Should a Biopsy be Conducted on Follow-up Visits in Prostate Cancer Active Surveillance?: A Personalized Biopsy Scheduling Approach

Medical Decision Making
XX(X):3–16

DOI: 10.1177/ToBeAssigned
www.sagepub.com/

SAGE

Au Thor[1] and Aut Hor[2]

Prepared using sagej.cls

Abstract

Background. Low-risk prostate cancer patients enrolled in active surveillance (AS) programs commonly undergo biopsies for examination of cancer progression. Biopsies are conducted as per a fixed and frequent schedule (e.g. annual biopsies), common for all patients. Such schedules may schedule unnecessary biopsies. Since biopsies are burdensome, patients do not always comply with the schedule, which increases the risk of delayed detection of cancer progression.

Objective. Motivated by the world's largest AS study, Prostate Cancer Research International Active Surveillance (PRIAS), our aim is to better balance the number of biopsies and the delay in detection of cancer progression. We intend to achieve this by personalizing the decision of conducting biopsies.

Methods. Using joint models for time-to-event and longitudinal data, we jointly model the historical prostate-specific antigen levels, digital rectal examination, and latest biopsy results of a patient at each follow-up visit. This results in a visit and patient-specific posterior predictive distribution of the time of cancer progression. Using this distribution we personalize the decision of conducting biopsy at a visit. We compare the personalized approach with the fixed schedules via a realistic and extensive simulation study based on an exact replica of the population of the patients from the PRIAS program.

Results. In comparison to the fixed schedules, the personalized approach saves one to seven biopsies per patient, depending upon the cancer progression speed of the patient. Despite a reduction in the number of biopsies, the delay in the detection of cancer progression for the personalized approach remains comparable with that of the schedule of PRIAS program.

Conclusions. We conclude that the personalized schedules better balance the number of biopsies per detected cancer progression.

Keywords

Active surveillance, biopsy, joint models, personalized medicine, prostate cancer

* *

Introduction

Prostate cancer is the second most frequently diagnosed cancer in men worldwide[1]. The increase in diagnosis of low-grade prostate cancer has been attributed to increase in life expectancy and increase in the number of screening programs[2]. An issue of prostate cancer screening programs is overdiagnosis. To avoid further over-treatment, patients diagnosed with low-grade prostate cancer are commonly advised to join active surveillance (AS) programs. In AS, serious treatments such as surgery, chemotherapy, or radiotherapy are delayed until necessary. Instead, cancer progression is routinely examined via serum prostate-specific antigen (PSA) levels: a protein biomarker, digital rectal examination (DRE) score: a measure of the size and location of the tumor, medical imaging, and biopsies etc.

Biopsies are the most reliable prostate cancer progression examination technique used in AS. When a patient's biopsy Gleason grading becomes larger than 6, AS is sto[illegible] and patient is advised for treatment of cancer progression[3]. ~~Since~~ biopsies are invasive they are conducted intermittently, until cancer progression is detected. Consequently, progression is always [illegible]ected with a delay, equal to the difference between the time of the ~~last~~ biopsy and the unobserved true time of progression. Biopsies are also painful, and prone to medical complications[4]. Hence, the decision of conducting a biopsy requires a fine compromise between the burden incurred due to biopsies and the delay in detection of progression. However, currently there is no consensus on the best time interval for subsequent repeat biopsies[5]. Many AS programs focus on minimizing only the delay, by scheduling biopsies annually[6]. Annual biopsies may work well for patients who progress

[1]Department of Biostatistics, Erasmus University Medical Center, the Netherlands
[2]Department of XXXX, XXX XXXX XXXX XXXX, the Netherlands

Corresponding author:
Anirudh Tomer, Erasmus MC, t.a.v. Anirudh Tomer / kamer Na-2701, PO Box 2040, 3000 CA Rotterdam, the Netherlands.
Email: a.tomer@erasmusmc.nl
*Financial support for this study was provided Netherlands Organization for Scientific Research's VIDI grant nr. 016.146.301, and Erasmus MC funding. The funding agreement ensured the authors independence in designing the study, interpreting the data, writing, and publishing the report.

fast, but for slowly progressing patients many unnecessary burdensome biopsies are scheduled. To improve the share of burden between fast and slow progressing patients, the world's largest AS program Prostate Cancer Research International Active Surveillance (PRIAS)[7], schedules annual biopsy only if at a follow-up visit a patient has a PSA doubling time between 0 and 10 years. PSA doubling time is measured as the inverse of the slope of the regression line through the base two logarithm of the observed PSA values. For everyone else, PRIAS schedules biopsies at following fixed follow-up times: year 1, 4, 7, and 10, and every 5 years thereafter. Despite this effort, in PRIAS over a period of 10 years a patient may get scheduled for 4 to 10 biopsies. Consequently, patients may not always comply with the schedule[3]. This can lead to the original problem of delayed detection of prostate cancer progression, and reduce the effectiveness of AS.

This article is motivated by the need to better balance the number of biopsies and the delay in detection of prostate cancer progression, than in practice currently. We intend to achieve this by personalizing the decision of conducting biopsies at follow-up visits (see Figure 1 for illustration). Personalized decision making has received much interest in the literature, especially for various cancers. For example, Markov decision process (MDP) models have been used to create personalized screening schedules for breast cancer[8], cervical cancer[9], and colorectal cancer[10]. In the specific case of prostate cancer, Zhang et al.[11] have used partially observable MDP models to personalize the decision of (not) deferring a biopsy to the next checkup time during the screening process. This decision is based on the baseline characteristics as well as a discretized PSA level of the patient at the current screening-visit time.

In comparison to the work referenced above, we do not base the decision of biopsy only on the current PSA value, but instead we utilize the entire history of PSA values, DRE scores, and results of the latest biopsy. To this end, we employ joint models for time-to-event and longitudinal data[12,13]. Since joint models use random effects[14] to model the between-patient heterogeneity, they are inherently patient-specific. Using joint models we first obtain a full specification of the joint distribution of the time of cancer progression, and PSA and DRE measurements. We then use it separately for each patient at each follow-up visit to define a visit and patient-specific posterior predictive distribution of the time of cancer progression, given

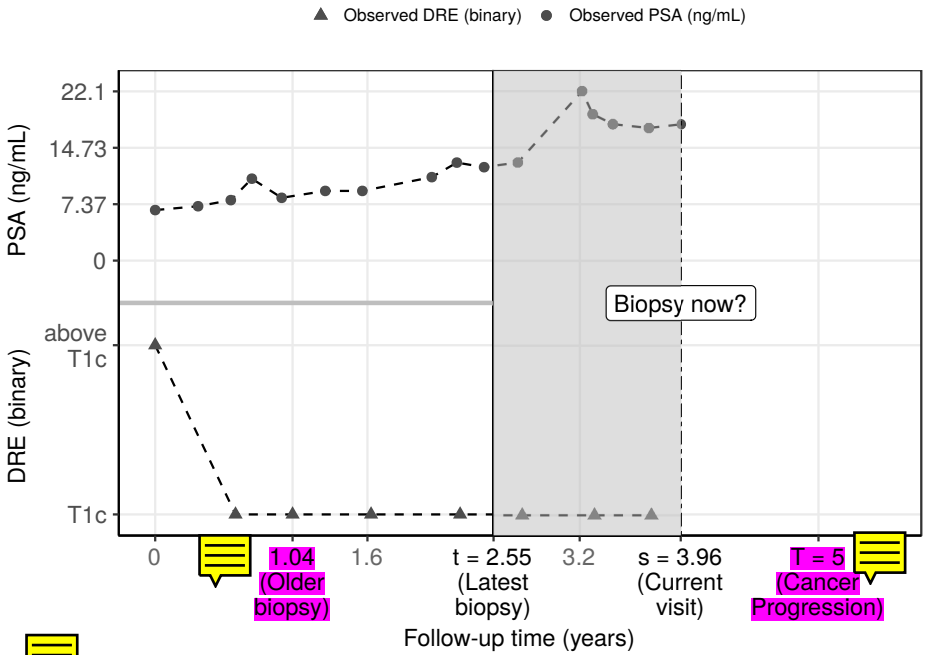

Figure 1. Illustration of the problem of making a decision for conducting biopsy at a follow-up visit ($s = 3.96$ years). The shaded region shows the time period between follow-up visit and the latest biopsy ($t = 2.55$ years), in which the patient is at a risk of cancer progression. In this work we utilize the entire history of PSA and DRE measurements along with the time of latest biopsy up to a follow-up visit to make the decision of biopsy at that visit.

the observed PSA and DRE measurements, and time of the latest biopsy up to that visit. We calculate the patient's risk of cancer progression at the follow-up visit using this distribution. If the risk is higher than a certain threshold our method schedules a biopsy at the same follow-up visit. Since there is no clear consensus on such risk thresholds we not only use fixed thresholds suggested by urologists, but also present a methodology to automate the choice of thresholds.

To develop our methodology we rely on the PRIAS dataset. However, biopsies are already conducted for PRIAS patients according to the PRIAS schedule. In addition, the cancer progression times of patients are observed with interval and right censoring. Given these reasons, and ~~the ethical issues with testing new methods of biopsies on real patients,~~ it is not possible to evaluate our methodology on ~~re~~ patients. Instead,

to reliably evaluate our methodology we conduct an extensive simulation study. In the simulation study we compare the personalized approach with the annual and PRIAS schedules. For a realistic comparison, we utilize an exact replica of the population of the PRIAS patients, generated using the model fitted to the PRIAS dataset.

The rest of the article is structured as follows: The details of the joint modeling framework and biopsy decision making methodology are presented in detail in the Methods section. The details of the simulation study and the corresponding results are presented in Methods and Results sections, respectively.

Methods

We start with a short introduction of the PRIAS dataset and the joint modeling framework we will use in our following developments. The PRIAS dataset consists of 5270 AS patients. For each patient, PSA measurements (ng/mL) are scheduled every 3 months for first 2 years and every 6 months thereafter. DRE measurements which are binary with two levels, namely DRE $>$ T1c and DRE $\leq$ T1c[15], are scheduled every 6 months. Larger values for PSA and/or larger score for DRE, may indicate cancer progression. Biopsies are scheduled as per the PRIAS protocol (see Introduction).

Joint Model for Time-to-Event and Longitudinal Data

Let T_i^* denote the true cancer progression time for the i-th AS patient. Let T_i^R and T_i^{R-1} denote the time of his latest and second latest biopsies, respectively. Since biopsies are conducted periodically, T_i^* cannot be observed directly and it is only known to fall in an interval $l_i < T_i^* \leq r_i$, where $l_i = T_i^{R-1}, r_i = T_i^R$ if progression is observed at the latest biopsy, and $l_i = T_i^R, r_i = \infty$ if progression is not observed yet. Further let $\boldsymbol{y}_{1i}, \boldsymbol{y}_{2i}$ denote the $n_{1i} \times 1$ and $n_{2i} \times 1$ vectors of the DRE and PSA longitudinal measurements, respectively. For a sample of n patients the observed data is denoted by $\mathcal{D}_n = \{l_i, r_i, \boldsymbol{y}_{1i}, \boldsymbol{y}_{2i}; i = 1, \ldots, n\}$.

The patient-specific PSA and DRE measurements over time are modeled using a generalized linear mixed effects model. For the i-th patient, the sub-model for DRE is given by:

$$\begin{aligned} \text{logit}\big[\Pr\{y_{1i}(t) > \text{T1c}\}\big] &= \beta_{01} + b_{01i} + (\beta_{11} + b_{11i})t \\ &\quad + \beta_{21}(\text{Age}_i - 70) + \beta_{31}(\text{Age}_i - 70)^2 \end{aligned} \tag{1}$$

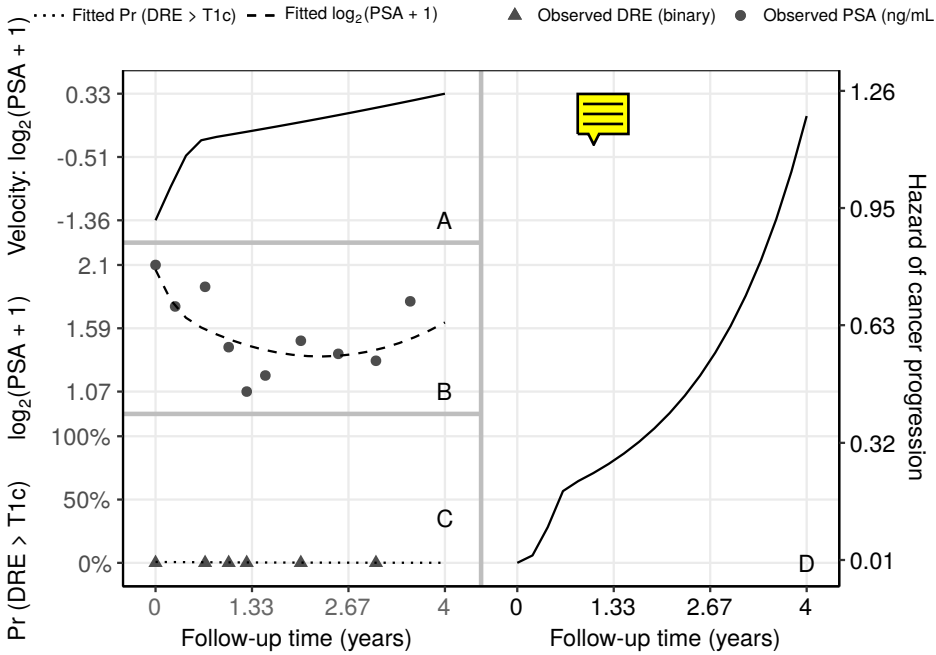

Figure 2. Illustration of the joint model fitted to the PRIAS dataset. Panel A shows the fitted $\log_2\{y_{2i}(t)+1\}$ velocity over time. Panel B shows the observed and underlying measurement error free (fitted) $\log_2\{y_{2i}(t)+1\}$ levels. Panel C shows the observed DRE scores and the fitted probability of obtaining a DRE score greater than T1c level. The hazard function shown in Panel D, depends on the fitted log odds of having DRE $>$ T1c, and the fitted $\log_2\{y_{2i}(t)+1\}$ value and velocity.

where, Age_i is the age of the i-th patient at the time of inclusion in AS, $\beta_{.1}$ are the fixed effect parameters, and $b_{.1i}$ are the random effect parameters. An example model fit for DRE is shown in Panel C of Figure 2. For the i-th patient, the sub-model for PSA is given by:

$$\begin{aligned}\log_2\left\{y_{2i}(t)+1\right\} = \beta_{02} + b_{02i} + \sum_{k=1}^{4}(\beta_{k2}+b_{k2i})B_k(t,\mathcal{K}) \\ + \beta_{52}(\text{Age}_i - 70) + \beta_{62}(\text{Age}_i-70)^2 + \varepsilon_{2i}(t),\end{aligned} \tag{2}$$

where, $B_k(t,\mathcal{K})$ denotes the k-th basis function of a B-spline with three internal knots at $\mathcal{K}=\{0.1, 0.7, 4\}$ years, and boundary knots at 0 and 5.42 years (0.95 quantile of the observed follow-up times). The fixed effects parameters are denoted by $\beta_{.2}$ and random effects are denoted by

$b_{.2i}$. The error $\varepsilon_{2i}(t)$ is assumed to be t-distributed with three degrees of freedom and scale σ (see Web Appendix C.1), and is independent of the random effects $b_{.2i}$. An example model fit for PSA is shown in Panel B of Figure 2. To account for the association between the DRE and PSA measurements, we link their corresponding random effects. More specifically, the complete vector of random effects $\boldsymbol{b}_i = (b_{.1i}, b_{.2i})^T$ 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 cancer progression, we use a relative risk sub-model. More specifically, the hazard of cancer progression $h_i(t)$ at a time t is given by:

$$\begin{aligned} h_i(t) = h_0(t) \exp \Big(& \alpha_{11} \text{logit}\big[\Pr\{y_{1i}(t) > \text{T1c}\}\big] \\ & + \alpha_{21} E\big[\log_2\big\{y_{2i}(t)+1\big\}\big] + \alpha_{22} \frac{\mathrm{d}E\big[\log_2\big\{y_{2i}(t)+1\big\}\big]}{\mathrm{d}t} \\ & + \gamma_1(\text{Age}_i - 70) + \gamma_2(\text{Age}_i - 70)^2 \Big), \end{aligned} \tag{3}$$

where, γ_1, γ_2 are the coefficients for the effect of Age. The parameter α_{11}models the impact of log odds of having DRE $>$ T1c on the hazard of cancer progression. The impact of PSA on cancer progression is modeled in two ways, namely at any time t the effect of the instantaneous underlying value of PSA $E[\log_2\{y_{2i}(t)+1\}]$ is given by α_{21}, and the effect of the instantaneous underlying velocity (panel A in Figure 2) of PSA $\mathrm{d}E[\log_2\{y_{2i}(t)+1\}]/\mathrm{d}t$ is given by α_{22}. Lastly, $h_0(t)$ is the baseline hazard at time t, and is modeled flexibly using P-splines. An example fitted hazard is shown in Panel D of Figure 2. The detailed specification of the baseline hazard, and parameter estimation using the Bayesian approach are presented in Web Appendix A of the supplementary material.

Personalized Decisions for Biopsy During Follow-up Visit

~~We intend to use the joint model fitted to the PRIAS dataset, to personalize the decision of conducting biopsies at follow-up visits. To this end,~~ let us assume that a decision of conducting a biopsy is to be made for a new patient j, who is not present in the PRIAS dataset. Let t be the time of his latest biopsy, and s denotes the current follow-up visit time. Let $\mathcal{Y}_{1j}(s)$ and $\mathcal{Y}_{2j}(s)$ denote the vector of all DRE and PSA measurements taken up to the current visit, respectively. We first combine all the observed

information to yield a posterior predictive distribution $g(T_j^*)$ of the time of cancer progression T_j^*. It is given by:

$$\begin{aligned} g(T_j^*) &= p\big\{T_j^* \mid T_j^* > t, \mathcal{Y}_{1j}(s), \mathcal{Y}_{2j}(s), \mathcal{D}_n\big\} \\ &= \int\int p\big(T_j^* \mid T_j^* > t, \boldsymbol{b}_j, \boldsymbol{\theta}\big) \\ &\quad \times p\big\{\boldsymbol{b}_j \mid T_j^* > t, \mathcal{Y}_{1j}(s), \mathcal{Y}_{2j}(s), \boldsymbol{\theta}\big\} p\big(\boldsymbol{\theta} \mid \mathcal{D}_n\big) \mathrm{d}\boldsymbol{b}_j \mathrm{d}\boldsymbol{\theta}. \end{aligned}$$

The distribution $g(T_j^*)$ is unique for each patient, and updates at every follow-up visit. It depends on the historical data of the patient via the posterior distribution of the random effects $\boldsymbol{b}_j$.

To make personalized decisions for biopsies, we utilize the personalized distribution $g(T_j^*)$. From a patient or doctor's perspective ~~the decision of conducting biopsy at a follow-up visit should account for the patient's risk of having a cancer progression $R_j(s \mid t, s)$ up to the current visit since the latest biopsy. We estimate this risk from the distribution $g(T_j^*)$ [16]. The risk is given~~ by:

$$R_j(s \mid t, s) = \Pr\big\{T_j^* \leq s \mid T_j^* > t, \mathcal{Y}_{1j}(s), \mathcal{Y}_{2j}(s)\mathcal{D}_n\big\}.$$

~~We propose to conduct a biopsy at a follow-up visit if this risk is higher than a certain threshold $0 \leq \kappa \leq 1$, as illustrated in Figure 3. While this approach is considerate about a patient's apprehensions for his risk of cancer progression, choosing too a small a threshold may lead to too many biopsies. To obtain an insight on the impact of thresholds, we evaluate a small risk of 5% a slightly higher risk of 15% in this work.~~

However, A drawback of ~~fixed thresholds is that they do not vary as per the general cancer progression rates of AS patients at a certain visit time.~~ In this regard, we propose to select a threshold on the basis of its ability to discriminate between patients who obtain cancer progression versus others. More specifically, given the time t of the latest biopsy we propose to choose a κ for which a binary classification accuracy measure[17], discriminating between patients obtaining progression versus others, is maximized. In joint models, a patient j is predicted to have progression in the time period up to current visit since last biopsy $(t, s]$, if $R_j(s \mid t, s) > \kappa$, or a control if $R_j(s \mid t, s) \leq \kappa$[18,19]. Since we are interested in detecting progression, we choose F_1 score as the classification accuracy measure. It combines both time dependent true positive rate (TPR) and

positive predictive value (PPV)[19], and is defined as:

$$\begin{aligned}
\mathrm{F}_1(t,s,\kappa) &= 2\frac{\mathrm{TPR}(t,s,\kappa)\ \mathrm{PPV}(t,s,\kappa)}{\mathrm{TPR}(t,s,\kappa)+\mathrm{PPV}(t,s,\kappa)},\\
\mathrm{TPR}(t,s,\kappa) &= \Pr\big\{R_j(s \mid t,s) > \kappa \mid t < T_j^* \le s\big\},\\
\mathrm{PPV}(t,s,\kappa) &= \Pr\big\{t < T_j^* \le s \mid R_j(s \mid t,s) > \kappa\big\}.
\end{aligned}$$

Since a high F_1 score is desired, the corresponding value of κ is $\arg\max_\kappa \mathrm{F}_1(t,s,\kappa)$.

Simulation Study

Although the personalized decision making approach is motivated by the PRIAS study, it is not possible to evaluate it on the PRIAS dataset. This is due to the fact that the PRIAS patients have already had their biopsies as per the PRIAS protocol. In addition, the true time of cancer progression is interval or right censored for all patients, making it impossible to correctly estimate the delay in detection of cancer progression due to a particular schedule. To this end, we conduct an extensive simulation study to compare pe alized, PRIAS and annual schedules. For a realistic comparison, we ~~utilize an exact replica of the population of the PRIAS patients. The simulation study follow-up period of 10 years also similar to the that of PRIAS study.~~

~~To simulate a replica of the population of the PRIAS patients, we generate the PSA and DRE profiles, and distribution of cancer progression times for hypothetical patients using the posterior distribution of the parameters of the joint model fitted to the PRIAS dataset.~~ From this population, we first sample 500 datasets with 1000 patients each. We generate a true cancer progression time for each of the patients, and then sample a set of PSA and DRE measurements at the same time points as given in PRIAS protocol. We then split ~~the~~ each dataset into a training (750 patients) and a test (250 patients) part, and generate a random and noninformative censoring time for the training patients. We next fit a joint model of the specification given in Equation (1), (2) and (3) to each of the 500 training datasets and obtain MCMC samples from the 500 sets of the posterior distribution of the parameters. Using these fitted joint models, we obtain the posterior predictive distribution of time of cancer progression for each of the 500×250 test patients at each of their visits. While maintaining a gap of 1 year between consecutive biopsies, individually

for each patient at each follow-up visit we make the decision of (not) conducting a biopsy ~~as per the methodology described in Personalized Decisions for Biopsy.~~ This results into an entire personalized schedule for each patient.

~~In this simulation study, for every test patient we conduct hypothetical biopsies.~~ One biopsy is conducted for all patients at the beginning of the AS program and another one is conducted at the end of the follow-up period at 10 years. The rest of the biopsies are scheduled using the following methods: (abbreviated names in parenthesis): biopsy every year (Annual), biopsy as per PRIAS protocol (PRIAS), personalized biopsy using a risk threshold of 5% (Risk: 5%), personalized biopsy using a risk threshold of 15% (Risk: 15%), and personalized biopsy using a risk threshold chosen on the ba of progression history of patients from the training dataset (Risk: F_1). We compare the resulting biopsy schedules on two measures, namely the number of biopsies they schedule and the delay in detection of cancer progression incurred due the schedule. We define the delay as the difference between the time of the biopsy on which cancer progression is detected and the true time of cancer progression. Ideal numbers for these two measures are 1 biopsy and 0 years of delay.

Results

We first discuss the results pertaining to the joint model fitted to the PRIAS dataset and then discuss results from the simulation study.

Model Fit

From the joint model fitted to the PRIAS dataset, we found that both $\log_2\{\text{PSA}+1\}$ velocity, and log odds of having DRE $>$ T1c were significantly associated with the hazard of cancer progression. For any patient, an increase in $\log_2\{\text{PSA}+1\}$ velocity from 0.03 to 0.16 (first and third quartiles of the fitted velocities, respectively) corresponds to a 1.92 fold increase in the hazard of cancer progression. Whereas, an increase in log odds of DRE $>$ T1c from -6.65 to -4.36 (first and third quartiles of the fitted log odds, respectively) corresponds to a 1.40 fold increase in the hazard of cancer progression. In terms of the predictive performance, we found that the area under the receiver operating characteristic curves (AUC)[19] was 0.59, 0.66, and 0.66 at years 1, 2, and 3 of ~~followup,~~ follow-up respectively. The AUC estimates for a joint model

ignoring DRE measurements were 0.58, 0.64 and 0.60 at years 1, 2 and 3 of follow-up. Parameter estimates are presented in detail in Web Appendix C.

Simulation Study

In the 500×750 training patients, we observed that for roughly 50% of the patients cancer progression did not take place in the 10 year follow-up. These could be seen as patients with a slow speed of cancer progression. Roughly 30% of the patients obtain cancer progression within first 3.5 years. These could be high risk patients who choose AS instead of immediate treatment, or patients with an initially misdiagnosed state of cancer[20]. In this work we consider these patients as the fast progressing patients. We consider that the remaining 20% patients with cancer progression times between 3.5 and 10 years have an intermediate speed of cancer progression.

For faster progressing patients (30% of the total patients), the boxplot in Figure 4 shows the variation in number of biopsies, and the delay in detection of cancer progression, in years (time of last biopsy - true time of cancer progression) due to various biopsy schedules. We can see that the personalized schedules conduct a median of one biopsy compared to two biopsies for PRIAS and annual schedule. The performance of personalized schedule with automatically chosen threshold is similar to that of PRIAS schedule. Thus with personalized approach, one biopsy may get saved for faster progressing patients.

For patients with intermediate progression speed (20% of the total patients), the boxplot in Figure 5 shows the variation in number of biopsies, and the delay in detection of cancer progression due to various biopsy schedules. Firstly, we can see that personalized schedules with a small risk threshold such as 5% risk conduct many more biopsies than other personalized schedules. Consequently, their performance with respect to the delay in detection of progression is similar to that of annual schedule. However, personalized schedule with slightly higher risk (15%) and risk chosen automatically, schedule a median of 3 and 4 biopsies, respectively. This is despite the fact that the delay in detection of cancer progression due to these schedules is similar to that of PRIAS. The PRIAS schedule also conducts more biopsies (median of 5 biopsies).

Thus, personalized approach may lead to one to two less biopsies for patients with intermediate speed of progression.

The patients who are at most advantage with personalized schedules are the patients who progress slowly (50% of the total patients). Figure 6 shows a boxplot of the number of biopsies conducted by various biopsy schedules for such patients. It can be seen that the annual schedule may lead to 10 unnecessary biopsies for everyone. The PRIAS schedule, schedules a median of 6 unnecessary biopsies. In comparison the personalized schedules using 15% and automatically chosen risk schedule only 3 and 4 biopsies, respectively.

Discussion

Prostate cancer active surveillance (AS) programs schedule biopsies for patients to detect cancer progression. Biopsies are burdensome and hence each biopsy counts. However, currently there is no consensus on the best time interval for subsequent repeat biopsies[5]. In order to reduce the delay in detection of cancer progression many AS programs schedule biopsies annually which leads to many unnecessary biopsies for slowly progressing patients. The world's largest AS program PRIAS attempts to identify these patients using their prostate-specific antigen (PSA) profile. However, despite their methodology compliance for biopsies is low. With an aim to better balance the burden of biopsies on patients and the delay in detection of cancer patients, in this article, we presented a methodology for personalizing the biopsy decision making process in AS programs.

Our methodology utilizes joint models for timetoevent and longitudinal data. Existing approaches for biopsy schedules either discard information from PSA and digital rectal examination (DRE), or use crude measures such as PSA doubling time. In contrast, our proposed methodology makes a separate decision for each patient at each follow-up visit on the basis of finer measures such as patient specific instantaneous PSA value, PSA velocity, probability of having DRE larger than level T1c, and results from the previous biopsies. Our method combines the aforementioned measures into a patient and visit specific cancer progression risk function. It schedules a biopsy if the risk of cancer crosses a certain threshold. We compared our approach with the existing annual and PRIAS schedules, by conducting a realistic and extensive simulation study, for a 10 year follow-up period.

In the simulation study we found that the patients who never obtain cancer progression in the 10 year follow-up incur much burden due to currently used schedules. The PRIAS schedule, despite its effort to identify such patients using PSA doubling time, schedules a minimum of 4 biopsies and a median of 6 unnecessary biopsies for such patients. The annual schedule performs even worse by scheduling 10 unnecessary biopsies. In contrast, the personalized schedules that we proposed reduce it to a median of 3 to 4 biopsies depending upon the choice of the risk threshold. The choice of the risk threshold is also important. A low risk threshold such as 5% risk may seem attractive to detect cancer progressions in time. However, it can work as worse as annual schedule in most situations, by scheduling unnecessary biopsies. A better idea is to either use a higher risk threshold or to automatically select it (see Methods). The personalized approach based on automatically chosen thresholds has the advantage of being generic for use in other active surveillance programs. In addition, it inherently accounts for the population specific cancer progression rate for the period in which patient has not been biopsied since last biopsy visit.

For patients with faster and intermediate speed of cancer progression, we found in our simulation study that the personalized approach with automatically chosen threshold conducts less biopsies than the PRIAS schedule but still leads to a similar delay in detection of cancer progression. For slow, intermediate, and fast progressing patients combined, the aforementioned personalized approach schedules 3.75 biopsies before detecting a cancer progression. The personalized approach using 15% risk schedules only 3 biopsies in total. These numbers are similar to the number of biopsies patients agree to undergo in PRIAS, if the non-compliance rates are also accounted in the PRIAS schedule.

A disadvantage of our approach is that it does not create an entire optimal biopsy schedule based on a certain number of biopsies that a patient agrees to undergo beforehand. However given the fact that cancer progression information obtained from PSA and DRE measurements updates on each visit, and patients may also disagree with more biopsies over time, creating an optimal schedule may not always work.

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, and Erasmus MC funding. 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. Lastly, we thank Frank-Jan H. Drost from the Department of Urology, Erasmus University Medical Center, for helping us in accessing the PRIAS data set.

Supplemental material

Supplementary material for this article are available in the document supplementary_material.pdf.

References

1. Torre LA, Bray F, Siegel RL et al. Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians* 2015; 65(2): 87–108.
2. Potosky AL, Miller BA, Albertsen PC et al. The role of increasing detection in the rising incidence of prostate cancer. *JAMA* 1995; 273(7): 548–552.
3. Bokhorst LP, Alberts AR, Rannikko A et al. Compliance rates with the Prostate Cancer Research International Active Surveillance (PRIAS) protocol and disease reclassification in noncompliers. *European Urology* 2015; 68(5): 814–821.
4. Ehdaie B, Vertosick E, Spaliviero M et al. The impact of repeat biopsies on infectious complications in men with prostate cancer on active surveillance. *The Journal of urology* 2014; 191(3): 660–664.
5. Loeb S, Carter HB, Schwartz M et al. Heterogeneity in active surveillance protocols worldwide. *Reviews in urology* 2014; 16(4): 202.
6. Welty CJ, Cowan JE, Nguyen H et al. Extended followup and risk factors for disease reclassification in a large active surveillance cohort for localized prostate cancer. *The Journal of Urology* 2015; 193(3): 807–811.
7. Bul M, Zhu X, Valdagni R et al. Active surveillance for low-risk prostate cancer worldwide: the prias study. *European urology* 2013; 63(4): 597–603.
8. Ayer T, Alagoz O and Stout NK. A POMDP approach to personalize mammography screening decisions. *Operations Research* 2012; 60(5): 1019–1034.
9. Akhavan-Tabatabaei R, Sánchez DM and Yeung TG. A Markov decision process model for cervical cancer screening policies in Colombia. *Medical Decision Making* 2017; 37(2): 196–211.
10. Erenay FS, Alagoz O and Said A. Optimizing colonoscopy screening for colorectal cancer prevention and surveillance. *Manufacturing & Service Operations Management* 2014; 16(3): 381–400.
11. Zhang J, Denton BT, Balasubramanian H et al. Optimization of prostate biopsy referral decisions. *Manufacturing & Service Operations Management* 2012; 14(4): 529–547.
12. Tsiatis AA and Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica* 2004; 14(3): 809–834.

13. Rizopoulos D. *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*. CRC Press, 2012. ISBN 9781439872864.
14. Laird NM and Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982; : 963–974.
15. Schröder F, Hermanek P, Denis L et al. The tnm classification of prostate cancer. *The Prostate* 1992; 21(S4): 129–138.
16. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* 2011; 67(3): 819–829.
17. López-Ratón M, Rodríguez-Álvarez MX, Cadarso-Suárez C et al. OptimalCutpoints: an R package for selecting optimal cutpoints in diagnostic tests. *Journal of Statistical Software* 2014; 61(8): 1–36.
18. Rizopoulos D. The R package JMbayes for fitting joint models for longitudinal and time-to-event data using MCMC. *Journal of Statistical Software* 2016; 72(7): 1–46.
19. Rizopoulos D, Molenberghs G and Lesaffre EM. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biometrical Journal* 2017; 59(6): 1261–1276.
20. Cooperberg MR, Cowan JE, Hilton JF et al. Outcomes of active surveillance for men with intermediate-risk prostate cancer. *Journal of Clinical Oncology* 2011; 29(2): 228.

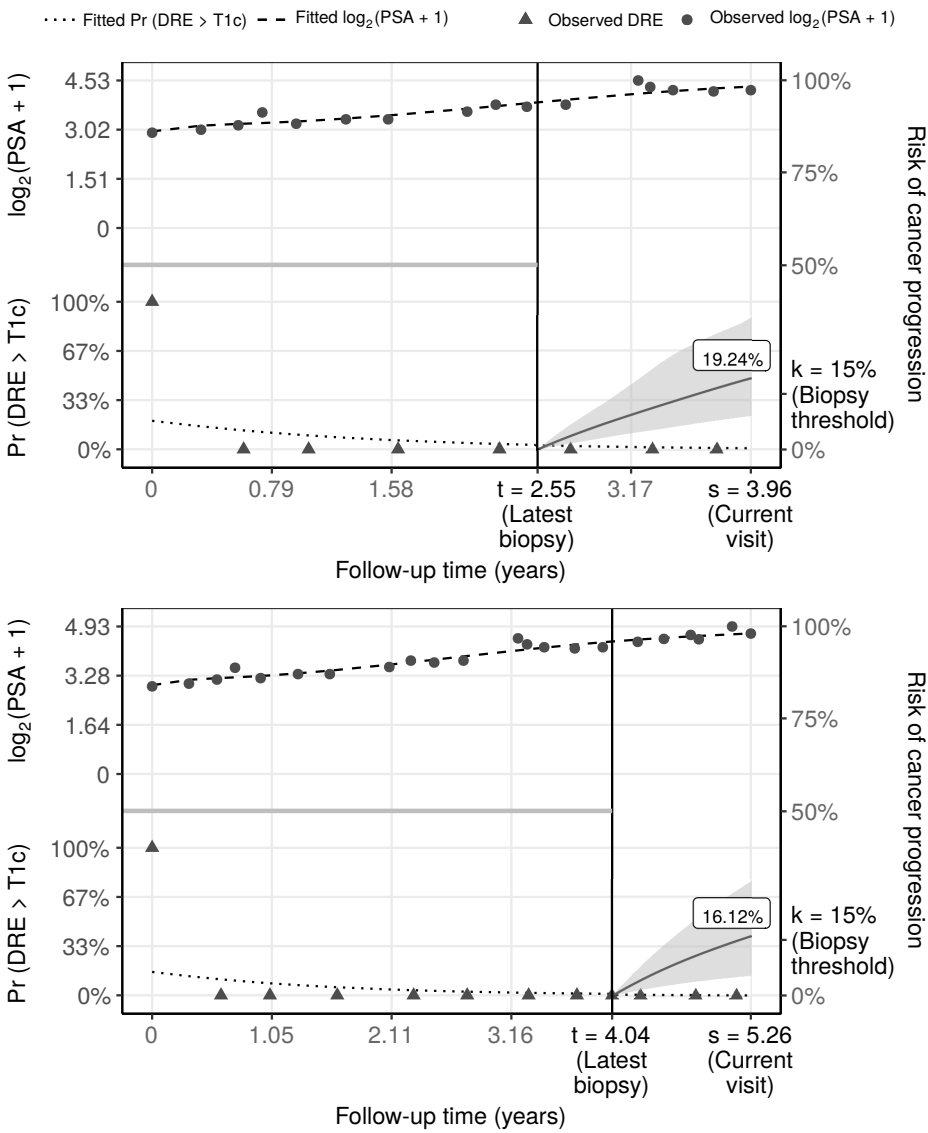

Figure 3. Illustration of personalized decision making for a patient at two different follow-up visits, namely $s = 3.96$ years in top panel and $s = 5.26$ years in bottom panel. The latest biopsy of the patient at which cancer progression was not detected is at $t = 2.55$ years in top panel and $t = 4.04$ in bottom panel. The risk of cancer progression, estimated on the basis of the fitted PSA and DRE profiles, and the latest biopsy, is shown in the right sub-panel with the 95% credible interval (shaded region). The risk at the current visit is estimated to be 19.24% in top panel and 16.12% in bottom panel. Since the risk is higher than the biopsy threshold $\kappa = 15\%$, a biopsy will be scheduled at the current visit in both panels.

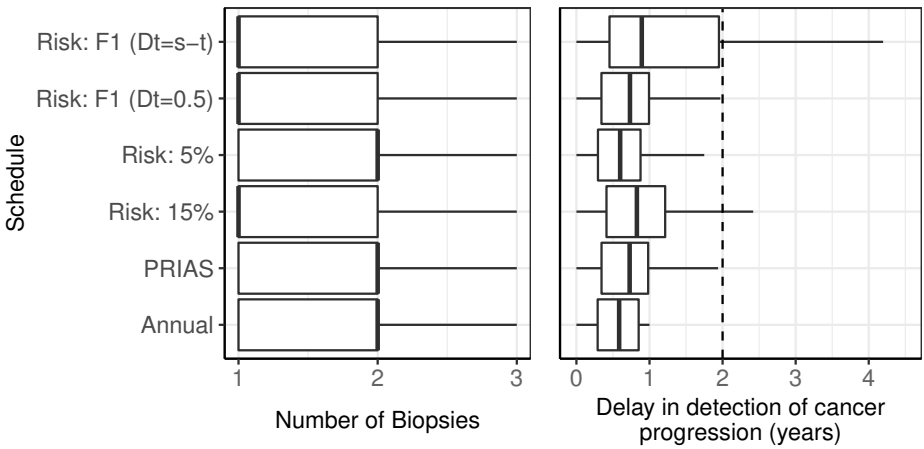

Figure 4. Boxplot showing variation in number of biopsies, and the delay in detection of cancer progression, in years (time of last biopsy - true time of cancer progression) for various biopsy schedules. The plot presents results for only those simulated test patients who had a faster speed of cancer progression, with progression times between 0 and 3.5 years. Biopsies are conducted until cancer progression is detected. Types of personalized schedules: Risk: 15% and Risk: 5% schedules, schedule biopsy if the risk of cancer progression at a visit is more than 15% and 5%, respectively. Risk: F1 (Dt=0.5) and Risk: F1 (Dt=s-t) work similar to Risk: 15% and Risk: 5%, except that the risk threshold for biopsy is chosen automatically by maximizing F_1 score. The term Dt represents the time window (years) in which F_1 score is maximized, and s is the time of the current follow-up visit and t is the time of the latest biopsy Annual corresponds to a schedule of yearly biopsies and PRIAS corresponds to biopsies as per PRIAS protocol.

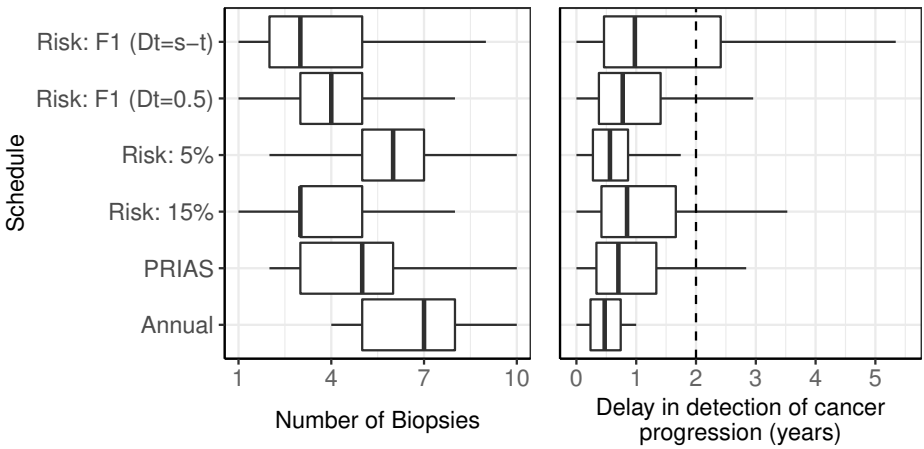

Figure 5. Boxplot showing variation in number of biopsies, and the delay in detection of cancer progression, in years (time of last biopsy - true time of cancer progression) for various biopsy schedules. The plot presents results for only those simulated test patients who had an intermediate speed of cancer progression, with progression times between 3.5 and 10 years. Biopsies are conducted until cancer progression is detected. Types of personalized schedules: Risk: 15% and Risk: 5% schedules, schedule biopsy if the risk of cancer progression at a visit is more than 15% and 5%, respectively. Risk: F1 (Dt=0.5) and Risk: F1 (Dt=s-t) work similar to Risk: 15% and Risk: 5%, except that the risk threshold for biopsy is chosen automatically by maximizing F_1 score. The term Dt represents the time window (years) in which F_1 score is maximized, and s is the time of the current follow-up visit and t is the time of the latest biopsy Annual corresponds to a schedule of yearly biopsies and PRIAS corresponds to biopsies as per PRIAS protocol.

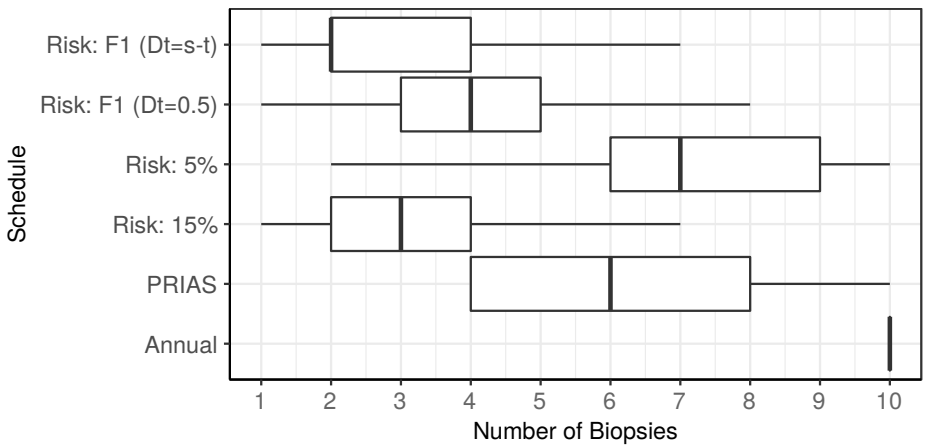

Figure 6. Boxplot showing variation in number of biopsies conducted by various biopsy schedules for those simulated test patients who did not have cancer progression over a period of 10 years. Biopsies are conducted until cancer progression is detected. Types of personalized schedules: Risk: 15% and Risk: 5% schedules, schedule biopsy if the risk of cancer progression at a visit is more than 15% and 5%, respectively. Risk: F1 (Dt=0.5) and Risk: F1 (Dt=s-t) work similar to Risk: 15% and Risk: 5%, except that the risk threshold for biopsy is chosen automatically by maximizing F_1 score. The term Dt represents the time window (years) in which F_1 score is maximized, and s is the time of the current follow-up visit and t is the time of the latest biopsy Annual corresponds to a schedule of yearly biopsies and PRIAS corresponds to biopsies as per PRIAS protocol.