Personalized Biopsy Schedules using Joint Models

Anirudh Tomer, Dimitris Rizopoulos, Ewout Steyerberg, Monique Roobol

October 25, 2017

Active Surveillance

- To avoid over-treatment men with low grade cancers are advised to enroll AS programs

- Decision to start active treatment is based on biopsies
  - reliable
  - painful $\Rightarrow$ cause anxiety
  - complications (inflammation, hematuria)

- High non-compliance rates $\Rightarrow$ compromised effectiveness of AS

Active Surveillance (cont'd)

- Currently, AS programs primarily use fixed schedules
  - yearly (aggressive)
  - every 3 years
  - …

- Unnecessary biopsies in patients who progress slowly

PRIAS Study

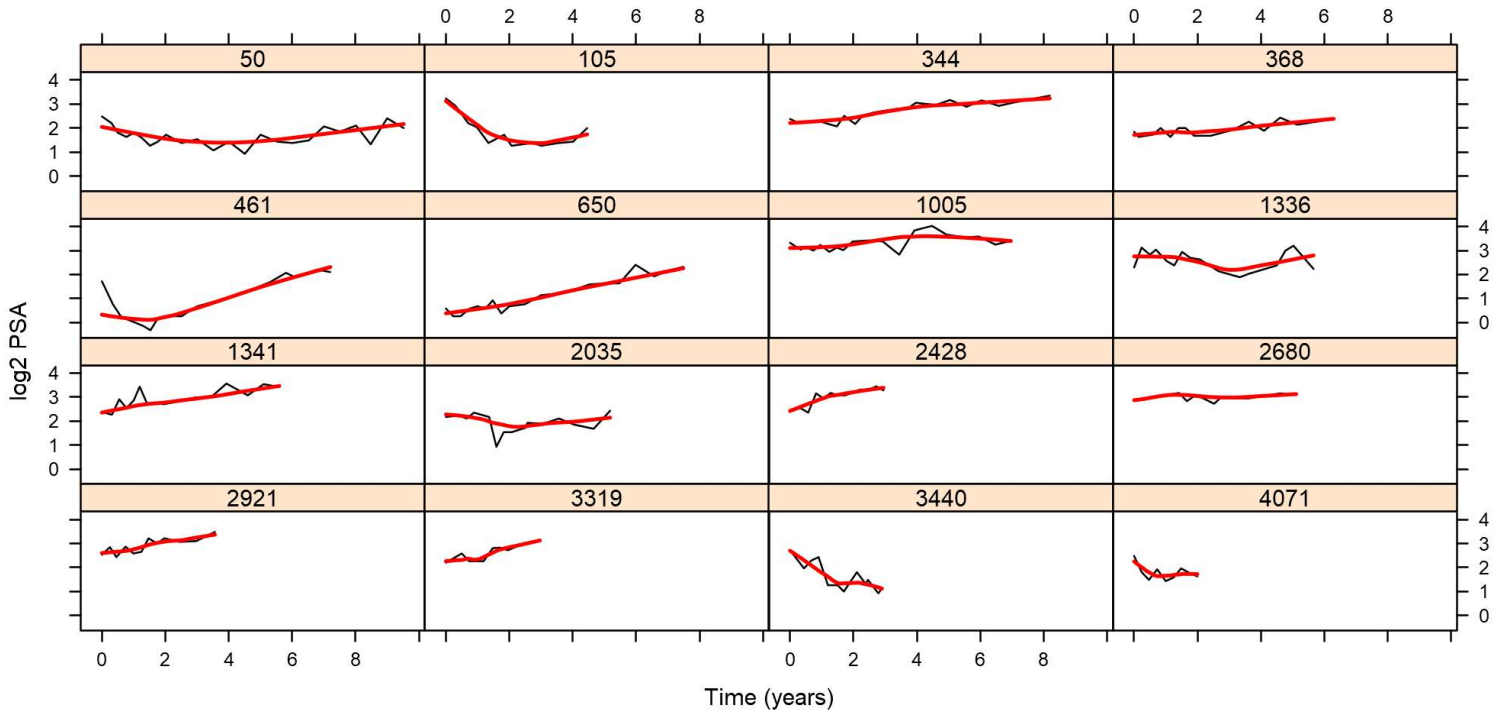
- AS program in which men with early prostate cancer are managed by a follow-up strategy
  - 5267 patients
  - biopsies at baseline, 1, 4, 7 and 10 years
  - or yearly after PSA doubling within a year

- Outcomes of interest:
  - time to Gleason score reclassification (from 6 to ≥ 7)
  - longitudinal PSA measurements

PRIAS Study (cont'd)

Outcome:

subject PSA



PRIAS Study (cont'd)

How to better plan biopsies?

- In steps:
  - *How the longitudinal PSA is related to Gleason reclassification?*
  - *How to combine previous PSA measurements and biopsies to predict reclassification?*
  - *When to plan the next biopsy?*

Time-varying Covariates

- To answer these questions we need to link
  - the time to Gleason reclassification (survival outcome)
  - the PSA measurements (longitudinal outcome)

- Biomarkers are *endogenous* time-varying covariates
  - their future path depends on previous events
  - standard time-varying Cox model not appropriate

Time-varying Covariates (cont'd)

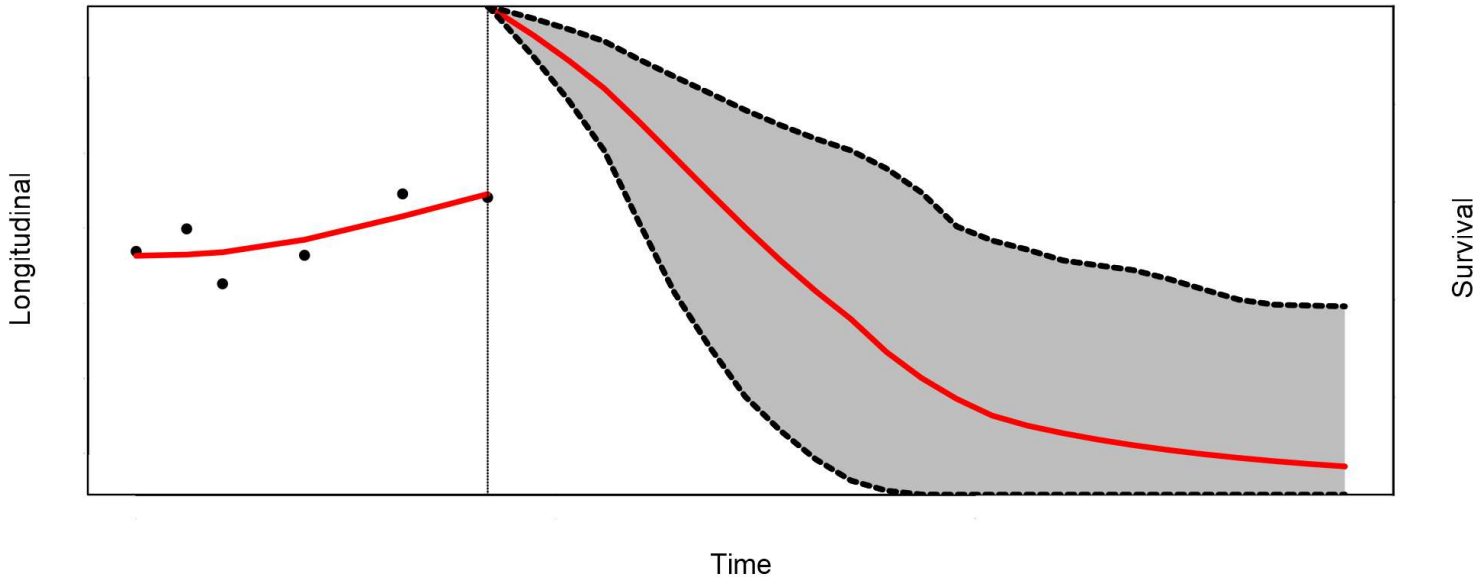
To account for endogeneity we use the framework of

Joint Models for Longitudinal & Survival Data

The Basic Joint Model

Measurements:

1 6 10



The Basic Joint Model (cont'd)

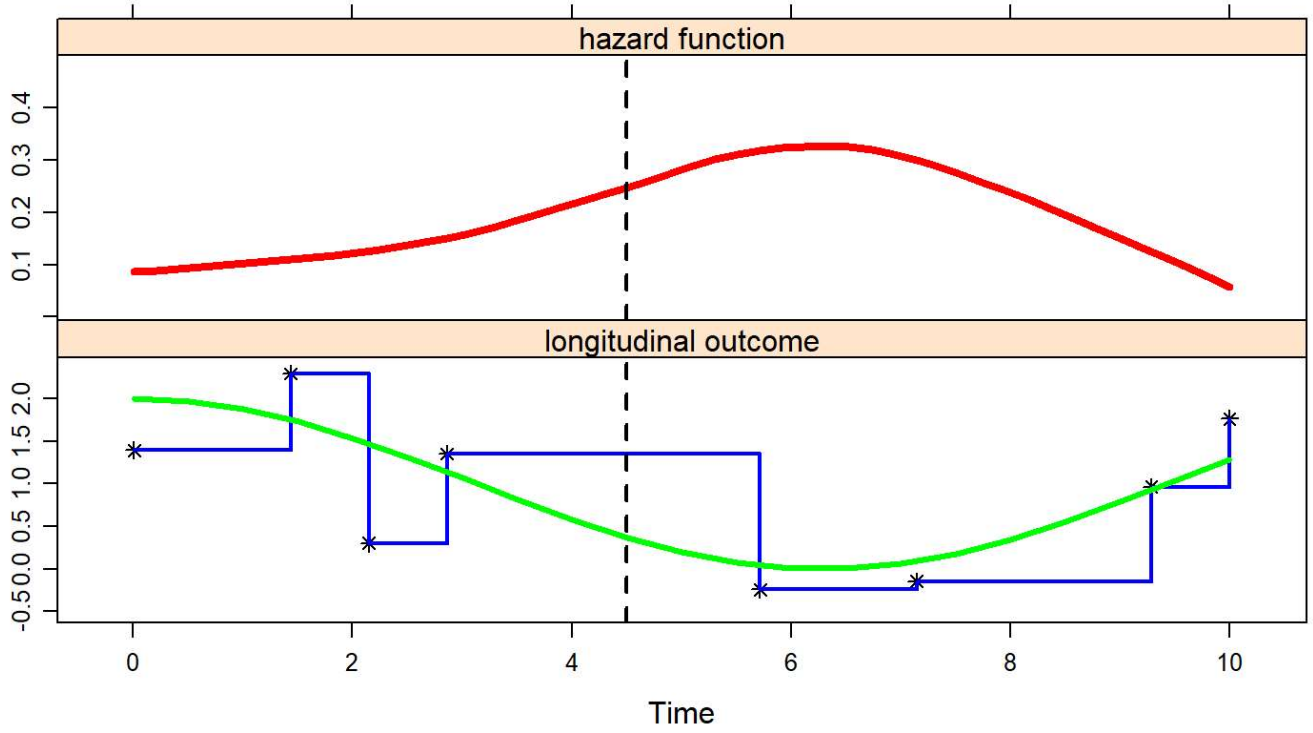
- We need some notation
  - T_i^* the true reclassification time
  - T_i^L last biopsy time point Gleason Score was < 7
  - T_i^R first biopsy time point Gleason Score was ≥ 7
  - $T_i^R = \infty$ for patients who haven't been reclassified yet
  - $\mathbf{y}_i$ vector of longitudinal PSA measurements
  - $\mathcal{Y}_i(t) = \{y_i(s), 0 \leq s < t\}$

The Basic Joint Model (cont'd)

- Formally, we have

$$
\begin{cases}
h_i(t) &= h_0(t)\exp\{\gamma^\top \mathbf{w}_i + \alpha\eta_i(t)\} \\
\\
y_i(t) &= \eta_i(t) + \varepsilon_i(t) \\
&= \mathbf{x}_i^\top(t)\beta + \mathbf{z}_i^\top(t)\mathbf{b}_i + \varepsilon_i(t) \\
\\
& \mathbf{b}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{D}), \quad \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2)
\end{cases}
$$

The Basic Joint Model (cont'd)



The Basic Joint Model (cont'd)

- The longitudinal and survival outcomes are jointly modeled

$$p(y_i, T_i^L, T_i^R) = \int p(y_i \mid b_i) \times \left\{ S(T_i^L \mid b_i) - S(T_i^R \mid b_i) \right\} \times p(b_i)\, db_i$$

  - the random effects b_i explain the interdependencies

The Basic Joint Model (cont'd)

- Estimation of joint models is based on either
  - Maximum likelihood (requires numerical integration)
  - Bayesian approaches (e.g., MCMC or HMC)

- Here, we follow a Bayesian approach
  - more on this later…

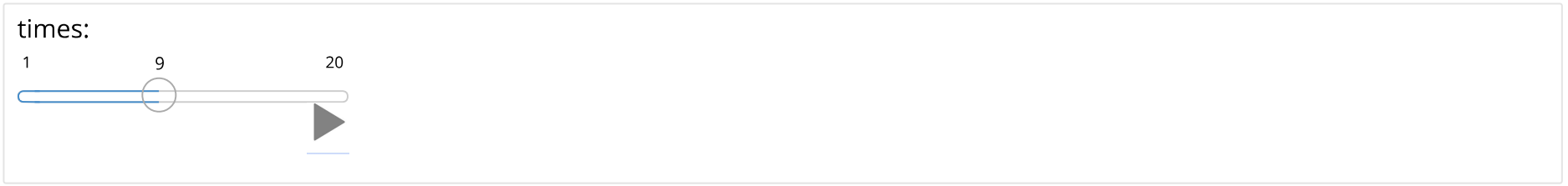
Functional Form

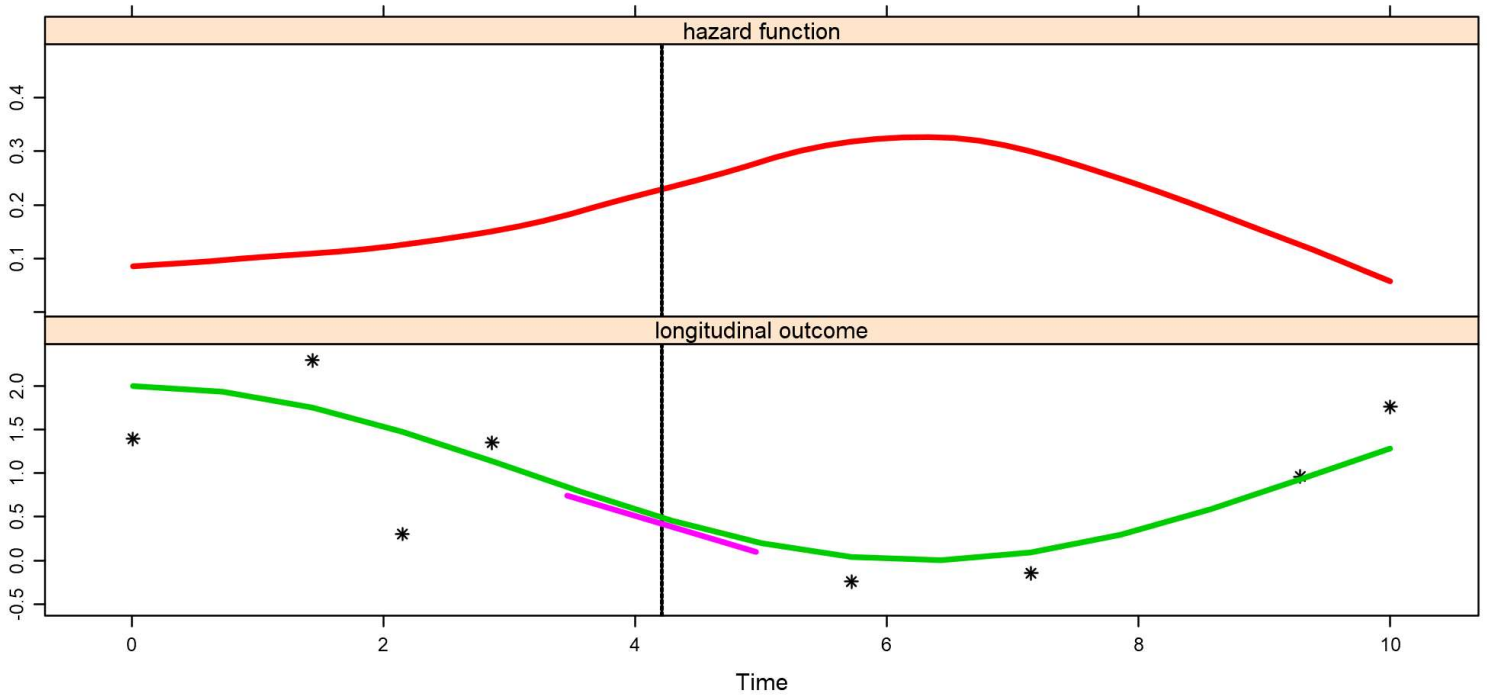
- Biomarker's rate of change
  - fast increasing PSA indicative of progression

$$h_i(t) = h_0(t) \exp\{\gamma^\top \mathbf{w}_i + \alpha_1 \eta_i(t) + \alpha_2 \eta_i'(t)\}$$

where $\eta_i'(t) = \frac{d}{dt}\eta_i(t)$

Functional Form (cont'd)





PRIAS Study Analysis

- PSA growth

$$
\begin{cases}
\log_2(PSA) &= \eta_i(t) + \varepsilon_i(t) \\
&= \beta_0 + \sum_{k=1}^{3} \beta_k \text{NS}_k(t, \nu) + \beta_4 \text{Age} + \beta_5 \text{Age}^2 \\
&\quad + b_{i0} + \sum_{k=1}^{2} b_{ik} \text{NS}_k(t, \nu) + \varepsilon_i(t) \\
\\
\mathbf{b}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{D}), & \quad \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2)
\end{cases}
$$

PRIAS Study Analysis (cont'd)

- Risk of reclassification

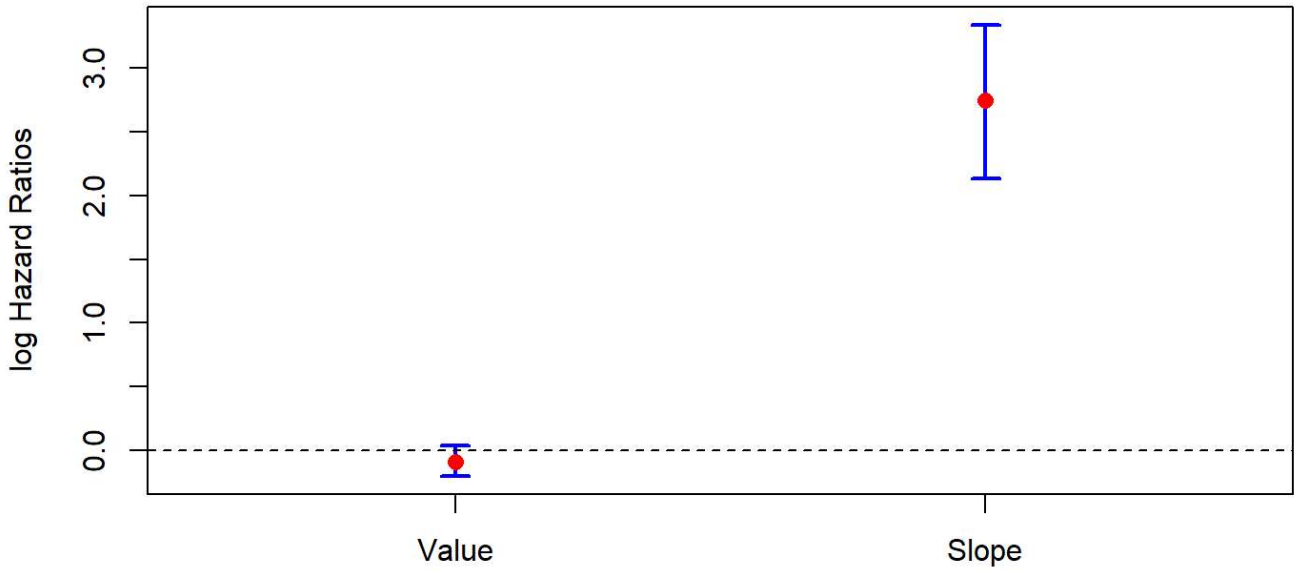
$$h_i(t) = h_0(t) \exp\left\{\gamma_1 \text{Age} + \gamma_2 \text{Age}^2 + \alpha_1 \eta_i(t) + \alpha_2 \frac{d\eta_i(t)}{dt}\right\}$$

where

- $\eta_i(t)$ log2(PSA) current value
- $\frac{d\eta_i(t)}{dt}$ log2(PSA) velocity

PRIAS Study Analysis (cont'd)

- Results



Info From a New Patient

- Suppose a new patient j comes in

- What is the available information
  - last biopsy: $T_j^* > t$
  - PSA measurements: $\mathcal{Y}_j(s) = \{y_j(t_{jl}); 0 \leq t_{jl} \leq s, l = 1, \ldots, n_j\}$
  - typically $s \geq t$
  - $\mathcal{D}_n$ sample on which the joint model was fitted

Info From a New Patient (cont'd)

- **Aim**: We want to combine all this information into a single tool to predict Gleason upgrade

- Posterior Predictive Distribution (PPD)

$$p\{T_j^* \mid T_j^* > t, \mathcal{Y}_j(s), \mathcal{D}_n\}$$

Two Avenues to Plan Biopsies

- What is the best guess on when Gleason upgrade will occur?

- **Option 1**: Theory and intuition suggest a measure of central tendency, e.g.,
  - $u =$ mean of PPD
  - $u =$ median of PPD

- **Option 2**: Risk-based approach, select $u > t$ such that

$$\Pr(T_j^* < u \mid T_j^* > t, \mathcal{Y}_j(s), \mathcal{D}_n) = \kappa, \quad 0 \leq \kappa \leq 1$$

Two Avenues to Plan Biopsies (cont'd)

- Selection of κ for **Option 2**

- Risk urologists are willing to take
  - e.g., risk of reclassification should be less than 10%
  - *often no clear consensus*

- Data-driven approach based on ROC analysis in $\mathcal{D}_n$
  - F1 score (https://en.wikipedia.org/wiki/F1_score), Youden's J (https://en.wikipedia.org/wiki/Youden%27s_J_statistic), …
  - time-varying thresholds $k(t), \quad 0 \leq t \leq t_{max}$

Simulations - Design

- We evaluated both approaches in simulations

- Data were simulated from the model fitted to PRIAS
  - value + velocity of PSA
  - 3 subgroups, early, mid and late progression

Simulations - Design (cont'd)

- Objective evaluation - each simulated dataset split in two parts
  - *training*: to fit the model
  - *test*: to apply the procedure for selecting biopsy times

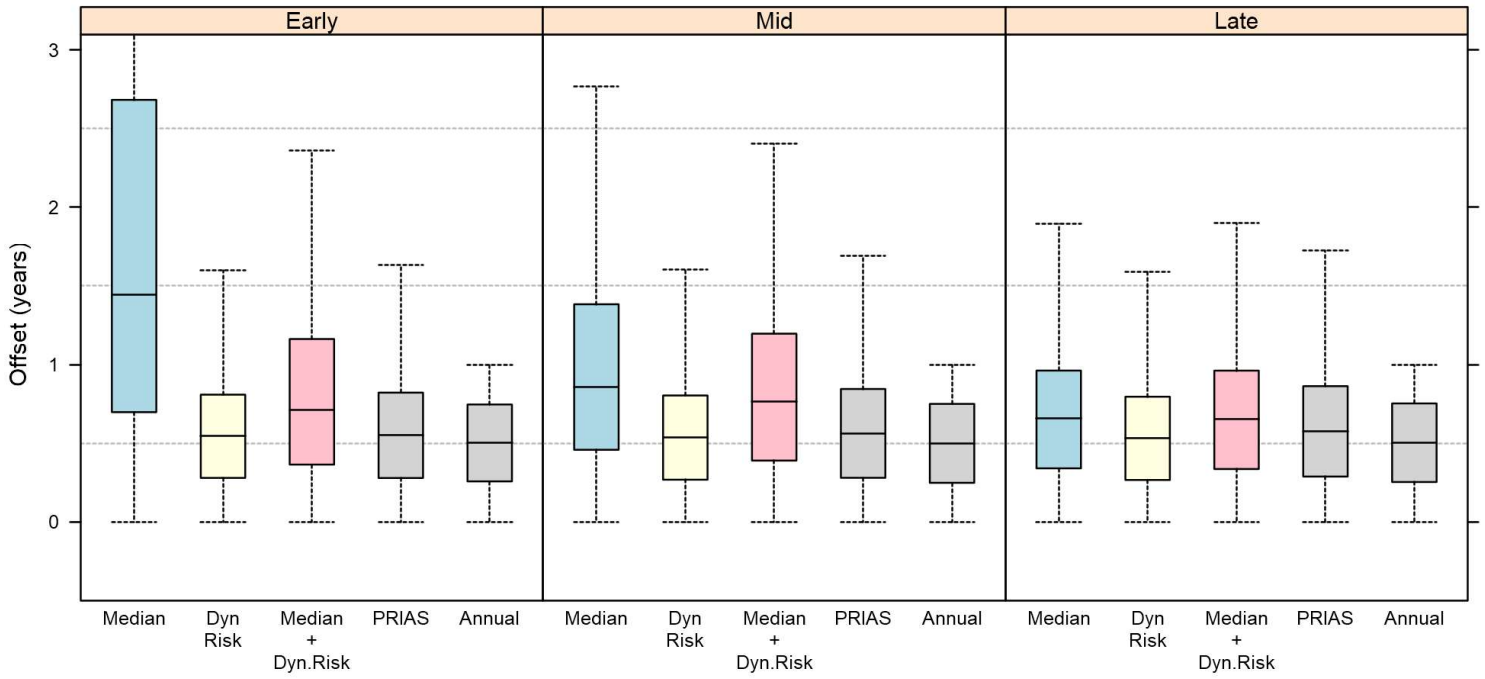
- Performance judged based on
  - effort required to find reclassification $\Rightarrow$ # Biopsies
  - how much we overshoot $\Rightarrow$ Offset

Simulation - Results

Statistic: Offset

Method: PRIAS Annual Dyn.Risk Median Median + Dyn.Risk

Period ☑ Zoom-in ☑



Online App

- We have created an online app to facilitate the whole procedure:

https://emcbiostatistics.shinyapps.io/Dynamic_Predictions/
(https://emcbiostatistics.shinyapps.io/Dynamic_Predictions/)

Discussion

- Things to improve
  - account for miss-classification
  - include DRE

- Software: available in **JMbayes** on CRAN & GitHub
  - https://cran.r-project.org/package=JMbayes (https://cran.r-project.org/package=JMbayes)
  - https://github.com/drizopoulos/JMbayes (https://github.com/drizopoulos/JMbayes)

Thank you for your attention!

http://www.drizopoulos.com/ (http://www.drizopoulos.com/)