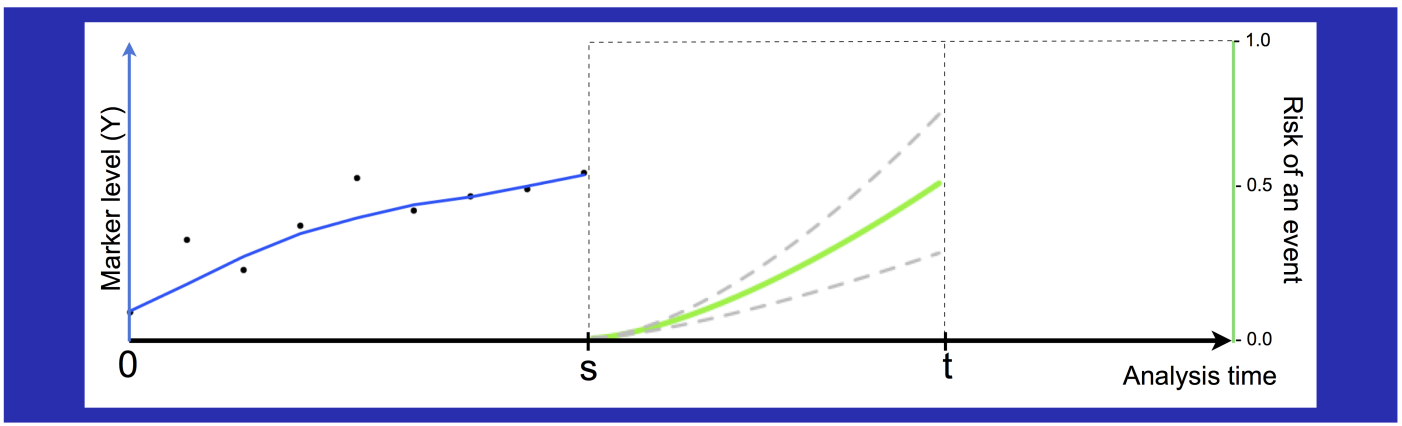
For settings where interest is in using **longitudinal measurements** to predict onset of **adverse outcome**, methods for linking (1) and (2) have been developed:

- Joint Models
- Partly Conditional Models

Motivating Example

- **End Stage Renal Disease (ESRD) Data**
  - **Cohort**: $n = 689$ subjects with severe non-dialysis requiring chronic kidney disease
  - **Longitudinal Marker**: estimated glomular filtration rate (eGFR) obtained approximately every 6 months
  - **Survival Outcome**: transition to ESRD or death (composite)
  - **Goal**: risk prediction to choose aggressive prevention strategies
    High-risk patients targeted for intervention





Objective: Given survival and covariate information up to **now** (time s), predict risk of adverse outcome in a given time frame, i.e. by time t.

Notation

For individual i, $i = 1, ...n$,

- T_i: event time; C_i: censoring time
- Observe $(X_i, \Delta_i) = (min(T_i, C_i), I(T_i \leq C_i))$

Notation

For individual i, $i = 1, \ldots n$,

- T_i: event time; C_i: censoring time
- Observe $(X_i, \Delta_i) = (min(T_i, C_i), I(T_i \leq C_i))$
- $\mathbf{s}_i = \{s_{i1}, \ldots, s_{im_i}; s_{im_i} < X_i\}$: vector of measurement times
- $Y_i(s_{ij})$: j^{th} marker measurement, $j = 1, \ldots, m_i$
  : marker measurement at time s_{ij}
- $\mathbf{Y}_i(u) = \{Y_i(s_{ij}) : 0 \leq s_{ij} \leq u,\ j = 1, \ldots, m_i,\ u < X_i\}$: vector of marker measurements up to time u

Notation

For individual i, $i = 1, ...n$,

- T_i: event time; C_i: censoring time
- Observe $(X_i, \Delta_i) = (min(T_i, C_i), I(T_i \leq C_i))$
- $\mathbf{s}_i = \{s_{i1}, \ldots, s_{im_i}; s_{im_i} < X_i\}$: vector of measurement times
- $Y_i(s_{ij})$: j^{th} marker measurement, $j = 1, \ldots, m_i$
  : marker measurement at time s_{ij}
- $\mathbf{Y}_i(u) = \{Y_i(s_{ij}) : 0 \leq s_{ij} \leq u,\ j = 1, \ldots, m_i,\ u < X_i\}$: vector of marker measurements up to time u
- $\mathbf{Z}_i$: a vector of baseline covariates

Notation

For individual i, $i = 1, \ldots n$,

- T_i: event time; C_i: censoring time
- Observe $(X_i, \Delta_i) = (min(T_i, C_i), I(T_i \leq C_i))$
- $\mathbf{s}_i = \{s_{i1}, \ldots, s_{im_i}; s_{im_i} < X_i\}$: vector of measurement times
- $Y_i(s_{ij})$: j^{th} marker measurement, $j = 1, \ldots, m_i$
  : marker measurement at time s_{ij}
- $\mathbf{Y}_i(u) = \{Y_i(s_{ij}) : 0 \leq s_{ij} \leq u,\ j = 1, \ldots, m_i,\ u < X_i\}$: vector of marker measurements up to time u
- $\mathbf{Z}_i$: a vector of baseline covariates
- s: time at which the prediction is made
- t: time for which the prediction is made
- $\tau_0 = t - s$: time of prediction since the conditioning time

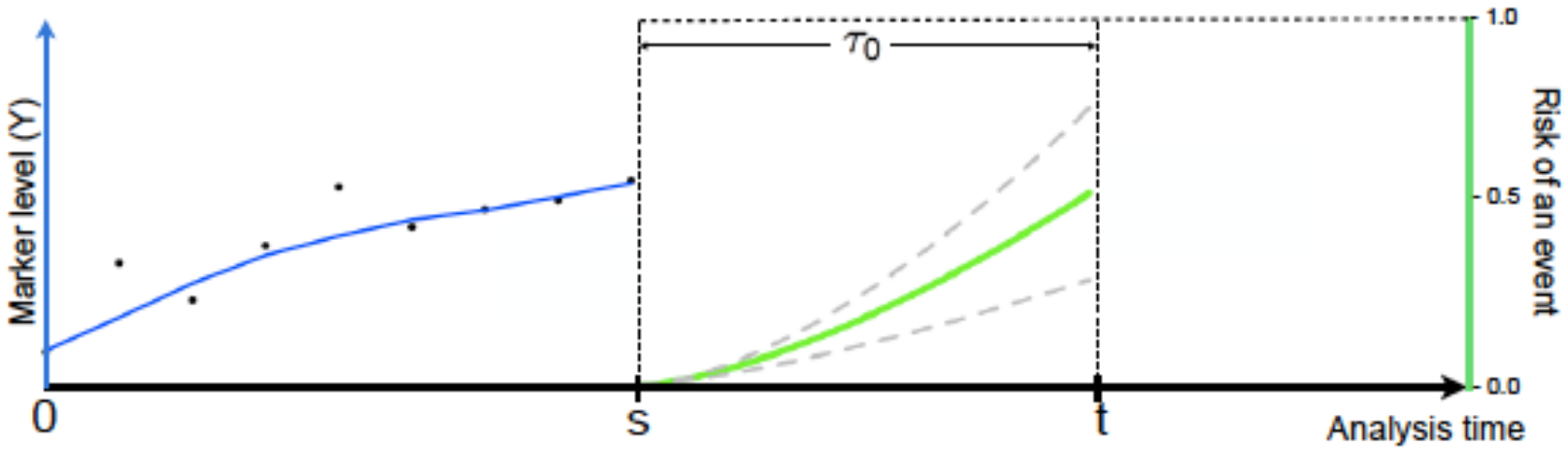
The key quantity in dynamic prediction

Goal: **Conditional risk prediction**

The probability of developing an adverse outcome in the τ_0 time interval from s, given survival up to time s, and covariate information available up to time s

$$R_i\{\tau_0|s, \mathbf{H}(s)\} = \mathrm{P}\{T_i \leq s + \tau_0 | T_i > s, \mathbf{H}_i(s)\}$$

where $\mathbf{H}_i(u) = \{\mathbf{Z}_i, \mathbf{Y}_i(u), \mathbf{s}_i(u)\}$ is the observed history of the covariate process at time $u \geq 0$



Modeling approaches

1. Joint models
2. Partly conditional survival models

Modeling approaches

1. Joint models
2. Partly conditional survival models

Q: Why not use a standard Cox model with time-dependent covariate?

Standard Cox Model

Baseline Measurements

$$\lambda\{t|Y\} = \lambda_0(t)\ exp\{\eta Y\}$$

- $\exp(\eta)$ is the instantaneous hazard ratio or multiplicative increase in the hazard of an event for a one-unit increase in marker Y
- $\int_0^t \lambda\{u|Y\}du = \Lambda(t)$

  $\rightarrow S(t) = exp\{-\Lambda(t)\}$
- Can get survival function, therefore can get predictions

Cox Model with Time-Dependent Covariates

Longitudinal Measurements: Our interest is in using longitudinally measured biomarker to predict onset of adverse outcome

$$\lambda\{t|\mathbf{Y}(\mathbf{t})\} = \lambda_0(t)\ exp\{\eta Y(\mathbf{t})\}$$

- $\int_0^t \lambda\{u|\mathbf{Y}(\mathbf{u})\}du$?
- Integration with unknown future marker path
  $\rightarrow$ Cannot get survival function or predictions
- Need to stop marker somehow to make prediction at time t

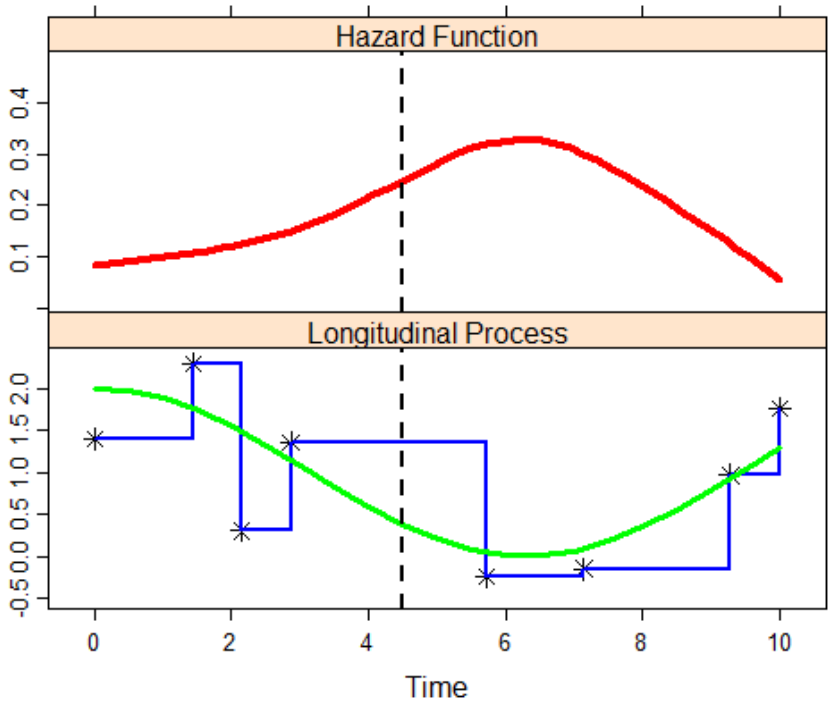
Cox Model with Time-Dependent Covariates

Longitudinal Measurements: Our interest is in using longitudinally measured biomarker to predict onset of adverse outcome

$$\lambda\{t|\mathbf{Y}(\mathbf{t})\} = \lambda_0(t)\ exp\{\eta Y(\mathbf{t})\}$$

- $\int_0^t \lambda\{u|\mathbf{Y}(\mathbf{u})\}du$?
- Integration with unknown future marker path
  $\rightarrow$ Cannot get survival function or predictions
- Need to stop marker somehow to make prediction at time t
- Furthermore, this approach does not handle:
  - Measurement error in biomarker measurements
    - using observed values can lead to attenuated effects
  - Intermittent measurement times
    - missing measurement at time of prediction t

Joint Models

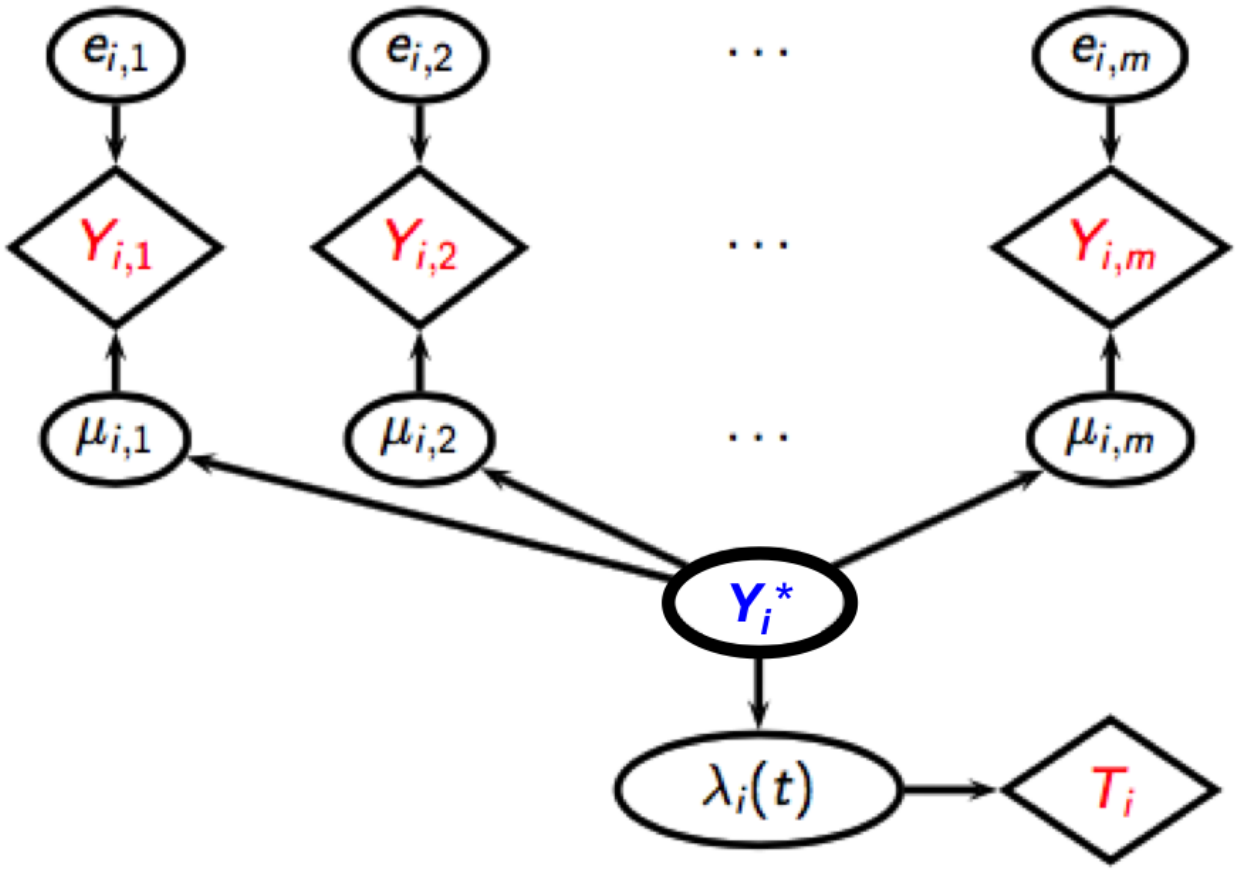


* = observed longitudinal measurements; – = underlying longitudinal process

General Idea: Hazard function at time point t (vertical dashed line) is associated with value of underlying longitudinal process at the same time point

(Rizopoulos 2014, *R-bloggers*)

Joint Models



Assumption: Association between the observed biomarker process and event-time process induced by **shared underlying latent process** $\mathbf{Y}_i^*$, i.e. **shared random effects** between the two processes

Joint Models

- Tsiatis et al. (1995); Faucett and Thomas (1996); Wulfsohn and Tsiatis (1997); Rizopoulos et al. (2011)
- **Two linked sub-models**:
  1. Survival model linking **event time** to underlying "true" biomarker process
  2. Model for recovering underlying **biomarker process** from observed data

Joint Models

Step 1: Time-varying covariate Cox model for **event time process**

$$\lambda_i\{t|Y_i^*(t)\} = \lambda_0(t)\ exp\{\eta Y_i^*(t) + \gamma' \mathbf{Z}_i(t)\},$$

where

- $\mathbf{Y}_i^*$: true unobserved marker value
- $\mathbf{Z}_i$: time-constant covariates

η characterizes the **association between the marker and risk of event**

Note: The functional form of the time-dependent covariate may be made more flexible by replacing $\eta Y_i^*(t)$ with $f\{Y_i^*(t), \eta, \mathbf{b}_i\}$, which specifies components of the longitudinal outcomes process (slope, area under the longitudinal trajectory) that are included in the linear predictor of the survival model (Brown 2009; Rizopoulos & Ghosh 2011, Rizopoulos 2012; Taylor et al. 2013; Rizopoulos et al. 2014; Rizopoulos et al. 2016)

Joint Models

Step 2: Linear mixed effects model for **biomarker process**

Letting $\mathbf{Y}_i$ represent the observed marker values and $\mathbf{Y}_i^*(s_{ij})$ represent the history of the hypothetical true longitudinal process (without measurement error) up to time s_{ij} for subject i,

$$\begin{aligned} Y_i(s_{ij}) &= Y_i^*(s_{ij}) + \epsilon_i(s_{ij}) \\ &= \mathbf{U_i}(s_{ij})\boldsymbol{\beta} + \mathbf{W_i}(s_{ij})\mathbf{b_i} + \epsilon_i(s_{ij}) \end{aligned}$$

where

- $\boldsymbol{\beta}$: Fixed effects vector
- $\mathbf{b_i}$: Random effects vector
- $\mathbf{U_i}$ and $\mathbf{W_i}$: Covariate matrices

Fixed effects: average longitudinal trajectory in time
Random effects: how each individual deviates from average trajectory

Standard assumptions: $\mathbf{b_i} \sim \mathcal{N}$, $\epsilon_i(t) \sim \mathcal{N}$

Joint Models

Linked Models: $[\text{vec}(Y_i),\ T_i]$

longitudinal $\quad \mathbf{Y_i}(t)|\, Y_i^*(t) \quad = Y_i^*(t) + \epsilon_i(t)$

survival $\quad \lambda_i\{t|\, Y_i^*(t)\} \quad = \lambda_0(t)\ exp\{\eta Y_i^*(t) + \gamma' \mathbf{Z_i}\}$

$$\mathbf{Y}_i^* \sim \mathcal{N} \qquad \epsilon_i \sim \mathcal{N}$$

Estimation of model based on joint distribution of the two outcomes. Since both model specifications involve unknown quantities, fitting involves **iteration between longitudinal and survival submodels**.

- **Goal**: Estimation of η, conditional risk prediction
- **Estimation**: (i) likelihood; (ii) Bayesian approach (Rizopoulos 2011)
- **Issues addressed**: measurement error; intermittent observation

Joint Models

Risk Prediction: For a future individual with $\mathbf{H}_o(s) = \{\mathbf{Z}_o, \mathbf{Y}_o(s), \mathbf{s}_o(s)\}$ and event-free at time s,

$$\begin{aligned} R^{JM}(\tau_0 \mid s, \mathbf{H}_o(s)) &= P(T_o \leq s + \tau_0 \mid T_o > s, \mathbf{H}(s) = \mathbf{H}_o(s), \mathcal{D}_n, \theta) \\ &= \text{conditional risk (recall: slide 75)} \end{aligned}$$

where

- $\mathcal{D}_n = X_i, \Delta_i, \mathbf{H}_i, i = 1, ..., n$ represents the full data used to fit JM
- θ = parameter vector of the joint model

Rizopoulos (2011):

$$\begin{aligned} R^{JM}(\tau_0 \mid s, \mathbf{H}_o(s)) &= 1 - P(T_o > s + \tau_0 \mid T_o > s, \mathbf{H}(s) = \mathbf{H}_o(s), \mathcal{D}_n, \theta) \\ &= 1 - \text{conditional survival} \\ &= 1 - \int \frac{S\{s + \tau_0 | \mathbf{Y}_o^*(s + \tau_0), \mathbf{Z}_o, \theta\}}{S\{s | \mathbf{Y}_o^*(s), \mathbf{Z}_o, \theta\}} p(\mathbf{b} | T_o > s, \mathbf{H}_o(s), \theta) d\mathbf{b} \end{aligned}$$

where

- $S(t) = exp\Big\{ \int_0^t \lambda(u | \mathbf{Y}_o^*(u), Z_o, \boldsymbol{\theta}) du \Big\}$

Joint Models

Risk Prediction: Joint model fitted to available data. Then, for a future individual, first-order **estimate**:

$$\widehat{R}_i^{JM}(\tau_0 \mid s) = 1 - \frac{\widehat{S}_i(s+\tau_0 | \mathbf{Y}_i^*(s+\tau_0, \widehat{b}_i, \mathbf{Z}_i, \widehat{\boldsymbol{\theta}}_{JM}), \widehat{\boldsymbol{\theta}}_{JM})}{\widehat{S}_i(s | \mathbf{Y}_i^*(s, \widehat{b}_i, \mathbf{Z}_i, \widehat{\boldsymbol{\theta}}_{JM}), \widehat{\boldsymbol{\theta}}_{JM})} + O\left(\frac{1}{m_i}\right)$$

where

- $\widehat{b}_i$: empirical Bayes estimate of b_i
- $\widehat{\boldsymbol{\theta}}_{JM}$: vector of the maximum likelihood estimates of the joint model
- $\widehat{S}(t) = exp\left\{ \int_0^t \widehat{\lambda}(u | \mathbf{Y}_o^*(u), Z_o, \widehat{\boldsymbol{\theta}}_{JM}) du \right\}$

(Rizopoulos, 2011)

Joint Models

Risk Prediction:

- **Baseline hazard function** must be specified **parametrically** - use splines for flexible models
- **Point and interval estimates** obtained using **Monte Carlo simulations** (Proust-Lima & Taylor 2009; Rizopoulos 2011)
- For some large number of simulations S (e.g. S=500),
  - Point estimate $\widehat{R}_i^{JM}(\tau_0 \mid s)$: Median over S Monte Carlo samples
  - 95% CI: (2.5^{th} percentile, 97.5^{th} percentile)
- Prediction for an individual requires **complex computation**

Joint Models

Software: Rizopoulos (2010)

- R package `JM`
- Examples at `http://jmr.r-forge.r-project.org`

Joint Models

Potential Issues:

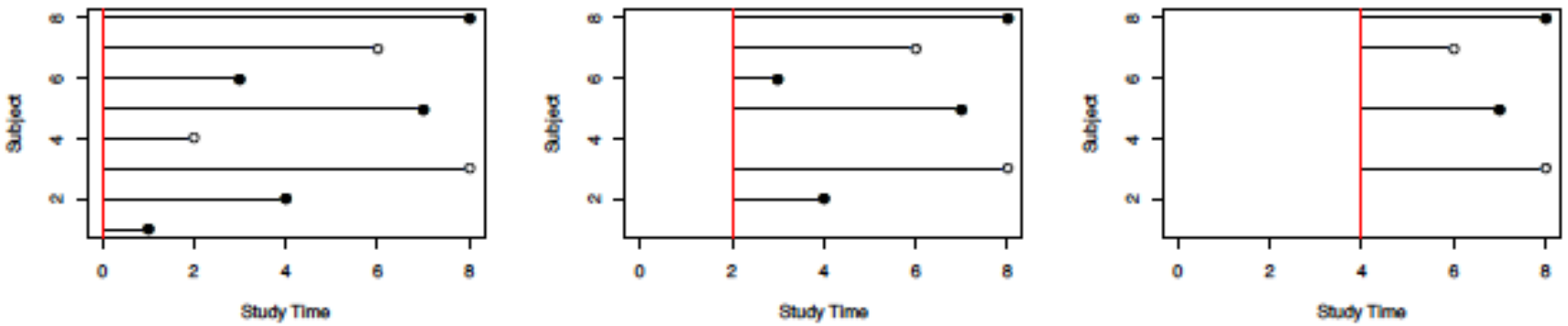
- Highly **parametric**
- Predictions can be sensitive to **assumption of latent process model**, which cannot be easily checked
- Prediction for an individual using Monte Carlo simulations is **computationally intensive**

Partly Conditional (PC) Models

Zheng & Heagerty (2005); Maziarz et al. (2017)

General Idea:

- Related to landmark approach (van Houwelingen & Putter (2012))
- Condition on survival up to some time s

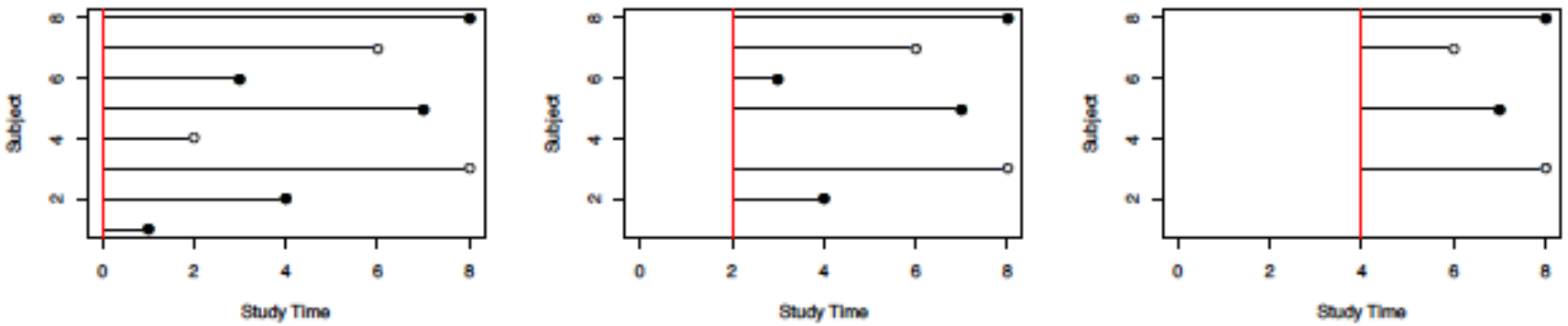


Partly Conditional (PC) Models

Zheng & Heagerty (2005); Maziarz et al. (2017)

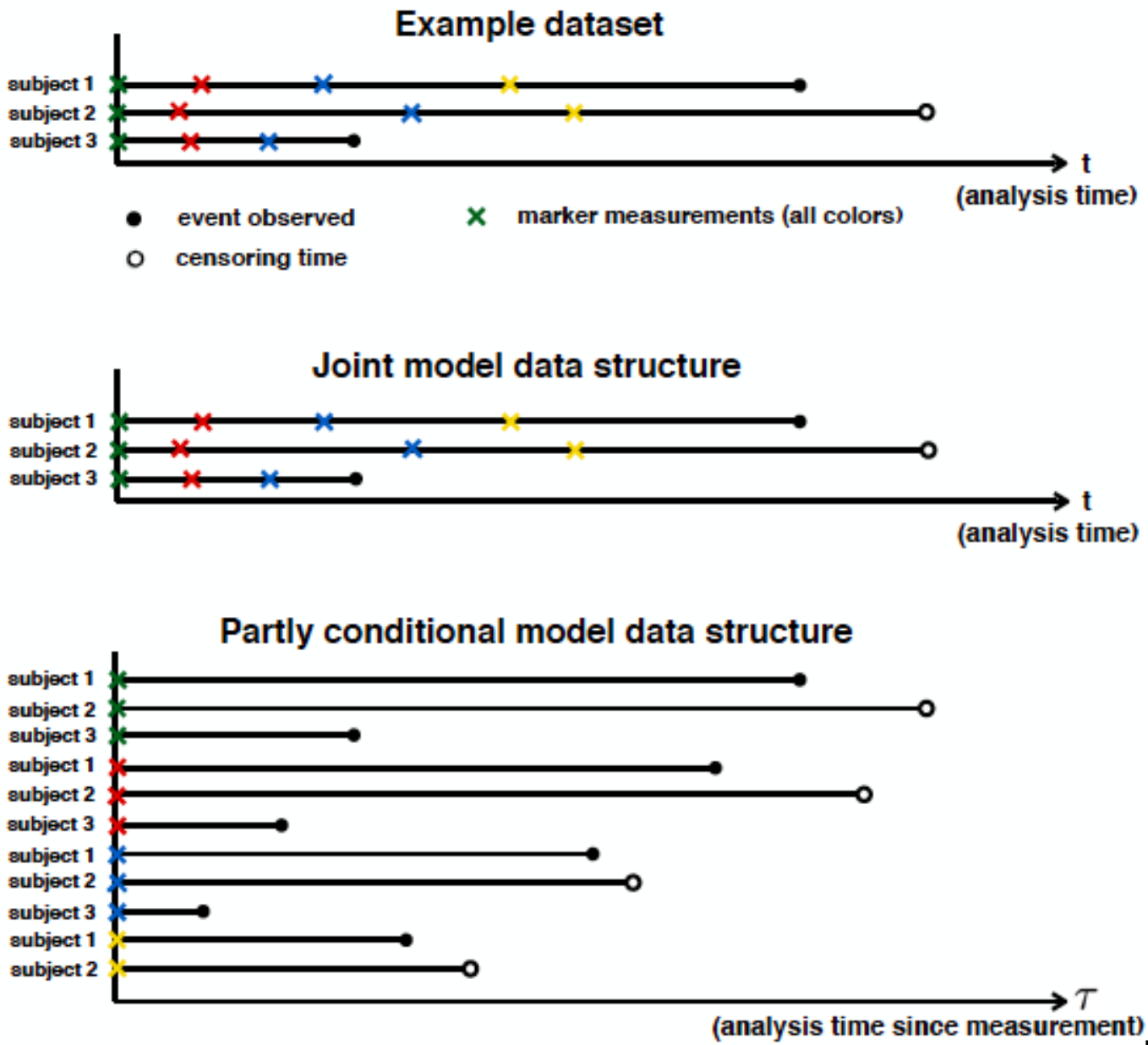
General Idea:

- Related to landmark approach (van Houwelingen & Putter (2012))
- Condition on survival up to some time s



- **Treat time s as new baseline**
- Predict residual life time from s: $\mathcal{T}_s = T - s | \mathcal{T}_s > 0$, using observed covariate history up to s, $\mathbf{H}(s)$

Joint Models vs Partly Conditional Models



Partly Conditional Models

Estimation:

- Semi-parametric
- PC model specifies relationship between $\mathcal{T}_{ij} = T_i - s_{ij}$ and $\mathbf{H}_i(s_{ij})$ **without having to specify full marker process**
- Two estimation approaches for survival outcomes:
  (a) Partly conditional Cox-type model ($\mathbf{PC_{Cox}}$)
  (Zheng & Heagerty, 2005)
  (b) Novel two-stage PC model ($\mathbf{PC_{Cox}}$ **BLUP**)
  (Maziarz et al, 2017)

(a) $\mathbf{PC}_{\mathbf{Cox}}$

PC models relate the covariate history **up to time** s to the residual survival time, or time since measurement, τ.

$$\begin{aligned}\lambda(\tau|\mathbf{H}_i(s_{ij})) &= \lambda_0(\tau)\exp\big(\boldsymbol{\alpha}'\mathbf{B}(s_{ij}) + \boldsymbol{\gamma}'\mathbf{Z}_i + \eta'\mathbf{B}(\tau)h(\mathbf{Y}_i(s_{ij}))\big) \\ &= \lambda_0(\tau)\exp(\boldsymbol{\theta}'_{\text{Cox}}\mathbf{H}_i^B(s_{ij},\tau))\end{aligned}$$

where

- $\lambda_0(\cdot)$ is an unknown baseline hazard
- $h(\mathbf{Y}_i(s_{ij}))$ is a functional of $\mathbf{Y}_i(s_{ij})$ - e.g. last observed value $Y_i(s_{ij})$
- $\boldsymbol{\theta}_{\text{Cox}} = [\boldsymbol{\alpha}', \boldsymbol{\gamma}', \eta']'_{P\times 1}$
- $\mathbf{H}_i^B(s_{ij},\tau) = [\mathbf{B}(s_{ij})', \mathbf{Z}_i', \mathbf{B}(\tau)h(\mathbf{Y}_i(s_{ij}))']'$
- $\mathbf{B}(\cdot)$ is a spline basis function of measurement time

(a) $\mathbf{PC}_{\mathbf{Cox}}$

Risk Prediction: For a future individual with covariate history up to time s, $\mathbf{H}_\circ(s) = \{\mathbf{s}_\circ(s), \mathbf{Z}_\circ, \mathbf{Y}_\circ(s)\}$ and survival up to time s, the probability of an event within time τ_0 from s can be estimated as

$$\begin{aligned}\widehat{R}^{\mathrm{PC}}_{\mathrm{Cox}}(\tau_0 \mid s, \mathbf{H}_\circ(s), \widehat{\boldsymbol{\theta}}_{\mathrm{Cox}}) &= \widehat{P}(T_o \le s+\tau_0 \mid \mathbf{H}(s) = \mathbf{H}_o(s), T_o > s, \widehat{\boldsymbol{\theta}}_{\mathrm{Cox}}) \\ &= 1 - \exp(-\widehat{\Lambda}(\tau_0 \mid s, \mathbf{H}_\circ(s), \widehat{\boldsymbol{\theta}}_{\mathrm{Cox}}))\end{aligned}$$

where

$$\widehat{\Lambda}(\tau_0|\mathbf{H}_\circ(s), \widehat{\boldsymbol{\theta}}_{\mathrm{Cox}}) = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{\tau_0} \frac{\exp(\widehat{\boldsymbol{\theta}}'_{\mathrm{Cox}}\mathbf{H}^B_\circ(s_{ij},u))}{\sum_{k=1}^{n}\sum_{l=1}^{m_k} I(X_{kl}\ge X_{ij})\exp(\widehat{\boldsymbol{\theta}}'_{\mathrm{Cox}}\mathbf{H}^B_k(s_{kl},u))}\, dN_{ij}(u)$$

(b) PC_{Cox} BLUP

- PC_{Cox} does not account for measurement error
- Maziarz et al. (2017) extended PC_{Cox} and proposed a two-stage estimator that:
  (1) Models the longitudinal process and calculates a fitted/smoothed $\widehat{Y}_i(s_{ij})$ based on the *best linear unbiased predictor* (**BLUP**) estimator
  (2) Obtains BLUP-based estimators of risk model parameters

(b) PC_{Cox} BLUP

Step 1: Best linear unbiased predictor (BLUP) smoothing of **longitudinal process**

Model for $\mathbf{Y}_i$:

$$\mathbf{Y}_i = \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i + \epsilon_i, i = 1, ..., n$$

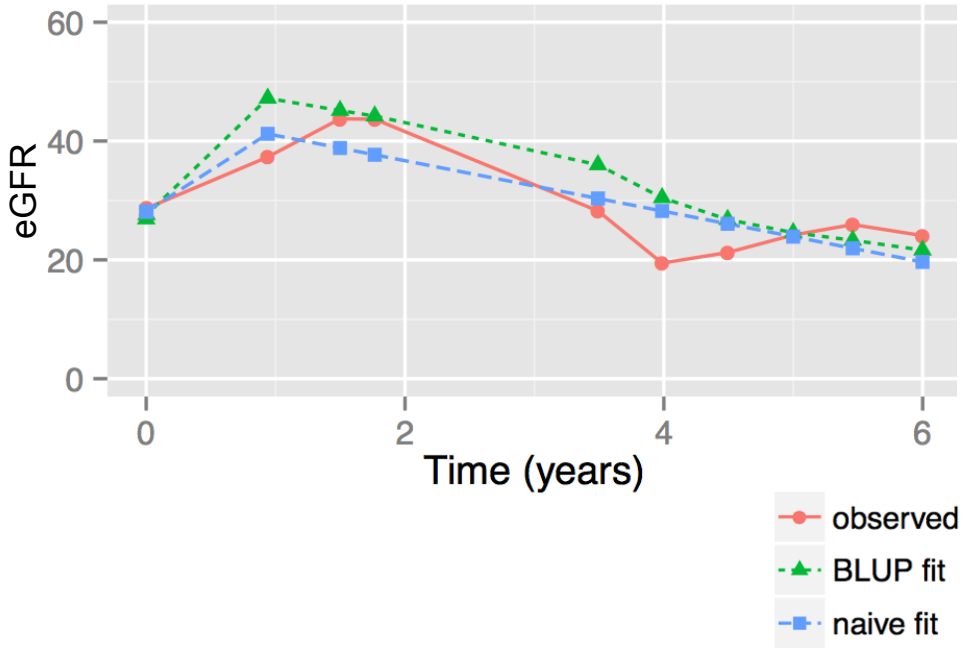
where

- $\epsilon_i \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}), \mathbf{b}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{D}(\phi)), \mathbf{Y}_i \sim \mathcal{N}(\mathbf{X}_i\beta, \mathbf{\Sigma}_i)$
- $\mathbf{U}_i$ and $\mathbf{W}_i$: Covariate matrices
- $\phi = (\phi_1, ..., \phi_q)^T$ is the parameter vector of the variance components of the random effects
- $\mathbf{\Sigma}_i = \sigma^2\mathbf{I} + \mathbf{W}_i\mathbf{D}(\phi)\mathbf{W}_i^T$

Then the BLUP estimator is:

$$\widehat{\mathbf{Y}}_i = \mathbf{U}_i\widehat{\beta} + \mathbf{W}_i\mathbf{D}(\widehat{\phi})\mathbf{W}_i^T(\mathbf{\Sigma}_i)^{-1}(\mathbf{Y}_i - \mathbf{U}_i\widehat{\beta})$$

(b) PC_{Cox} BLUP



- **BLUP fit**: Uses **only past information** to obtain BLUP values at a given observation time, provides smoothing to individual marker data by *shrinking* individual marker trajectory toward population-averaged trajectory
- **Naïve fit**: Linear mixed effects (LME) model uses **past and future information** of new individual to obtain fitted marker trajectory

(b) PC_{Cox} BLUP

Step 2: Cox model for **event time process** (same as PC_{Cox})

$$\begin{aligned}\lambda(\tau|\mathbf{H}_i(s_{ij})) &= \lambda_0(\tau)\exp\left(\boldsymbol{\alpha}'\mathbf{B}(s_{ij}) + \boldsymbol{\gamma}'\mathbf{Z}_i + \eta'\mathbf{B}(\tau)h(\mathbf{Y}_i(s_{ij}))\right) \\ &= \lambda_0(\tau)\exp(\boldsymbol{\theta}'_{\text{Cox}}\mathbf{H}_i^B(s_{ij},\tau)),\end{aligned}$$

where

- $\lambda_0(\cdot)$ is an unknown baseline hazard
- $h(\mathbf{Y}_i(s_{ij}))$ is a functional of $\mathbf{Y}_i(s_{ij})$ - **BLUP estimator** $\widehat{Y}_i(s_{ij})$
- $\boldsymbol{\theta}_{\text{Cox}} = [\boldsymbol{\alpha}', \boldsymbol{\gamma}', \eta']'_{P\times 1}$
- $\mathbf{H}_i^B(s_{ij},\tau) = [\mathbf{B}(s_{ij})', \mathbf{Z}_i', \mathbf{B}(\tau)h(\mathbf{Y}_i(s_{ij}))']'$
- $\mathbf{B}(\cdot)$ is a spline basis function of measurement time

Obtain BLUP-based estimators of risk model parameters $\widehat{\boldsymbol{\theta}}_{\text{Cox}}^{\text{BLUP}}$

Note: The two-stage approach described here uses a partly conditional model, different from JM two-stage approach which uses a time-varying covariate survival model.

(b) PC_{Cox} BLUP

Prediction: For a future individual with $\mathbf{H}_o(s) = \{\mathbf{Z}_o, \mathbf{Y}_o(s), \mathbf{s}_o(s)\}$ and event-free at time s,

1. The predicted random effect is

$$\widehat{\mathbf{b}}_o|s_{oj} = \mathbf{D}(\widehat{\phi})\mathbf{W}'_o(s_{oj})(\widehat{\sigma}^2\mathbf{I} + \mathbf{W}_o(s_{oj})\mathbf{D}(\widehat{\phi})\mathbf{W}'_o(s_{oj}))^{-1}(\mathbf{Y}_o(s_{oj}) - \mathbf{U}_o(s_{oj})\widehat{\boldsymbol{\beta}}),$$

2. The fitted marker value based on covariate data only up to time s_{oj} is

$$\widehat{Y}_o(s_{oj}) = \mathbf{U}_o(s_{oj})\widehat{\boldsymbol{\beta}} + \mathbf{W}_o(s_{oj})(\widehat{\mathbf{b}}_o|s_{oj}).$$

Iterating through each marker measurement of each subject, one can obtain vectors of BLUP fitted marker values for each subject, $\widehat{\mathbf{Y}}_o$

Note: $\widehat{\mathbf{Y}}_o$ is not the same as $\mathbf{Y}^*_o$, obtained for a joint model.

Then, the predicted risk for a new subject using PC_{Cox} BLUP:

$$\widehat{R}^{\text{BLUP}}_{\text{Cox}}(\tau_0 \mid s, \mathbf{H}_o(s), \widehat{\boldsymbol{\theta}}^{\text{BLUP}}_{\text{Cox}}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Phi}}) = 1 - \exp(-\widehat{\Lambda}(\tau_0 \mid s, \widehat{\mathbf{H}}_o(s, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Phi}}), \widehat{\boldsymbol{\theta}}^{\text{BLUP}}_{\text{Cox}}))$$

where

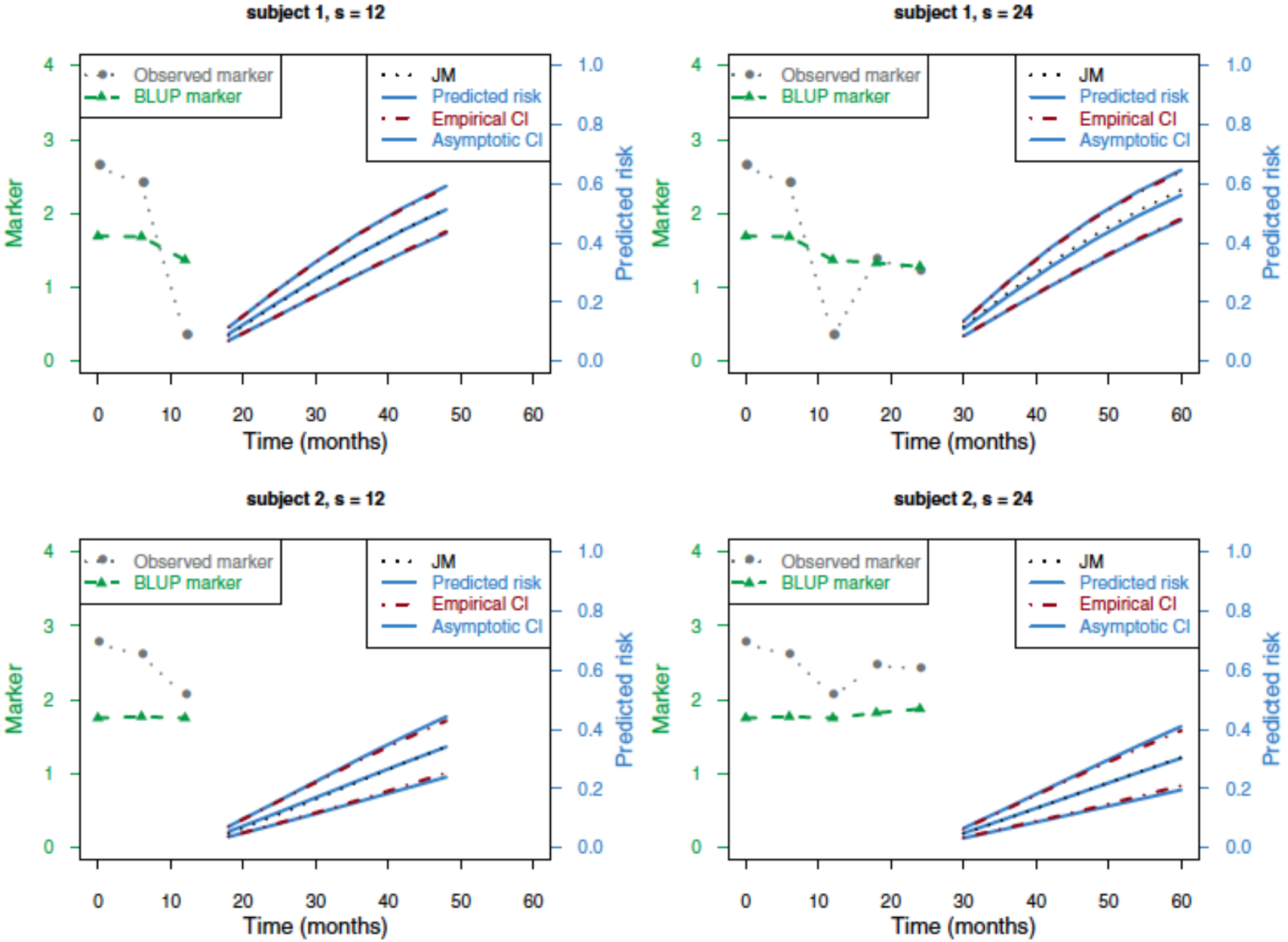
- $\widehat{\mathbf{H}}_o(s, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Phi}}) = \{\mathbf{s}_o(s), \mathbf{Z}_o, \widehat{\mathbf{Y}}^{BLUP}_o(s)\}$
- $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\Phi}}$ are vectors of parameter estimates of the fixed effects and the variance components, respectively

Partly Conditional Model

Software: Available as supplementary web-based material for:

Maziarz M, Heagerty PJ, Cai T, Zheng Y (2017). On Longitudinal Prediction with Time-to-Event Outcomes: Comparison of Modeling Options. *Biometrics*, 73(1): 83-93.

JM vs PC: Simulation of Individual Risk Prediction



Simulation

Goal: To compare the **calibration** and **discrimination** performance of the different modeling approaches: PC_{Cox}, PC_{Cox} BLUP, JM

- **Calibration / Overall performance**: Prediction error (PE) or Brier Score (Brier, 1950) extended to survival outcomes (Schoop et al., 2008)

$$\text{PE} = \text{E}\{\text{I}(s < T_i \leq s + \tau_0) - R(\tau_0|s, \mathbf{H}(s))^2\}$$

  $\rightarrow$ Distance between observed and predicted outcomes

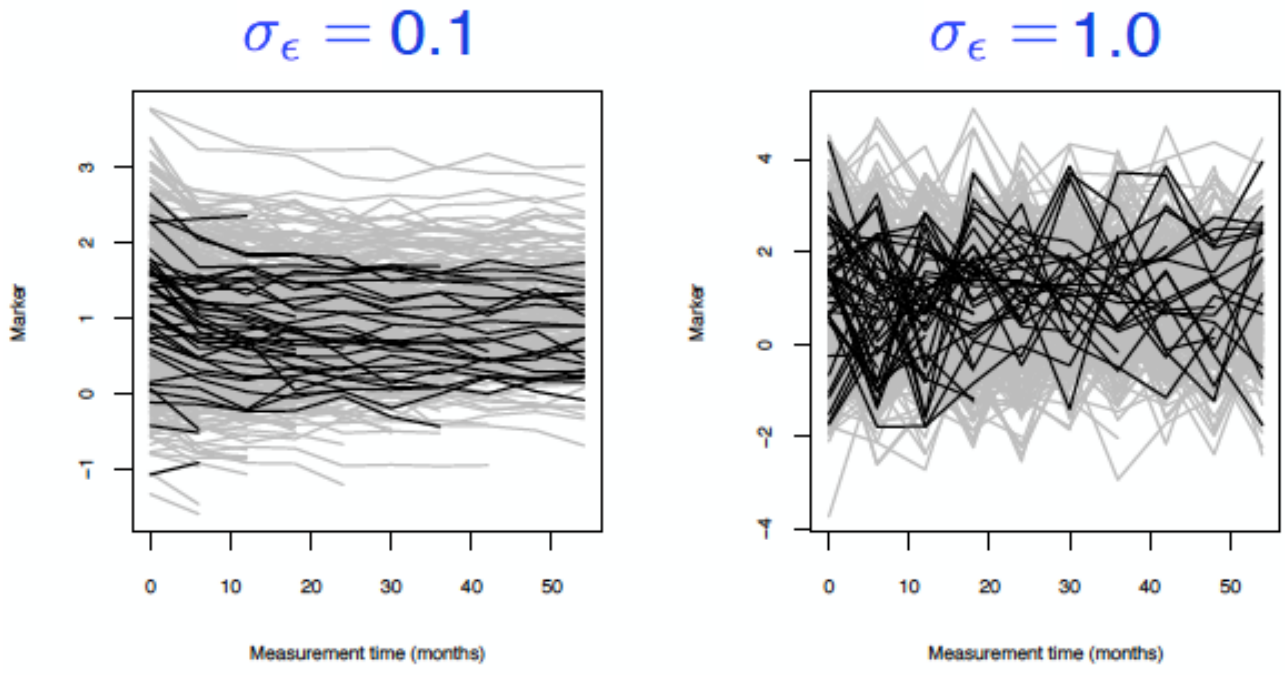
- **Discrimination accuracy**: $\text{AUC}^{\text{C/D}}$ based on TP_t^{C} and FP_t^{D} estimated over full range of risk thresholds $c \in (0, 1)$:

$$\begin{aligned} \text{TP}^{\text{C}}_{s,\tau_0}(c) &= P\{R_i(\tau_0 \mid s) \geq c \mid s < T_i \leq s + \tau_0\} \\ \text{FP}^{\text{D}}_{s,\tau_0}(c) &= P\{R_i(\tau_0 \mid s) \geq c \mid T_i > s + \tau_0\} \end{aligned}$$

  $\rightarrow$ As covered in Part 1

Simulation

Setup



- Measurement error: $\sigma_\epsilon = 0.1$ and 1.0
- (s, t) = (24, 36), (48, 60), (24, 48), (48, 72) in months
- Further details of simulation setup in manuscript (Maziarz et al., 2017)

Simulation Results (500 replications, n = 500)

$\sigma_e = 0.1$

	$\tau_0 = 12$		$\tau_0 = 24$	
	s = 24	s = 48	s = 24	s = 48
	EST (ESD)	EST (ESD)	EST (ESD)	EST (ESD)
PC_{Cox}				
PE	**0.113** (0.014)	0.101 (0.018)	0.165 (0.013)	**0.155** (0.015)
AUC	0.761 (0.032)	**0.740** (0.051)	0.774 (0.031)	0.759 (0.038)
PC_{Cox}*BLUP*				
PE	**0.113** (0.014)	0.101 (0.018)	0.164 (0.013)	**0.154** (0.015)
AUC	0.763 (0.032)	**0.742** (0.048)	0.775 (0.032)	0.762 (0.038)
JM				
PE	**0.111** (0.012)	0.097 (0.015)	0.162 (0.012)	**0.156** (0.016)
AUC	0.773 (0.036)	**0.754** (0.051)	0.786 (0.028)	0.761 (0.037)

EST = Estimate, ESD = empirical standard error

Models are comparable for small measurement error.

Simulation Results (500 replications, n = 500)

$\sigma_e = 1.0$

	$\tau_0 = 12$		$\tau_0 = 24$	
	s = 24	s = 48	s = 24	s = 48
	EST (ESD)	EST (ESD)	EST (ESD)	EST (ESD)
PC_{Cox}				
PE	**0.126** (0.015)	0.110 (0.019)	0.194 (0.014)	**0.178** (0.015)
AUC	0.646 (0.044)	**0.632** (0.061)	0.655 (0.034)	**0.639** (0.045)
PC_{Cox} *BLUP*				
PE	**0.128** (0.015)	0.108 (0.017)	0.196 (0.015)	**0.168** (0.015)
AUC	0.696 (0.043)	**0.701** (0.054)	0.702 (0.033)	**0.719** (0.039)
JM				
PE	**0.118** (0.012)	0.102 (0.015)	0.177 (0.012)	**0.166** (0.015)
AUC	0.730 (0.039)	**0.719** (0.055)	0.732 (0.029)	**0.723** (0.041)

EST = Estimate, ESD = empirical standard error

PC_{Cox} BLUP provides improvement over PC_{Cox} when nonsystematic error is present

Illustration: ESRD Data

- **End Stage Renal Disease (ESRDS) Data**
  - **Cohort**: $n = 689$ subjects with severe non-dialysis requiring chronic kidney disease
  - **Longitudinal Marker**: estimated glomular filtration rate (eGFR) obtained approximately every 6 months
  - **Survival Outcome**: transition to ESRD or death (composite)
  - **Goal**: risk prediction to choose aggressive prevention strategies
    High-risk patients targeted for intervention

Illustration: observed biomarker trajectories

Individual observed eGFR trajectories stratified by age and eGFR at baseline

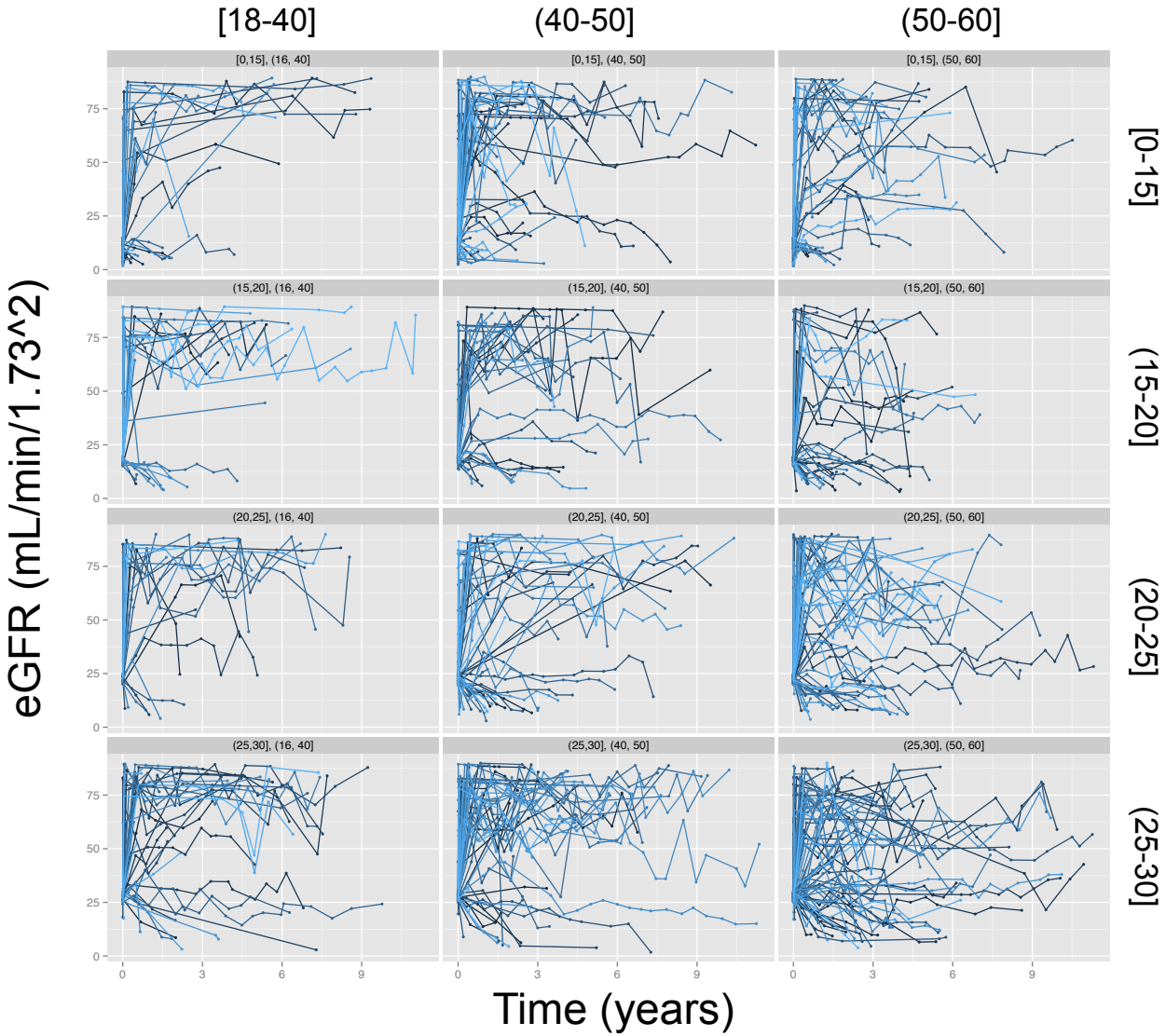


Illustration: modeled biomarker trajectories

Individual fitted eGFR trajectories stratified by age and eGFR at baseline

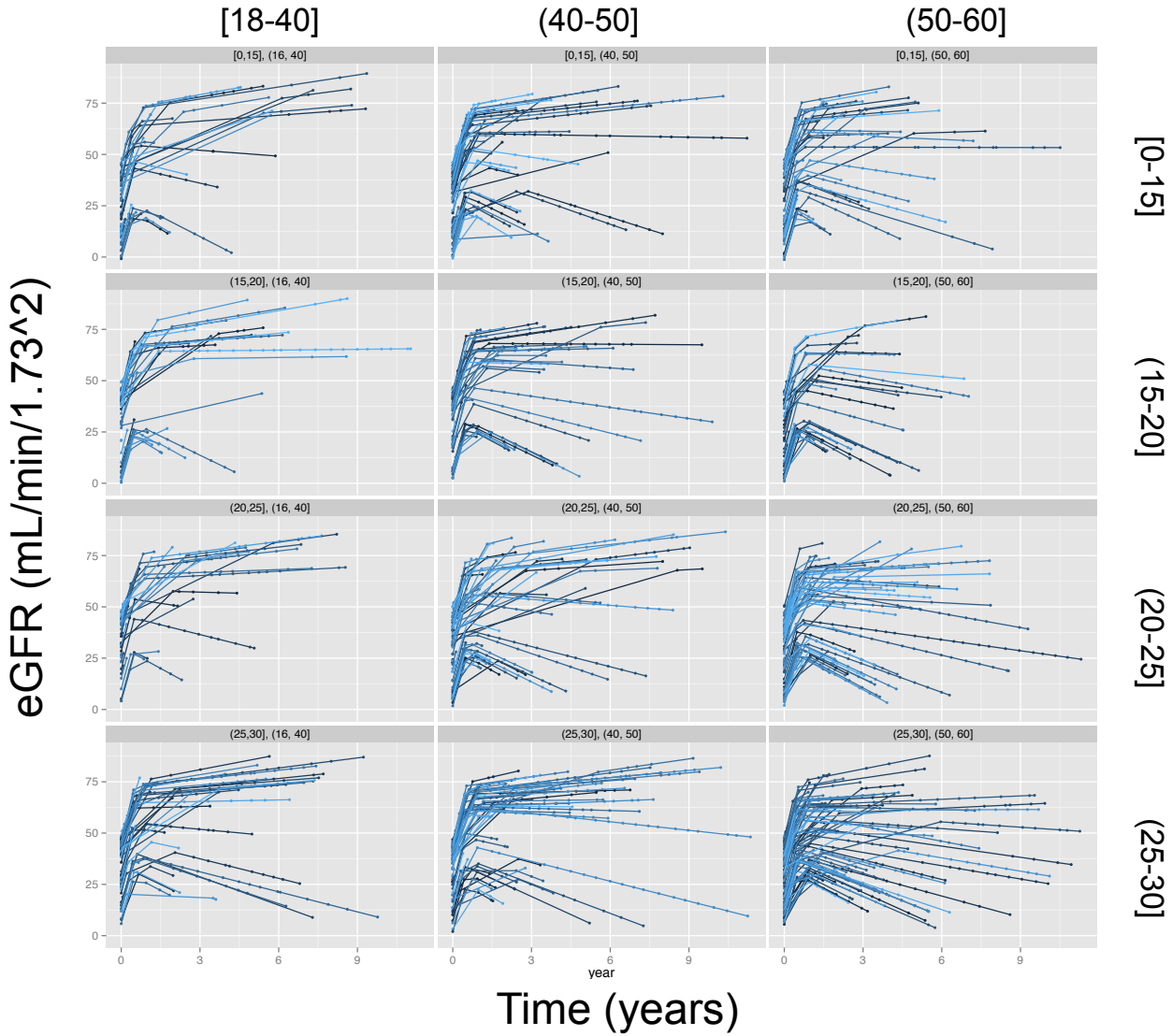
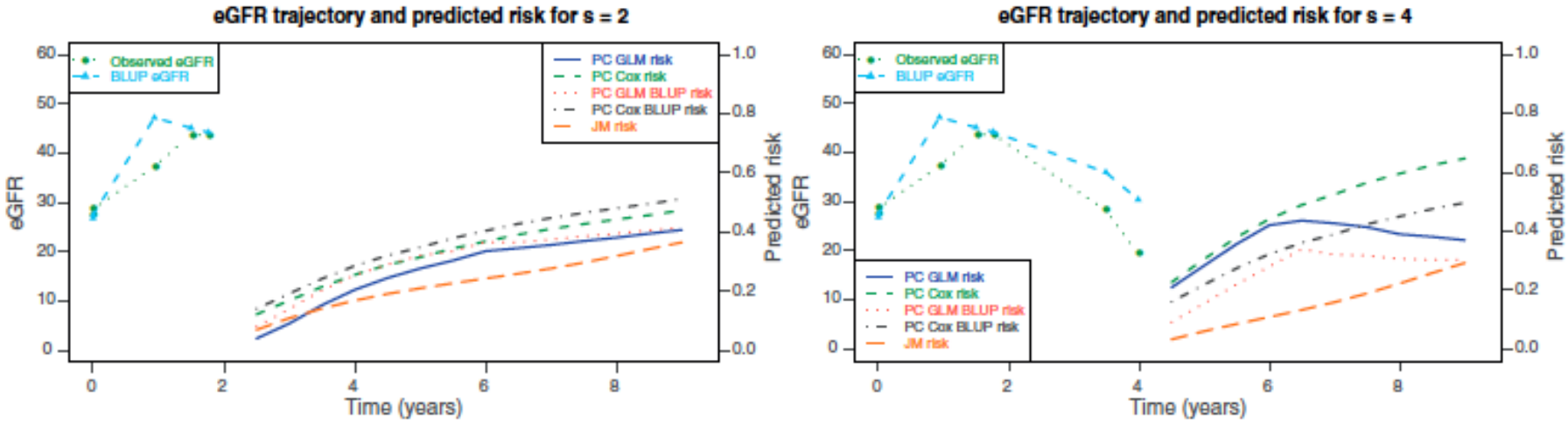


Illustration: individual risk prediction



..... Observed eGFR

– – – BLUP eGFR

– – – PC_{Cox} risk

–.– PC_{Cox} BLUP risk

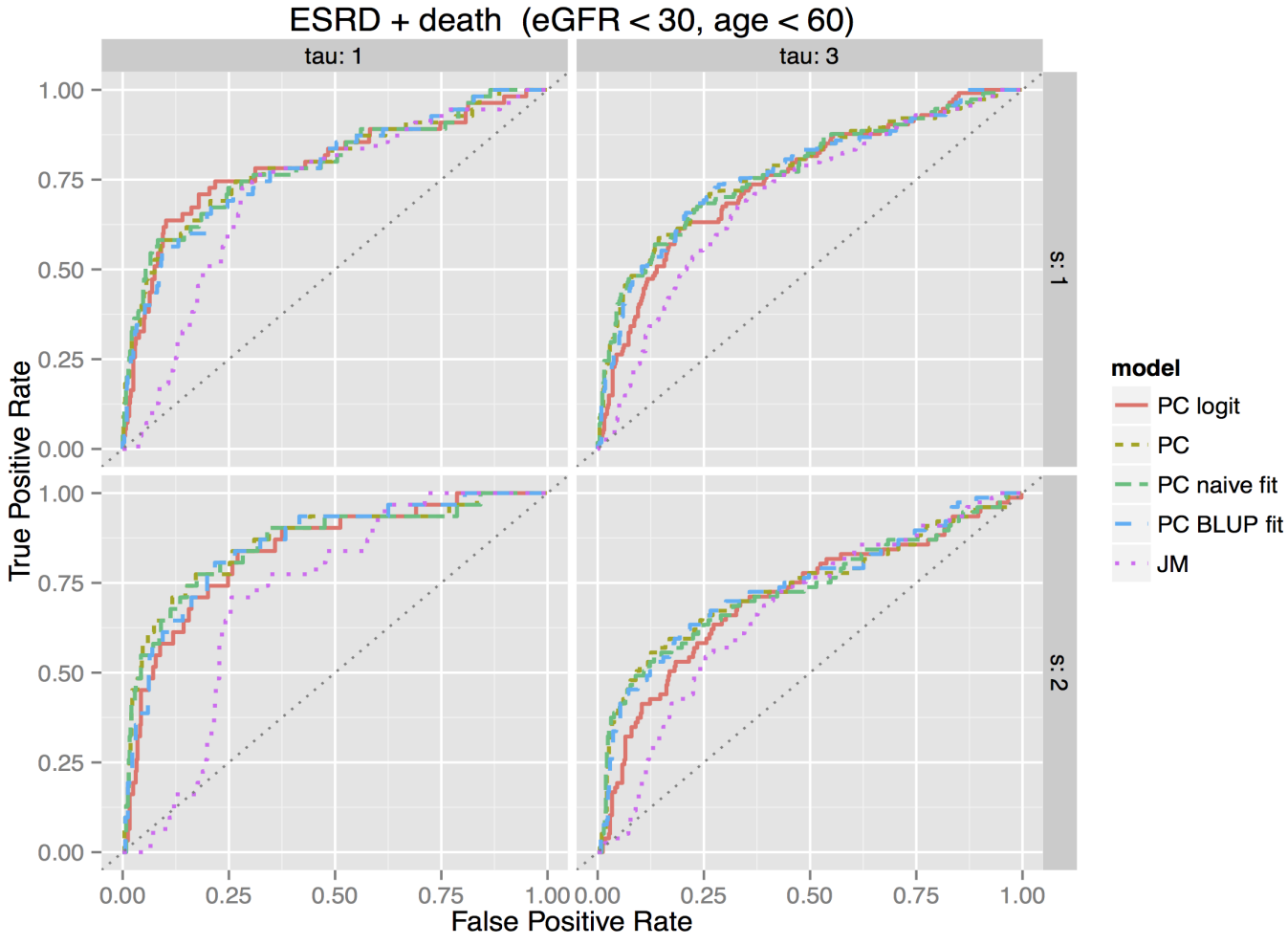
– – – JM risk

Illustration: ESRD Data Analysis Results

	τ_0 = 1 year		τ_0 = 3 years	
	s = 1 years	s = 2 years	s = 1 years	s = 2 years
	n_e/n=55/574	n_e/n=31/519	n_e/n=114/574	n_e/n=77/519
	EST (ESD)	EST (ESD)	EST (ESD)	EST (ESD)
PE	**0.075** (0.009)	0.053 (0.011)	0.132 (0.024)	0.130 (0.032)
AUC	0.791 (0.033)	**0.861** (0.041)	0.772 (0.024)	**0.735** (0.029)
$PC_{Cox}BLUP$				
PE	**0.079** (0.007)	0.056 (0.010)	0.139 (0.020)	0.131 (0.027)
AUC	0.782 (0.039)	**0.852** (0.047)	0.771 (0.028)	**0.738** (0.032)
JM				
PE	**0.087** (0.008)	0.073 (0.014)	0.154 (0.025)	0.149 (0.027)
AUC	0.714 (0.033)	**0.712** (0.038)	0.702 (0.023)	**0.671** (0.025)

EST = Estimate, ESD = empirical standard error

Illustration: ESRD Data Analysis Results



Summary: Partly Conditional Models vs Joint Models

PC models

- provide a flexible, robust and practical alternative to JM for dynamic prediction (Simulation computation time: 6 hours for PC versus 20 hours for JM)
- require no modeling assumptions for the longitudinal biomarker trajectory
- are relatively simple to implement, easy to modify and extend. Can easily be scaled to include multiple biomarkers, by simply including their BLUP fits in the Cox model. JM would get analytically and computationally complex for multiple biomarkers.

Based on simulations,

- PC_{Cox} BLUP performs comparably to JM
- PC_{Cox} BLUP provides improvement over PC_{Cox} when nonsystematic error is present and marker trajectories can be modeled well
- PC_{Cox} outperforms PC_{Cox} BLUP and JM when marker trajectory is complex and is difficult to model well with a linear mixed effects model

Some Future Considerations

- Multiple longitudinal markers
- Competing risks / cause-specific transitions
- Development targeted at performance measures
    - C-index
    - Population yield