Biometrics 00, 1–21 DOI: 10.1111/j.1541-0420.2005.00454.x
December 0000

Web-based Supplementary Materials for "Personalized Schedules for Surveillance of Cancer Patients"

Anirudh Tomer[1,*], Daan Nieboer[2], Monique J. Roobol[3], Ewout W. Steyerberg[2], 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
email: atomer@erasmusmc.nl

This paper has been submitted for consideration for publication in *Biometrics*

Web Appendix A. Joint Model for Time to Event and Longitudinal Outcomes

We start with the definition of the joint modeling framework that will be used to fit a model to the available dataset, and then to plan biopsies for future patients. Let T_i^* denote the true Gleason reclassification (referred to as GR hereafter) time for the i-th patient enrolled in an AS program. Let the vector of times at which biopsies are conducted for this patient be denoted by $T_i^b = \{T_{i0}^b, T_{i1}^b, \ldots, T_{iN_i^b}^b; T_{ij}^b < T_{ik}^b, \forall j < k\}$, where N_i^b are the total number of biopsies conducted. Because of the periodical nature of biopsy schedules, T_i^* cannot be observed directly and it is only known to fall in an interval $(l_i, r_i]$, where $l_i = T_{iN_i^b-1}^b, r_i = T_{iN_i^b}^b$ if GR is observed, and $l_i = T_{iN_i^b}^b, r_i = \infty$ if patient drops out of AS before GR is observed. Further let $\boldsymbol{y}_i$ denote the $n_i \times 1$ vector of PSA levels for the i-th patient. For a sample of n patients the observed data is denoted by $\mathcal{D}_n = \{l_i, r_i, \boldsymbol{y}_i; i = 1, \ldots, n\}$.

The longitudinal outcome of interest, namely PSA level, is continuous in nature and thus to model it the joint model utilizes a linear mixed effects model (LMM) of the form:

$$\begin{aligned} y_i(t) &= m_i(t) + \varepsilon_i(t) \\ &= \boldsymbol{x}_i^T(t)\boldsymbol{\beta} + \boldsymbol{z}_i^T(t)\boldsymbol{b}_i + \varepsilon_i(t), \end{aligned}$$

where $\boldsymbol{x}_i(t)$ denotes the row vector of the design matrix for fixed effects and $\boldsymbol{z}_i(t)$ denotes the same for random effects. Correspondingly the fixed effects are denoted by $\boldsymbol{\beta}$ and random effects by $\boldsymbol{b}_i$. The random effects are assumed to be normally distributed with mean zero and $q \times q$ covariance matrix $\boldsymbol{D}$. The true and unobserved PSA level at time t is denoted by $m_i(t)$. Unlike $y_i(t)$, the former is not contaminated with the measurement error $\varepsilon_i(t)$. The error is assumed to be normally distributed with mean zero and variance σ^2, and is independent of the random effects $\boldsymbol{b}_i$.

To model the effect of PSA on hazard of GR, joint models utilize a relative risk sub-model. The hazard of GR for patient i at any time point t, denoted by $h_i(t)$, depends on a function

of subject specific linear predictor $m_i(t)$ and/or the random effects:

$$h_i(t \mid \mathcal{M}_i(t), \boldsymbol{w}_i) = \lim_{\Delta t \to 0} \frac{\Pr\{T_i^* \epsilon [t, t+\Delta t) \mid T_i^* \geqslant t, \mathcal{M}_i(t), \boldsymbol{w}_i\}}{\Delta t}$$
$$= h_0(t) \exp\left[\boldsymbol{\gamma}^T \boldsymbol{w}_i + f\{M_i(t), \boldsymbol{b}_i, \boldsymbol{\alpha}\}\right], \quad t > 0,$$

where $\mathcal{M}_i(t) = \{m_i(v), 0 \leqslant v \leqslant t\}$ denotes the history of the underlying PSA levels up to time t. The vector of baseline covariates is denoted by $\boldsymbol{w}_i$, and $\boldsymbol{\gamma}$ are the corresponding parameters. The function $f(\cdot)$ parametrized by vector $\boldsymbol{\alpha}$ specifies the functional form of PSA levels (Brown, 2009; Rizopoulos, 2012; Taylor et al., 2013; Rizopoulos et al., 2014) that is used in the linear predictor of the relative risk model. Some functional forms relevant to the problem at hand are the following:

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

These formulations of $f(\cdot)$ postulate that the hazard of GR at time t may be associated with the underlying level $m_i(t)$ of the PSA at t, or with both the level and velocity $m_i'(t)$ of the PSA at t. Lastly, $h_0(t)$ is the baseline hazard at time t, and is modeled flexibly using P-splines. 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 parameters of the joint model using Markov chain Monte Carlo (MCMC) methods under the Bayesian framework. Let $\boldsymbol{\theta}$ denote the vector of the parameters of the

joint model. The joint model postulates that given the random effects, time to GR and longitudinal responses taken over time are all mutually independent. Under this assumption the posterior distribution of the parameters is given by:

$$\begin{aligned} p(\boldsymbol{\theta}, \boldsymbol{b} \mid \mathcal{D}_n) &\propto \prod_{i=1}^{n} p(l_i, r_i, \boldsymbol{y}_i \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) \\ &\propto \prod_{i=1}^{n} p(l_i, r_i \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{y_i} \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}), \\ p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) &= \frac{1}{\sqrt{(2\pi)^q \det(\boldsymbol{D})}} \exp(\boldsymbol{b}_i^T \boldsymbol{D}^{-1} \boldsymbol{b}_i), \end{aligned}$$

where the likelihood contribution of longitudinal outcome conditional on random effects is:

$$\begin{aligned} p(\boldsymbol{y_i} \mid \boldsymbol{b}_i, \boldsymbol{\theta}) &= \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^{n_i}} \exp\left(- \frac{\|\boldsymbol{y_i} - \boldsymbol{X}_i\boldsymbol{\beta} - \boldsymbol{Z}_i\boldsymbol{b}_i\|^2}{\sigma^2} \right), \\ \boldsymbol{X}_i &= \{\boldsymbol{x}_i(t_{i1})^T, \ldots, \boldsymbol{x}_i(t_{in_i})^T\}^T, \\ \boldsymbol{Z}_i &= \{\boldsymbol{z}_i(t_{i1})^T, \ldots, \boldsymbol{z}_i(t_{in_i})^T\}^T. \end{aligned}$$

The likelihood contribution of the time to GR outcome is given by:

$$p(l_i, r_i \mid \boldsymbol{b}_i, \boldsymbol{\theta}) = \exp\Big\{ - \int_0^{l_i} h_i(s \mid \mathcal{M}_i(s), \boldsymbol{w}_i) \mathrm{d}s \Big\} - \exp\Big\{ - \int_0^{r_i} h_i(s \mid \mathcal{M}_i(s), \boldsymbol{w}_i) \mathrm{d}s \Big\}. \quad (1)$$

The integral in (1) 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 effects $\boldsymbol{\beta}$, and inverse Gamma prior with shape and rate both equal to 0.01 for the parameter σ^2. 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 number q of the random effects. For the relative risk model's parameters $\boldsymbol{\gamma}$ and the association parameters $\boldsymbol{\alpha}$, we use a a global-local ridge-type shrinkage prior. For example, for the s-th element of $\boldsymbol{\alpha}$ we assume (similarly for $\boldsymbol{\gamma}$):

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

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. For the penalized version of the B-spline approximation to the baseline hazard, we use the following prior for parameters γ_{h_0} (Lang and Brezger, 2004):

$$p(\gamma_{h_0} \mid \tau_h) \propto \tau_h^{\rho(\boldsymbol{K})/2} \exp\left(-\frac{\tau_h}{2} \gamma_{h_0}^T \boldsymbol{K} \gamma_{h_0} \right),$$

where τ_h is the smoothing parameter that takes a Gamma(1, 0.005) hyper-prior in order to ensure a proper posterior for γ_{h_0}, $\boldsymbol{K} = \Delta_r^T \Delta_r + 10^{-6} \boldsymbol{I}$, where Δ_r denotes the r-th difference penalty matrix, and $\rho(\boldsymbol{K})$ denotes the rank of $\boldsymbol{K}$.

Web Appendix B. Derivations for Equation 6 and 7 of the Main Manuscript

In this section we present the derivations for Equation 6 and 7 of the main manuscript. To this end, we first expand the formula for dynamic survival probability presented in Equation 4 of the main manuscript.

$$\begin{aligned}\pi_j(u \mid t,s) &= \Pr\big\{T_j^* \geqslant u \mid T_j^* > t, \mathcal{Y}_j(s\big), D_n\big) \\ &= \int\int \Pr\big(T_j^* \geqslant u \mid T_j^* > t, \boldsymbol{b}_j, \boldsymbol{\theta}\big) p\big\{\boldsymbol{b}_j \mid T_j^* > t, \mathcal{Y}_j(s), \boldsymbol{\theta}\big\} p(\boldsymbol{\theta} \mid \mathcal{D}_n) \mathrm{d}\boldsymbol{b}_j \mathrm{d}\boldsymbol{\theta} \\ &= \int\int \frac{\exp\big\{-H_j(u|\boldsymbol{b}_j,\boldsymbol{\theta})\big\}}{\exp\big\{-H_j(t|\boldsymbol{b}_j,\boldsymbol{\theta})\big\}} p\big\{\boldsymbol{b}_j \mid T_j^* > t, \mathcal{Y}_j(s), \boldsymbol{\theta}\big\} p(\boldsymbol{\theta} \mid \mathcal{D}_n) \mathrm{d}\boldsymbol{b}_j \mathrm{d}\boldsymbol{\theta},\end{aligned} \tag{2}$$

where $H_j(u|\boldsymbol{b}_j,\boldsymbol{\theta}) = \int_0^u h_i(s \mid \boldsymbol{b}_j,\boldsymbol{\theta})\mathrm{d}s$ is the cumulative hazard up to time point u.

Web Appendix B.1 *Derivation of Equation 6 of the Main Manuscript*

$$E_g(T_j^*) = \int_t^\infty T_j^* g(T_j^*) \mathrm{d}T_j^*.$$

Using integration by parts, wherein $\mathrm{d}\big\{-\pi_j(T_j^* \mid t,s)\big\}/\mathrm{d}T_j^* = g(T_j^*)$,

$$\begin{aligned}E_g(T_j^*) &= \Big[-T_j^* \pi_j(T_j^* \mid t,s)\Big]_t^\infty + \int_t^\infty \pi_j(T_j^* \mid t,s) \frac{\mathrm{d}(T_j^*)}{\mathrm{d}T_j^*} \mathrm{d}T_j^* \\ &= t\pi_j(t \mid t,s) - \lim_{T_j^* \to \infty} T_j^* \pi_j(T_j^* \mid t,s) \\ &\quad + \int_t^\infty \pi_j(T_j^* \mid t,s) \mathrm{d}T_j^*,\end{aligned}$$

where $\pi_j(t \mid t,s) = \Pr\big\{T_j^* \geqslant t \mid T_j^* > t, \mathcal{Y}_j(s), D_n\big\} = 1$. As for $\lim_{T_j^* \to \infty} T_j^* \pi_j(T_j^* \mid t,s)$, the limit can be interchanged with the integral in Equation 2, because as $T_j^* \to \infty$ the integrand in the equation converges uniformly on the domain of $(\boldsymbol{b}_j,\boldsymbol{\theta})$. Thus,

$$\begin{aligned}\lim_{T_j^* \to \infty} T_j^* \pi_j(T_j^* \mid t,s) = \int\int \lim_{T_j^* \to \infty} &\frac{T_j^*}{\exp\big\{H_j(T_j^*|\boldsymbol{b}_j,\boldsymbol{\theta})\big\}} \\ &\times \frac{p\big\{\boldsymbol{b}_j \mid T_j^* > t, \mathcal{Y}_j(s), \boldsymbol{\theta}\big\} p(\boldsymbol{\theta} \mid \mathcal{D}_n)}{\exp\big\{-H_j(t|\boldsymbol{b}_j,\boldsymbol{\theta})\big\}} \mathrm{d}\boldsymbol{b}_j \mathrm{d}\boldsymbol{\theta}.\end{aligned}$$

Using L'Hospital's rule,

$$\begin{aligned}\lim_{T_j^*\to\infty} T_j^*\pi_j(T_j^* \mid t,s) &= \int\int \frac{1}{\lim_{T_j^*\to\infty}\exp\big\{H_j(T_j^*|\boldsymbol{b}_j,\boldsymbol{\theta})\big\}H_j'(T_j^*|\boldsymbol{b}_j,\boldsymbol{\theta})} \\ &\quad\times \frac{p\big\{\boldsymbol{b}_j \mid T_j^*>t,\mathcal{Y}_j(s),\boldsymbol{\theta}\big\}p(\boldsymbol{\theta}\mid\mathcal{D}_n)}{\exp\big\{-H_j(t|\boldsymbol{b}_j,\boldsymbol{\theta})\big\}}\mathrm{d}\boldsymbol{b}_j\mathrm{d}\boldsymbol{\theta} \\ &= \int\int 0\times\frac{p\big\{\boldsymbol{b}_j \mid T_j^*>t,\mathcal{Y}_j(s),\boldsymbol{\theta}\big\}p(\boldsymbol{\theta}\mid\mathcal{D}_n)}{\exp\big\{-H_j(t|\boldsymbol{b}_j,\boldsymbol{\theta})\big\}}\mathrm{d}\boldsymbol{b}_j\mathrm{d}\boldsymbol{\theta} \\ &= 0.\end{aligned}$$

In light of these results, we obtain:

$$E_g(T_j^*) = t + \int_t^\infty \pi_j(T_j^* \mid t,s)\mathrm{d}T_j^*.$$

Web Appendix B.2 *Derivation of Equation 7 of the Main Manuscript*

Since $\mathrm{var}_g(T_j^*) = E_g\{(T_j^*)^2\} - E_g(T_j^*)^2$, we first show the derivation for $E_g\{(T_j^*)^2\}$.

$$E_g\{(T_j^*)^2\} = \int_t^\infty (T_j^*)^2 g(T_j^*)\mathrm{d}T_j^*.$$

Using integration by parts, wherein $\mathrm{d}\big\{-\pi_j(T_j^* \mid t,s)\big\}/\mathrm{d}T_j^* = g(T_j^*)$,

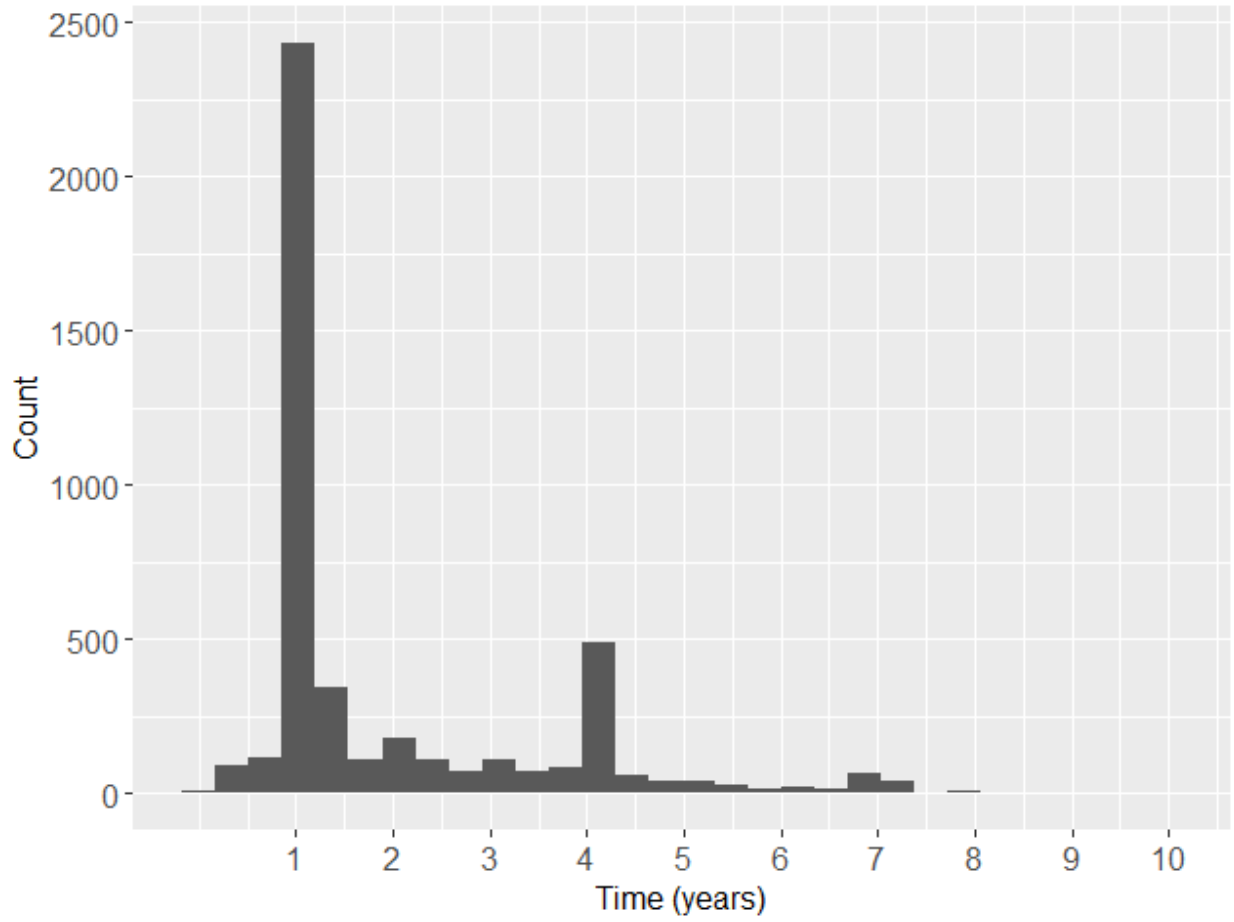
$$\begin{aligned}E_g\big\{(T_j^*)^2\big\} &= \Big[-(T_j^*)^2\pi_j(T_j^* \mid t,s)\Big]_t^\infty + \int_t^\infty \pi_j(T_j^* \mid t,s)\frac{\mathrm{d}(T_j^*)^2}{\mathrm{d}T_j^*}\mathrm{d}T_j^* \\ &= t^2\pi_j(t \mid t,s) - \lim_{T_j^*\to\infty}(T_j^*)^2\pi_j(T_j^* \mid t,s) \\ &\quad + 2\int_t^\infty T_j^*\pi_j(T_j^* \mid t,s)\mathrm{d}T_j^* \\ &= t^2 + 2\int_t^\infty T_j^*\pi_j(T_j^* \mid t,s)\mathrm{d}T_j^*.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{var}_g(T_j^*) &= t^2 + 2\int_t^\infty T_j^*\pi_j(T_j^* \mid t,s)\mathrm{d}T_j^* \\ &\quad - \left[t^2 + \Big\{\int_t^\infty \pi_j(T_j^* \mid t,s)\mathrm{d}T_j^*\Big\}^2 + 2t\int_t^\infty \pi_j(T_j^* \mid t,s)\mathrm{d}T_j^*\right] \\ &= 2\int_t^\infty (T_j^* - t)\pi_j(T_j^* \mid t,s)\mathrm{d}T_j^* - \Big\{\int_t^\infty \pi_j(T_j^* \mid t,s)\mathrm{d}T_j^*\Big\}^2.\end{aligned}$$

Web Appendix C. Non-compliance in PRIAS

The PRIAS schedule for biopsies is: one biopsy each at year one, year four, year seven and year ten (and every five years thereafter) after induction in AS. If at any time point a patient has PSA-DT less than 10 years, then biopsies are conducted annually. However, it has been observed by Bokhorst et al. (2015) that patients/doctors do not always comply with this biopsy schedule. Some of the reasons given by patients/doctors for non-compliance are: 'patient does not want biopsy', 'PSA stable', 'complications on last biopsy' and 'no signs of disease progression on previous biopsy'. Such non-compliance can lead to delays in detection of prostate cancer progression. To elucidate the issue of non-compliance in PRIAS, we show in Web Figure 1 a histogram of the repeat biopsies conducted for the PRIAS patients who never obtained GR (4560 out 5267 patients). It can be seen that the compliance for the first repeat biopsy around year one is much higher than the same at later years.



Web Figure 1. Histogram of repeat biopsies conducted for the 4560 out of 5267 PRIAS patients who did not obtain GR.

Web Appendix D. Parameter estimates for PRIAS dataset

The posterior parameter estimates for the joint model we fitted to the PRIAS dataset are shown in Web Table Web Table 1 (longitudinal sub-model) and Web Table Web Table 2 (relative risk sub-model), and parameter estimates for the variance-covariance matrix from the longitudinal sub-model are the following:

$$\boldsymbol{D} = \begin{bmatrix} 0.409 & 0.105 & -0.140 \\ 0.105 & 1.725 & 0.431 \\ -0.140 & 0.431 & 1.326 \end{bmatrix}$$

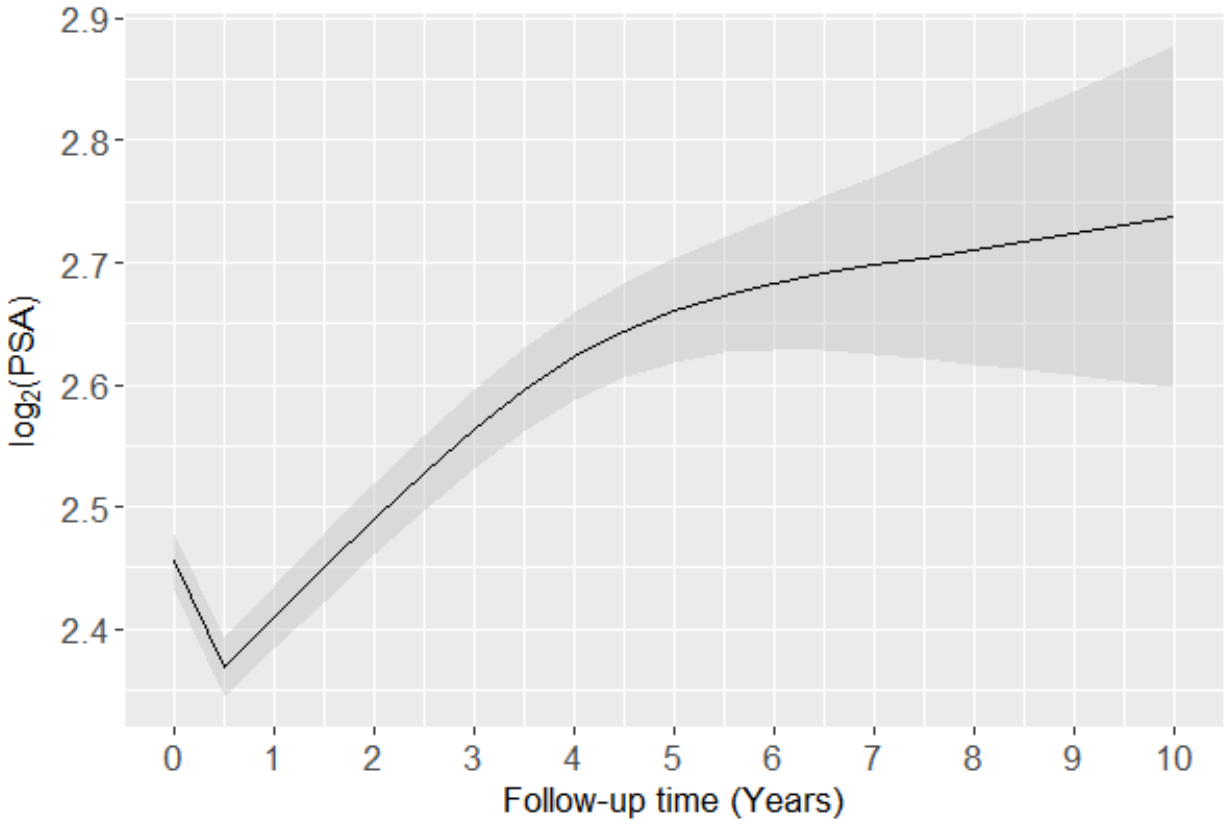
The effect of age only affects the baseline $\log_2$ PSA score. However it is so small that it can be ignored for all practical purposes. Since the longitudinal evolution of $\log_2$ PSA is modeled with non-linear terms, the interpretation of the coefficients corresponding to time is not straightforward. In lieu of the interpretation we present the fitted evolution of PSA (Web Figure 2) over a period of 10 years for a hypothetical patient.

Table Web Table 1

Longitudinal sub-model estimates for mean and 95% credible interval, for the joint model fitted to the PRIAS dataset.

	Mean	Std. Dev	2.5%	97.5%	P
Intercept	2.455	0.012	2.433	2.480	<0.000
(Age − 70)	0.003	0.001	4.9×10^{-4}	0.006	0.032
$(\text{Age} - 70)^2$	-0.001	1.4×10^{-4}	-0.001	-3.5×10^{-4}	<0.000
Spline: visitTimeYears[0.0, 0.1]	-0.006	0.012	-0.031	0.017	0.674
Spline: visitTimeYears[0.1, 0.5]	0.228	0.019	0.192	0.265	<0.000
Spline: visitTimeYears[0.5, 4.0]	0.140	0.029	0.088	0.197	<0.000
Spline: visitTimeYears[4.0, 7.0]	0.303	0.039	0.227	0.379	<0.000
σ	0.324	0.001	0.321	0.326	

For the relative risk sub-model, the parameter estimates in Web Table 2 show that only $\log_2$ PSA velocity is strongly associated with hazard of GR. For any patient, a unit increase in $\log_2$ PSA velocity corresponds to a 11 time increase in hazard of GR. The effect of $\log_2$ PSA value and effect of age on hazard of GR are small enough to be safely ignored for all practical purposes.



Web Figure 2. Fitted evolution of $\log_2$ PSA over a period of 10 years with 95% credible interval, for a patient who was inducted in AS at the Age of 70 years.

Table Web Table 2

Relative risk sub-model estimates for mean and 95% credible interval, for the joint model fitted to the PRIAS dataset.

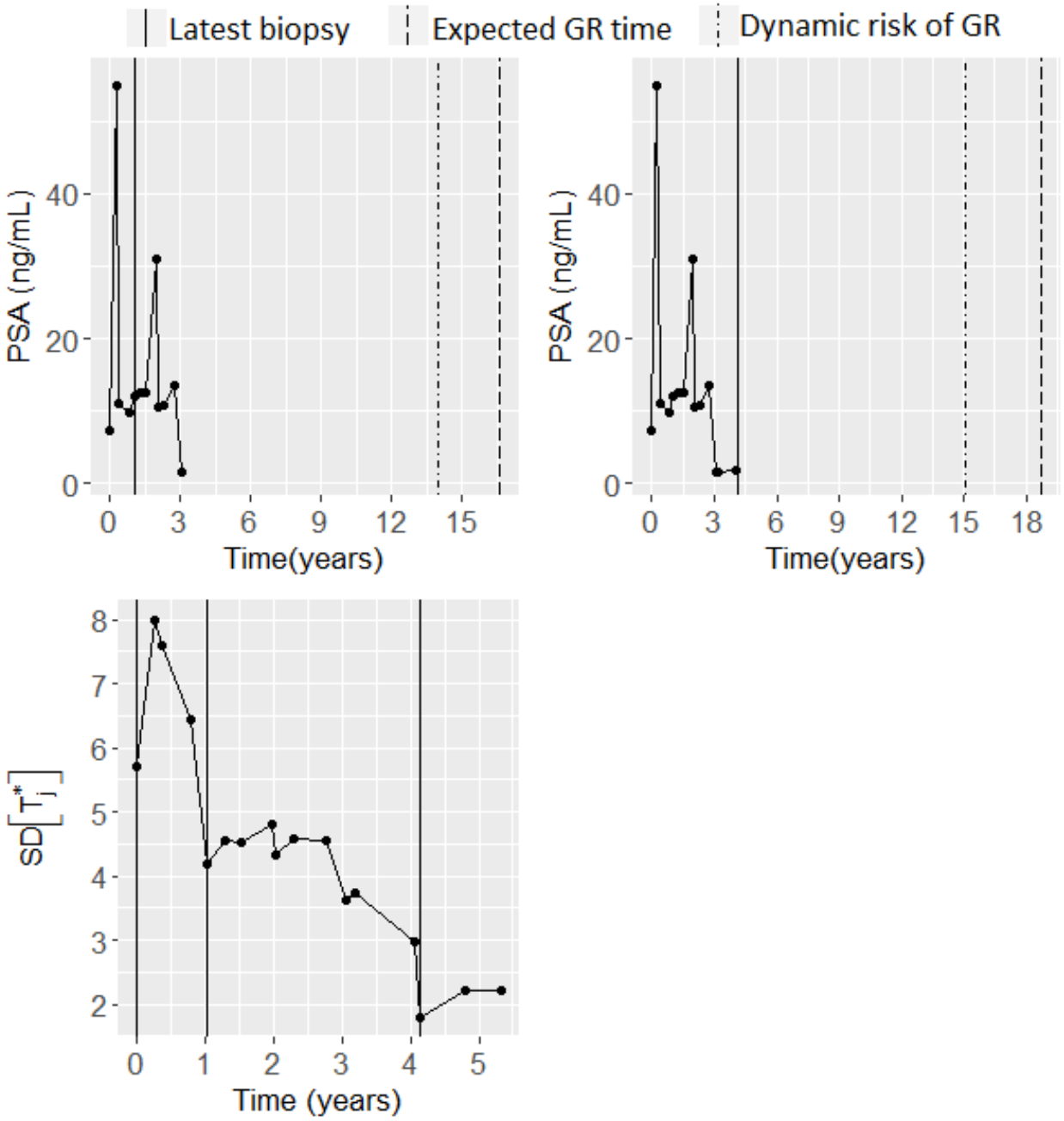
Variable	Mean	Std. Dev	2.5%	97.5%	P
$(\text{Age} - 70)$	0.037	0.006	0.025	0.0490	<0.000
$(\text{Age} - 70)^2$	-0.001	0.001	-0.003	1.8×10^{-4}	0.104
$\log_2$ PSA	-0.049	0.064	-0.172	0.078	0.414
Slope($\log_2$ PSA)	2.407	0.319	1.791	3.069	<0.000

Web Appendix E. Demonstration of Personalized Schedules for Patient 911 and Patient 2340

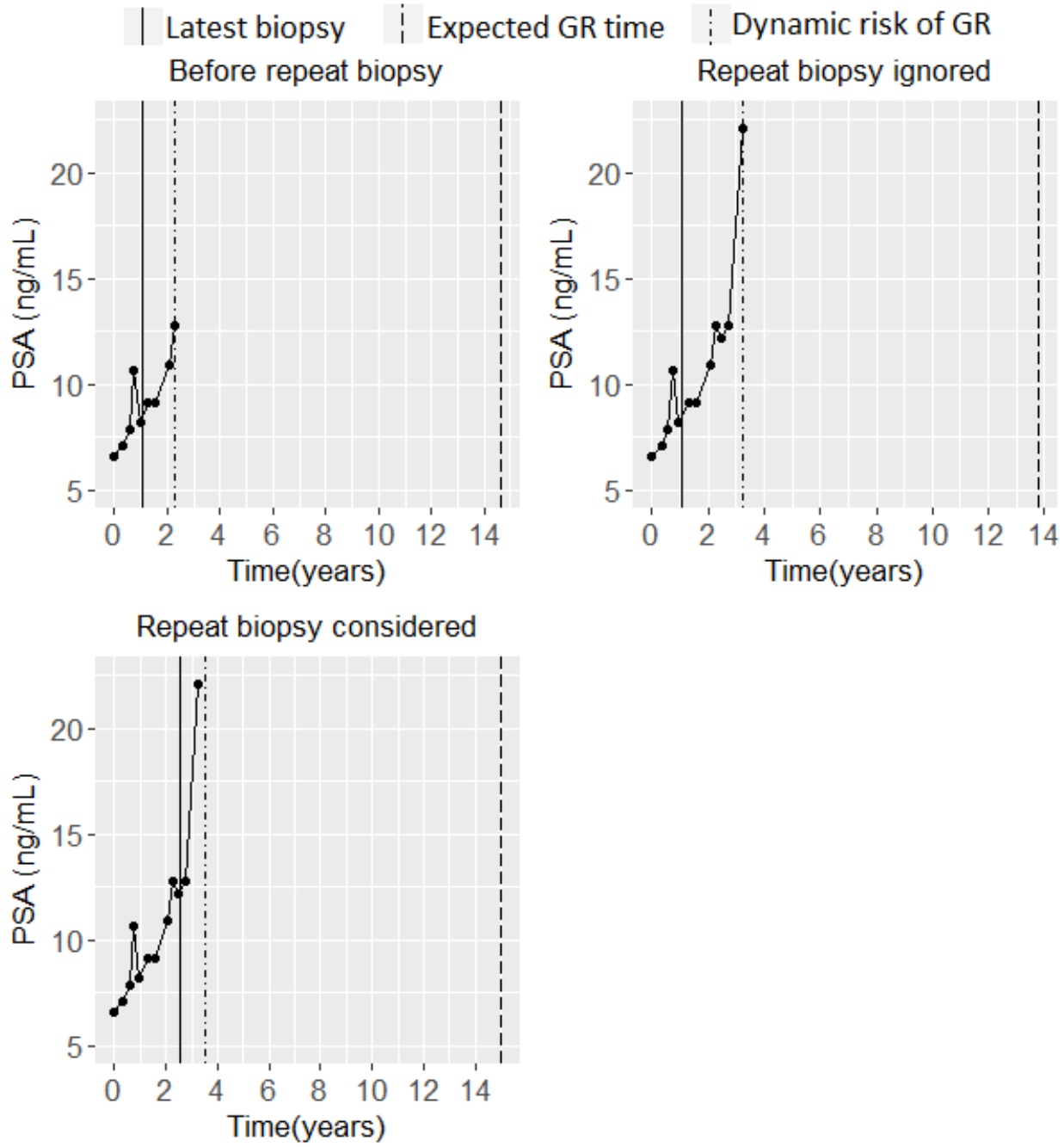
In this section we demonstrate the application of personalized schedules on patients from PRIAS. In Section 5.2 of the main manuscript we demonstrated personalized schedules for patient 3174. Here we demonstrate them for the remaining two patients.

The second patient of the demonstration data set is patient 911, for whom the evolution of PSA, time of last biopsy and proposed biopsy times are shown in the top panel of Web Figure 3. We can see the combined effect of decreasing PSA levels and a negative repeat biopsy on personalized schedules, between year three and year 4.5 for this patient. In accordance with the many negative repeat biopsies and consistently decreasing PSA, the proposed time of biopsy based on expected time of GR increases from 16.6 years to 18.8 years in this period. We can also see that after each repeat biopsy, $\text{SD}[T_j^*] = \sqrt{\text{var}_g(T_j^*)}$ decreases sharply (bottom panel of Web Figure 3), thus in turn reducing the offset as well. As for schedules based on dynamic risk of GR, in accordance with the many negative repeat biopsies and consistently decreasing PSA, the proposed time of biopsy increases from 14 years to 15 years during the time period of interest.

Patient 2340 presents a case where information from PSA levels and repeat biopsies is conflicting. In Web Figure 4 we can see that the PSA for this patient becomes twice between year two and year 3.2. If only information from PSA is considered, then we can see that proposed time of biopsy based on expected time of GR is preponed from 14.6 to 13.0 years during this period. However, if we also take into account the negative result from the repeat biopsy at year 2.5, then the proposed time of biopsy is postponed from 14.6 years to 15 years. Thus more weight is given to a recent negative biopsy result than PSA, which is in accordance with the clinical practice. The proposed time of biopsy based on dynamic risk of GR is also postponed from 2.3 to 3.6 years in light of the negative biopsy result.



Web Figure 3. Top panel: Evolution of PSA, history of repeat biopsies and corresponding personalized schedules for patient 911. Bottom Panel: History of repeat biopsies and $\mathrm{SD}_g(T_j^*) = \sqrt{\mathrm{var}_g(T_j^*)}$ over time for patient 911.



Web Figure 4. Evolution of PSA, history of repeat biopsies and corresponding personalized schedules for patient 2340.

Web Appendix F. Simulation study

Web Appendix F.1 *Results with κ chosen on the basis of Youden's J*

In the main manuscript, for the personalized schedules based on dynamic risk of GR we chose κ on the basis of F_1 score. However while conducting the simulation study, we also tried choosing it on the basis of Youden's J. Unlike F_1 score, Youden's J, is a measure of classification accuracy for both cases and controls. It is defined as:

$$\mathrm{J}(t, \Delta t, s) = \mathrm{TPR}(t, \Delta t, s) - \mathrm{FPR}(t, \Delta t, s), \quad \mathrm{J}\epsilon[-1, 1],$$

$$\mathrm{TPR}(t, \Delta t, s) = \Pr\big\{\pi_j(t + \Delta t \mid t, s) \leqslant \kappa \mid T_j^* \epsilon (t, t + \Delta t]\big\},$$

$$\mathrm{FPR}(t, \Delta t, s) = \Pr\big\{\pi_j(t + \Delta t \mid t, s) > \kappa \mid T_j^* > t + \Delta t\big\}.$$

where $\mathrm{TPR}(\cdot)$ and $\mathrm{FPR}(\cdot)$ denote time dependent true positive rate (sensitivity) and false positive rate (1−Specificity). The estimation for both is similar to the estimation of $\mathrm{AUC}(t, \Delta t, s)$ given by Rizopoulos et al. (2017). The optimal value of κ is $\arg\max_{\kappa} \mathrm{J}(t, \Delta t, s)$.

The simulation study results for Youden's J were not presented in the main manuscript for brevity and are presented here in Web Table 3. In addition results for a hybrid approach between median time of GR and dynamic risk of GR based on Youden's J are also presented.

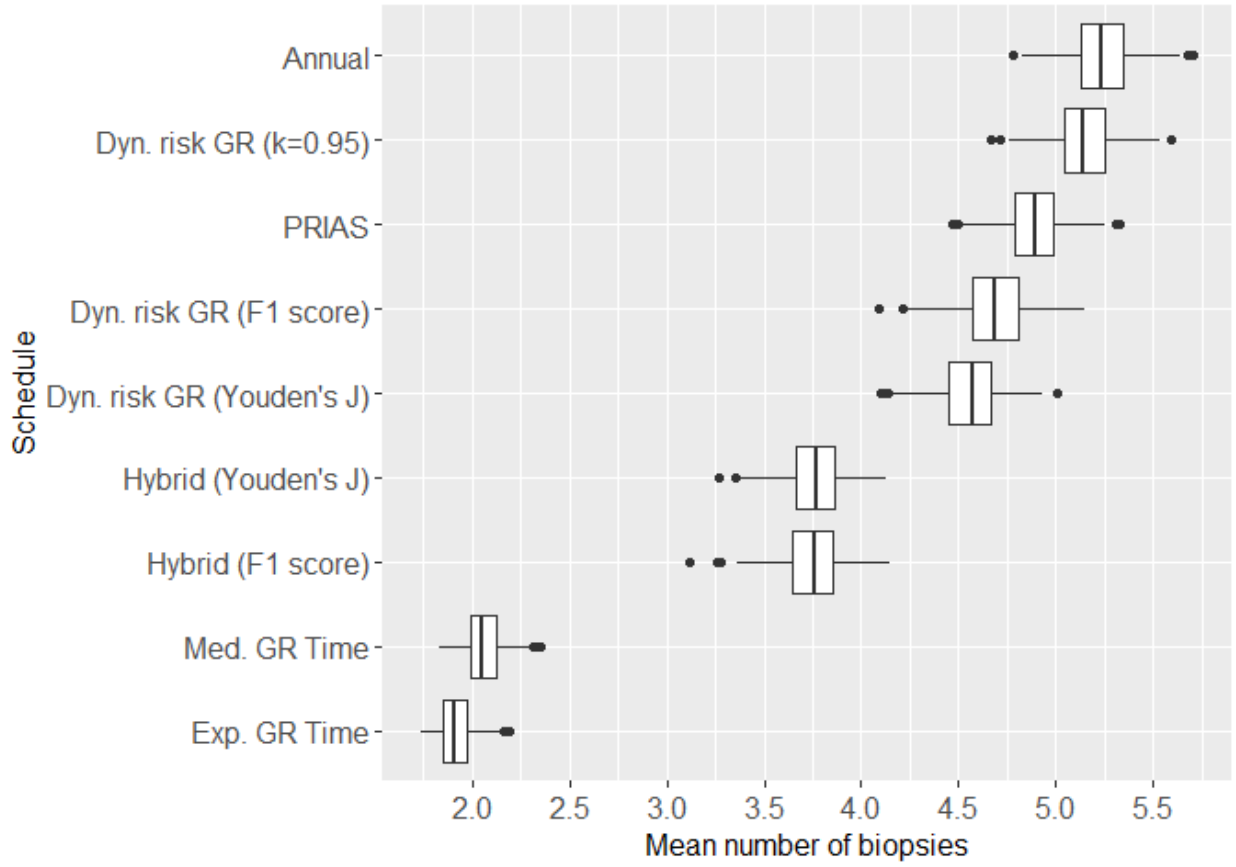
Web Appendix F.2 *Results for a fixed $\kappa = 0.95$*

We have also presented results for a fixed κ of 0.95, which means that the next biopsy is scheduled at a time point where risk of GR is 5%. The results for this approach are presented in Web Table 3.

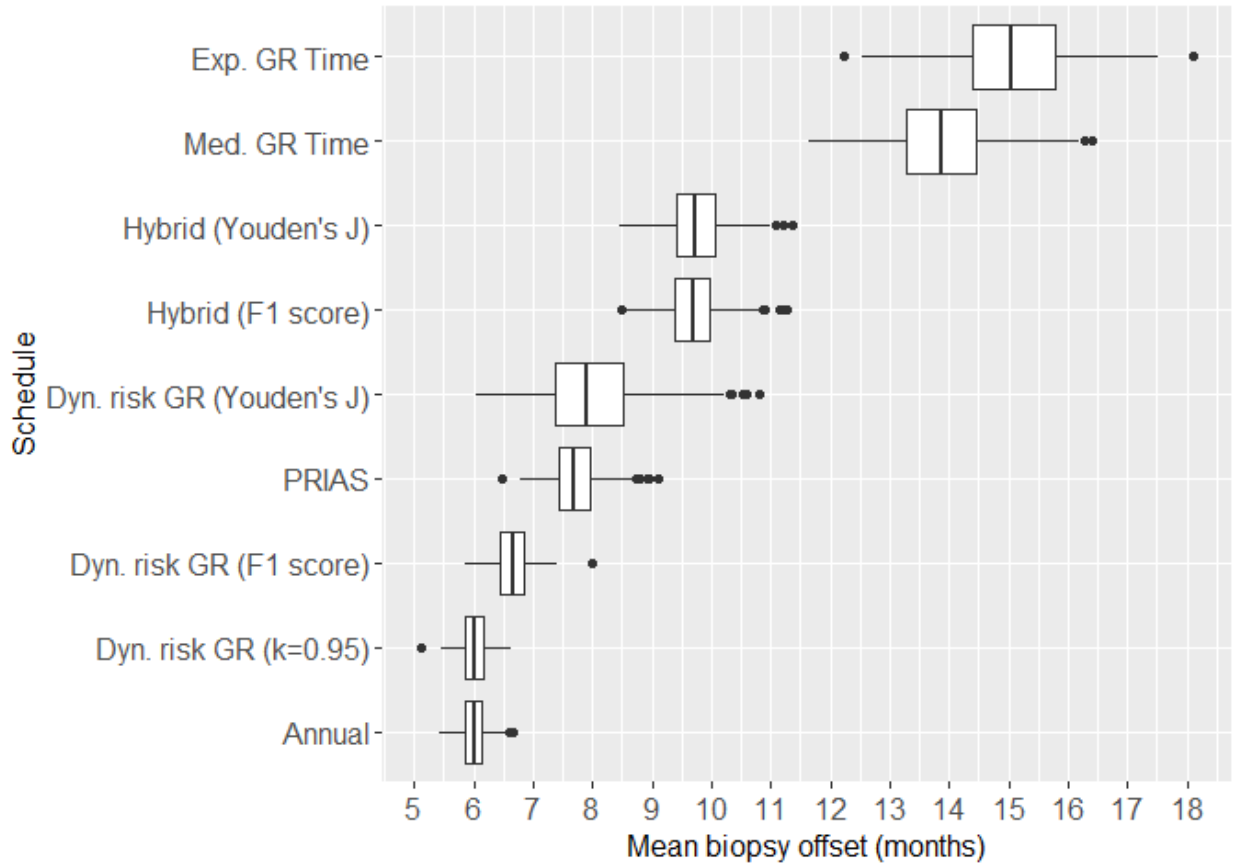
Web Appendix F.3 *Plots for the results of simulation study*

In this section we present figures related to the simulation study results discussed in Section 6 of main manuscript. The figures we present next are population specific, i.e. subgroup level differentiation is not done.

- Variation in estimated mean for number of biopsies and offset (months) for different methods is shown in Web Figure 5 and Web Figure 6.
- Variation in estimated standard deviation of number of biopsies and offset (months) for different methods is shown in Web Figure 7 and Web Figure 8.



Web Figure 5. Boxplot showing variation in estimated mean number of biopsies across the simulations, for different methods. Patients from all subgroups are considered.

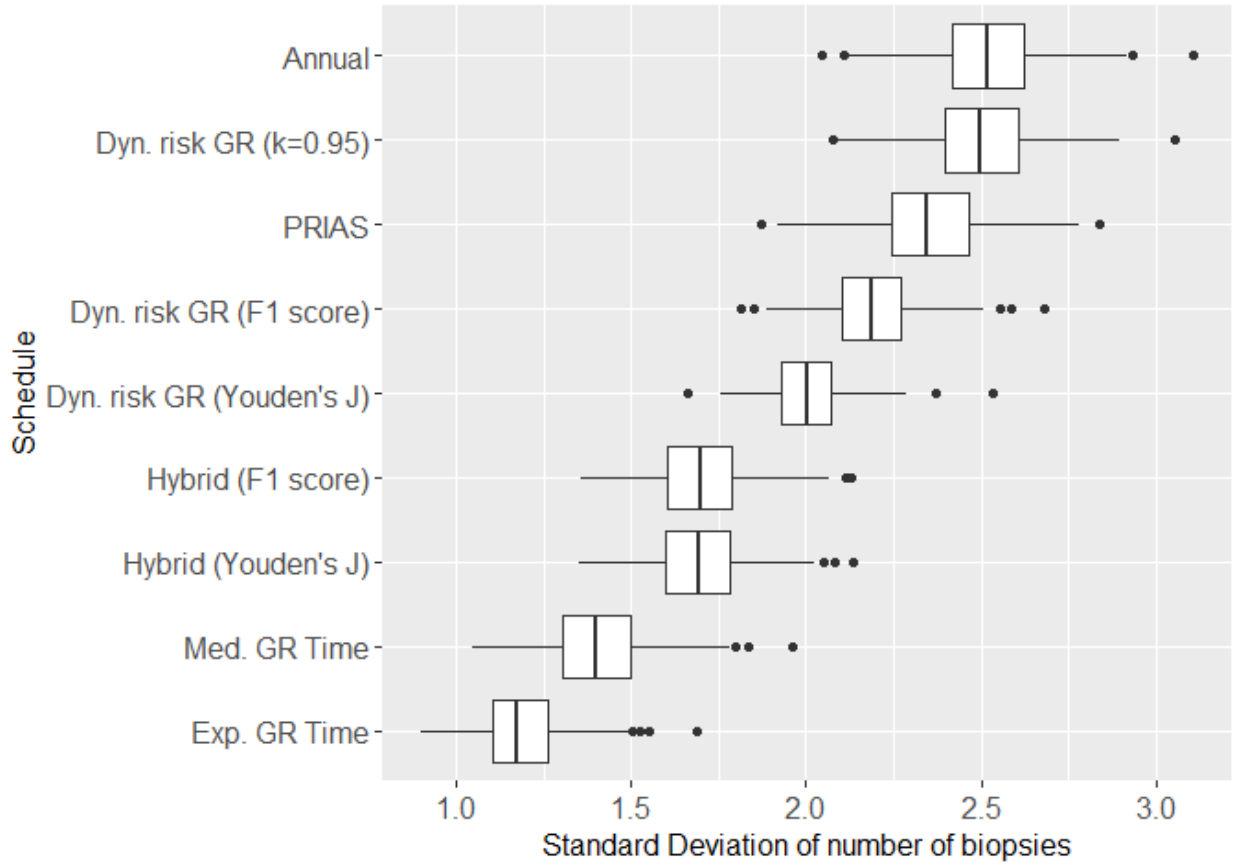


Web Figure 6. Boxplot showing variation in estimated mean of biopsy offset (months) across the simulations, for different methods. Patients from all subgroups are considered.

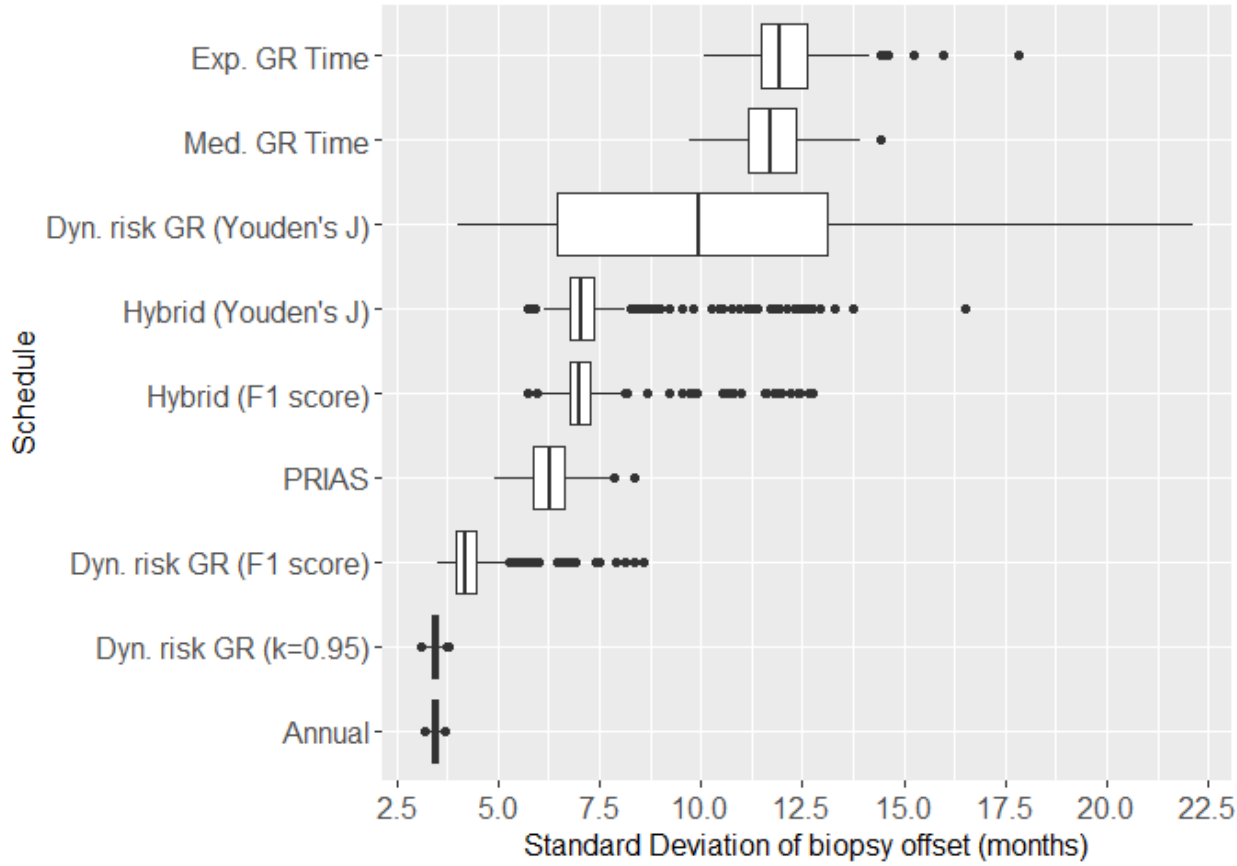
Table Web Table 3

Estimated mean and standard deviation of the number of biopsies and offset (months). Method names are abbreviated for consistency with Figure 3.

a) All subgroups: 124781 patients				
Schedule	$E[N^{bS}]$	$E(O_j^S)$	SD$[N^{bS}]$	SD$[O^S]$
Annual	5.24	6.01	2.53	3.46
PRIAS	4.90	7.71	2.36	6.31
Exp. GR time	1.92	15.08	1.19	12.11
Med. GR time	2.06	13.88	1.41	11.80
Dyn. risk GR (F_1 score)	4.69	6.66	2.19	4.38
Hybrid (F_1 score)	3.75	9.70	1.71	7.25
Dyn. risk GR (Youden's J)	4.56	8.00	2.00	11.00
Hybrid (Youden's J)	3.76	9.74	1.70	7.46
Dyn. risk GR ($\kappa = 0.95$)	5.15	6.02	2.51	3.47
b) Subgroup G_1: 41484 patients				
Schedule	$E[N^{bS}]$	$E(O_j^S)$	SD$[N^{bS}]$	SD$[O^S]$
Annual	4.32	6.02	3.13	3.44
PRIAS	4.07	7.44	2.88	6.11
Exp. GR time	1.72	21.65	1.47	14.75
Med. GR time	1.84	20.66	1.76	14.62
Dyn. risk GR (F_1 score)	3.85	6.75	2.69	4.44
Hybrid (F_1 score)	3.25	10.25	2.16	8.07
Dyn. risk GR (Youden's J)	3.73	8.93	2.41	14.56
Hybrid (Youden's J)	3.25	10.31	2.15	8.36
Dyn. risk GR ($\kappa = 0.95$)	4.23	6.05	3.10	3.46
c) Subgroup G_2: 41423 patients				
Schedule	$E[N^{bS}]$	$E(O_j^S)$	SD$[N^{bS}]$	SD$[O^S]$
Annual	5.18	5.98	2.13	3.47
PRIAS	4.85	7.70	2.00	6.29
Exp. GR time	1.77	13.54	0.98	9.83
Med. GR time	1.89	12.33	1.16	9.44
Dyn. risk GR (F_1 score)	4.63	6.66	1.82	4.37
Hybrid (F_1 score)	3.68	10.32	1.37	7.45
Dyn. risk GR (Youden's J)	4.52	7.67	1.67	9.57
Hybrid (Youden's J)	3.69	10.36	1.37	7.56
Dyn. risk GR ($\kappa = 0.95$)	5.09	5.99	2.11	3.47
d) Subgroup G_3: 41874 patients				
Schedule	$E[N^{bS}]$	$E(O_j^S)$	SD$[N^{bS}]$	SD$[O^S]$
Annual	6.20	6.02	1.76	3.46
PRIAS	5.76	7.98	1.71	6.51
Exp. GR time	2.27	10.09	0.99	7.47
Med. GR time	2.45	8.70	1.15	6.32
Dyn. risk GR (F_1 score)	5.58	6.58	1.56	4.33
Hybrid (F_1 score)	4.32	8.55	1.26	5.91
Dyn. risk GR (Youden's J)	5.42	7.40	1.44	7.61
Hybrid (Youden's J)	4.32	8.56	1.26	6.16
Dyn. risk GR ($\kappa = 0.95$)	6.11	6.01	1.76	3.46



Web Figure 7. Boxplot showing variation in estimated standard deviation of number of biopsies across the simulations, for different methods. Patients from all subgroups are considered.



Web Figure 8. Boxplot showing variation in estimated standard deviation of biopsy offset (months) across the simulations, for different methods. Patients from all subgroups are considered.

Web Appendix G. Source code

The source code for the joint model fitted to the PRIAS data set can be found at:

`https://goo.gl/phQkxG`

The source code for the simulation study can be found at:

`https://goo.gl/TpLTM8`.

References

Bokhorst, L. P., Alberts, A. R., Rannikko, A., Valdagni, R., Pickles, T., Kakehi, Y., Bangma, C. H., Roobol, M. J., study group PRIAS, et al. (2015). Compliance rates with the Prostate Cancer Research International Active Surveillance (PRIAS) protocol and disease reclassification in noncompliers. *European Urology* **68,** 814–821.

Brown, E. R. (2009). Assessing the association between trends in a biomarker and risk of event with an application in pediatric HIV/AIDS. *The Annals of Applied Statistics* **3,** 1163–1182.

Eilers, P. H. and Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science* **11,** 89–121.

Lang, S. and Brezger, A. (2004). Bayesian P-splines. *Journal of Computational and Graphical Statistics* **13,** 183–212.

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

Rizopoulos, D., Hatfield, L. A., Carlin, B. P., and Takkenberg, J. J. (2014). Combining dynamic predictions from joint models for longitudinal and time-to-event data using Bayesian model averaging. *Journal of the American Statistical Association* **109,** 1385–1397.

Rizopoulos, D., Molenberghs, G., and Lesaffre, E. M. (2017). Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biometrical Journal* doi:10.1002/bimj.201600238.

Taylor, J. M., Park, Y., Ankerst, D. P., Proust-Lima, C., Williams, S., Kestin, L., Bae, K., Pickles, T., and Sandler, H. (2013). Real-time individual predictions of prostate cancer recurrence using joint models. *Biometrics* **69,** 206–213.

Received October 0000. *Revised February* 0000. *Accepted March* 0000.