

Personalized Biopsies in Prostate Cancer Active Surveillance[☆]

Anirudh Tomer, MSc^{a,*}, Dimitris Rizopoulos, PhD^a, Daan Nieboer, MSc^b,
Monique J. Roobol, PhD^c, Au Thor^d, Aut Hor^d, Auth Or^d

^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 xxxx, xxxx University Medical Center, City, Country*

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

Outcome Measurements, and Statistical Analysis: Prostate-specific anti-

[☆]Word Count Abstract: 300; Word Count Text: 2800

*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.rizopoulos@erasmusmc.nl (Dimitris Rizopoulos, PhD), d.nieboer@erasmusmc.nl
(Daan Nieboer, MSc), m.roobol@erasmusmc.nl (Monique J. Roobol, PhD)

gen (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 with Hazard Ratio (increase from 1st to 3rd quartile): 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 (PCa) are usually advised active surveillance (AS) instead of immediate radical treatment [1]. In AS, PCa progression is routinely monitored via prostate-specific antigen (PSA), digital rectal examination, and the Gleason score from repeat prostate biopsies. Since Gleason score is the strongest indicator of cancer-related outcomes, patients are commonly advised curative treatment when Gleason ≥ 7 (GS7) is detected [2].

Biopsies are conducted periodically, and hence GS7 is always detected with a time delay. The smaller this delay, the larger is the window of opportunity for curative treatment. For timely detection of GS7, many AS programs use fixed biopsy schedules (e.g., annual or biennial biopsies) for all patients [3, 4]. However, many AS patients do not require any biopsy in the first ten years of follow-up (see PRIAS and John's Hopkins AS cohorts in Figure 1). Biopsies are also invasive, painful and prone to medical complications. Biopsy burden combined with patient non-compliance [5] to frequent biopsies, has raised concerns regarding the optimal biopsy schedule [6, 7].

A simple alternative to frequent biopsies is infrequent biopsies. However, studies suggest not reducing biopsy frequency beyond 24 months, to have a sufficient window of opportunity for curative treatment [6, 9]. Although, biopsy every 24 months still leads to five unnecessary biopsies over ten years for 50% patients (Figure 1). A promising alternative to fixed biopsies is biopsies based on personalized risk of GS7. For example, among the two patients shown in Figure 2, patient B is a more suitable candidate for biopsy than patient A, because his risk of GS7 is much higher. Simulation studies



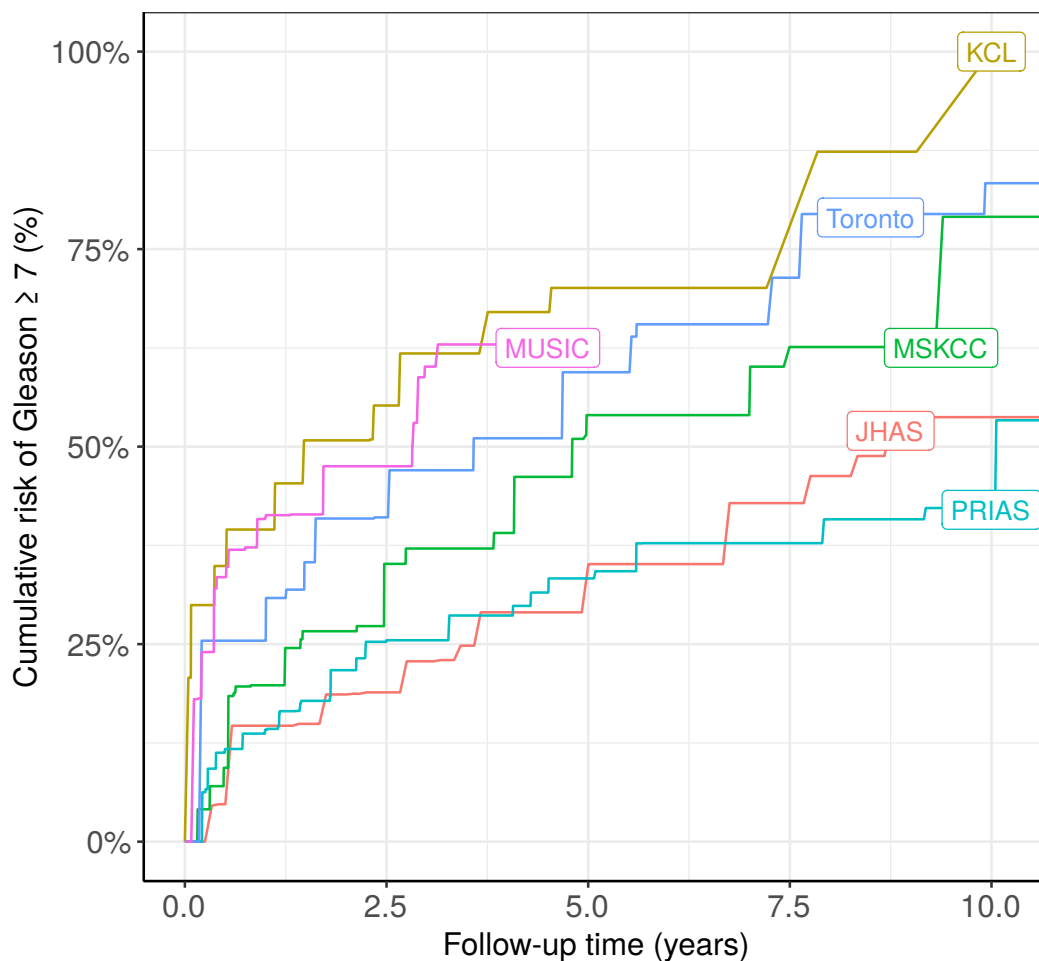



Figure 1: **Estimated cumulative risk of having Gleason ≥ 7 (GS7)** in the world's largest AS cohort PRIAS, and five of the largest AS cohorts part of the GAP3 database [8]. Abbreviations are *JHAS*: Johns Hopkins Active Surveillance, *PRIAS*: Prostate Cancer International Active Surveillance, *Toronto*: University of Toronto Active Surveillance, *MSKCC*: Memorial Sloan Kettering Cancer Center Active Surveillance, *KCL*: King's College London Active Surveillance. In the world's largest AS cohort PRIAS and in JHAS, roughly 50% of patients do not obtain GS7 in the first ten years. That is, ideally no biopsies should be prescribed to 50% of the patients in first ten years of AS.

26 have shown that personalized schedules may better balance the number of
 27 biopsies per detected GS7 than fixed schedules [10].

28 The first challenge in developing risk-based schedules is consolidating ob-
 29 served patient data (e.g., PSA, previous biopsy results) into GS7 risk esti-
 30 mates. For this previous studies have employed joint models for time-to-event 

31 and longitudinal data [10, 11, 12]. However, translating risk estimates into
 32 clinical decisions is challenging. For example, a 10% risk can be perceived as
 33 high/low depending upon the patient’s age. Patients may also weigh the risk
 34 of GS7 with the potential consequences of another biopsy. Two important

aspects

~~consequences~~ are the timing and total number of biopsies (burden), and the
 36 time delay in the detection of GS7 (smaller is better). These consequences
 37 vary between the patients, and also over the follow-up period for the same
 38 patient.

39 The goal of this work was to assist patients and doctors in making bet-
 40 ter decisions of biopsies than fixed and frequent biopsies. We intended to
 41 achieve this by providing the patients risk based personalized schedules of
 42 biopsies, and to allow them to compare the consequences of each schedule
 43 before making a decision. To this end, we took three steps. First we fitted
 44 a prediction (joint) model to the world’s largest AS dataset, PRIAS [2]. We
 45 then externally validated the model predictions in five largest AS cohorts
 46 that are part of the GAP3 database. Lastly, we utilized the personalized
 47 GS7 risk predictions to calculate the timing and total number of biopsies,
 48 and the time delay in the detection of GS7 for risk-based and fixed biopsy
 49 schedules.

relevant

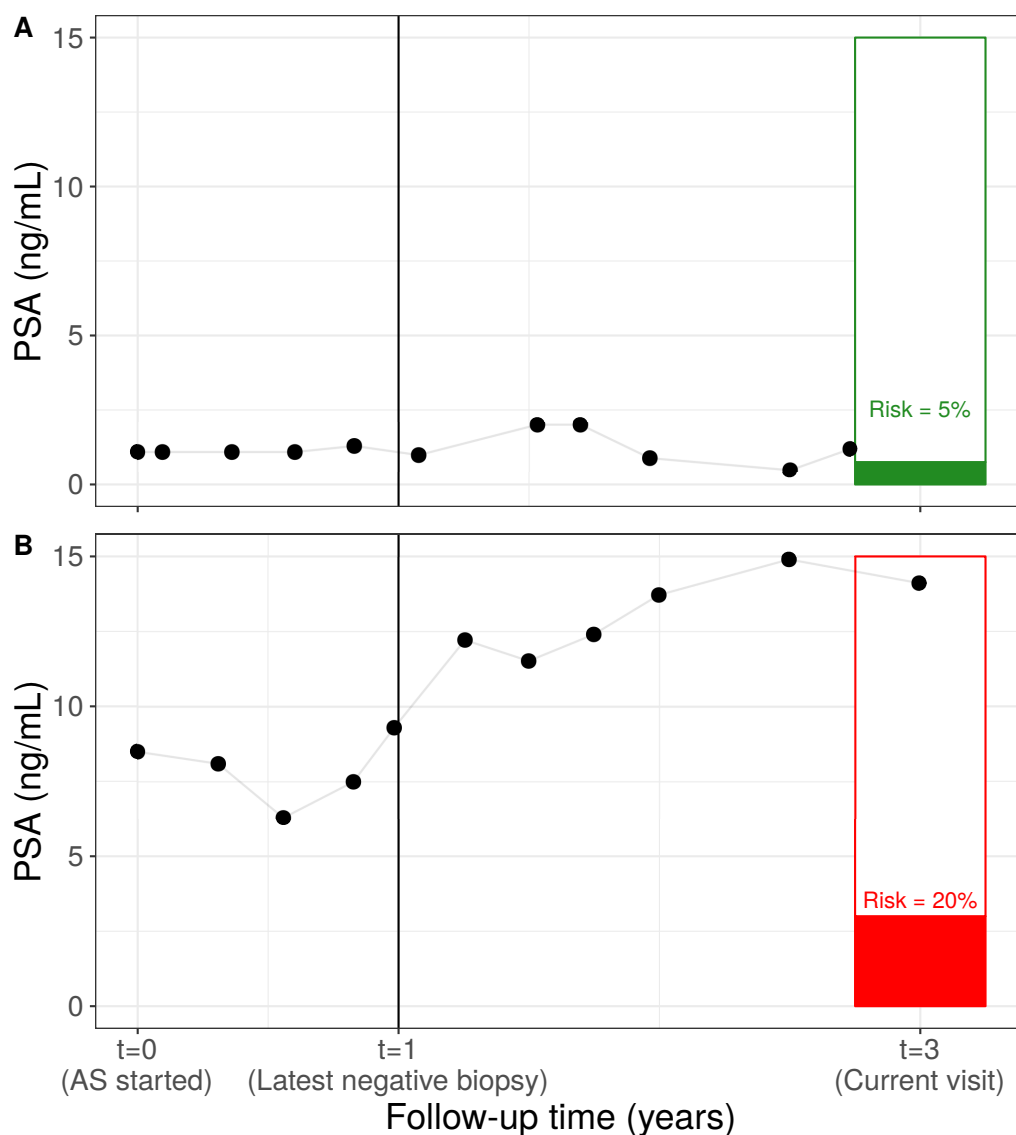


Figure 2: **Motivation for personalized risk-based decisions of biopsy:** Patient A and B had their latest biopsy at year one of follow-up. Patient A's prostate-specific antigen profile remained stable until the current visit time at year three, whereas patient B's profile has shown a rise. Consequently, patient B's estimated cumulative risk of GS7 at year three is higher than that of patient A, making patient B a more suitable candidate for biopsy than Patient A. Risk estimates in this figure are only illustrative.

50 2. Patients and Methods

51 2.1. Study Cohort

52 Prostate Cancer International Active Surveillance (PRIAS) is an ongoing
 53 prospective cohort study of men with low- and very-low risk PCa diagnoses
 54 [2]. More than 100 medical centers from 17 countries contribute
 55 in PRIAS, using a common study protocol (www.prias-project.org). We
 56 used the data collected between December 2006 (beginning of PRIAS study)
 57 and May 2019. The PSA was measured every three months until year two
 58 of follow-up and every six months thereafter. Biopsy schedule was year one,
 59 four, seven, and ten, and additional yearly biopsies when PSA doubling time
 60 was between zero and ten years. The primary event used in this work is
 61 Gleason ≥ 7 (GS7) because it is commonly used as a trigger for treatment
 62 advice. It was observed in 1134 patients. However, 2250 patients were provided
 63 treatment (see Table 1). Treatment in absence of GS7 may have been
 64 advised on the basis of PSA, number of biopsy cores with cancer, anxiety,
 65 or other reasons. We focused only on GS7 because of its strong association
 66 with cancer-related outcomes. Due to the periodical nature of biopsies, the
 67 time of GS7 was only available as a time interval in which GS7 occurred.

Table 1: **Patient characteristics for the PRIAS dataset.** The primary event of interest is Gleason ≥ 7 . IQR: interquartile range, PSA: prostate-specific antigen.

Characteristic	Value
Total patients	7813
Gleason ≥ 7 (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)	61 (IQR: 66–71)
Median follow-up period per patient (years)	1.8 (IQR: 0.9–3.99)
Total PSA measurements	65798
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	15563
Median number of biopsies per patient	2 (IQR: 1–3)

2.2. Statistical Methods

Our aim was to develop a model for predicting the time of GS7. The available data for each patient were, age at the start of AS, all observed PSA measurements, and the history of biopsies. We wanted to account for the correlation between the PSA measurements of the same patient, and also their correlation with the time of GS7. An additional complication was that the PSA values were missing once a patient obtained GS7. A commonly used model to handle these issues is the joint model for time-to-event and longitudinal data [12, 10, 11].

The joint model we utilized, exploited patient-specific random effects [13] to act as a common source of correlation between the PSA and time of GS7 outcomes (see Figure 3 and Appendix A.2's Figure 1). Random effects also represented the underlying state of PCa, and were included in both the linear mixed effects sub-model for $\log_2\{\text{PSA} + 1\}$ transformed measurements (see Appendix A.5), and the relative risk sub-model (similar to `cox` model) for time of GS7. In the sub-model for PSA, random effects non-linearly modeled the evolution of PSA over time. Simultaneously, in the relative risk model random effects were used indirectly by including fitted $\log_2\{\text{PSA} + 1\}$ value and velocity as time dependent covariates. ~~This established the correlation between PSA and time of GS7. Unlike observed $\log_2\{\text{PSA} + 1\}$ values, the fitted values were free of measurement errors.~~ The $\log_2\{\text{PSA} + 1\}$ velocity was mathematically derived from fitted $\log_2\{\text{PSA} + 1\}$ values. Consequently, the $\log_2\{\text{PSA} + 1\}$ velocity also changed non-linearly over follow-up.



The parameters of the two sub-models were estimated jointly using the

93 R package **JMbayes** [14]. This package utilizes the Bayesian methodology
 94 to estimate model parameters.


95 *2.3. Assessment of Predictions of GS7*

96 We validated the risk predictions of GS7 from our model within the
 97 PRIAS dataset (internal validation), as well as in five of the largest AS
 98 cohorts part of the GAP3 database [8] (external validation). The external
 99 cohorts were University of Toronto AS (Toronto), Johns Hopkins AS (JHAS),
 100 Memorial Sloan Kettering Cancer Center AS (MSKCC), King’s College Lon-
 101 don AS (KCL), and Michigan Urological Surgery Improvement Collaborative
 102 AS (MUSIC). For validation, we utilized the area under the receiver operat-
 103 ing characteristic curve or AUC [15] as a measure of discrimination, and root
 104 mean squared prediction error or RMSPE [15] as a measure of calibration.
 105 Since AS studies are longitudinal in nature, we computed AUC and RMSPE
 106 in a time dependent manner, at a gap of every six months (follow-up schedule
 107 of PRIAS) until year five (95-percentile of the observed GS7 times in PRIAS)
 108 of follow-up.

109 *2.4. Personalized Schedule of Biopsies, and Its Consequences*

110  Consider a new patient shown in Figure 4 for whom we intended to make
 111 a personalized schedule of biopsies. First, using the joint model fitted to the
 112 PRIAS dataset we obtained his cumulative risk of GS7 over the entire follow-
 113 up period (Panel B of Figure 4). A biopsy at his current visit may be then
 114 suggested if the cumulative risk of GS7 at the current visit is above a certain
 115 threshold (e.g., 10% risk). By repeatedly applying this 10% threshold rule
 116 over the whole follow-up, we obtained his personalized schedule of biopsies 

117 (see Appendix C). Similar schedules can be made with another risk threshold
118 such as 5% or 15% risk. These personalized risk based biopsy schedules are
119 updated at each follow-up visit based on newly gathered patient data.

120 To assist patients in making an informed choice for a schedule, be it
121 personalized or fixed, we provided them patient-specific consequences of fol-
122 lowing each schedule. To this end, we first calculated the probability of 
123 occurrence of GS7 in time gaps between successive biopsies of each schedule.
124 Using these probabilities we then obtained the expected time delay in de-
125 tection of GS7 for following that schedule (see Appendix C). Thus, patients
126 had a method to compare across various schedules in terms of the personal-
127 ized burden (time and total biopsies), and personalized benefit (less delay in
128 detection of GS7 is beneficial). Lastly, we implemented this approach in a
129 web-application.

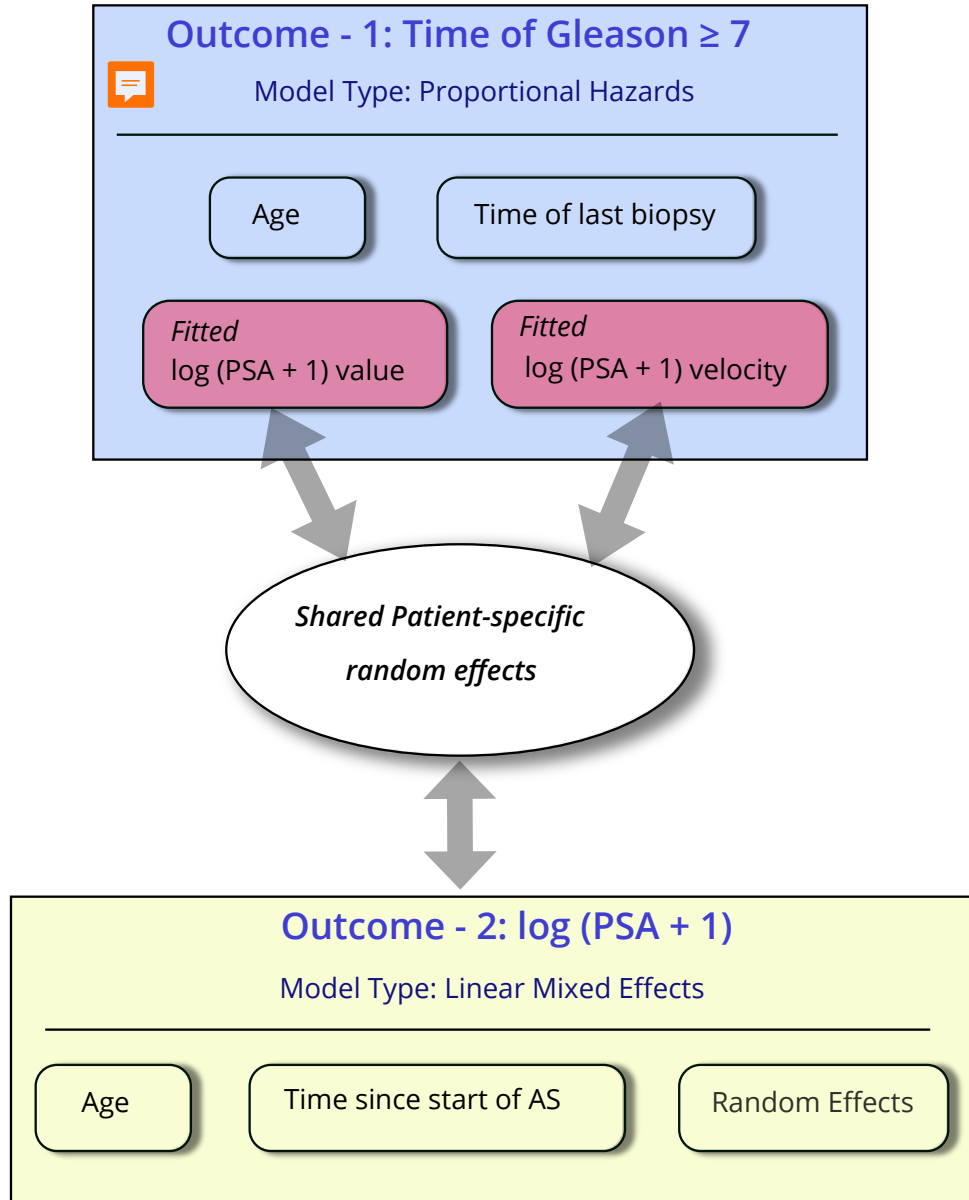


Figure 3: **Diagram of the joint model:** Patient-specific random effects (ellipse in center) are shared between sub-models for all outcomes pertaining to prostate cancer progression. The random effects model the correlation between the outcomes. In the linear mixed effects sub-model for $\log_2\{\text{PSA} + 1\}$ transformed PSA outcome (bottom rectangle), random effects are used as covariates. In the relative-risk sub-model (similar to Cox model) for time of Gleason ≥ 7 (GS7), random effects are utilized indirectly, by including fitted $\log_2\{\text{PSA} + 1\}$ value and velocities as covariates. Age of patient at baseline is included in both sub-models. Parameters of both sub-models are estimated jointly.

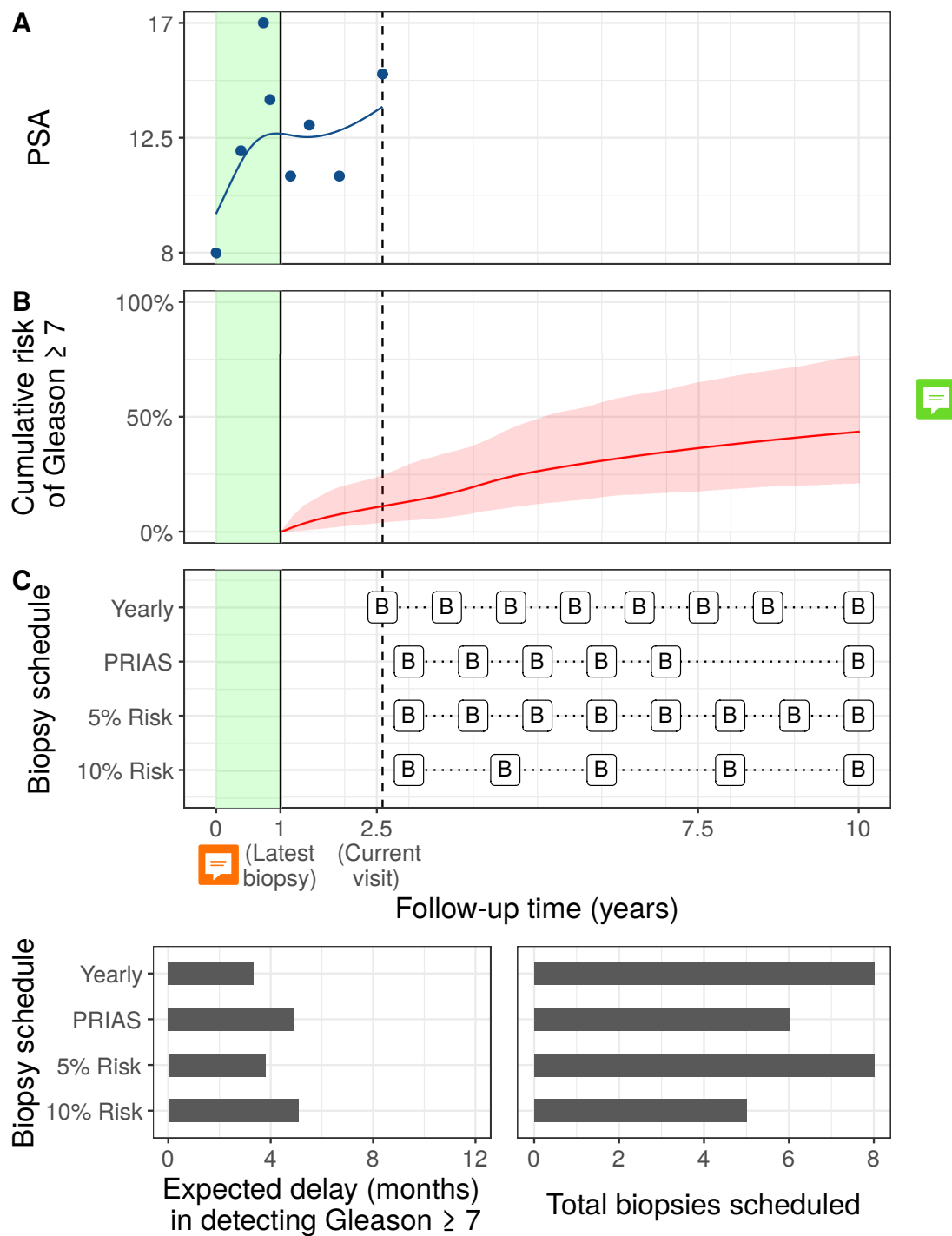


Figure 4: **Personalized and fixed schedules of biopsies for new patient.** **Panel A,B:** show the observed and fitted $\log_2(\text{PSA} + 1)$ measurements, and the dynamic cumulative risk of Gleason ≥ 7 over the follow-up period. **Panel C** shows the personalized and fixed schedules of biopsies with a 'B' indicating the time of biopsy. In the bottom two panels, the various schedules are compared in terms of the number of biopsies they schedule, and the expected delay in detection of Gleason ≥ 7 if they are followed.

130 3. Results

131 For patients in the PRIAS dataset, the probability of obtaining reclassi-
 132 fication within the first five and ten years is 33% and 42%, respectively (see
 133 Figure 1). That is, ideally more than 50% of the patients do not require any
 134 biopsy in the first ten years. We next discuss the results from the joint model
 135 fitted to the PRIAS dataset. For every ten years increase in a patient age the
 136 corresponding adjusted hazard ratio of reclassification is 1.45 (95%CI: 1.30–
 137 1.63). For an increase in fitted $\log_2\{\text{PSA} + 1\}$ value from the first quartile
 138 of fitted values (2.36) to the third quartile (3.07), the corresponding adjusted
 139 hazard ratio of reclassification is 0.99 (95%CI: 0.89–1.11). On the other hand
 140 an increase in fitted $\log_2\{\text{PSA} + 1\}$ velocity from the first quartile of fitted
 141 velocity (-0.09) to the third quartile (0.31), the corresponding adjusted haz-
 142 ard ratio of reclassification is 2.47 (95%CI: 1.93–2.99). These results indicate
 143 that the velocity of $\log_2\{\text{PSA} + 1\}$ measurements is a stronger predictor of
 144 hazard of reclassification than the $\log_2\{\text{PSA} + 1\}$ value. Detailed parameter
 145 estimates are presented in Tables 2, 3 and 5 of Appendix A.4.

146 Using the joint model fitted to the PRIAS dataset we made risk predic-
 147 tions for GS7 in real PRIAS patients. As shown in Figure 4 of Appendix
 148 B, these risk estimates become more accurate as more data is gathered over
 149 follow-up. The risk estimates also increase when time last biopsy increases.
 150 For these risk predictions we calculated the time dependent area under the
 151 receiver operating characteristic curves (AUC) and the root mean squared
 152 prediction error (RMSPE). These are shown in Figure 5. For predictions
 153 within PRIAS (internal validation), the time-dependent AUC was between
 154 0.62 and 0.69, and RMSPE between 0.23 and 0.37 over the whole follow-up

155 period. For validation in external cohorts, the AUC was similar to the AUC of
156 PRIAS for all cohorts during the first three years of follow-up. The RMPSE
157 however differed much more during the same period. The AS cohorts closest
158 to PRIAS in terms of RMSPE were Johns Hopkins Active Surveillance
159 and Memorial Sloan Kettering Cancer Center Active Surveillance. Detailed
160 AUC and RMSPE results for all cohorts with 95% bootstrapped confidence
161 intervals are presented in Table 6 to Table 11 of Appendix B.

162 Using the risk predictions for GS7, we developed personalized schedules
163 of biopsy for real PRIAS patients (see Figure 4 and Appendix C's Figure
164 6, 7, 8 and 9). In all of these patients the biopsies denoted by 'B' show
165 that personalized schedules schedule fewer biopsies than fixed schedules. At
166 the same time the expected time delay in detection of GS7 is less than an
167 year for personalized schedules. We have implemented this approach in a
168 web-application (www.ourapp.url) for practical use.

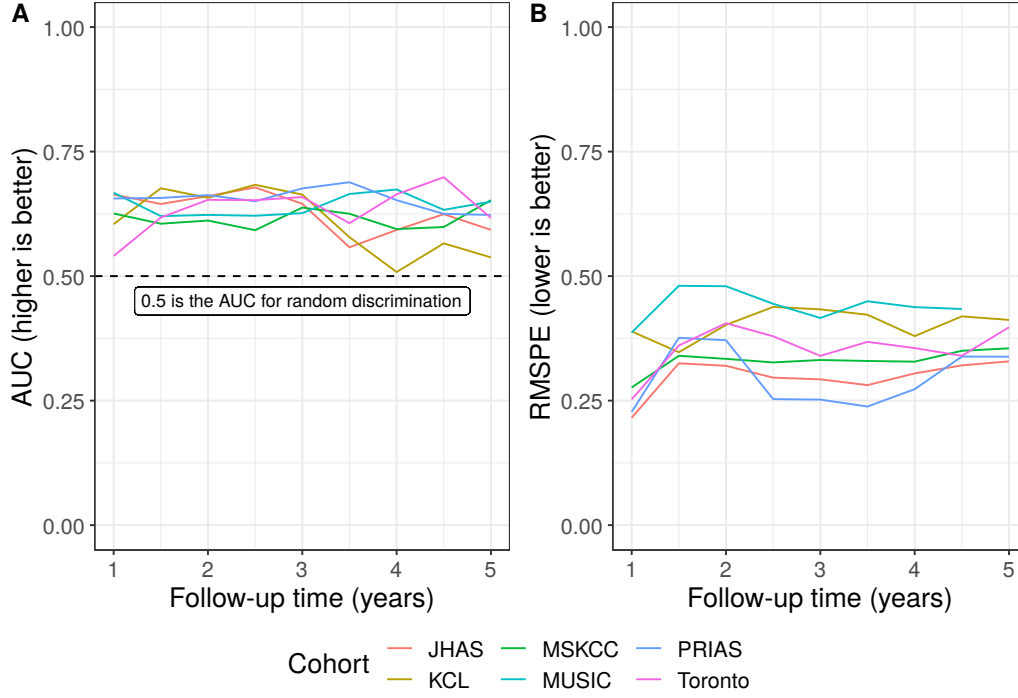


Figure 5: **Validation of predictions of Gleason ≥ 7 (GS7).** In **Panel A** we can see that the time dependent area under the receiver operating characteristic curve or AUC (measure of discrimination) is above 0.5 in PRIAS (internal validation), and in Toronto, JHAS, MSKCC, KCL, and MUSIC AS cohorts (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.

169 4. Discussion

170 We developed personalized schedules for repeat biopsies in PCa patients
 171 enrolled in AS programs. These schedules were based on a patient’s risk for
 172 having a Gleason ≥ 7 (GS7). Patient- and visit-specific risks of GS7 were
 173 estimated using their entire history of PSA and repeat biopsies, and baseline
 174 characteristics. Consequently, the personalized schedules were updated as
 175 more data was gathered over follow-up. Risk calculators for GS7 are not new
 176 [11, 16]. However, the novelty of our work is that we developed a methodology
 177 for scheduling personalized biopsies using those risks, and also a methodology
 178 to compare schedules, be it personalized or fixed, in simple terms of burden
 179 and benefit. More specifically, for each schedule we provided patients the
 180 times of biopsy and total biopsies (burden), and the time delay in detection of
 181 GS7 (less is beneficial) expected due to that schedule. We also implemented
 182 our methodology in a web-application.

183 The proposed joint model accounted for the complex correlation struc-
 184 ture that exists between longitudinal PSA measurements and time of GS7
 185 of a patient. It also accounted for PSA measurements that were missing in
 186 patients who obtained GS7. This model adjusts the risks of GS7 upon a neg-
 187 ative repeat biopsy. Thus complete patient information is consolidated into
 188 a single patient risk profile. Our model is fitted to the world’s largest PCa
 189 AS program, PRIAS. We also externally validated our model predictions for
 190 GS7 in five largest AS cohorts that are part of the GAP3 database [8]. We
 191 found that the AUC for predictions of GS7 over the follow-up period (Fig-
 192 ure 5) was similar in external cohorts and PRIAS (internal validation). The
 193 RMSPE however was similar to PRIAS only for Memorial Sloan Kettering

194 Cancer Center and Johns Hopkins cohorts. Given the large size of the latter
195 two cohorts, we expect that our model and the methodology will be useful
196 to a large number of AS patients. Extending our model and methodology in
197 other cohorts only requires fitting the model to their AS dataset.

198 The clinical implications of our work are as follows. The median survival
199 time for GS7 is more than ten years in PRIAS, and in some other cohorts
200 (Figure 1). That is, more than 50% of AS patients do not require any biopsy
201 during the first ten years of follow-up. We hope that our work will address
202 patient apprehensions regarding adverse outcomes in AS, in a more objective
203 manner. Many AS programs still utilize a rigorous schedule of yearly biopsies
204 [3]. However, with concerns about non-compliance and burden of biopsies
205 [5], the availability of our web based tool may encourage patients and doctors
206 to consider personalized schedules.

207 Our work has certain limitations. The proposed model is valid only for
208 the first thirteen years of follow-up in PRIAS, whereas GS7 may occur much
209 later in many patients. Due to this issue, the calibration and discrimination
210 measures of predictions were also less accurate in later follow-up periods.
211 These issue can be mitigated by refitting the model as more follow-up data
212 is gathered in PRIAS. While we focused only on GS7, it is susceptible to
213 inter-observer variation. Models which account for this variation [11, 17] will
214 be interesting to investigate further. However, the methodology to schedule
215 biopsies, and to estimate the consequences of following a schedule need not
216 change. There is also a potential for including diagnostic information from
217 novel biomarkers, quality of life measures, and magnetic resonance imaging
218 (MRI). Currently, this data is very sparsely available in the PRIAS dataset.

219 However, in future, adding this information in our model is trivial. This is
220 because modeling correlation for extra outcomes, mainly entails connecting
221 sub-models for the outcomes to shared random effects (see Figure 3). Our
222 model can also be used to schedule MRI scans, since they are expensive in
223 developing countries.

224 5. Conclusions

225 We developed a novel methodology for scheduling biopsies to detect Glea-
226 son ≥ 7 (GS7) in PCa patients enrolled in AS. Our methodology consolidates
227 a patient's entire history of PSA and repeat biopsies, and baseline character-
228 istics into risk profile of GS7 over his follow-up period. It then utilizes this
229 risk profile to schedule biopsies in a personalized manner. The personalized
230 schedule is updated as more patient data is gathered over follow-up. To as-
231 sist patients in making the choice of the best biopsy schedule, we provided
232 them personalized burden (time and total biopsies), and personalized bene-
233 fit (less time delay in detection of GS7 is beneficial), for both personalized
234 and currently used schedules. Lastly, we implemented this approach in a
235 web-application.

236 References

- 237 1. Briganti A, Fossati N, Catto JW, Cornford P, Montorsi F, Mottet N,
 238 Wirth M, Van Poppel H. Active surveillance for low-risk prostate cancer:
 239 the European Association of Urology position in 2018. *European urology*
 240 2018;74(3):357–68.
- 241 2. Bul M, Zhu X, Valdagni R, Pickles T, Kakehi Y, Rannikko A, Bjartell A,
 242 Van Der Schoot DK, Cornel EB, Conti GN, et al. Active surveillance for
 243 low-risk prostate cancer worldwide: the prias study. *European urology*
 244 2013;63(4):597–603.
- 245 3. Nieboer D, Tomer A, Rizopoulos D, Roobol MJ, Steyerberg EW. Active
 246 surveillance: a review of risk-based, dynamic monitoring. *Translational*
 247 *andrology and urology* 2018;7(1):106–15.
- 248 4. Loeb S, Carter HB, Schwartz M, Fagerlin A, Braithwaite RS, Lepor H.
 249 Heterogeneity in active surveillance protocols worldwide. *Reviews in*
 250 *urology* 2014;16(4):202–3.
- 251 5. Bokhorst LP, Alberts AR, Rannikko A, Valdagni R, Pickles T, Kakehi Y,
 252 Bangma CH, Roobol MJ, PRIAS study group . Compliance rates with
 253 the Prostate Cancer Research International Active Surveillance (PRIAS)
 254 protocol and disease reclassification in noncompliers. *European Urology*
 255 2015;68(5):814–21.
- 256 6. Inoue LY, Lin DW, Newcomb LF, Leonardson AS, Ankerst D, Gulati R,
 257 Carter HB, Trock BJ, Carroll PR, Cooperberg MR, et al. Comparative

- analysis of biopsy upgrading in four prostate cancer active surveillance cohorts. *Annals of internal medicine* 2018;168(1):1–9.
7. Bratt O, Carlsson S, Holmberg E, Holmberg L, Johansson E, Josefsson A, Nilsson A, Nyberg M, Robinson D, Sandberg J, et al. The study of active monitoring in sweden (sams): a randomized study comparing two different follow-up schedules for active surveillance of low-risk prostate cancer. *Scandinavian journal of urology* 2013;47(5):347–55.
 8. Bruinsma SM, Zhang L, Roobol MJ, Bangma CH, Steyerberg EW, Nieboer D, Van Hemelrijck M, consortium MFGAPPCASG, Trock B, Ehdaie B, et al. The movember foundation’s gap3 cohort: a profile of the largest global prostate cancer active surveillance database to date. *BJU international* 2018;121(5):737–44.
 9. de Carvalho TM, Heijnsdijk EA, de Koning HJ. Estimating the risks and benefits of active surveillance protocols for prostate cancer: a microsimulation study. *BJU international* 2017;119(4):560–6.
 10. Tomer A, Nieboer D, Roobol MJ, Steyerberg EW, Rizopoulos D. Personalized schedules for surveillance of low-risk prostate cancer patients. *Biometrics* 2019;75(1):153–62.
 11. Coley RY, Zeger SL, Mamawala M, Pienta KJ, Carter HB. Prediction of the pathologic gleason score to inform a personalized management program for prostate cancer. *European urology* 2017;72(1):135–41.
 12. Rizopoulos D. Joint Models for Longitudinal and Time-to-Event Data: With Applications in R. CRC Press; 2012. ISBN 9781439872864.

- 281 13. Laird NM, Ware JH, et al. Random-effects models for longitudinal data.
282 *Biometrics* 1982;38(4):963–74.
- 283 14. Rizopoulos D. The R package JMBayes for fitting joint models for lon-
284 gitudinal and time-to-event data using MCMC. *Journal of Statistical*
285 *Software* 2016;72(7):1–46.
- 286 15. Rizopoulos D, Molenberghs G, Lesaffre EM. Dynamic predictions with
287 time-dependent covariates in survival analysis using joint modeling and
288 landmarking. *Biometrical Journal* 2017;59(6):1261–76.
- 289 16. Ankerst DP, Xia J, Thompson Jr IM, Hoeffler J, Newcomb LF, Brooks
290 JD, Carroll PR, Ellis WJ, Gleave ME, Lance RS, et al. Precision
291 medicine in active surveillance for prostate cancer: development of the
292 canary–early detection research network active surveillance biopsy risk
293 calculator. *European urology* 2015;68(6):1083–8.
- 294 17. Balasubramanian R, Lagakos SW. Estimation of a failure time distribu-
295 tion based on imperfect diagnostic tests. *Biometrika* 2003;90(1):171–82.