Personalized Biopsies in Prostate Cancer Active Surveillance[☆]

Anirudh Tomer, MSc[a,*], Daan Nieboer, MSc[b], Monique J. Roobol, PhD[c], Anders Bjartell, PhD[d], Dimitris Rizopoulos, PhD[a], Movember Foundations Global Action Plan Prostate Cancer Active Surveillance (GAP3) consortium[e]

[a] *Department of Biostatistics, Erasmus University Medical Center, Rotterdam, the Netherlands*
[b] *Department of Public Health, Erasmus University Medical Center, Rotterdam, the Netherlands*
[c] *Department of Urology, Erasmus University Medical Center, Rotterdam, the Netherlands*
[d] *Department of Urology, Skåne University Hospital, Malmö, Sweden*
[e] *The Movember Foundations Global Action Plan Prostate Cancer Active Surveillance (GAP3) consortium members presented in Appendix A*

Abstract

`Background`: Low-risk prostate cancer patients enrolled in active surveillance (AS) undergo repeat biopsies. Treatment often advised when biopsy Gleason ≥ 7 (GS7). Many patients never obtain GS7, yet undergo biopsies frequently.

`Objective`: Better balance the number of biopsies and time delay in detection of GS7.

`Design, Setting, and Participants`: World's largest AS study PRIAS,

[☆] Word Count Abstract: 300; Word Count Text: 2509

[*] Corresponding author (Anirudh Tomer): Erasmus MC, kamer flex Na-2823, PO Box 2040, 3000 CA Rotterdam, the Netherlands. Tel: +31 10 70 43393

Email addresses: a.tomer@erasmusmc.nl (Anirudh Tomer, MSc), d.nieboer@erasmusmc.nl (Daan Nieboer, MSc), m.roobol@erasmusmc.nl (Monique J. Roobol, PhD), anders.bjartell@med.lu.se (Anders Bjartell, PhD), d.rizopoulos@erasmusmc.nl (Dimitris Rizopoulos, PhD)

7813 patients, 1134 obtained GS7, 100 medical centers in 17 countries, common study protocol.

`Outcome Measurements, and Statistical Analysis`: Prostate-specific antigen (PSA) measured every six months, repeat biopsies every one to three years. Bayesian joint model fitted to the observed time of GS7 and longitudinal PSA measurements. Risk predictions for GS7 externally validated in five largest AS cohorts from GAP3 database. Personalized risk based biopsy schedules developed using GS7 predictions. Total biopsies, time of biopsies and expected time delay in detection of GS7 calculated for various schedules to compare them.

`Results and Limitations`: Roughly 50% patients did not obtain GS7 in first 10 years in PRIAS. PSA velocity was a stronger predictor of GS7 (Hazard Ratio: 2.47, 95%CI: 1.93–2.99), than PSA value (Hazard Ratio: 0.99, 95%CI: 0.89–1.11). Internal validation: Time varying area under ROC curve for GS7 prediction between 0.62 and 0.69, and prediction error between 0.23 and 0.37. External validation: Results similar to internal validation only for Toronto, Memorial Sloan Kettering, and Johns Hopkins AS cohorts.

`Conclusions`: We developed personalized risk based biopsy schedules as alternative to fixed schedules. To assist patients in biopsy decisions we provided total and time of biopsies, and expected time delay in detection of GS7, for fixed and personalized schedules. Personalized schedules update with more patient data over follow-up.

`Patient Summary`: Personalized biopsy schedules are a novel alternative to fixed biopsy schedules (e.g., yearly biopsies). They utilize a patient's PSA and biopsy history to decide best time of biopsies. Personalized schedules

can offer a better balance between number of biopsies and delay in detection of Gleason upgrade than yearly biopsies.

Keywords: Active Surveillance, Biopsies, Personalized Medicine, Prostate Cancer, Shared Decision Making

1. Introduction

Patients with low- and very low-risk screening-detected localized prostate cancer are usually advised active surveillance (AS) instead of immediate radical treatment [1]. In AS, cancer progression is routinely monitored via prostate-specific antigen (PSA), digital rectal examination, and repeat biopsies. Among these, biopsy Gleason grade is the strongest indicator of cancer-related outcomes [2]. When the Gleason grade increases from grade 1 [Gleason score $\leq$ 6 (3+3)] to grade 2 [Gleason score 7 (3+4)] or higher (reclassification), the patient is commonly advised curative treatment [3].

Biopsies are conducted periodically, and hence reclassification is always detected with a time delay (Figure 1). The smaller this time delay, the larger is the window of opportunity for curative treatment. For detecting reclassification timely, many AS programs use fixed and frequent biopsy schedules (e.g., annual biopsies) for all patients [4, 5]. However, such schedules also lead to many unnecessary biopsies in slow/non-progressing patients. Biopsies are invasive, painful and prone to medical complications. Thus, biopsy burden combined with patient non-compliance to frequent biopsies [6] has raised concerns regarding the optimal biopsy schedule [7, 8].

A simple alternative to fixed and frequent biopsies is infrequent biopsies. For example, scheduling biopsies biennially instead of annually, may not substantially increase the risk of adverse downstream outcomes [7, 9]. Although, scheduling biopsies biennially may still lead to five unnecessary biopsies over ten years (current study period of large AS programs) for slow/non-progressing patients. A promising alternative to fixed biopsies is personalized schedule of biopsies. Personalized schedules utilize a patient's

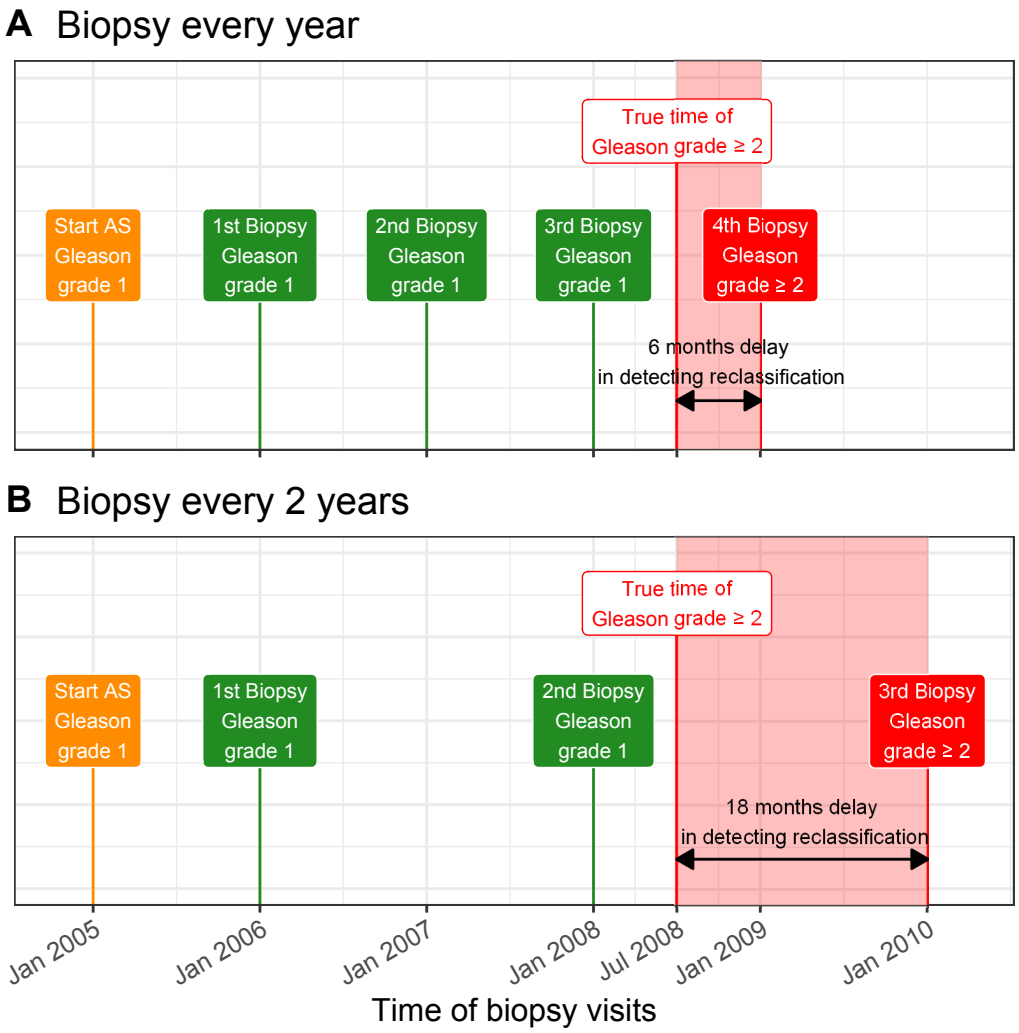

Figure 1: **Trade-off between the number of biopsies and time delay in detecting reclassification (Increase in Gleason grade from 1 to 2):** The true time of reclassification for the patient in this figure is July 2008. When biopsies are scheduled annually (**Panel A**), reclassification is detected in January 2009 with a time delay of six months, and a total of four biopsies are scheduled. When biopsies are scheduled biennially (**Panel B**) reclassification is detected in January 2010 with a time delay of 18 months, and a total of three biopsies are scheduled. Since biopsies are conducted periodically, the time of reclassification is observed as an interval. For example, Jan 2008–Jan 2009 in **Panel A** and Jan 2008–Jan 2010 in **Panel B**.

risk of reclassification to make decisions for biopsies (Figure 2).

The first challenge in developing risk of reclassification based personalized schedules is consolidating observed patient data (e.g., PSA, previous biopsy results) into risk estimates for reclassification. Previous studies have utilized the latest value of PSA to predict the Gleason grade [10, 11]. However, this approach ignores historical PSA measurements. To combine such longitudinal PSA data with historical biopsy results, a suitable model is the joint model for time-to-event and longitudinal data [12, 13, 14]. This model provides personalized predictions for risk of reclassification. Although, the subsequent challenge is to translate risk predictions into clinical decisions. For example, a 10% risk of reclassification can be perceived as high/low depending upon the patient's age. Patients may also weigh these risks with the potential *consequences* of another biopsy. Two relevant *consequences* of biopsies (Figure 1) are the timing and total number of biopsies (burden), and the time delay in detecting reclassification (smaller is beneficial). The relative importance of these *consequences* can vary between the patients, and also over the follow-up period for the same patient.

The goal of this work was to assist patients and doctors in making better decisions of biopsies than fixed and frequent biopsies. For this purpose, we created a tool that provides patients their current and future risk of having a reclassification, and personalized schedules of biopsies based on these risks. The tool also provides expected *consequences* of both fixed and personalized schedules. Thus, patients can compare them before making a decision. We took three steps for our goal. First, we fitted a prediction (joint) model to the world's largest AS dataset, PRIAS [3]. We then externally validated it

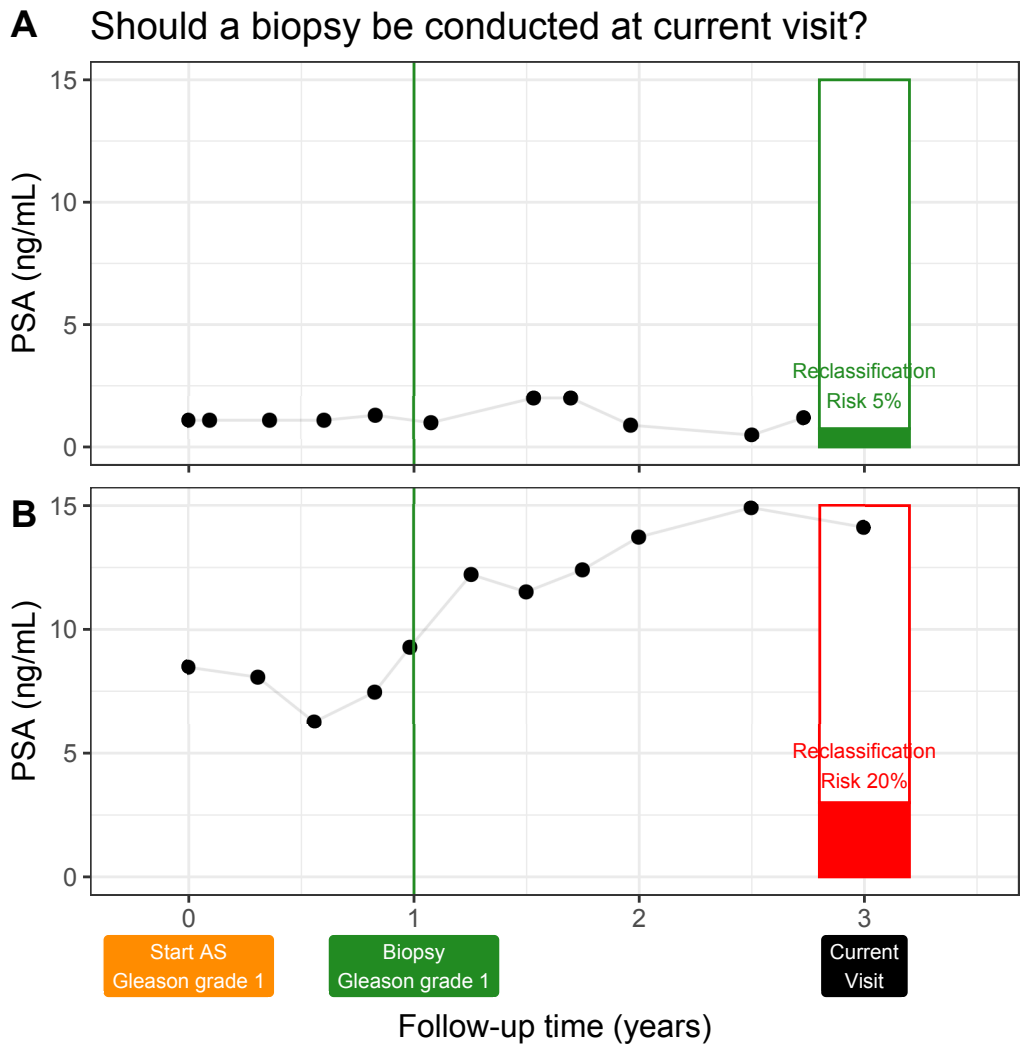

Figure 2: **Motivation for personalized risk-based decisions of biopsy**: Patients A (**Panel A**) and B (**Panel B**) had their latest biopsy at year one of follow-up (green vertical line). Patient A's prostate-specific antigen profile remained stable until his current visit at year three, whereas patient B's profile has shown a rise. Consequently, patient B's estimated cumulative risk of reclassification at the current visit (year three) is higher than that of patient A. This makes patient B a more suitable candidate for biopsy than Patient A. *Risk estimates in this figure are only illustrative.*

in five largest AS cohorts from the GAP3 database [15]. Lastly, we utilized our model to develop personalized schedule of biopsies, and to estimate the *consequences* of following both personalized and fixed biopsy schedules.

2. Patients and Methods

2.1. Study Cohort

Prostate Cancer International Active Surveillance (PRIAS) is an ongoing (December 2006 – to date) prospective cohort study of men with low- and very-low risk prostate cancer diagnoses [3]. More than 100 medical centers from 17 countries contributed to PRIAS, using a common study protocol (`www.prias-project.org`). Cancer progression was monitored by measuring PSA every three months for the first two years of follow-up and every six months thereafter. Biopsies were scheduled at year one, four, seven, and ten. Additional yearly biopsies were scheduled when PSA doubling time fell between zero and ten years.

We selected all 7813 patients between December 2006 and April 2019 who had Gleason grade 1 [2] at the time of inclusion in AS (Table 1). Our primary event of interest is reclassification (1134 patients), defined as, increase in Gleason grade on repeat biopsy from grade 1 to 2 or higher. Reclassification is a trigger for treatment advice in PRIAS. Although, 2250 patients were also advised treatment on the basis of PSA, the number of biopsy cores with cancer, anxiety, or other reasons. We focused solely on reclassification because of its strong association with cancer-related outcomes. Due to the periodical nature of biopsies, the true time of reclassification was interval censored (Figure 1).

Table 1: **Summary of the PRIAS dataset**. The primary event of interest is reclassification, that is, increase in Gleason grade from grade 1 to 2 or higher. IQR: interquartile range, PSA: prostate-specific antigen.

Characteristic	**Value**
Total patients	7813
Reclassification (primary event)	1134
Treatment	2250
Watchful waiting	334
Loss to follow-up	250
Death (unrelated to prostate cancer)	95
Death (related to prostate cancer)	2
Median age at diagnosis (years)	66 (IQR: 61–71)
Median follow-up period per patient (years)	1.8 (IQR: 0.9–4.0)
Total PSA measurements	67578
Median number of PSA measurements per patient	6 (IQR: 4–12)
Median PSA value (ng/mL)	5.7 (IQR: 4.1–7.7)
Total biopsies	15686
Median number of biopsies per patient	2 (IQR: 1–2)

2.2. Statistical Methods

For our aim to create risk of reclassification based personalized biopsy schedules, we required a risk prediction model. Available data was each patient's age at the start of AS, PSA measurements over follow-up, the timing of repeat biopsies and corresponding Gleason grades, and the time of reclassification (Figure 1). Analysis of this data required modeling the correlation between the PSA measurements of the same patient, the correlation between the Gleason grade and the PSA profile, and to account for PSA measurements missing after a patient experienced reclassification. To this end, a commonly used model is the joint model for time-to-event and longitudinal data [12, 13, 14].

The joint model we utilized consisted of two sub-models. First, a linear mixed model [16] for the longitudinally measured PSA (log-transformed). Second, a relative-risk model (similar to Cox model) for modeling the risk of reclassification using the observed times of reclassification. In the sub-model for PSA, we modeled PSA profiles non-linearly (Panel A, Figure 3). From the fitted PSA profiles, we extracted the instantaneous PSA velocity for each patient. This instantaneous PSA velocity varies as the PSA profile varies over time (Panel B, Figure 3). Thus it is more precise than the currently used PSA velocity/doubling-time [17]. We connected the two sub-models by using the fitted PSA profile and instantaneous PSA velocity as predictors for risk of reclassification in the relative-risk sub-model (Panel C, Figure 3). Patient age was included as a predictor in both sub-models. The parameters of the two sub-models were estimated jointly using the R package **JMbayes** [18]. This package utilizes the Bayesian methodology to estimate model parameters

(Appendix A).

2.3. Risk Predictions for Reclassification

The key component in personalized schedules is the cumulative-risk of reclassification. For new patients, we predicted this risk over their whole follow-up period. To this end, we inputted their PSA measurements and Gleason grade results into the joint model fitted to the PRIAS dataset (e.g., Panel C, Figure 3). When new PSA measurements or biopsy results become available over follow-up, the cumulative-risk profile is also updated.

We validated our prediction joint model using the PRIAS dataset (internal validation), as well using five largest AS cohorts from the GAP3 database [15] (external validation). These were University of Toronto AS (Toronto), Johns Hopkins AS (JHAS), Memorial Sloan Kettering Cancer Center AS (MSKCC), King's College London AS (KCL), and Michigan Urological Surgery Improvement Collaborative AS (MUSIC). We calculated the area under the receiver operating characteristic curve or AUC [19] as a measure of discrimination, and root mean squared prediction error or RMSPE [19] as a measure of calibration. Since AS studies are longitudinal in nature, we computed AUC and RMSPE in a time-dependent manner, at a gap of every six months (follow-up schedule of PRIAS) until year five (95-percentile of the observed reclassification times in PRIAS) of follow-up.

2.4. Personalized Schedule of Biopsies, and Their Consequences

Patients in PRIAS already have a follow-up schedule for testing PSA (Section 2.1). We reused this schedule to create personalized biopsy schedules out of it. Specifically, we scheduled biopsy on a PSA visit if the predicted

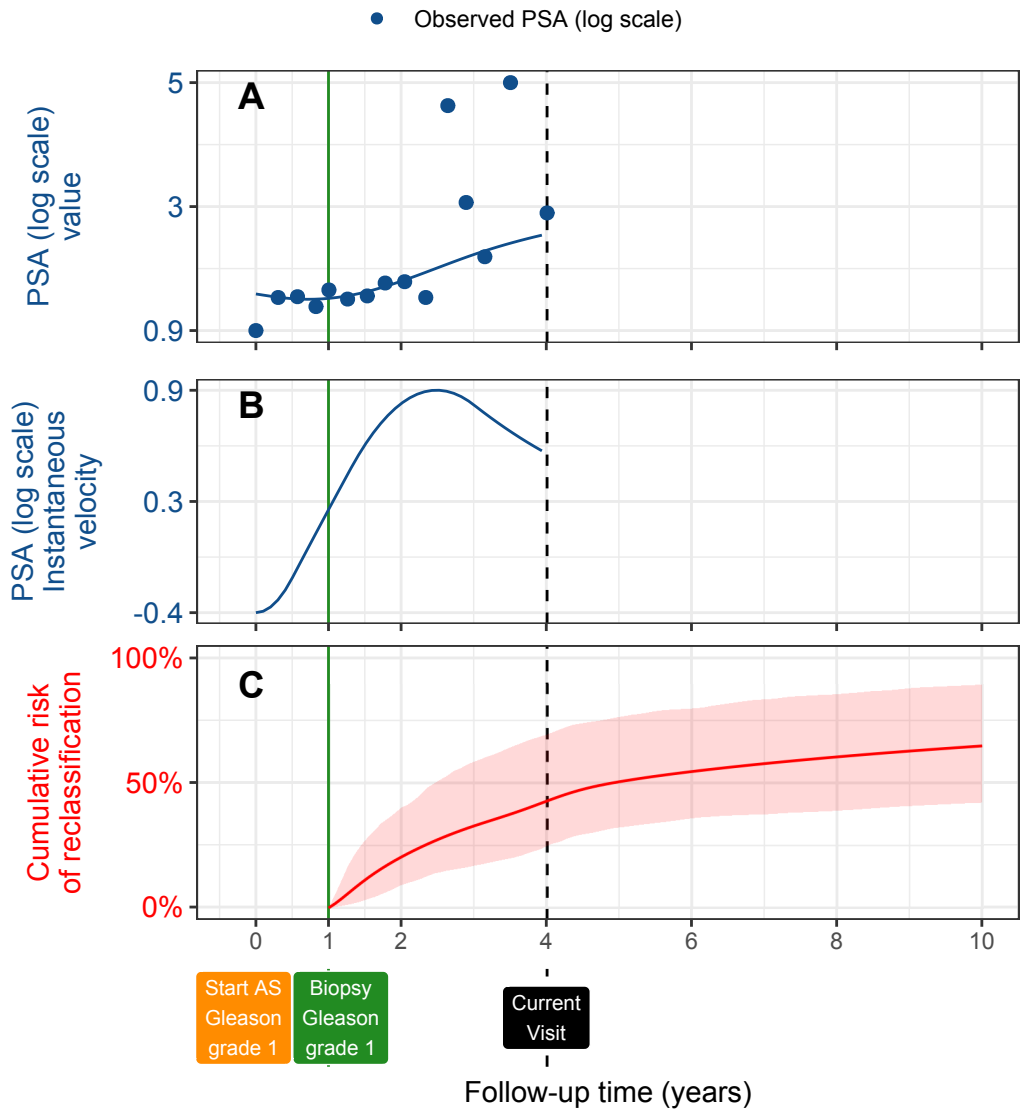

Figure 3: **Illustration of the joint model on a real PRIAS dataset patient**. **Panel A:** Observed (blue dots) and fitted PSA (solid blue line) measurements, log-transformed. **Panel B:** Estimated instantaneous velocity of PSA (log-transformed). **Panel C**: Predicted cumulative-risk of reclassification (95% credible interval shaded). Reclassification is defined as increase in Gleason grade [2] from grade 1 to 2 or higher. This risk of reclassification is available starting from the time of the latest negative biopsy (vertical green line at year 1 of follow-up). Joint model estimated it by combining the fitted PSA value and velocity (both on log scale of PSA) and time of latest negative biopsy. Black dashed line at year 4 denotes time of current visit.

cumulative-risk of reclassification on that visit was more than 10% (example threshold). We applied this rule of biopsy iteratively, starting from the current visit of a patient. After scheduling each subsequent biopsy, we updated the cumulative-risk of reclassification. This was done to account for the probability of not observing reclassification on the latest biopsy. We also kept a minimum gap of one year between consecutive biopsies (PRIAS recommendation). Personalized schedule using various risk thresholds are shown for an example patient in Figure 4.

The choice of the risk threshold in the personalized schedule dictates the *consequences* of that schedule. *Consequences* are, the timing and total number of biopsies, and the expected delay in detecting reclassification if that schedule is followed. A larger risk threshold will lead to infrequent biopsies. However, it may also lead to a longer delay in detecting reclassification. To assist patients in choosing an appropriate risk threshold for personalized schedule, and to compare personalized and fixed schedules, we provided an estimate of these *consequences*. These *consequences* were also personalized (Appendix C). That is, two patients may follow the same schedule of biopsies, but their expected delay in detection of reclassification will be different.

2.5. Web-Application

We implemented our methodology in a web-application `https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/`. This web-application uses the joint model that is fitted to the PRIAS dataset. Consequently, predictions for risk of reclassification can only be made for ten years of follow-up (follow-up period of PRIAS). Patient data can be uploaded in Microsoft Excel format, or entered manually on the web-application. Patients can check the evolution

of their cumulative-risk of reclassification over the future follow-up visits. In addition, the web-application allows comparison of the following schedules: personalized schedules based on 5%, 10%, and 15% risk threshold, annual and biennial biopsies, and the PRIAS schedule.

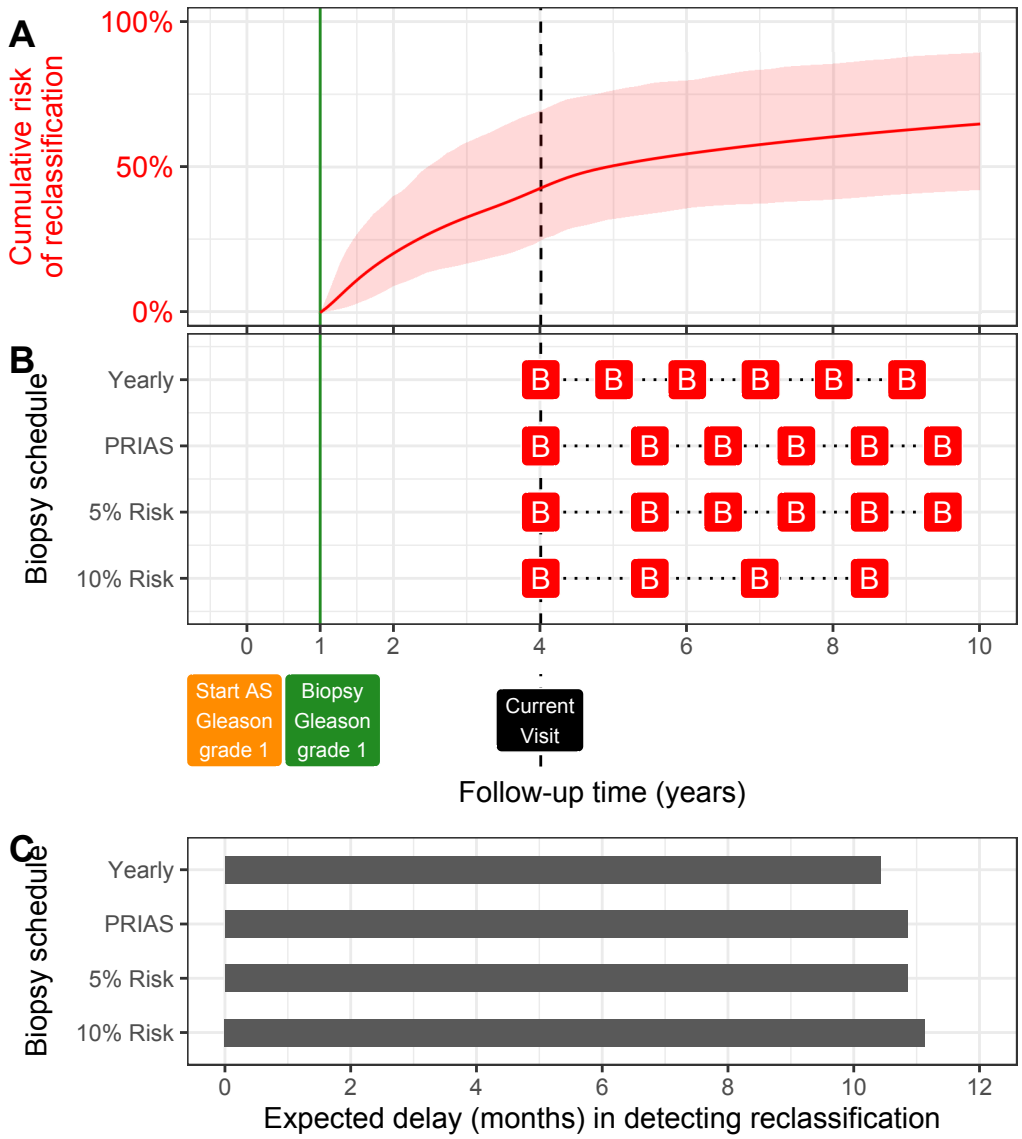

Figure 4: **Illustration of personalized and fixed schedules of biopsies**. The PSA profile of this patient is shown in Figure 3. **Panel A:** Predicted cumulative-risk of reclassification (95% credible interval shaded). **Panel B:** Personalized and fixed schedules of biopsies, with a red 'B' indicating a scheduled biopsy. **Panel C:** Expected delay in detecting reclassification for different schedules. Green vertical line at year 1 denotes time of latest negative biopsy. Black dashed line at year 4 denotes time of current visit.

3. Results

3.1. General Results

In PRIAS, the probability of experiencing reclassification within the first five and ten years was 33% and 42%, respectively (cumulative-risk plot in Appendix A). That is, ideally more than 50% of the patients may not require any biopsy in the first ten years.

3.2. Joint Model Results

For every ten years increase in a patient age at the time diagnosis, the adjusted hazard ratio of reclassification is 1.45 (95%CI: 1.30–1.63). When fitted PSA value (log scale) increases from 2.36 (25-th percentile of fitted PSA) to 3.07 (75-th percentile of fitted PSA), the adjusted hazard ratio of reclassification is 0.99 (95%CI: 0.89–1.11). When estimated instantaneous PSA (log scale) velocity increases from -0.09 (25-th percentile of estimated velocity) to 0.31 (75-th percentile of estimated velocity), the adjusted hazard ratio of reclassification is 2.47 (95%CI: 1.93–2.99). Hence, the instantaneous PSA (log scale) velocity is a stronger predictor for hazard of reclassification than the PSA value (log scale). Detailed parameter estimates are in Appendix A.4.

3.3. Validation Results

Using the joint model fitted to the PRIAS dataset we made risk predictions for GS7 in real PRIAS patients. As shown in Figure 4 of Appendix B, these risk estimates become more accurate as more data is gathered over follow-up. To check the accuracy of these risk predictions, we calculated the time dependent area under the receiver operating characteristic curves (AUC) as a measure of discrimination, and the root mean squared prediction

error (RMSPE) as a measure of calibration. These are shown in Figure 5. For predictions within PRIAS (internal validation), the time-dependent AUC was between 0.62 and 0.69, and RMSPE between 0.23 and 0.37 over the whole follow-up period. For validation in external cohorts, the AUC was similar to the AUC of PRIAS for all cohorts during the first three years of follow-up. The RMPSE however differed much more during the same period. The AS cohorts closest to PRIAS in terms of RMSPE were Johns Hopkins Active Surveillance and Memorial Sloan Kettering Cancer Center Active Surveillance. Detailed AUC and RMSPE results for all cohorts with 95% bootstrapped confidence intervals are presented in Table 6 to Table 11 of Appendix B.

3.4. Personalized Schedule Results

Using the risk predictions for GS7, we developed personalized schedules of biopsy for real PRIAS patients. We maintained a minimum gap of one year between biopsies as advised by the PRIAS protocol. In addition, we scheduled biopsies only for the first ten years follow-up because of limited follow-up period of the training dataset PRIAS. A compulsory biopsy was done scheduled year ten of follow-up in all schedules for meaningful comparison of their expected delays in detection of GS7. Various personalized and fixed biopsy schedules for demo patients are shown in Figure 4 and Appendix C's Figure 6, 7, 8 and 9. The biopsies denoted by 'B' show that personalized schedules schedule fewer biopsies than fixed schedules. At the same time the expected time delay in detection of GS7 is less than an year for personalized schedules. We have implemented this approach in a web-application (`https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/`, and Ap-

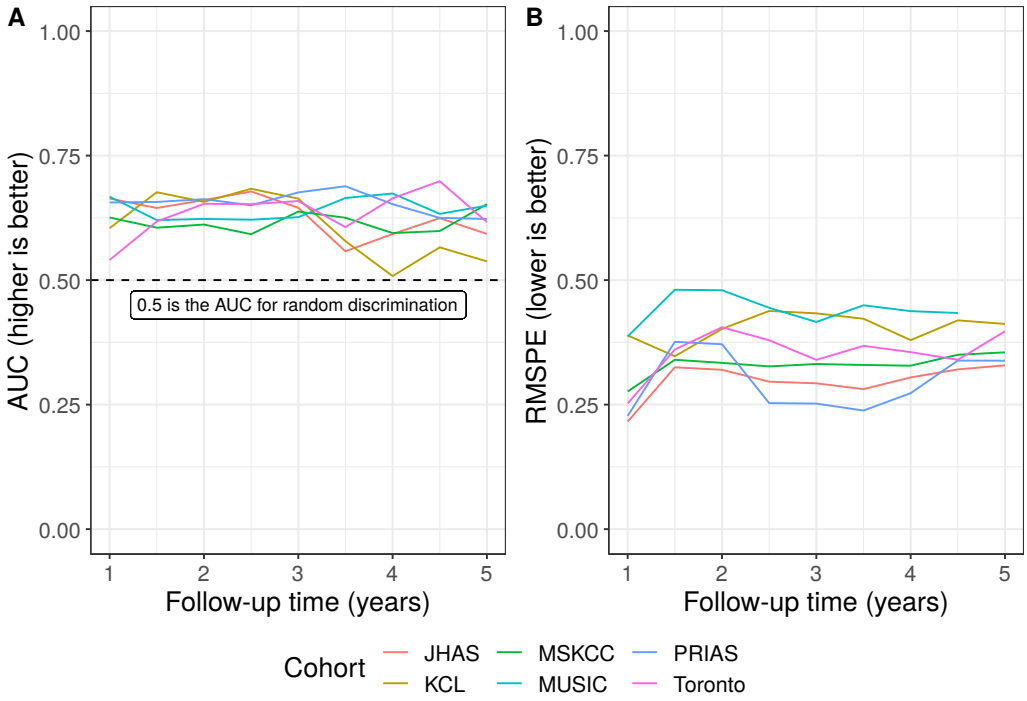

Figure 5: **Validation of predictions of Gleason $\geq$ 7 (GS7)**. In **Panel A** we can see that the time dependent area under the receiver operating characteristic curve or AUC (measure of discrimination) is above 0.5 in PRIAS (internal validation), and in Toronto, JHAS, MSKCC, KCL, and MUSIC AS cohorts (external validation). In **Panel B** we can see that the time dependent root mean squared prediction error or RMSPE (measure of calibration) is similar for PRIAS, and JHAS and Toronto cohorts. The bootstrapped 95% confidence interval for these estimates are presented in Table 6 to Table 11 of Appendix B. Full names of Cohorts are *PRIAS*: Prostate Cancer International Active Surveillance, *Toronto*: University of Toronto Active Surveillance, *JHAS*: Johns Hopkins Active Surveillance, *MSKCC*: Memorial Sloan Kettering Cancer Center Active Surveillance, *KCL*: King's College London Active Surveillance, *MUSIC*: Michigan Urological Surgery Improvement Collaborative Active Surveillance.

pendix D) for practical use.

4. Discussion

We developed personalized schedules of biopsies for active surveillance (AS) prostate cancer patients. These schedules utilize a patient's risk of reclassification (increase in biopsy Gleason grade [2] from grade 1 to 2 or higher). Reclassification risk calculators are not new [13, 20]. However, our work has three novel features. First, we personalized the risk of reclassification and it to schedule biopsies in a personalized manner. Second, we developed a methodology to assist patients and doctors to compare and choose between personalized and fixed schedules. Specifically, for each schedule we provided the timing and total number of biopsies (burden), and the expected time delay in detecting reclassification (smaller is beneficial). Third, we implemented our methodology in a web-application: `https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/`.

Our risk prediction model was a joint model [14**?**], fitted to the world's largest AS dataset, PRIAS. This model consolidated all available patient data, that is, historical PSA measurements and biopsy results, and baseline characteristics, into a single reclassification risk profile. This personalized risk profile and corresponding personalized schedules were updated as more patient data became available over follow-up.

We externally validated our model predictions for reclassification in five largest AS cohorts that are part of the GAP3 database [15]. We found that the AUC for predictions of GS7 over the follow-up period (Figure 5) was similar in external cohorts and PRIAS (internal validation). The RMSPE however was similar to PRIAS only for Memorial Sloan Kettering Cancer Center and Johns Hopkins cohorts. Given the large size of the latter two

cohorts, we expect that our model and the methodology will be useful to a large number of AS patients. Extending our model and methodology in other cohorts only requires fitting the model to their AS dataset.

The clinical implications of our work are as follows. The median survival time for reclassification is more than ten years in PRIAS, and in some other cohorts (Figure **??**). That is, more than 50% of AS patients do not require any biopsy during the first ten years of follow-up. We hope that our work will address patient apprehensions regarding adverse outcomes in AS, in a more objective manner. Many AS programs still utilize a rigorous schedule of yearly biopsies [4]. However, with concerns about non-compliance and burden of biopsies [6], the availability of our web based tool may encourage patients and doctors to consider personalized schedules.

Our work has certain limitations. The proposed model is valid only for the first ten years of follow-up in PRIAS, whereas reclassification may occur much later in many patients. In addition, model predictions were less accurate in later follow-up period due to lack of training data. These problems can be mitigated by refitting the model with new follow-up data in future. Although, we focused only on reclassification, it is susceptible to inter-observer variation. Models which account for this variation [13, 21] will be interesting to investigate further. However, the methodology for personalized scheduling, and for comparison of various schedules need not change. There is also a potential for including diagnostic information from novel biomarkers, quality of life measures, and magnetic resonance imaging (MRI). Currently, this data is very sparsely available in the PRIAS dataset. However, in future, our model can be extended by adding the new biomarkers as predictors.

5. Conclusions

We developed a novel methodology for personalized scheduling of biopsies, to detect reclassification (increase in Gleason grade from grade 1 to 2 or higher) in active surveillance prostate cancer patients. Personalized schedules utilize personalized risk of reclassification. Our model predicts this risk using a patient's whole history of PSA, repeat biopsy results and baseline characteristics. Our methodology is time-dynamic because it updates the personalized schedules as more patient data is gathered over follow-up. To assist patients and doctors in making the choice of a suitable biopsy schedule, we provide them estimates of personalized burden (time and total biopsies), and personalized benefit (time delay in detection of reclassification; lesser is beneficial), for both personalized and currently used schedules. Our methodology is available for use in practice at `https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/`.

Acknowledgments

This work was supported by the Movember Foundation. The funder did not play any role in the study design, collection, analysis or interpretation of data, or in the drafting of this paper. The first and last Authors would like to acknowledge support by Nederlandse Organisatie voor Wetenschappelijk Onderzoek (the national research council of the Netherlands) VIDI grant nr. 016.146.301, and Erasmus University Medical Center funding.

1. Briganti A, Fossati N, Catto JW, Cornford P, Montorsi F, Mottet N, Wirth M, Van Poppel H. Active surveillance for low-risk prostate cancer: the European Association of Urology position in 2018. *European urology* 2018;74(3):357–68.

2. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma. *The American journal of surgical pathology* 2016;40(2):244–52.

3. Bul M, Zhu X, Valdagni R, Pickles T, Kakehi Y, Rannikko A, Bjartell A, Van Der Schoot DK, Cornel EB, Conti GN, et al. Active surveillance for low-risk prostate cancer worldwide: the prias study. *European urology* 2013;63(4):597–603.

4. Nieboer D, Tomer A, Rizopoulos D, Roobol MJ, Steyerberg EW. Active surveillance: a review of risk-based, dynamic monitoring. *Translational andrology and urology* 2018;7(1):106–15.

289 5. Loeb S, Carter HB, Schwartz M, Fagerlin A, Braithwaite RS, Lepor H.
290 Heterogeneity in active surveillance protocols worldwide. *Reviews in*
291 *urology* 2014;16(4):202–3.

292 6. Bokhorst LP, Alberts AR, Rannikko A, Valdagni R, Pickles T, Kakehi Y,
293 Bangma CH, Roobol MJ, PRIAS study group . Compliance rates with
294 the Prostate Cancer Research International Active Surveillance (PRIAS)
295 protocol and disease reclassification in noncompliers. *European Urology*
296 2015;68(5):814–21.

297 7. Inoue LY, Lin DW, Newcomb LF, Leonardson AS, Ankerst D, Gulati R,
298 Carter HB, Trock BJ, Carroll PR, Cooperberg MR, et al. Comparative
299 analysis of biopsy upgrading in four prostate cancer active surveillance
300 cohorts. *Annals of internal medicine* 2018;168(1):1–9.

301 8. Bratt O, Carlsson S, Holmberg E, Holmberg L, Johansson E, Josefsson
302 A, Nilsson A, Nyberg M, Robinsson D, Sandberg J, et al. The study of
303 active monitoring in sweden (sams): a randomized study comparing two
304 different follow-up schedules for active surveillance of low-risk prostate
305 cancer. *Scandinavian journal of urology* 2013;47(5):347–55.

306 9. de Carvalho TM, Heijnsdijk EA, de Koning HJ. Estimating the risks
307 and benefits of active surveillance protocols for prostate cancer: a mi-
308 crosimulation study. *BJU international* 2017;119(4):560–6.

309 10. Partin AW, Yoo J, Carter HB, Pearson JD, Chan DW, Epstein JI, Walsh
310 PC. The use of prostate specific antigen, clinical stage and gleason score

311 to predict pathological stage in men with localized prostate cancer. *The*
312 *Journal of urology* 1993;150(1):110–4.

313 11. Makarov DV, Trock BJ, Humphreys EB, Mangold LA, Walsh PC, Ep-
314 stein JI, Partin AW. Updated nomogram to predict pathologic stage of
315 prostate cancer given prostate-specific antigen level, clinical stage, and
316 biopsy gleason score (partin tables) based on cases from 2000 to 2005.
317 *Urology* 2007;69(6):1095–101.

318 12. Tomer A, Nieboer D, Roobol MJ, Steyerberg EW, Rizopoulos D. Per-
319 sonalized schedules for surveillance of low-risk prostate cancer patients.
320 *Biometrics* 2019;75(1):153–62.

321 13. Coley RY, Zeger SL, Mamawala M, Pienta KJ, Carter HB. Prediction
322 of the pathologic gleason score to inform a personalized management
323 program for prostate cancer. *European urology* 2017;72(1):135–41.

324 14. Rizopoulos D. Joint Models for Longitudinal and Time-to-Event Data:
325 With Applications in R. CRC Press; 2012. ISBN 9781439872864.

326 15. Bruinsma SM, Zhang L, Roobol MJ, Bangma CH, Steyerberg EW,
327 Nieboer D, Van Hemelrijck M, consortium MFGAPPCASG, Trock B,
328 Ehdaie B, et al. The movember foundation's gap3 cohort: a profile of
329 the largest global prostate cancer active surveillance database to date.
330 *BJU international* 2018;121(5):737–44.

331 16. Laird NM, Ware JH, et al. Random-effects models for longitudinal data.
332 *Biometrics* 1982;38(4):963–74.

17. Vickers AJ, Savage C, O'Brien MF, Lilja H. Systematic review of pre-treatment prostate-specific antigen velocity and doubling time as predictors for prostate cancer. *Journal of Clinical Oncology* 2009;27(3):398.

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, Lesaffre EM. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biometrical Journal* 2017;59(6):1261–76.

20. Ankerst DP, Xia J, Thompson Jr IM, Hoefler J, Newcomb LF, Brooks JD, Carroll PR, Ellis WJ, Gleave ME, Lance RS, et al. Precision medicine in active surveillance for prostate cancer: development of the canary–early detection research network active surveillance biopsy risk calculator. *European urology* 2015;68(6):1083–8.

21. Balasubramanian R, Lagakos SW. Estimation of a failure time distribution based on imperfect diagnostic tests. *Biometrika* 2003;90(1):171–82.