BIOMETRICS 00, 1–10
December 0000

DOI: 10.1111/j.1541-0420.2005.00454.x

# Personalized Schedules for Prostate Cancer Biopsies

**Aniruddh 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

SUMMARY: Low risk prostate cancer patients enrolled in active surveillance (AS) programs have to undergo biopsies on a frequent basis for examination of disease progression. Majority of the AS programs worldwide employ fixed schedules of biopsies for all patients. It has been found that such fixed and frequent schedules discourage patients to receive biopsies, and also bring a financial burden on the healthcare systems. Motivated by the world's largest AS program PRIAS, in this paper we present personalized schedules for biopsies to counter these problems. Using joint models for time to event and longitudinal data, our methods combine information from previous biopsy results and historical prostate-specific antigen (PSA) levels of a patient, to schedule the next biopsy. We also present criteria to compare the efficacy of personalized schedules with that of existing biopsy schedules, and a method to select the optimal schedule.

KEY WORDS: Active surveillance; Joint models; Personalized medicine; Prostate cancer;.

## 1. Introduction

In this decade prostate cancer is the second most frequently diagnosed cancer (14% of all cancers) in males worldwide, with nearly 67% of all prostate cancer cases reported in developed countries (Torre et al., 2015). The increase in diagnosis of low grade prostate cancers has been attributed to increase in life expectancy and increase in number of screening programs (Potosky et al., 1995). A major issue of screening programs that also has been established in other types of cancers (e.g., breast cancer) is over-diagnosis. To avoid overtreatment, patients diagnosed with low grade prostate cancer are often motivated to join active surveillance (AS) programs. The goal of AS is to routinely examine the progression of prostate cancer and avoid serious treatments such as surgery or chemotherapy as long as they are not needed.

Currently the largest AS program worldwide is Prostate Cancer Research International Active Surveillance, also known as PRIAS (Bokhorst et al., 2016). Patients enrolled in PRIAS are closely monitored using serum prostate-specific antigen (PSA) levels, digital rectal examination (DRE) and repeat prostate biopsies. Results from biopsies are graded on a scale called Gleason, which takes values between 2 and 10, with 10 corresponding to a serious state of the disease. At the time of induction in PRIAS, patients must have a Gleason score of six or less, DRE score of cT2c or less and a PSA of 10 ng/mL or less. In PRIAS, a PSA doubling time (PSA-DT) of less than three years indicates prostate cancer progression, where PSA-DT is measured as the inverse of the slope of regression line through the base two logarithm of PSA values. However until either DRE or Gleason are observed to be higher than the aforementioned thresholds, patients are not removed from AS for curative treatment (Bokhorst et al., 2016). When the Gleason score becomes greater than six, it is also known as Gleason reclassification (referred to as GR hereafter).

Biopsies are reliable but they are also difficult to conduct, cause pain and have serious side effects such as hematuria and sepsis (Loeb et al., 2013). Due to these reasons, majority of the AS programs worldwide strongly advise at most one biopsy per year. Conducting a biopsy annually (we refer to it as annual schedule hereafter) has the advantage that GR can be detected within one year of its occurrence. This may work well for patients with a faster progressing disease, however for patients with a slower progressing disease many unnecessary biopsies are scheduled. The drawbacks of unnecessary biopsies are not only medical but also financial. Keegan et al. (2012) have shown that annual schedules can cost more than the treatment (brachytherapy or prostatectomy) to AS programs, and if biopsies were to be conducted every other year, then up to 28% increase in savings per head from AS over treatment could be achieved. Despite this, several AS programs employ the annual schedule (Tosoian et al., 2011; Welty et al., 2015).

Scheduling frequent biopsies for patients has also lead to a high non-compliance rate for repeat biopsies (Bokhorst et al., 2015), which reduces the effectiveness of AS programs because progression is detected late. In PRIAS some of the reasons reported for non-compliance were: 'patient does not want biopsy', 'complications on last biopsy' and 'no signs of disease progression on previous biopsy'. The PRIAS schedule and the compliance rates are: one biopsy each at year one (81%), year four (60%), year seven (53%) and year 10 (33%). After year 10 biopsies are conducted every five years. An exception is made, if at any time a patient has PSA-DT less than 10 years, wherein the annual schedule is prescribed. It is important to note that, unlike biopsies, measurement of PSA has a high compliance rate of 91% in PRIAS. This is because PSA is easy to obtain and measuring PSA does not lead to any side effects. The use of PSA-DT in the scheduling process is justified, because it was found to be indicative of GR in PRIAS (Bokhorst et al., 2015).

This paper is motivated by the need to reduce the burden of biopsies and most optimally find the onset of GR. To this end, we intend to create personalized schedules for biopsies which improve upon the PRIAS and annual schedule. That is, a different schedule for every patient utilizing his recorded information. Personalized schedules for screening have received much interest in the literature, especially in the medical decision making context. For diabetic retinopathy, cost optimized personalized schedules based on Markov models have been developed by Bebu and Lachin (2017). For breast cancer, personalized mammography screening policy based on the prior screening history and personal risk characteristics of women, using partially observable Markov decision process (MDP) models have been proposed by Ayer, Alagoz, and Stout (2012). MDP models have also been used to develop personalized screening policies for cervical cancer (Akhavan-Tabatabaci, Sánchez, and Yeung, 2017) and colorectal cancer (Erenay, Alagoz, and Said, 2014). Another type of model called joint model for time to event and longitudinal data (Tsiatis and Davidian, 2004; Rizopoulos, 2012) has also been used to create personalized schedules, albeit for longitudinal biomarkers (Rizopoulos et al., 2016). In context of prostate cancer, Zhang et al. (2012) have used partially observable MDP models to personalize the decision of (not) deferring a biopsy to the next checkup time during the screening process. The decision is based on the baseline characteristics as well as a discretized PSA level of the patient at the current check up time.

Our work differs from the above referenced work in certain aspects. Firstly, the schedules we propose in this paper, account for the latent between-patient heterogeneity. We achieve this using joint models, which are inherently patient-specific because they utilize random effects. Secondly, joint models allow a continuous time scale and utilize the entire history of PSA levels. Lastly, instead of making a binary decision of (not) deferring a biopsy to the next check up time, we schedule biopsies at a per patient optimal future time, utilizing the historical PSA measurements and repeat biopsy results of the patient up to the current check up time. To this end, using joint models we first obtain a full specification of the joint distribution of PSA levels and time of GR. We then use it to define a patient-specific posterior predictive distribution of the time of GR given the observed PSA measurements and previous biopsies. Using the general framework of Bayesian decision theory, we propose a set of loss functions which are minimized to find the optimal time of conducting a biopsy. These loss functions yield us two categories of personalized schedules, those based on expected time of GR and those based on the risk of GR. We also analyze an approach where the two types of schedules are combined. To compare the proposed personalized schedules with the PRIAS and annual schedule we conduct a simulation study, and then discuss various criteria for evaluating the efficacy of each schedule, and a method to choose the most suitable one.

The rest of the paper is organized as follows. Section 2 briefly covers the joint modeling framework. Section 3 details the personalized scheduling approaches we have proposed in this paper. In Section 4 we discuss criteria for evaluation of the efficacy of a schedule and the choice of the optimal schedule. In Section 5 we demonstrate the personalized schedules by employing them for the patients from the PRIAS program. Lastly, in Section 6, we present the results from a simulation study we conducted to compare personalized schedules with PRIAS and annual schedule.





## 2. 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 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 $\sigma^2$, 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:

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

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.

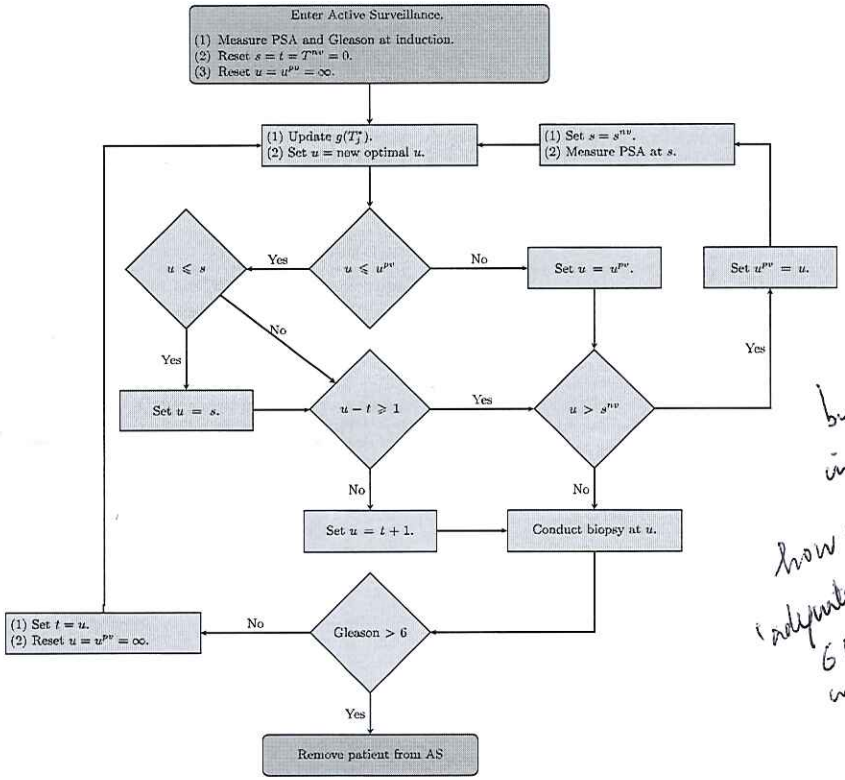


**Figure 1.** Algorithm for creating a personalized schedule for patient $j$. $t$ denotes the time of the latest biopsy. $s$ denotes the time of the latest available PSA measurement. $u$ denotes the proposed personalized time of biopsy. $u^{pv}$ denotes the time at which a repeat biopsy was proposed on the last visit to the hospital. $T^{nv}$ denotes the time of the next visit for measurement of PSA.

### 4.1 *Evaluation of Schedules*

We measure the efficacy of a schedule $S$ using two criteria, namely, for the $j$-th patient the number of biopsies $N_j^{bS} \epsilon [1, \infty]$ a schedule conducts before GR is detected, and the offset $O_j^S \epsilon [0, \infty]$ by which it overshoots the true GR time $T_j^*$. The offset $O_j^S$ is defined as $O_j^S = T^S_{jN_j^{bS}} - T_j^*$, where $T^S_{jN_j^{bS}} \geqslant T_j^*$ is the time at which GR is detected. Our interest lies in the joint distribution $p(N_j^{bS}, O_j^S)$ of the number of biopsies and the offset. Given the medical and financial burden associated with biopsies, ideally only one biopsy which leads to a zero offset should be conducted. That is, a method with a low mean number of biopsies $E(N_j^{bS})$ as well a low mean offset $E(O_j^S)$ is desired. It is also desired that a method has low variance of the number of biopsies $\text{var}(N_j^{bS})$, as well as low variance of the offset $\text{var}(O_j^S)$, so that the method works similarly for most patients. Quantiles of $p(N_j^{bS})$ may also be of interest. For example, a schedule which conducts less than two biopsies in 95% of the cases may be preferred.

### 4.2 *Finding the Optimal Schedule*

Given the multiple measures of efficacy of a schedule, the next step is to find the optimal schedule. Using principles from compound optimal designs (Läuter, 1976) we propose to choose a schedule $S$ which minimizes a loss function of the



following form:

$$L(S) = \sum_{r=1}^{R} \eta_r \mathcal{R}_r(N_j^{bS}). \quad (8)$$

where $\mathcal{R}_r(\cdot)$ is a measure of efficacy of either the number of biopsies or the offset (in the equation above, only $N_j^{bS}$ is used for brevity of notation). Some examples of $\mathcal{R}_r(\cdot)$ are mean, median, variance and quantile function. Constants $\eta_1, \ldots, \eta_R$, where $\eta_r \epsilon [0, 1]$ and $\sum_{r=1}^{R} \eta_r = 1$, are weights to differentially weigh-in the contribution of each of the $R$ measures of efficacy. An example loss function is:

$$L(S) = \eta_1 E(N_j^{bS}) + \eta_2 E(O_j^S). \quad (9)$$

The choice of $\eta_1$ and $\eta_2$ is not easy, because biopsies have serious medical side effects and consequently the cost of an extra biopsy cannot be quantified or compared to a unit increase in offset easily. To obviate this problem we utilize the equivalence between compound and constrained optimal designs (Cook and Wong, 1994). More specifically, it can be shown that for any $\eta_1$ and $\eta_2$ there exists a constant $C > 0$ for which minimization of loss function in (9) is equivalent to minimization of the loss function subject to the constraint that $E(O_j^S) < C$. That is, the optimal schedule is the one with the least number of biopsies and an offset less than $C$. The choice of $C$ now can be based on the protocol of AS program. In the more generic case in (8), the optimal solution can be found by minimizing $\mathcal{R}_R(\cdot)$ under the constraint $\mathcal{R}_r(\cdot) < C_r; r = 1, \ldots, R-1$.

## 5. Personalized Schedules for Patients in PRIAS

To demonstrate how the personalized schedules work, we apply them to the patients enrolled in PRIAS. To this end, we divide the PRIAS dataset into a training dataset with 5264 patients and a demonstration dataset with three patients who never experienced GR. We fit a joint model to the training dataset and then use it to create personalized schedules for patients in demonstration dataset. We fit the joint model using the R package JMbayes (Rizopoulos, 2016), which uses the Bayesian methodology to estimate the model parameters.

### 5.1 *Fitting the Joint Model to PRIAS Dataset*

The training dataset contains age at the time of induction in PRIAS, PSA levels and the time interval in which GR is detected, for 5264 prostate cancer patients. PSA was measured at every three months for first two years and every six months thereafter. To detect GR, biopsies were conducted as per the PRIAS schedule (Section 1). For the longitudinal analysis of PSA we use $\log_2$ PSA measurements instead of the raw data. This because the PSA scores take very large values around the time of disease progression, indicating that the underlying distribution for PSA is right skewed. The longitudinal sub-model of the joint model we fit is given by:

$$\log_2 \text{PSA}(t) = \beta_0 + \beta_1 (Age - 70) + \beta_2 (Age - 70)^2 + \sum_{k=1}^{4} \beta_{k+2} B_k(t, \mathcal{K}) + b_{i0} + b_{i1} B_7(t, 0.1) + b_{i2} B_8(t, 0.1) + \varepsilon_i(t). \quad (10)$$

where $B_k(t, \mathcal{K})$ denotes the $k$-th basis function of a B-spline with three internal knots at $\mathcal{K} = \{0.1, 0.5, 4\}$ years, and boundary knots at zero and seven years. The spline for the random effects consists of one internal knot at 0.1 years and boundary knots at zero and seven years. The choice of knots was based on exploratory analysis as well as on model selection criteria AIC and BIC. Age of patients was median centered to avoid numerical instabilities during parameter estimation. For the relative risk sub-model the hazard function we fit is given by:

$$h_i(t) = h_0(t) \exp\{\gamma_1 (Age - 70) + \gamma_2 (Age - 70)^2 + \alpha_1 m_i(t) + \alpha_2 m_i'(t)\}. \quad (11)$$

where $\alpha_1$ and $\alpha_2$ are measures of strength of the association between hazard of GR and $\log_2$ PSA value $m_i(t)$ and $\log_2$ PSA velocity $m_i'(t)$, respectively. Since the PRIAS schedule depends only on the observed PSA values (via PSA-DT), the interval censoring observed in PRIAS is independent and non informative of the underlying health of the patient.

From the joint model fitted to the PRIAS dataset we found that only $\log_2$ PSA velocity was strongly associated with hazard of GR. For any patient, a unit increase in $\log_2$ PSA velocity led to an 11 times increase in the hazard of GR. The parameter estimates for the fitted joint model are presented in detail in Web Appendix C of the supplementary material.

### 5.2 *Demonstration of Personalized Schedules*

Using the demonstration dataset, we next present the functioning of personalized schedules based on expected time of GR and dynamic risk of GR. The first patient of interest is patient 3174. The evolution of PSA, repeat biopsy history and proposed times of biopsies for this patient are shown in the top panel of Figure 2. It can be seen that the schedule of biopsy based on expected time of GR adjusts the times of biopsy according to the rise in hazard, which increases due steep rise in $\log_2$ PSA velocity. More specifically, at year two the proposed biopsy time is 12.5 years whereas at year four it decreases to 5.3 years. On average, a biopsy scheduled using expected time of GR at year two should have a larger offset $O_j^S$ compared to the same at year four. This is because the standard deviation of $g(T_j^*)$, given by $\text{SD}_g(T_j^*) = \sqrt{\text{var}_g(T_j^*)}$, is considerably lower at year four as shown in the bottom panel of Figure 2. In the figure it can be seen that the standard deviation also strongly depends on $\log_2$ PSA velocity. As for the schedules based on dynamic risk of GR, the optimal $1 - \kappa$ value was found to be between 0 and 0.1 at all time points, because of the sharp rise in PSA values. This value of $\kappa$ corresponds to a time very close to the time of latest biopsy ($t = 0$). Hence the biopsies are scheduled much earlier than those based on expected time of GR.

The demonstration of personalized schedules for the two other patients from the demonstration data set is discussed in Web Appendix D of the supplementary material.

## 6. Simulation Study

The application of personalized schedules for patients from PRIAS demonstrated that the schedules adapt according to the historical data of each patient. However we could not

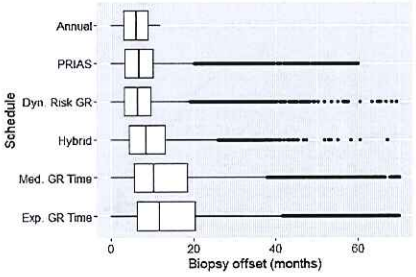


**Figure 5.** Boxplot showing variation in biopsy offset (months) for different methods, using all simulated patients. X-axis is trimmed at 70 months for visual clarity.

on variance of number of biopsies or offset they are not too high either for the hybrid approach.

## 7. Discussion

In this paper we presented personalized biopsy schedules for patients enrolled in AS programs. The problem at hand was that low risk prostate cancer patients enrolled in AS have to undergo repeat biopsies on a frequent basis for examination of disease progression, which causes medical side effects to patients and also brings financial burden on healthcare systems. Because of these issues, AS programs such as PRIAS are observing a high non-compliance for repeat biopsies, which further increases the chances of delayed detection of disease progression. To approach these problems we proposed personalized schedules based on joint model for time to event and longitudinal data. At any given point in time, the personalized schedules we proposed utilize a patient's information from historical PSA measurements and repeat biopsies conducted up to that time. We proposed two different classes of personalized schedules for individual patients. They are schedules based on the central tendency of the distribution of time of GR of a patient, and schedules based on dynamic risk of GR. In addition we also proposed a combination (hybrid approach) of these two approaches, which is effective even in scenarios where variance of time of GR for a patient is high. We then proposed criteria for evaluation of various schedules and a method to select the most optimal schedule.

We demonstrated using PRIAS dataset that the personalized schedules adjust the time of biopsy on the basis of results from historical biopsies and PSA, even when the two are not in concordance with each other (Web Appendix D). Secondly, we conducted a simulation study to compare various schedules. We observed that personalized schedules based on dynamic risk of GR performed better than PRIAS schedule in terms of both mean and variance of number of biopsies and offset. We also observed that the PRIAS schedule conducted more biopsies and had higher offset for patients having higher mean GR time. We prefer the schedule based on dynamic risk of GR over PRIAS schedule on the basis of these results. The schedules based on expected and median time of GR conducted only two biopsies on average, which is very promising compared to PRIAS and annual schedule which conducted 4.9 and 5.2 biopsies on average, respectively. In addition, for the former two schedules, at least 90% of the patients had an offset less than 36 months, which is the maximum possible offset in PRIAS during the first 10 years of AS. If a stronger restriction is prescribed for the offset, then we propose that the hybrid approach be used since it neither conducts too many biopsies nor it has a very high offset.

While each of the personalized methods has their own disadvantages and advantages, they also offer multiple choices to the AS programs to choose one as per their requirements, instead of choosing a common fixed schedule for all patients. In this regard, there is potential to develop more personalized schedules. For example, using loss functions which asymmetrically penalize overshooting/undershooting the target GR time can be interesting. Depending upon the requirements it is also possible to choose $\kappa$ on the basis of binary classification accuracy measures which focus on non-cases as well (Web Appendix E). Although in this work we assumed that GR time was interval censored, in reality the Gleason scores are susceptible to inter-observer variation (Carlson et al., 1998). Models and schedules which account for error in measurement of time of GR, will be interesting to investigate further. Lastly, there is potential for including diagnostic information from Magnetic resonance imaging (MRI) or DRE. Unlike PSA levels, such information may not always be continuous in nature, in which case our proposed methodology can be extended by utilizing the framework of generalized linear mixed models.


## Acknowledgements

The first and last authors would like to acknowledge support by the Netherlands Organization for Scientific Research's VIDI grant nr. 016.146.301. The authors also thank the Erasmus MC Cancer Computational Biology Center for giving access to their IT-infrastructure and software that was used for the computations and data analysis in this study.


## Supplementary Materials

Web Appendix A, C, D and E referenced in Section 2, Section 5, and Section 7, and the derivation of Equation (6) and (7) in Web Appendix B, are available in the document supplementary_material.pdf.

## References


Akhavan-Tabatabaei, R., Sánchez, D. M., and Yeung, T. G. (2017). A Markov decision process model for cervical cancer screening policies in colombia. *Medical Decision Making* **37**, 196–211.

Ayer, T., Alagoz, O., and Stout, N. K. (2012). A POMDP approach to personalize mammography screening decisions. *Operations Research* **60**, 1019–1034.

Bebu, I. and Lachin, J. M. (2017). Optimal screening schedules for disease progression with application to diabetic retinopathy. *Biostatistics* doi:10.1093/biostatistics/kxx009.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer Science & Business Media.

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.

Bokhorst, L. P., Valdagni, R., Rannikko, A., Kakehi, Y., Pickles, T., Bangma, C. H., Roobol, M. J., study group PRIAS, 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. *European Urology* **70**, 954–960.

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.

Carlson, G. D., Calvanese, C. B., Kahane, H., and Epstein, J. I. (1998). Accuracy of biopsy Gleason scores from a large uropathology laboratory: use of a diagnostic protocol to minimize observer variability. *Urology* **51**, 525–529.

Cook, R. D. and Wong, W. K. (1994). On the equivalence of constrained and compound optimal designs. *Journal of the American Statistical Association* **89**, 687–692.

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

Erenay, F. S., Alagoz, O., and Said, A. (2014). Optimizing colonoscopy screening for colorectal cancer prevention and surveillance. *Manufacturing & Service Operations Management* **16**, 381–400.

Keegan, K. A., Dall'Era, M. A., Durbin-Johnson, B., and Evans, C. P. (2012). Active surveillance for prostate cancer compared with immediate treatment. *Cancer* **118**, 3512–3518.

Läuter, E. (1976). Optimal multipurpose designs for regression models. *Mathematische Operationsforschung und Statistik* **7**, 51–68.

Loeb, S., Vellekoop, A., Ahmed, H. U., Catto, J., Emberton, M., Nam, R., Rosario, D. J., Scattoni, V., and Lotan, Y. (2013). Systematic review of complications of prostate biopsy. *European Urology* **64**, 876–892.

López-Ratón, M., Rodríguez-Álvarez, M. X., Cadarso-Suárez, C., Gude-Sampedro, F., et al. (2014). OptimalCutpoints: an R package for selecting optimal cutpoints in diagnostic tests. *Journal of Statistical Software* **61**, 1–36.

Potosky, A. L., Miller, B. A., Albertsen, P. C., and Kramer, B. S. (1995). The role of increasing detection in the rising incidence of prostate cancer. *JAMA* **273**, 548–552.

Rizopoulos, D. (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* **67**, 819–829.

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.

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.

Rizopoulos, D., Taylor, J. M. G., Van Rosmalen, J., Steyerberg, E. W., and Takkenberg, J. J. M. (2016). Personalized screening intervals for biomarkers using joint models for longitudinal and survival data. *Biostatistics* **17**, 149–164.

Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media.

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.

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians* **65**, 87–108.

Tosoian, J. J., Trock, B. J., Landis, P., Feng, Z., Epstein, J. I., Partin, A. W., Walsh, P. C., and Carter, H. B. (2011). Active surveillance program for prostate cancer: an update of the Johns Hopkins experience. *Journal of Clinical Oncology* **29**, 2185–2190.

Tsiatis, A. A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica* **14**, 809–834.

Welty, C. J., Cowan, J. E., Nguyen, H., Shinohara, K., Perez, N., Greene, K. L., Chan, J. M., Meng, M. V., Simko, J. P., Cooperberg, M. R., et al. (2015). Extended followup and risk factors for disease reclassification in a large active surveillance cohort for localized prostate cancer. *The Journal of Urology* **193**, 807–811.

Zhang, J., Denton, B. T., Balasubramanian, H., Shah, N. D., and Inman, B. A. (2012). Optimization of prostate biopsy referral decisions. *Manufacturing & Service Operations Management* **14**, 529–547.


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