Supplementary Materials for "Personalized Schedules for Burdensome Surveillance Tests"

Anirudh Tomer[1,*], Daan Nieboer[2], Monique J. Roobol[3], Ewout W. Steyerberg[2,4], and Dimitris Rizopoulos[1]

[1]Department of Biostatistics, Erasmus University Medical Center, the Netherlands
[2]Department of Public Health, Erasmus University Medical Center, the Netherlands
[3]Department of Urology, Erasmus University Medical Center, the Netherlands
[4]Department of Biomedical Data Sciences, Leiden University Medical Center, the Netherlands
*email: a.tomer@erasmusmc.nl

Web Appendix A Joint Model for Time-to-Progression and Longitudinal Outcomes

Let the true time of disease progression for the i-th patient be T_i^*. Progression is always observed with interval censoring $l_i < T_i^* \leq r_i$ (Figure **??**). In patients who obtain progression, r_i and l_i denote the time of their latest and second latest invasive tests. Otherwise, l_i denotes the time of their latest test and $r_i = \infty$. Assuming K auxiliary longitudinal outcomes, let $\boldsymbol{y}_{ki}$ denote the $n_{ki} \times 1$ longitudinal response vector of the k-th outcome, $k\epsilon\{1, \ldots, K\}$. The observed data of all n patients is given by $\mathcal{A}_n = \{l_i, r_i, \boldsymbol{y}_{1i}, \ldots \boldsymbol{y}_{Ki}; i = 1, \ldots, n\}$.

To accommodate longitudinal outcomes of different types in a unified framework, the joint model consists of a generalized linear mixed-effects sub-model (Laird and Ware, 1982). In particular, the conditional distribution of $\boldsymbol{y}_{ki}$ given a vector of patient-specific random effects $\boldsymbol{b}_{ki}$ is assumed to be a member of the exponential family, with linear predictor given by,

$$g_k\big[E\{y_{ki}(t) \mid \boldsymbol{b}_{ki}\}\big] = m_{ki}(t) = \boldsymbol{x}_{ki}^\top(t)\boldsymbol{\beta}_k + \boldsymbol{z}_{ki}^\top(t)\boldsymbol{b}_{ki}, \tag{1}$$

where $g_k(\cdot)$ denotes a known one-to-one monotonic link function, $y_{ki}(t)$ denotes the value of the k-th longitudinal outcome for the i-th patient at time t, and $\boldsymbol{x}_{ki}(t)$ and $\boldsymbol{z}_{ki}(t)$ denote the time-dependent design vectors for the fixed $\boldsymbol{\beta}_k$ and random effects $\boldsymbol{b}_{ki}$, respectively. To account for the association between the different longitudinal outcomes, we link their corresponding random effects. More specifically, the complete vector of random effects $\boldsymbol{b}_i = (\boldsymbol{b}_{1i}^\top, \ldots, \boldsymbol{b}_{Ki}^\top)^\top$ is assumed to follow a multivariate normal distribution with mean zero and variance-covariance matrix W.

For the survival process, we assume that the hazard of progression $h_i(t)$ at a time t depends on a function of the patient and outcome-specific linear predictors

$m_{ki}(t)$ and/or the random effects. More specifically,

$$h_i\{t \mid \mathcal{M}_i(t), \boldsymbol{w}_i(t)\} = h_0(t) \exp\left[\boldsymbol{\gamma}^\top \boldsymbol{w}_i(t) + \sum_{k=1}^{K}\sum_{l=1}^{L_k} f_{kl}\{\mathcal{M}_{ki}(t), \boldsymbol{w}_i(t), \boldsymbol{b}_{ki}, \boldsymbol{\alpha}_{kl}\}\right], \quad t > 0, \tag{2}$$

where $h_0(\cdot)$ denotes the baseline hazard function, $\mathcal{M}_{ki}(t) = \{m_{ki}(s) \mid 0 \leq s < t\}$ denotes the history of the k-th longitudinal process up to t, and $\boldsymbol{w}_i(t)$ is a vector of exogenous, possibly time-varying, covariates with corresponding regression coefficients $\boldsymbol{\gamma}$. Functions $f_{kl}(\cdot)$, parameterized by vector of coefficients $\boldsymbol{\alpha_{kl}}$, specify which features of each longitudinal outcome are included in the linear predictor of the relative-risk model (Brown, 2009; Rizopoulos, 2012; Taylor et al., 2013). Some examples, motivated by the literature (subscripts k and l dropped for brevity), are:

$$\begin{cases} f\{\mathcal{M}_i(t), \boldsymbol{w}_i(t), \boldsymbol{b}_i, \boldsymbol{\alpha}\} = \alpha m_i(t), \\ f\{\mathcal{M}_i(t), \boldsymbol{w}_i(t), \boldsymbol{b}_i, \boldsymbol{\alpha}\} = \alpha_1 m_i(t) + \alpha_2 m_i'(t), \quad \text{with } m_i'(t) = \frac{\mathrm{d}m_i(t)}{\mathrm{d}t}. \end{cases}$$

These formulations of $f(\cdot)$ postulate that the hazard of progression at time t may be associated with the underlying level $m_i(t)$ of the longitudinal outcome at t, or with both the level and velocity $m_i'(t)$ (e.g., PSA value and velocity in prostate cancer) of the outcome at t. Lastly, $h_0(t)$ is the baseline hazard at time t, and is modeled flexibly using P-splines (Eilers and Marx, 1996). More specifically:

$$\log h_0(t) = \gamma_{h_0,0} + \sum_{q=1}^{Q} \gamma_{h_0,q} B_q(t, \boldsymbol{v}),$$

where $B_q(t, \boldsymbol{v})$ denotes the q-th basis function of a B-spline with knots $\boldsymbol{v} = v_1, \ldots, v_Q$ and vector of spline coefficients γ_{h_0}. To avoid choosing the number and position of knots in the spline, a relatively high number of knots (e.g., 15 to 20) are chosen and the corresponding B-spline regression coefficients γ_{h_0} are penalized using a differences penalty (Eilers and Marx, 1996).

Web Appendix A.1 Parameter Estimation

We estimate the parameters of the joint model using Markov chain Monte Carlo (MCMC) methods under the Bayesian framework. Let $\boldsymbol{\theta}$ denote the vector of all of the parameters of the joint model. The joint model postulates that given the random effects, the time to progression, and all of the longitudinal measurements taken over time are all mutually independent. Under this assumption the posterior distribution of the parameters is given by:

$$\begin{aligned} p(\boldsymbol{\theta}, \boldsymbol{b} \mid \mathcal{D}_n) &\propto \prod_{i=1}^{n} p(l_i, r_i, \boldsymbol{y}_{1i}, \ldots \boldsymbol{y}_{Ki}, \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) \\ &\propto \prod_{i=1}^{n}\prod_{k=1}^{K} p(l_i, r_i \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{y}_{ki} \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}), \\ p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) &= \frac{1}{\sqrt{(2\pi)^{|W|}\det(\boldsymbol{D})}} \exp(\boldsymbol{b}_i^\top \boldsymbol{D}^{-1} \boldsymbol{b}_i), \end{aligned}$$

where, the likelihood contribution of the k-th longitudinal outcome vector $\boldsymbol{y}_{ki}$ for the i-th patient, conditional on the random effects is:

$$p(\boldsymbol{y}_{ki} \mid \boldsymbol{b}_i, \boldsymbol{\theta}) = \prod_{j=1}^{n_{ki}} \exp\left[\frac{y_{kij}\psi_{kij}(\boldsymbol{b}_{ki}) - c_k\{\psi_{kij}(\boldsymbol{b}_{ki})\}}{a_k(\varphi)} - d_k(y_{kij}, \varphi)\right],$$

where n_{ki} are the total number of longitudinal measurements of type k for patient i. The natural and dispersion parameters of the exponential family are denoted by $\psi_{kij}(\boldsymbol{b}_{ki}$ and φ, respectively. In addition, $c_k(\cdot), a_k(\cdot), d_k(\cdot)$ are known functions specifying the member of the exponential family. The likelihood contribution of the time to progression outcome is given by:

$$p(l_i, r_i \mid \boldsymbol{b}_i, \boldsymbol{\theta}) = \exp\Big[-\int_0^{l_i} h_i\big\{s \mid \mathcal{M}_i(t), \boldsymbol{w}_i(t)\big\}\mathrm{d}s\Big] - \exp\Big[-\int_0^{r_i} h_i\big\{s \mid \mathcal{M}_i(t), \boldsymbol{w}_i(t)\big\}\mathrm{d}s\Big]. \tag{3}$$

The integral in Equation (3) does not have a closed-form solution, and therefore we use a 15-point Gauss-Kronrod quadrature rule to approximate it.

We use independent normal priors with zero mean and variance 100 for the fixed effect parameters of the longitudinal model. For scale parameters we inverse Gamma priors. For the variance-covariance matrix $\boldsymbol{D}$ of the random effects we take inverse Wishart prior with an identity scale matrix and degrees of freedom equal to the total number of random effects. For the relative risk model's parameters $\boldsymbol{\gamma}$ and the association parameters $\boldsymbol{\alpha}$, we use independent normal priors with zero mean and variance 100. However, when $\boldsymbol{\alpha}$ becomes high dimensional (e.g., when several functional forms are considered per longitudinal outcome), we opt for a global-local ridge-type shrinkage prior, i.e., for the s-th element of $\boldsymbol{\alpha}$ we assume:

$$\alpha_s \sim \mathcal{N}(0, \tau\psi_s), \quad \tau^{-1} \sim \text{Gamma}(0.1, 0.1), \quad \psi_s^{-1} \sim \text{Gamma}(1, 0.01). \tag{4}$$

The global smoothing parameter τ has sufficiently mass near zero to ensure shrinkage, while the local smoothing parameter ψ_s allows individual coefficients to attain large values. Other options of shrinkage or variable-selection priors could be used as well (Andrinopoulou and Rizopoulos, 2016). Finally, the penalized version of the B-spline approximation to the baseline hazard is specified using the following hierarchical prior for γ_{h_0} (Lang and Brezger, 2004):

$$p(\gamma_{h_0} \mid \tau_h) \propto \tau_h^{\rho(\boldsymbol{K})/2} \exp\Big(-\frac{\tau_h}{2}\gamma_{h_0}^\top \boldsymbol{K}\gamma_{h_0}\Big) \tag{5}$$

where τ_h is the smoothing parameter that takes a Gamma$(1, \tau_{h\delta})$ prior distribution, with a hyper-prior $\tau_{h\delta} \sim$ Gamma$(10^{-3}, 10^{-3})$, which ensures a proper posterior distribution for γ_{h_0}(Jullion and Lambert, 2007), $\boldsymbol{K} = \Delta_r^\top \Delta_r + 10^{-6}\boldsymbol{I}$, with Δ_r denoting the r-th difference penalty matrix, and $\rho(\boldsymbol{K})$ denotes the rank of $\boldsymbol{K}$.

Web Appendix B Joint Model for the PRIAS Dataset

Web Appendix B.1 Dataset

We first provide an introduction to the world's largest active surveillance (AS) program called Prostate Cancer Research International Active Surveillance, abbreviated as PRIAS (Bokhorst et al., 2016), that we use to illustrate our methodology. We then outline the joint model that we fit to the PRIAS dataset. We then present the parameter estimates of our model.

Web Appendix B.2 PRIAS Dataset

We use the data of prostate cancer patients (see Web Table 1) from the world's largest AS study called PRIAS (Bokhorst et al., 2016). More than 100 medical centers from 17 countries worldwide contribute to the collection of data, utilizing a common study protocol and a web-based tool, both available at `www.prias-project.org`. We use data collected between December 2006 (beginning of PRIAS study) and May 2019. The primary event of interest is cancer progression detected upon a positive biopsy. Biopsies are scheduled at the following fixed follow-up times (measured since inclusion in AS): year 1, 4, 7, and 10, and every 5 years thereafter. An annual schedule of biopsies is prescribed to those patients who have a PSA (prostate-specific antigen) doubling time between 0 and 10 years. The PSA doubling time at any point during follow-up is measured as the inverse of the slope of the regression line through the base two logarithm of the observed PSA values. The time of progression is interval censored because biopsies are scheduled periodically. There are three types of competing events, namely death, removal of patients from AS on the basis of their observed DRE (digital rectal examination) and PSA measurements, and loss to follow-up. We assume these three types of events to be censored observations. Under this assumption of censoring, Figure 1 shows the cumulative-risk of progression over the study follow-up period.

Web Appendix B.3 Model Specification

Let T_i^* denote the true progression time of the i-th patient included in PRIAS. Since biopsies are conducted periodically, T_i^* is observed with interval censoring $l_i < T_i^* \leq r_i$. When progression is observed for the patient at his latest biopsy time r_i, then l_i denotes the time of the second latest biopsy. Otherwise, l_i denotes the time of the latest biopsy and $r_i = \infty$. Let $\boldsymbol{y}_{di}$ and $\boldsymbol{y}_{pi}$ denote his observed DRE (digital rectal examination) and PSA (prostate-specific antigen) longitudinal measurements, respectively. The observed data of all n patients is denoted by $\mathcal{D}_n = \{l_i, r_i, \boldsymbol{y}_{di}, \boldsymbol{y}_{pi}; i = 1, \ldots, n\}$.

In our joint model, the patient-specific DRE and PSA measurements over time are modeled using a bivariate generalized linear mixed effects sub-model. The sub-model for DRE is given by (see Panel A, Figure **??**):

$$\begin{aligned} \text{logit}\big[\Pr\{y_{di}(t) > \text{T1c}\}\big] &= \beta_{0d} + b_{0di} + (\beta_{1d} + b_{1di})t \\ &\quad + \beta_{2d}(\text{Age}_i - 65) + \beta_{3d}(\text{Age}_i - 65)^2 \end{aligned} \tag{6}$$

where, t denotes the follow-up visit time, and Age_i is the age of the i-th patient at the time of inclusion in AS. The fixed effect parameters are denoted by $\{\beta_{0d}, \ldots, \beta_{3d}\}$, and $\{b_{0di}, b_{1di}\}$ are the patient specific random effects. With this definition, we

Table Web Table 1: **Summary of the PRIAS dataset**. The primary event of interest is cancer progression (increase in biopsy Gleason grade group from grade group 1 to 2 or higher). Abbreviations: PSA is prostate-specific antigen; DRE is digital rectal examination, with level T1c (Schröder et al., 1992) indicating a clinically inapparent tumor which is not palpable or visible by imaging, whereas tumors with DRE > T1c are palpable; IQR is interquartile range.

Characteristic	**Value**
Total patients	7813
progression (primary event)	1134
Treatment	2250
Watchful waiting	334
Lost to follow-up	203
Discontinued on request	46
Death (other)	95
Death (prostate cancer)	2
Total DRE measurements	37326
Total PSA measurements	67578
Total biopsies	15686
Median follow-up period per patient (years)	1.8 (IQR: 0.9–4.0)
Median age at diagnosis (years)	66 (IQR: 61–71)
Median PSA (ng/mL)	5.7 (IQR: 4.1–7.7)
DRE = T1c (%)	34883/37326 (94%)
Median number of PSA per patient	6 (IQR: 4–12)
Median number of DRE per patient	4 (IQR: 2–7)
Median number of biopsies per patient	2 (IQR: 1–2)

assume that the patient-specific log odds of obtaining a DRE measurement larger than T1c (palpable tumor) remain linear over time.

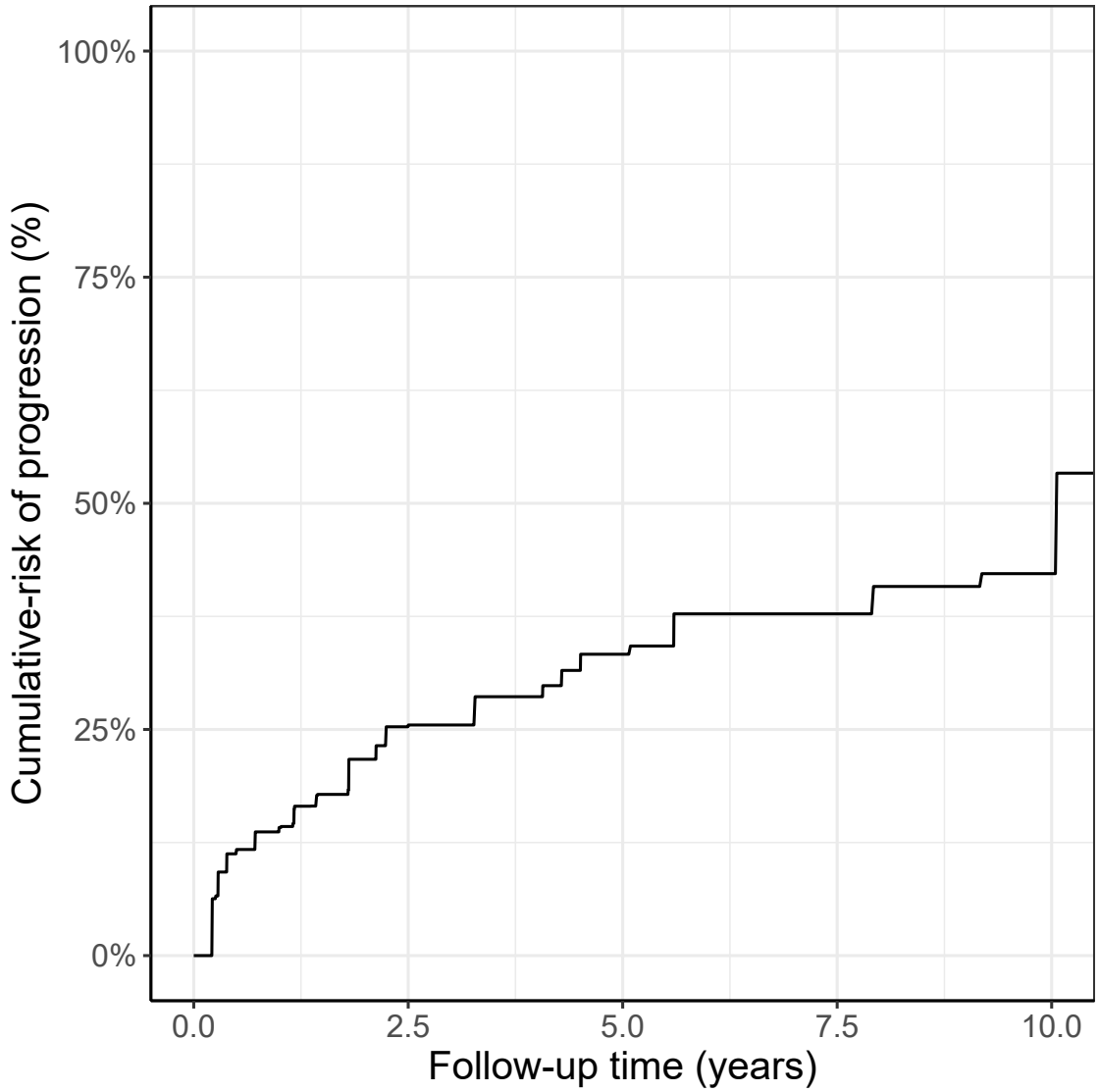
The mixed effects sub-model for PSA is given by (see Panel B, Figure **??**):

$$\begin{aligned}\log_2\left\{y_{pi}(t)+1\right\} &= m_{pi}(t)+\varepsilon_{pi}(t),\\ m_{pi}(t) &= \beta_{0p}+b_{0pi}+\sum_{k=1}^{3}(\beta_{kp}+b_{kpi})B_k(t,\mathcal{K})\\ &\quad+\beta_{4p}(\text{Age}_i-65)+\beta_{5p}(\text{Age}_i-65)^2,\end{aligned} \tag{7}$$

where, $m_{pi}(t)$ denotes the underlying measurement error free value of $\log_2(\text{PSA}+1)$ transformed (Lin et al., 2000; Pearson et al., 1994) measurements at time t. We model it non-linearly over time using B-splines (De Boor, 1978). To this end, our B-spline basis function $B_k(t,\mathcal{K})$ has 3 internal knots at $\mathcal{K}=\{0.75, 2.12\}$ years (33-rd and 66-th percentile of observed follow-up times), and boundary knots at 0 and 6.4 years (95-th percentile of the observed follow-up times). The fixed effect parameters are denoted by $\{\beta_{0p},\ldots,\beta_{5p}\}$, and $\{b_{0pi},\ldots,b_{3pi}\}$ are the patient specific random effects. The error $\varepsilon_{pi}(t)$ is assumed to be t-distributed with three degrees of freedom (see Web Appendix B.5) and scale σ, and is independent of the random effects.

To account for the correlation between the DRE and PSA measurements of a patient, we link their corresponding random effects. More specifically, the complete vector of random effects $\boldsymbol{b}_i=(b_{0di},b_{1di},b_{0pi},\ldots,b_{3pi})^\top$ is assumed to follow a multivariate normal distribution with mean zero and variance-covariance matrix $\boldsymbol{D}$.

To model the impact of DRE and PSA measurements on the risk of progression, our joint model uses a relative risk sub-model. More specifically, the hazard of



Web Figure 1: **Estimated cumulative-risk of cancer progression** for patients in the Prostate Cancer Research International Active Surveillance (PRIAS) dataset. Nearly 50% patients (*slow progressing*) do not progress in the ten year follow-up period. Cumulative risk is estimated using nonparametric maximum likelihood estimation (Turnbull, 1976), to account for interval censored progression times observed in the PRIAS dataset. Censoring includes death, removal from AS on the basis of observed longitudinal data, and patient dropout.

progression $h_i(t)$ at a time t is given by (see Panel D, Figure **??**):

$$h_i(t) = h_0(t) \exp\Big(\gamma_1(\text{Age}_i - 65) + \gamma_2(\text{Age}_i - 65)^2 + \alpha_{1d}\text{logit}\big[\Pr\{y_{di}(t) > \text{T1c}\}\big] + \alpha_{1p} m_{pi}(t) + \alpha_{2p}\frac{\partial m_{pi}(t)}{\partial t}\Big), \tag{8}$$

where, γ_1, γ_2 are the parameters for the effect of age. The parameter α_{1d} models the impact of log odds of obtaining a DRE > T1c on the hazard of progression. The impact of PSA on the hazard of progression is modeled in two ways: a) the impact of the error free underlying PSA value $m_{pi}(t)$ (see Panel B, Figure **??**), and b) the impact of the underlying PSA velocity $\partial m_{pi}(t)/\partial t$ (see Panel C, Figure **??**). The corresponding parameters are α_{1p} and α_{2p}, respectively. Lastly, $h_0(t)$ is the baseline

hazard at time t, and is modeled flexibly using P-splines (Eilers and Marx, 1996).

Web Appendix B.4 Parameter Estimates

We fit a joint model to the PRIAS dataset using the R package **JMbayes** (Rizopoulos, 2016). The corresponding posterior parameter estimates are shown in Web Table 3 (longitudinal sub-model for DRE outcome), Web Table 4 (longitudinal sub-model for PSA outcome) and Web Table 5 (relative risk sub-model). The parameter estimates for the variance-covariance matrix $\boldsymbol{D}$ from the longitudinal sub-model are shown in the following Web Table 2:

Table Web Table 2: Estimated variance-covariance matrix $\boldsymbol{D}$ of the random effects $\boldsymbol{b} = (b_{0d}, b_{1d}, b_{0p}, b_{1p}, b_{2p}, b_{3p})$ from the joint model fitted to the PRIAS dataset. The variances of the random effects are highlighted along the diagonal of the variance-covariance matrix.

Random Effects	b_{0d}	b_{1d}	b_{0p}	b_{1p}	b_{2p}	b_{3p}
b_{0d}	9.233	-0.183	-0.213	0.082	0.058	0.023
b_{1d}	-0.183	1.259	0.091	0.079	0.145	0.109
b_{0p}	-0.213	0.091	0.247	0.007	0.067	0.018
b_{1p}	0.082	0.079	0.007	0.248	0.264	0.189
b_{2p}	0.058	0.145	0.067	0.264	0.511	0.327
b_{3p}	0.023	0.109	0.018	0.189	0.327	0.380

Table Web Table 3: Estimated mean and 95% credible interval for the parameters of the longitudinal sub-model (see Equation 6) for the DRE outcome.

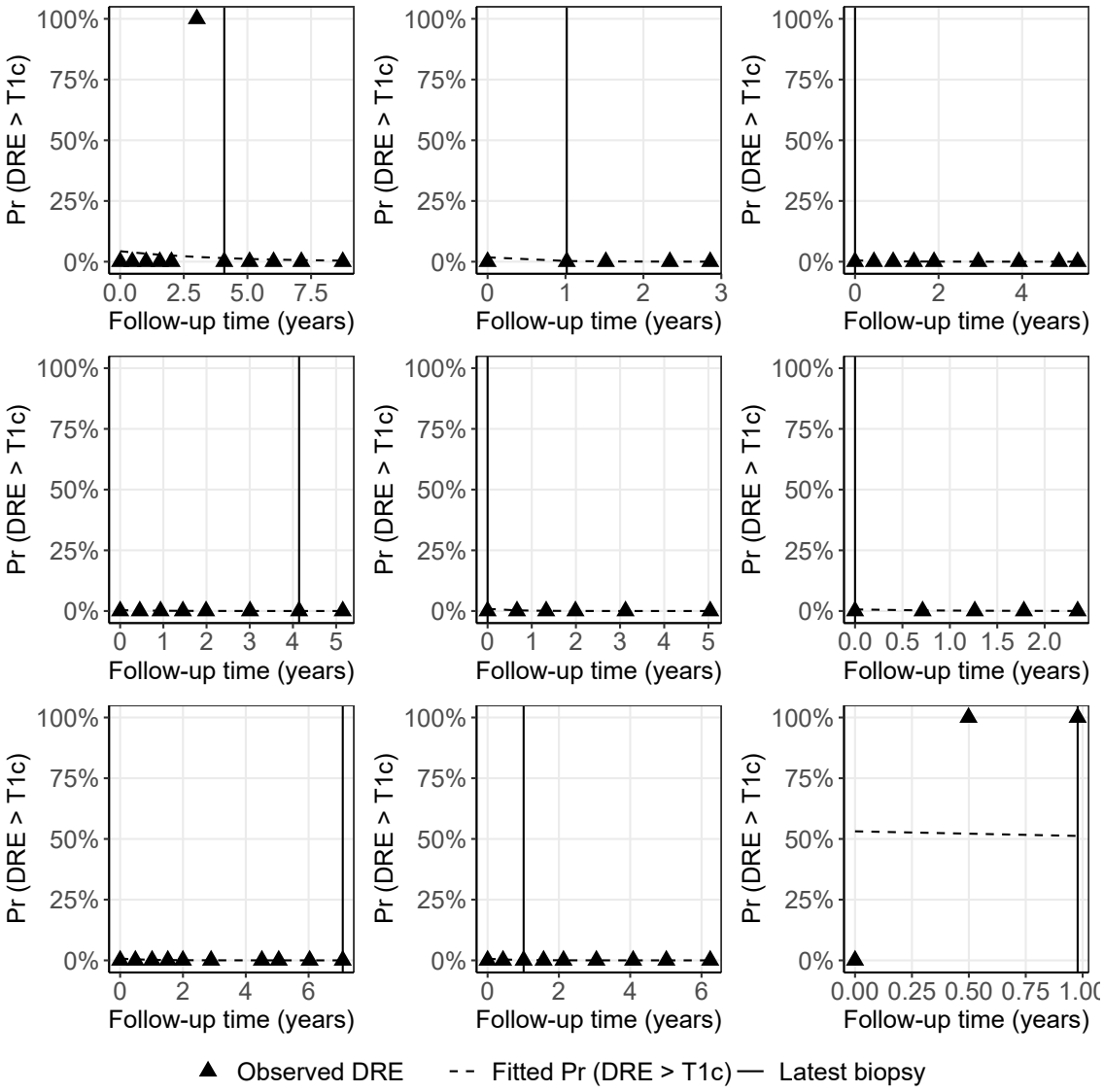
Variable	Mean	Std. Dev	2.5%	97.5%
(Intercept)	-4.407	0.151	-4.716	-4.113
(Age − 65)	0.057	0.009	0.039	0.075
$(\text{Age} - 65)^2$	-0.002	0.001	-0.004	0.000
visitTimeYears	-1.089	0.113	-1.292	-0.866

Table Web Table 4: Estimated mean and 95% credible interval for the parameters of the longitudinal sub-model (see Equation 7) for the PSA outcome.

Variable	Mean	Std. Dev	2.5%	97.5%
(Intercept)	2.687	0.007	2.674	2.701
(Age − 65)	0.008	0.001	0.006	0.010
$(\text{Age} - 65)^2$	-0.001	0.000	-0.001	0.000
Spline: [0.00, 0.75] years	0.199	0.009	0.181	0.217
Spline: [0.75, 2.12] years	0.293	0.012	0.269	0.316
Spline: [2.12, 6.4] years	0.379	0.014	0.352	0.406
σ	0.144	0.001	0.142	0.145

We present plots of observed DRE versus fitted probabilities of obtaining a DRE measurement larger than T1c, for nine randomly selected patients in Figure 2. Sim-

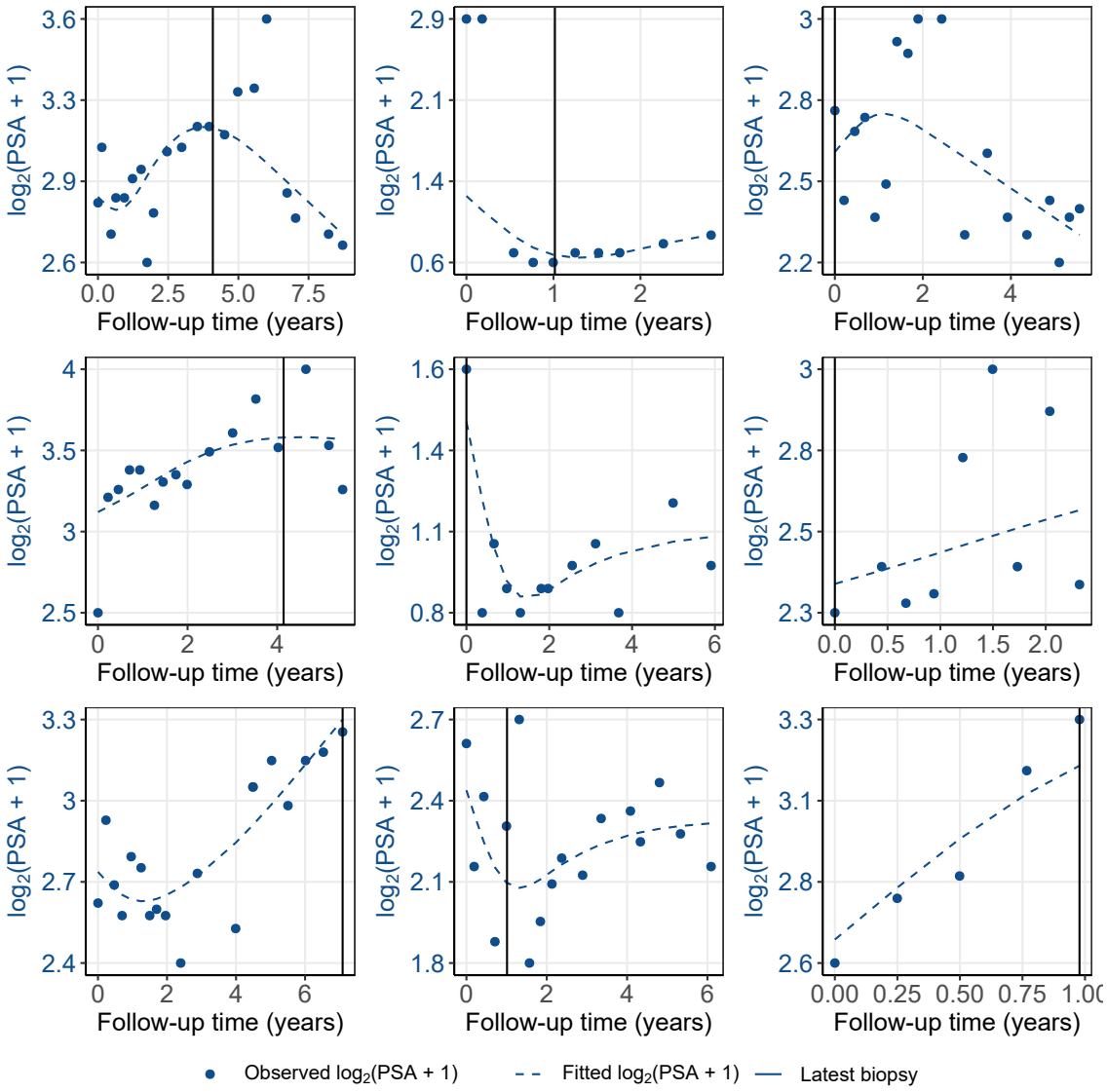
ilarly observed versus fitted PSA profiles for nine randomly selected patients are shown in Figure 3.



Web Figure 2: Observed DRE versus fitted probabilities of obtaining a DRE measurement larger than T1c, for nine randomly selected PRIAS patients. The fitted profiles utilize information from the observed DRE measurements, PSA measurements, and time of the latest biopsy. Observed DRE measurements plotted against 0% probability are equal to T1c. Observed DRE measurements plotted against 100% probability are larger than T1c.

For the relative risk sub-model (see Equation 8), the parameter estimates in Web Table 5 show that both $\log_2(\text{PSA}+1)$ velocity, and the log odds of having DRE $>$ T1c were significantly associated with the hazard of progression.

It is important to note that since age, $\log_2(\text{PSA}+1)$ value and velocity, and log odds of DRE $>$ T1c are all measured on different scales, a comparison between the corresponding parameter estimates is not easy. To this end, in Web Table 6, we present the hazard (of progression) ratio, for an increase in the aforementioned variables from their first to the third quartile. For example, an increase in log odds of DRE $>$ T1c, from -6.650 to -4.356 (fitted first and third quartiles) corresponds to a hazard ratio of 1.402. The interpretation for the rest is similar.



Web Figure 3: Fitted versus observed $\log_2(\text{PSA}+1)$ profiles for nine randomly selected PRIAS patients. The fitted profiles utilize information from the observed PSA measurements, DRE measurements, and time of the latest biopsy.

Table Web Table 5: Estimated mean and 95% credible interval for the parameters of the relative risk sub-model (see Equation 8) of the joint model fitted to the PRIAS dataset.

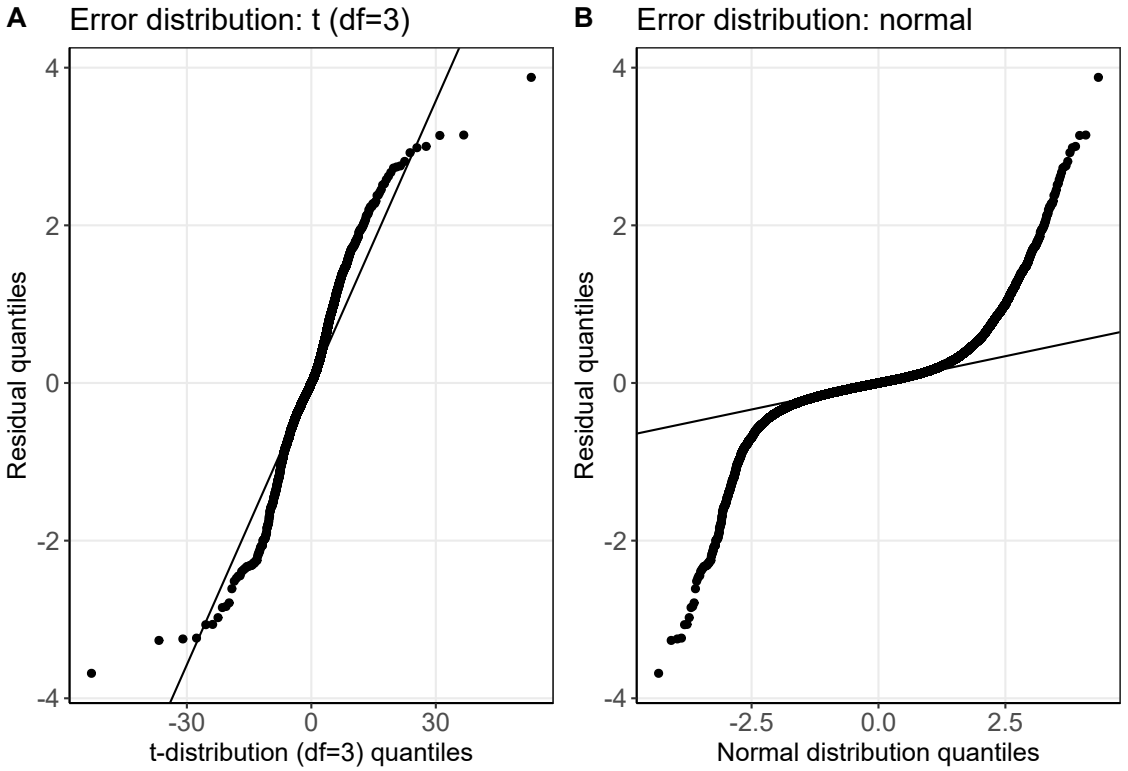
Variable	Mean	Std. Dev	2.5%	97.5%
$(\text{Age}-65)$	0.034	0.005	0.025	0.043
$(\text{Age}-65)^2$	0.000	0.001	-0.001	0.001
$\text{logit}\{\Pr(\text{DRE} > \text{T1c})\}$	0.047	0.014	0.018	0.073
Fitted $\log_2(\text{PSA}+1)$ value	0.024	0.076	-0.125	0.170
Fitted $\log_2(\text{PSA}+1)$ velocity	2.656	0.291	2.090	3.236

Table Web Table 6: Hazard (of progression) ratio and 95% credible interval (CI), for an increase in the variables of relative risk sub-model, from their first quartile (Q_1) to their third quartile (Q_3). Except for age, quartiles for all other variables are based on their fitted values obtained from the joint model fitted to the PRIAS dataset.

Variable	Q_1	Q_3	Hazard ratio [95% CI]
Age	61	71	1.402 [1.282, 1.539]
$\text{logit}\{\Pr(\text{DRE} > \text{T1c})\}$	-8.278	-5.155	1.157 [1.057, 1.255]
$\log_2(\text{PSA} + 1)$ value	2.362	3.076	1.017 [0.914, 1.129]
$\log_2(\text{PSA} + 1)$ velocity	-0.033	0.145	1.603 [1.450, 1.778]

Web Appendix B.5 Assumption of t-distributed (df=3) Error Terms

With regards to the choice of the distribution for the error term ε_p for the PSA measurements (see Equation 7), we attempted fitting multiple joint models differing in error distribution, namely t-distribution with three, and four degrees of freedom, and a normal distribution for the error term. However, the model assumption for the error term were best met by the model with t-distribution having three degrees of freedom. The quantile-quantile plot of subject-specific residuals for the corresponding model in Panel A of Figure 4, shows that the assumption of t-distributed (df=3) errors is reasonably met by the fitted model.



Web Figure 4: Quantile-quantile plot of subject-specific residuals from the joint models fitted to the PRIAS dataset. **Panel A**: model assuming a t-distribution (df=3) for the error term ε_p. **Panel B**: model assuming a normal distribution for the error term ε_p.

Web Appendix B.6 Predictive Performance of the Joint Model Fitted to the PRIAS dataset

The performance of personalized schedules depends on the performance of the joint model. In this regard, we evaluate the predictive performance of the model that we fitted to the PRIAS dataset, using two measures: the area under the receiver operating characteristic curve (AUC) as a measure of discrimination, and the mean absolute prediction error (MAPE) as a measure of calibration. Given the longitudinal nature of the data at hand, in a joint model time dependent AUC and MAPE are more relevant. More specifically, in PRIAS, given the time of latest biopsy t, and history of DRE and PSA measurements up to time s, we are interested in a medically relevant time frame $(t, s]$, within which the occurrence of progression is of interest. In the case of prostate cancer, at any point in time it is of interest to identify patients who may have obtained progression in the last one year $(s - t = 1)$. Using data of the patients from the PRIAS study, we calculate the AUC and MAPE (see Rizopoulos, Molenberghs, and Lesaffre (2017) for estimation) at every six months (schedule of PSA/DRE measurement follow-up in PRIAS) between year one and year six (roughly the 95-percentile of observed follow-up times). The resulting AUC, and MAPE are presented in Web Table 7. We can see that the our model has a moderate AUC and MAPE. Thus, we expect even better performance of personalized schedules than in our simulation study, in joint models with better AUC and MAPE.

Table Web Table 7: Follow-up time dependent, area under the receiver operating characteristic curves (AUC), and mean absolute prediction error (MAPE), with 95% confidence interval in brackets.

Follow-up period (years)	AUC (95% CI)	MAPE (95%CI)
0.0 to 1.0	0.658 [0.620, 0.693]	0.234 [0.229, 0.240]
0.5 to 1.5	0.648 [0.631, 0.663]	0.220 [0.213, 0.226]
1.0 to 2.0	0.624 [0.600, 0.644]	0.151 [0.147, 0.155]
1.5 to 2.5	0.649 [0.604, 0.704]	0.127 [0.118, 0.134]
2.0 to 3.0	0.683 [0.629, 0.729]	0.134 [0.121, 0.143]
2.5 to 3.5	0.681 [0.604, 0.739]	0.115 [0.105, 0.128]
3.0 to 4.0	0.647 [0.600, 0.710]	0.079 [0.073, 0.087]
3.5 to 4.5	0.630 [0.583, 0.668]	0.095 [0.089, 0.101]
4.0 to 5.0	0.614 [0.557, 0.659]	0.104 [0.098, 0.111]
4.5 to 5.5	0.615 [0.541, 0.702]	0.101 [0.088, 0.114]
5.0 to 6.0	0.617 [0.550, 0.713]	0.102 [0.086, 0.121]

Web Appendix C Additional Examples of Personalized Schedules

Web Appendix D Simulation Study Results

In the simulation study, we evaluated the following biopsy schedules (Inoue et al., 2018; Loeb et al., 2014): biopsy every year (annual), biopsy according to the PRIAS schedule (PRIAS), personalized biopsy schedules based on two fixed risk thresholds, namely, $\kappa^* = 5\%$ and $\kappa^* = 10\%$, and an automatically chosen $\kappa^*(v)$ (Section 3 of main manuscript). We compare all the aforementioned schedules on two criteria, namely the number of biopsies they schedule and the corresponding time delay in detection of cancer progression, in years (time of positive biopsy - true time of cancer progression). The corresponding results, using 500×250 test patients are presented in Table Web Table 8. Since the simulated cohorts are based on PRIAS, roughly only 50% of the patients progress in the ten year study period. While, we are able to calculate total number of biopsies scheduled in all 500×250 test patients, but the time delay in detection of progression is available only for those patients who progress in ten years (*progressing*). Hence, we show the simulation results separately for *progressing* and *non-progressing* patients.

Table Web Table 8: **Simulation study results for all patients**: Estimated first, second (median), and third quartiles for number of biopsies (Q_1^{nb}, Q_2^{nb}, Q_3^{nb}) and for the delay in detection of cancer progression (Q_1^{delay}, Q_2^{delay}, Q_3^{delay}), in years, for various biopsy schedules. The delay is equal to the difference between the time of the positive biopsy and the unobserved true time of progression. Types of personalized schedules: Risk: 10% and Risk: 5% approaches, schedule a biopsy if the cumulative-risk of cancer progression at a visit is more than 10% and 5%, respectively. Risk: $\kappa^*(v)$ works similar as previous, except that a visit-specific risk threshold is chosen automatically (Section 3 of main manuscript).

Progressing patients (50%)	Q_1^{nb}	Q_2^{nb}	Q_3^{nb}	Q_1^{delay}	Q_2^{delay}	Q_3^{delay}
Annual	1	3	6	0.29	0.57	0.82
PRIAS	1	2	4	0.38	0.74	1.00
Risk: 5%	1	3	5	0.33	0.65	0.91
Risk: 10%	1	2	4	0.45	0.85	1.34
Risk: $\kappa^*(v)$	1	2	3	0.54	0.96	1.74
Non-progressing patients (50%)						
Annual	10	10	10	-	-	-
PRIAS	4	6	8	-	-	-
Risk: 5%	6	7	9	-	-	-
Risk: 10%	4	5	6	-	-	-
Risk: $\kappa^*(v)$	3	4	4	-	-	-

References

Andrinopoulou, Eleni-Rosalina and Dimitris Rizopoulos (2016). "Bayesian shrinkage approach for a joint model of longitudinal and survival outcomes assuming different association structures". In: *Statistics in medicine* 35.26, pp. 4813–4823.

Bokhorst, Leonard P et al. (2016). "A decade of active surveillance in the PRIAS study: an update and evaluation of the criteria used to recommend a switch to active treatment". In: *European Urology* 70.6, pp. 954–960.

Brown, Elizabeth R. (2009). "Assessing the association between trends in a biomarker and risk of event with an application in pediatric HIV/AIDS". In: *The Annals of Applied Statistics* 3.3, pp. 1163–1182.

De Boor, Carl (1978). *A practical guide to splines*. Vol. 27. Springer-Verlag New York.

Eilers, Paul HC and Brian D Marx (1996). "Flexible smoothing with B-splines and penalties". In: *Statistical Science* 11.2, pp. 89–121.

Inoue, Lurdes YT et al. (2018). "Comparative Analysis of Biopsy Upgrading in Four Prostate Cancer Active Surveillance Cohorts". In: *Annals of internal medicine* 168.1, pp. 1–9.

Jullion, Astrid and Philippe Lambert (2007). "Robust specification of the roughness penalty prior distribution in spatially adaptive Bayesian P-splines models". In: *Computational statistics & data analysis* 51.5, pp. 2542–2558.

Laird, Nan M and James H Ware (1982). "Random-effects models for longitudinal data". In: *Biometrics*, pp. 963–974.

Lang, Stefan and Andreas Brezger (2004). "Bayesian P-splines". In: *Journal of computational and graphical statistics* 13.1, pp. 183–212.

Lin, Haiqun et al. (2000). "A latent class mixed model for analysing biomarker trajectories with irregularly scheduled observations". In: *Statistics in Medicine* 19.10, pp. 1303–1318.

Loeb, Stacy et al. (2014). "Heterogeneity in active surveillance protocols worldwide". In: *Reviews in urology* 16.4, pp. 202–203.

Pearson, Jay D et al. (1994). "Mixed-effects regression models for studying the natural history of prostate disease". In: *Statistics in Medicine* 13.5-7, pp. 587–601.

Rizopoulos, Dimitris (2012). *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*. CRC Press.

— (2016). "The R Package JMbayes for Fitting Joint Models for Longitudinal and Time-to-Event Data Using MCMC". In: *Journal of Statistical Software* 72.7, pp. 1–46.

Rizopoulos, Dimitris, Geert Molenberghs, and Emmanuel MEH Lesaffre (2017). "Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking". In: *Biometrical Journal* 59.6, pp. 1261–1276.

Schröder, FH et al. (1992). "The TNM classification of prostate cancer". In: *The Prostate* 21.S4, pp. 129–138.

Taylor, Jeremy M.G. et al. (2013). "Real-time individual predictions of prostate cancer recurrence using joint models". In: *Biometrics* 69.1, pp. 206–213.

Turnbull, Bruce W (1976). "The empirical distribution function with arbitrarily grouped, censored and truncated data". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 38.3, pp. 290–295.