Biometrics 00, 1–17 DOI: 10.1111/j.1541-0420.2005.00454.x
December 0000

Web-based Supplementary Materials for "Personalized schedules for biopsies in prostate cancer patients"

John Author*, Jane Author, and Dick Author
Department of Statistics, University of Latex, Coventry CV4 7AL, U.K
**email:* author@address.edu

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 GR 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 that it falls 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 $\boldsymbol{D}$. $m_i(t)$ denotes the true and unobserved value of longitudinal outcome at time t, i.e. unlike $y_i(t)$ it is not contaminated with 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 longitudinal outcome 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:

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

where $\mathcal{M}_i(t) = \{m_i(v), 0 \leqslant v \leqslant t\}$ denotes the history of the underlying longitudinal process up to time t. $\boldsymbol{w}_i$ is a vector of baseline covariates and $\boldsymbol{\gamma}$ are the corresponding parameters. The function $f(\cdot)$ parametrized by vector $\boldsymbol{\alpha}$ specifies the functional form of longitudinal outcome (Brown, 2009; Rizopoulos, 2012; Taylor et al., 2013; Rizopoulos et al., 2014; Rizopoulos, 2016) that is used in the linear predictor of the relative risk model. Some functional forms relevant to the problem at hand and their interpretation 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 of the biomarker $m_i(t)$, or with both the level and slope of the longitudinal profile $m_i'(t)$ at time 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 Gleason reclassification (referred to as GR hereafter) and longitudinal responses taken over time are all mutually independent. Under this assumption the posterior distribution of the parameters is given by:

$$\begin{aligned} p(\boldsymbol{\theta}, \boldsymbol{b} \mid \mathcal{D}_n) &\propto \prod_{i=1}^{n} p(l_i, r_i, \boldsymbol{y}_i \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) \\ &\propto \prod_{i=1}^{n} p(l_i, r_i \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{y_i} \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) \end{aligned}$$

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

$$p(\boldsymbol{y_i} \mid \boldsymbol{b}_i, \boldsymbol{\theta}) = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^{n_i}} \exp\left\{ - \frac{\|\boldsymbol{y_i} - \boldsymbol{X}_i\boldsymbol{\beta} - \boldsymbol{Z}_i\boldsymbol{b}_i\|^2}{\sigma^2} \right\},$$

$$\boldsymbol{X}_i = \{\boldsymbol{x}_i(t_{i1})^T, \ldots, \boldsymbol{x}_i(t_{in_i})^T\}^T,$$

$$\boldsymbol{Z}_i = \{\boldsymbol{z}_i(t_{i1})^T, \ldots, \boldsymbol{z}_i(t_{in_i})^T\}^T$$

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

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

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 e.g. 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.

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

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 $\dfrac{\mathrm{d}\big\{-\pi_j(T_j^* \mid t,s)\big\}}{\mathrm{d}T_j^*} = g(T_j^*)$

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

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,

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

Using L'Hospital's rule

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

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 $\dfrac{\mathrm{d}\big\{-\pi_j(T_j^* \mid t, s)\big\}}{\mathrm{d}T_j^*} = g(T_j^*)$

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

Therefore,

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

Web Appendix C. Parameter estimates for PRIAS dataset

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

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

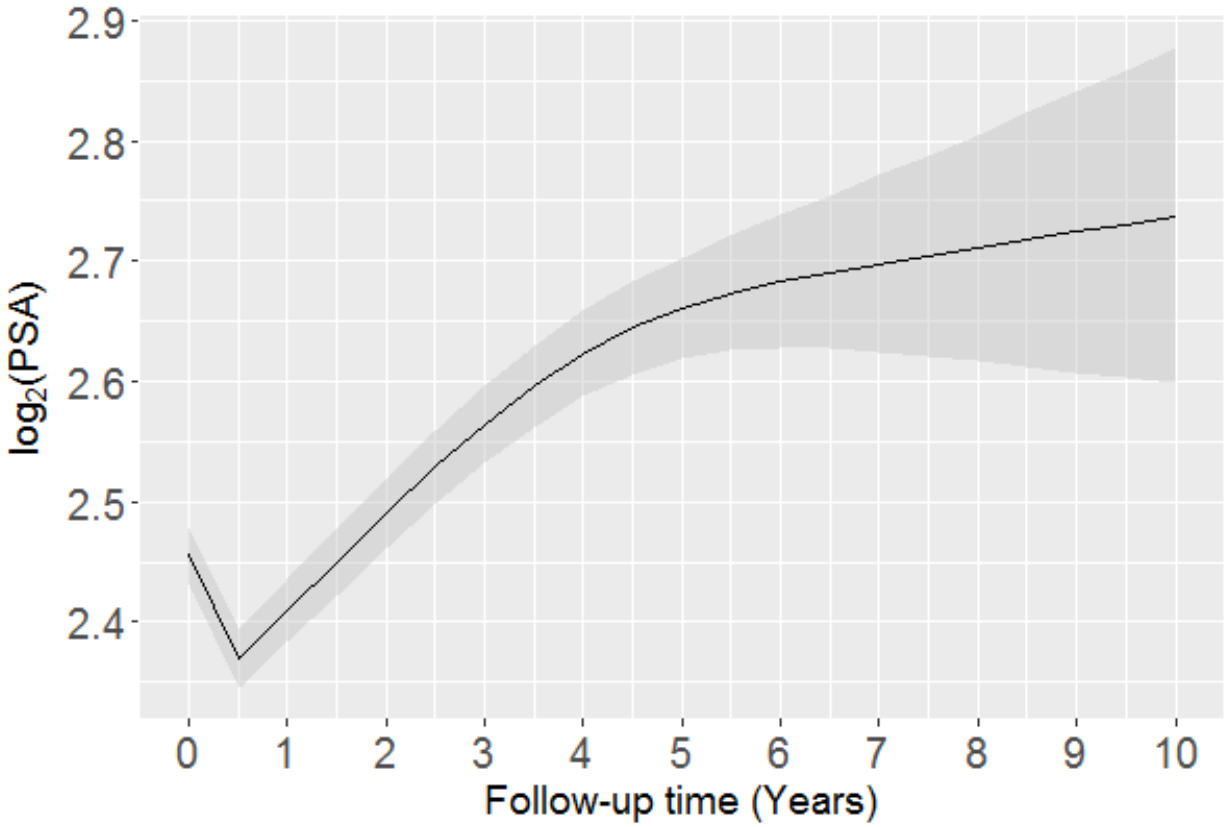
The effect of age only affects the baseline $\log_2$ PSA score. However it is so small that it can be ignored for all practical purposes. Since the longitudinal evolution of $\log_2$ PSA is modeled with non-linear terms, the interpretation of the coefficients corresponding to time is not straightforward. In lieu of the interpretation we present the fitted evolution of PSA (Web Figure 1) over a period of 10 years for a hypothetical patient.

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 1. 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 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
(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 D. Simulation study

Web Appendix D.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:

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

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

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

where TPR$(\cdot)$ and FPR$(\cdot)$ denote time dependent true positive rate (sensitivity) and false positive rate (1 − Specificity). The estimation for both proceeds as in (Rizopoulos, 2016). The optimal value of κ is $\arg\max_\kappa J(t, \Delta t, s)$.

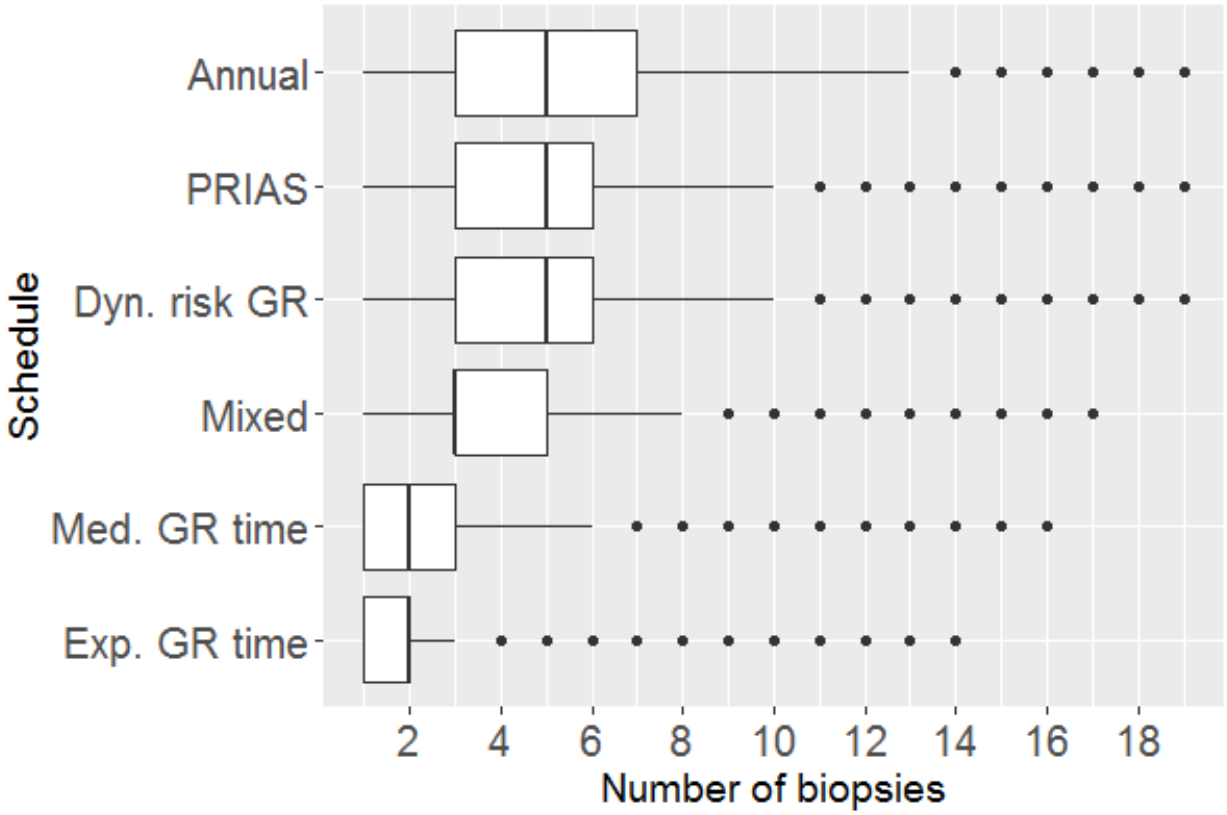
The simulation study results for Youden's J were not presented in the main manuscript for brevity and are presented here in Web Table 3. In addition results for a mixed approach between median time of GR and dynamic risk of GR based on Youden's J are also presented.

Web Appendix D.2 *Plots for the results of simulation study*

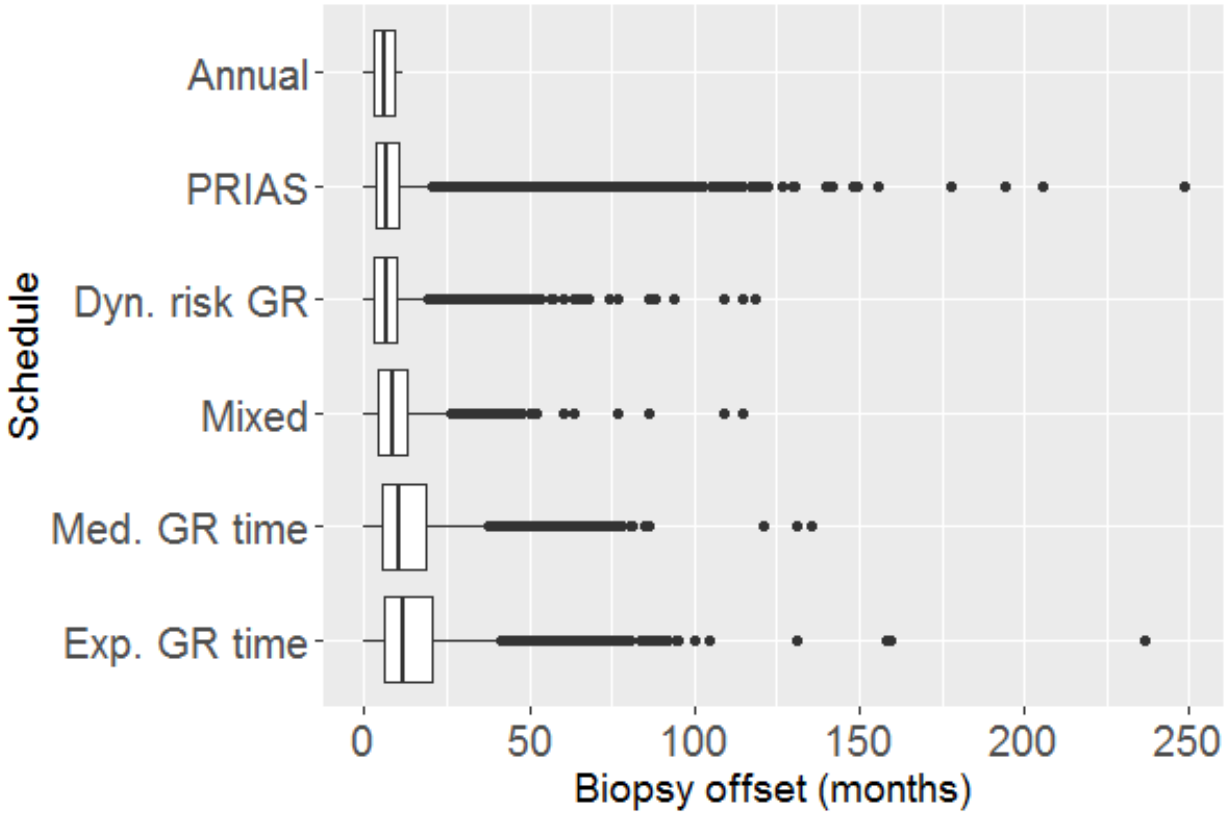
In this section we present figures related to the simulation study results discussed in Section 6 of main manuscript. The figures we present next are population specific, i.e. subgroup level differentiation is not done.

- Web Figure 2 and Web Figure 3 show the variation in number of biopsies and biopsy offset (months) for different methods.
- Variation in estimated mean for number of biopsies and offset (months) for different methods is shown in Web Figure 4 and Web Figure 6.

- Variation in estimated variance for number of biopsies and offset (months) for different methods is shown in Web Figure 5 and Web Figure 7.



Web Figure 2. Boxplot showing variation in number of biopsies conducted by different methods. Patients from all subgroups are considered.

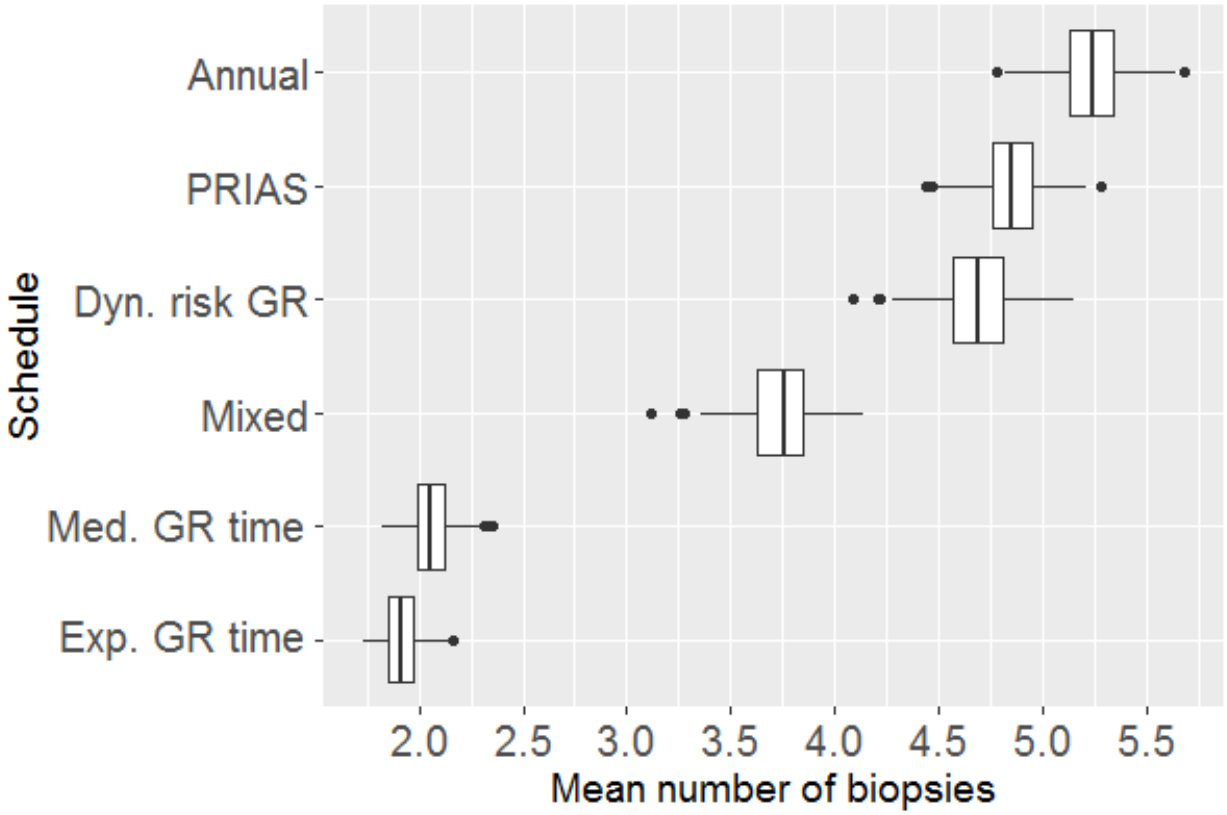


Web Figure 3. Boxplot showing variation in biopsy offset (months) for different methods. Patients from all subgroups are considered.

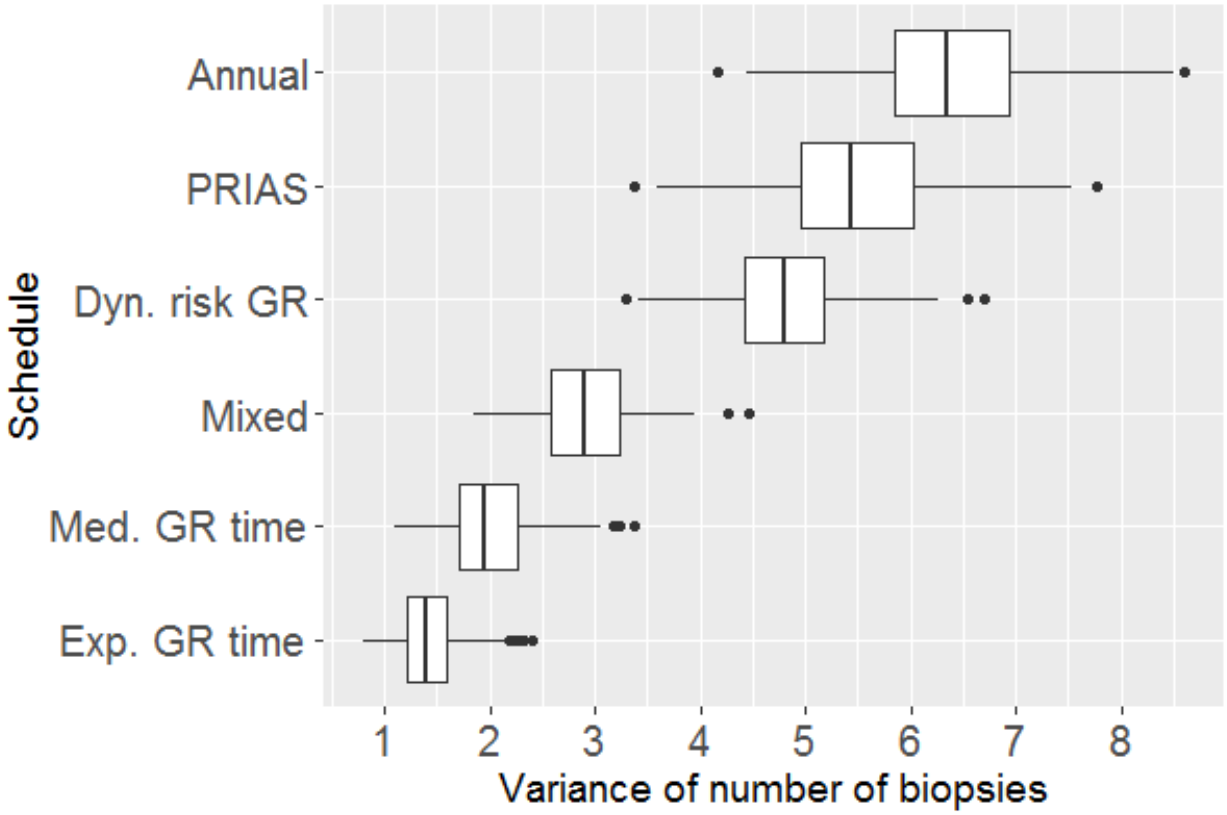
Table 3

Estimated mean and standard deviation of the number of biopsies and offset (months). Method names are abbreviated for consistency with Figure 6.

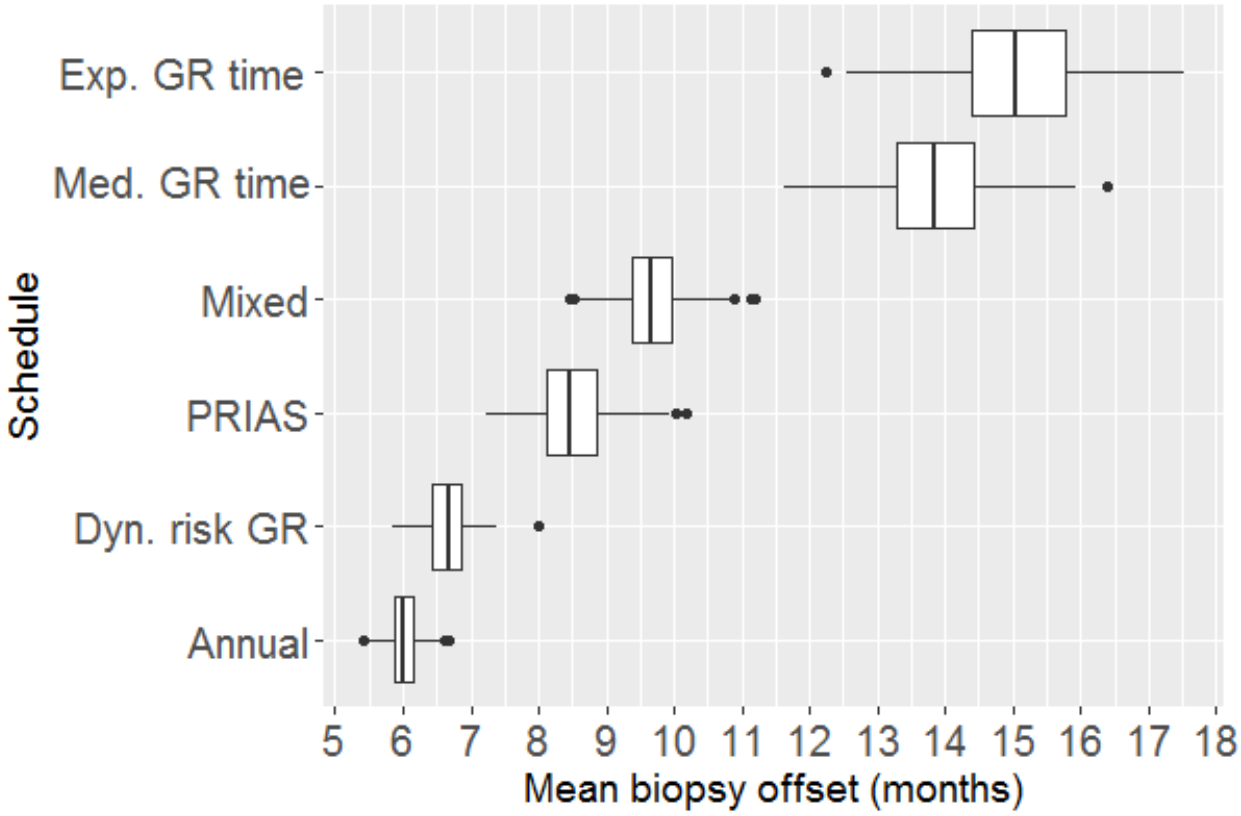
Schedule	$E[N^{bS}]$	$E[O^S]$	SD$[N^{bS}]$	SD$[O^S]$
a) All subgroups: 101823 patients				
Annual	5.24	6.01	2.53	3.45
PRIAS	4.86	8.49	2.35	8.69
Exp. GR time	1.92	15.06	1.19	12.11
Med. GR time	2.07	13.87	1.42	11.80
Dyn. risk GR (F_1 score)	4.69	6.66	2.20	4.37
Mixed (F_1 score)	3.75	9.69	1.71	7.02
Dyn. risk GR (Youden's J)	4.56	8.04	2.00	11.08
Mixed (Youden's J)	3.76	9.78	1.70	7.71
b) Subgroup G_1: 33680 patients				
Schedule	$E[N^{bS}]$	$E[O^S]$	SD$[N^{bS}]$	SD$[O^S]$
Annual	4.33	6.02	3.14	3.44
PRIAS	4.05	7.98	2.87	8.08
Exp. GR time	1.72	21.65	1.47	14.77
Med. GR time	1.85	20.67	1.77	14.64
Dyn. risk GR (F_1 score)	3.85	6.76	2.69	4.45
Mixed (F_1 score)	3.24	10.24	2.17	7.73
Dyn. risk GR (Youden's J)	3.73	9.02	2.41	14.78
Mixed (Youden's J)	3.26	10.4	2.16	8.80
c) Subgroup G_2: 33907 patients				
Schedule	$E[N^{bS}]$	$E[O^S]$	SD$[N^{bS}]$	SD$[O^S]$
Annual	5.18	5.99	2.13	3.48
PRIAS	4.82	8.57	1.99	8.65
Exp. GR time	1.78	13.53	0.98	9.82
Med. GR time	1.90	12.31	1.16	9.43
Dyn. risk GR (F_1 score)	4.63	6.66	1.82	4.34
Mixed (F_1 score)	3.68	10.30	1.38	7.17
Dyn. risk GR (Youden's J)	4.52	7.67	1.67	9.42
Mixed (Youden's J)	3.70	10.42	1.36	7.77
d) Subgroup G_3: 34236 patients				
Schedule	$E[N^{bS}]$	$E[O^S]$	SD$[N^{bS}]$	SD$[O^S]$
Annual	6.20	6.01	1.77	3.46
PRIAS	5.70	8.92	1.73	9.27
Exp. GR time	2.27	10.11	0.99	7.53
Med. GR time	2.45	8.71	1.15	6.36
Dyn. risk GR (F_1 score)	5.57	6.58	1.56	4.32
Mixed (F_1 score)	4.31	8.54	1.27	5.91
Dyn. risk GR (Youden's J)	5.42	7.45	1.44	7.81
Mixed (Youden's J)	4.32	8.55	1.25	6.19



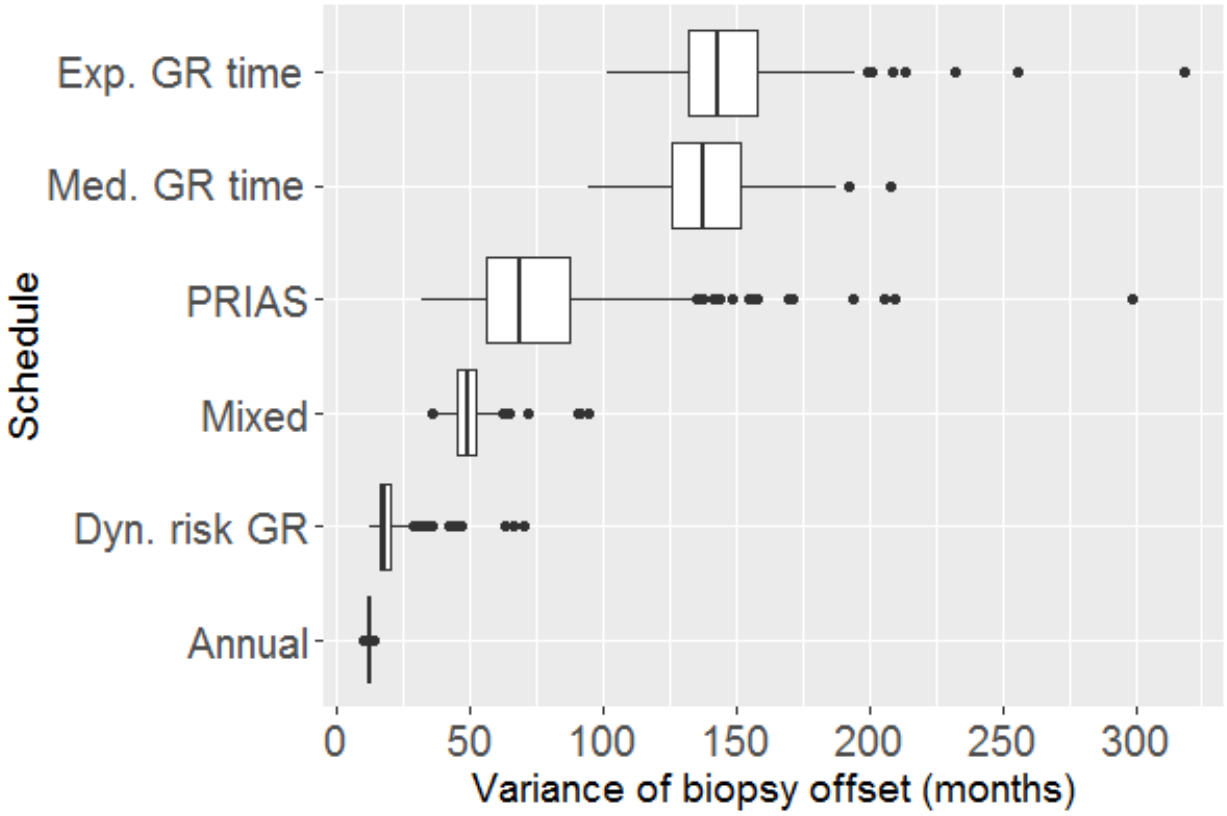
Web Figure 4. Boxplot showing variation in estimated mean number of biopsies across the simulations, for different methods. Patients from all subgroups are considered.



Web Figure 5. Boxplot showing variation in estimated variance of 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.



Web Figure 7. Boxplot showing variation in estimated variance of biopsy offset (months) across the simulations, for different methods. Patients from all subgroups are considered.

References

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. (2016). The r package jmbayes for fitting joint models for longitudinal and time-to-event data using mcmc. *Journal of Statistical Software* **72,** 1–46.

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.

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.