

# Personalized Biopsy Schedules Based on Risk of Gleason Upgrading for Low-Risk Prostate Cancer Active Surveillance Patients\*

Anirudh Tomer, MSc<sup>a,\*</sup>, Daan Nieboer, MSc<sup>b,c</sup>, Monique J. Roobol, PhD<sup>c</sup>, Anders Bjartell, MD, PhD<sup>d</sup>, Ewout W. Steyerberg, PhD<sup>b,e</sup>, Dimitris Rizopoulos, PhD<sup>a</sup>, Movember Foundation's Global Action Plan Prostate Cancer Active Surveillance (GAP3) consortium<sup>f</sup>

<sup>a</sup>*Department of Biostatistics, Erasmus University Medical Center, Rotterdam, the Netherlands*

<sup>b</sup>*Department of Public Health, Erasmus University Medical Center, Rotterdam, the Netherlands*

<sup>c</sup>*Department of Urology, Erasmus University Medical Center, Rotterdam, the Netherlands*

<sup>d</sup>*Department of Urology, Skåne University Hospital, Malmö, Sweden*

<sup>e</sup>*Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands*

<sup>f</sup>*The Movember Foundation's Global Action Plan Prostate Cancer Active Surveillance (GAP3) consortium members presented in Appendix A*

---

## Abstract

**Objective:** To develop a model and methodology for predicting the risk of Gleason *upgrading* in prostate cancer active surveillance (AS) patients, and using the predicted risks to create risk-based *personalized* biopsy schedules as an alternative to one-size-fits-all schedules (e.g., annually). Furthermore,

---

\*Word count: 3636

\*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](mailto:a.tomer@erasmusmc.nl) (Anirudh Tomer, MSc), [d.nieboer@erasmusmc.nl](mailto:d.nieboer@erasmusmc.nl) (Daan Nieboer, MSc), [m.roobol@erasmusmc.nl](mailto:m.roobol@erasmusmc.nl) (Monique J. Roobol, PhD), [anders.bjartell@med.lu.se](mailto:anders.bjartell@med.lu.se) (Anders Bjartell, MD, PhD), [e.w.steyerberg@lumc.nl](mailto:e.w.steyerberg@lumc.nl) (Ewout W. Steyerberg, PhD), [d.rizopoulos@erasmusmc.nl](mailto:d.rizopoulos@erasmusmc.nl) (Dimitris Rizopoulos, PhD)

to assist patients and doctors in making shared decisions of biopsy schedules, by providing them quantitative estimates of the *burden* and *benefit* of opting for personalized versus any other schedule in AS. Last, to externally validate our model and implement it along with personalized schedules in a ready to use web-application.

**Materials and Methods:** Repeat prostate-specific antigen (PSA) measurements, timing and results of previous biopsies, and age at baseline from the world’s largest AS study, Prostate Cancer Research International Active Surveillance or PRIAS (7813 patients, 1134 experienced upgrading). We fitted a Bayesian joint model for time-to-event and longitudinal data to this dataset. We then validated our model externally in the largest six AS cohorts of the Movember Foundation’s Global Action Plan (GAP3) database ( $> 20,000$  patients, 27 centers worldwide). Using the model predicted upgrading-risks, we scheduled biopsies whenever a patient’s upgrading-risk was above a certain threshold. To assist patients/doctors in choice of this threshold, and to compare the resulting personalized schedule with currently practiced schedules, along with the timing and the total number of biopsies (burden) planned, for each schedule we provided them the time delay expected in detecting upgrading (shorter is better).

**Results:** The cause-specific cumulative upgrading-risk at year five of follow-up was 35% in PRIAS, and at most 50% in GAP3 cohorts. In the PRIAS based model, PSA velocity was a stronger predictor of upgrading (Hazard Ratio: 2.47, 95%CI: 1.93–2.99) than PSA value (Hazard Ratio: 0.99,

95%CI: 0.89–1.11). Our model had a moderate area under the receiver operating characteristic curve (0.6–0.7) in validation cohorts. The prediction error was moderate (0.1–0.2) in validation cohorts where the impact of PSA value and velocity on upgrading-risk was similar to PRIAS, but large (0.2–0.3) otherwise. Our model required recalibration of baseline upgrading-risk in validation cohorts. We implemented the validated models and the methodology for personalized schedules in a web-application (<http://tiny.cc/biopsy>).

**Conclusions:** We successfully developed and validated a model for predicting upgrading-risk, and providing risk-based personalized biopsy decisions, in prostate cancer AS. Personalized prostate biopsies are a novel alternative to fixed one-size-fits-all schedules that may help to reduce unnecessary prostate biopsies while maintaining cancer control. The model and schedules made available via a web-application enable shared decision making of biopsy schedules by comparing fixed and personalized schedules on total biopsies and expected time delay in detecting upgrading.

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

---

## 1. Introduction

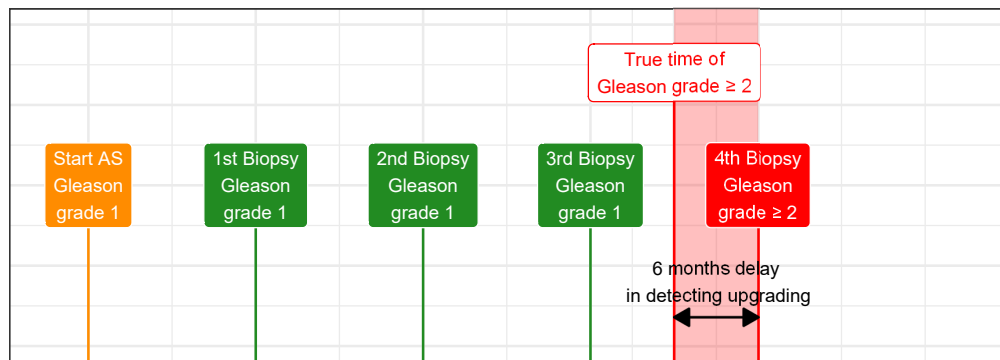
Patients with low- and very low-risk screening-detected localized prostate cancer are recommended active surveillance (AS) usually, instead of immediate radical treatment [1]. In AS, cancer progression is monitored routinely via prostate-specific antigen (PSA), digital rectal examination (DRE), repeat

6 biopsies, and recently, magnetic resonance imaging (MRI). Among these, the  
 7 strongest indicator of cancer-related outcomes is the biopsy Gleason grade  
 8 group [2]. When it increases from group 1 (Gleason 3+3) to 2 (Gleason 3+4)  
 9 or higher, it is called *upgrading* [3]. Upgrading is an important endpoint in  
 10 AS upon which patients are commonly advised curative treatment [4].

11 Biopsies in AS are always conducted with a time gap between them.  
 12 Consequently, upgrading is always detected with a time delay (Figure 1)  
 13 that cannot be measured directly. In this regard, to detect upgrading timely,  
 14 many patients are prescribed fixed and frequent biopsies, most often annu-  
 15 ally [6]. However, such one-size-fits-all schedules lead to unnecessary biopsies  
 16 in slow/non-progressing patients. Biopsies are invasive, may be painful, and  
 17 are prone to medical complications such as bleeding and septicemia[7]. Thus,  
 18 biopsy burden and patient non-compliance to frequent biopsies [8] have raised  
 19 concerns regarding the optimal biopsy schedule [9, 10] in AS.

20 Except for the confirmatory biopsy at year one of AS [8], opinions and  
 21 practice regarding the timing of remaining biopsies lack agreement [5]. Some  
 22 AS programs utilize patients' observed PSA, DRE, previous biopsy Glea-  
 23 son grade, and lately, MRI results to decide biopsies [? 4, 5]. In contrast,  
 24 others discourage schedules based on clinical data and MRI results [? 6],  
 25 and instead support periodical one-size-fits-all biopsy schedules. Further-  
 26 more, some suggest replacing frequent periodical schedules with infrequent  
 27 ones (e.g., biennially) [9, 11]. Each of these approaches has limitations. For  
 28 example, one-size-fits-all schedules can lead to many unnecessary biopsies  
 29 because of differences in baseline *upgrading-risk* across cohorts [9]. Whereas,  
 30 since observed clinical data has measurement error (e.g., PSA fluctuations),

### A Biopsy every year



### B Biopsy every 2 years

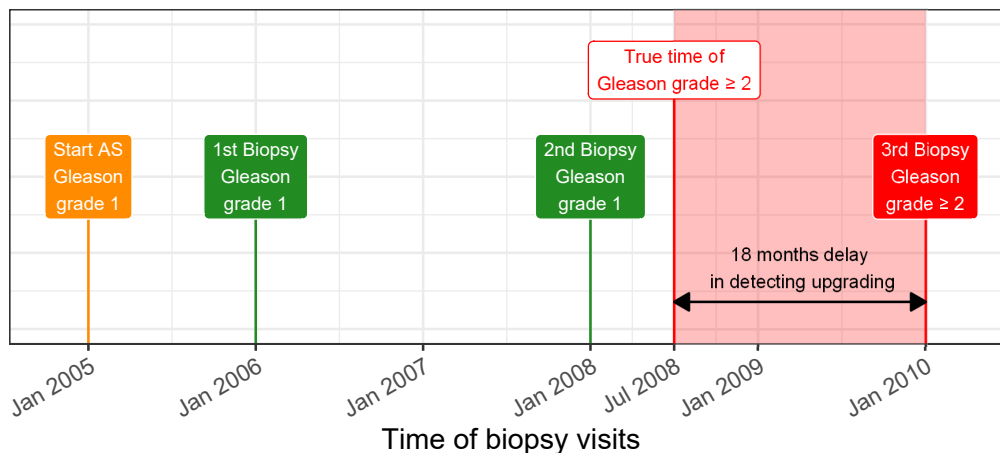
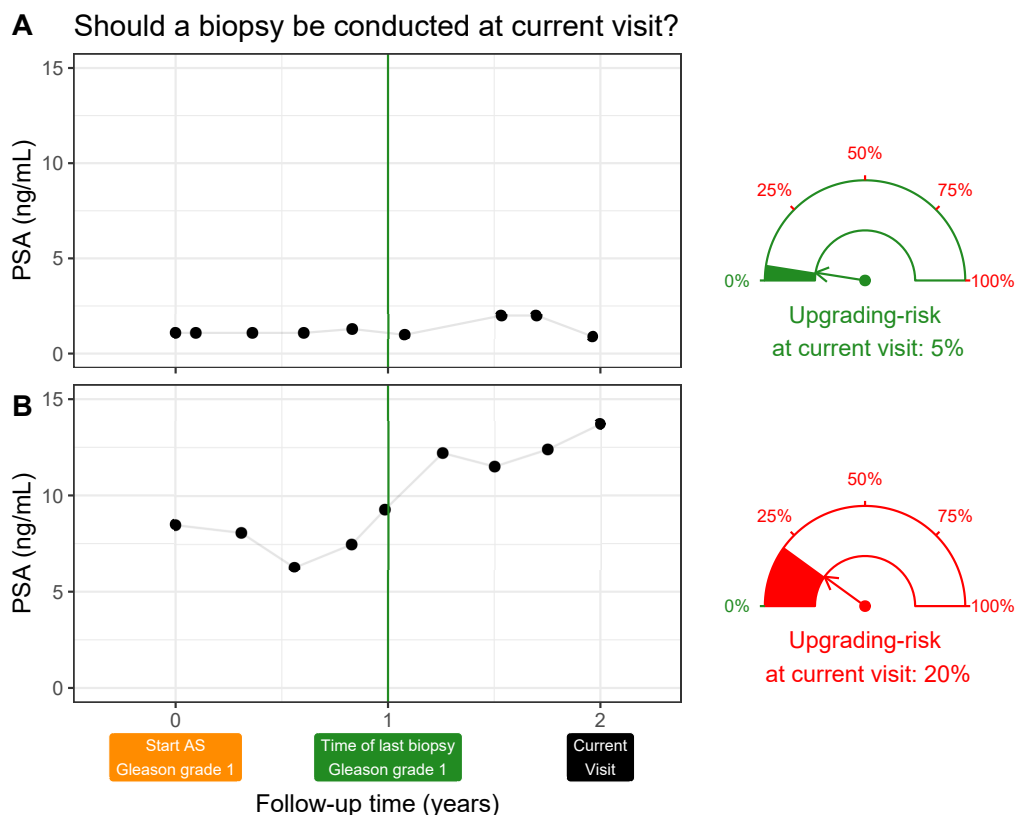


Figure 1: **Trade-off between the timing and number of biopsies (burden) and time delay in detecting Gleason upgrading (shorter is better):** The true time of Gleason upgrading (increase in Gleason grade group from group 1 to 2 or higher) for the patient in this figure is July 2008. When biopsies are scheduled annually (**Panel A**), upgrading 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**), upgrading 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 upgrading is observed as an interval. For example, between Jan 2008–Jan 2009 in **Panel A** and between Jan 2008–Jan 2010 in **Panel B**. The phrase ‘Gleason grade group’ is shortened to ‘Gleason grade’ for brevity.

31 a flaw of using it directly is that it may lead to poor decisions. Also, deci-  
 32 sions based on clinical data typically rely only on the latest data point and  
 33 ignore previous repeated measurements. A novel alternative that counters  
 34 these drawbacks is first processing patient data via a statistical model, and  
 35 subsequently using model predicted upgrading-risks to create *personalized*  
 36 biopsy schedules [5] (Figure 2). While, upgrading-risk calculators are not  
 37 new [15, 25, 12, 13], not all are personalized either. Besides, they do not  
 38 specify how risk predictions can be exploited to create a schedule.

39 This work is motivated by the problem of scheduling biopsies in AS. We  
 40 have two goals. First, we want to assist practitioners in using clinical data  
 41 in biopsy decisions in a statistically sound manner. To this end, we plan to  
 42 develop a robust, generalizable statistical model that provides reliable indi-  
 43 vidual upgrading-risk in AS. Subsequently, we will employ these predictions  
 44 to derive risk-based personalized biopsy schedules. Our second goal is to en-  
 45 able shared decision making of biopsy schedules. We intend to achieve this by  
 46 allowing patients and doctors to compare the *burden* and *benefit* (Figure 1) of  
 47 opting for personalized schedules versus periodical schedules versus schedules  
 48 based on clinical data. Specifically, we propose timing and number of planned  
 49 biopsies (more/frequent are burdensome), and the expected time delay in  
 50 detecting upgrading (shorter is beneficial) for any given schedule. While  
 51 fulfilling our goals, we want to capture the maximum possible information  
 52 from the available data. Hence, we will use all repeated PSA measurements  
 53 of patients, previous biopsy results, and baseline characteristics. To fit this  
 54 model, we will utilize data of the world’s largest AS study, Prostate Cancer  
 55 Research International Active Surveillance (PRIAS). To evaluate our model,



**Figure 2: Motivation for upgrading-risk based personalized biopsy decisions:** To utilize patients' complete longitudinal data and results from previous biopsies in making biopsy decisions. For this purpose, we first process data using a statistical model and then utilize the patient-specific predictions for risk of Gleason upgrading to schedule biopsies. For example, Patient 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 (PSA) profile remained stable until his current visit at year two, whereas patient B's profile has shown a rise. Consequently, patient B's upgrading-risk at the current visit (year two) 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.

we will externally validate it in the largest six AS cohorts from the Movember Foundation’s Global Action Plan (GAP3) database [17]. Last, we aim to implement the validated model and methodology in a web-application.

## 2. Patients and Methods

### 2.1. Study Cohort

For developing a statistical model to predict upgrading-risk, we used the world’s largest AS dataset, Prostate Cancer International Active Surveillance or PRIAS [4], dated April 2019 (Table 1). In PRIAS, biopsies were scheduled at year one, four, seven, ten, and additional yearly biopsies were scheduled when PSA doubling time was between zero and ten years. We selected all 7813 patients who had Gleason grade group 1 at inclusion in AS. Our primary event of interest is an increase in this Gleason grade group observed upon repeat biopsy, called *upgrading* (1134 patients). Upgrading is a trigger for treatment advice in PRIAS. Some examples of treatment options in active surveillance are radical prostatectomy, brachytherapy, definitive radiation therapy, and other alternative local treatments such as cryosurgery, High Intensity Focused Ultrasound, and External Beam Radiation Therapy. Comprehensive details on treatment options and their side effects are available in EAU-ESTRO-SIOG guidelines on prostate cancer [? ]. In PRIAS 2250 patients were provided treatment based on their PSA, the number of biopsy cores with cancer, or anxiety/other reasons. However, our reasons for focusing solely on upgrading are that upgrading is strongly associated with cancer-related outcomes, and other treatment triggers vary between cohorts [5].



For externally validating our model’s predictions, we selected the following largest (by the number of repeated measurements) six cohorts from Movember Foundation’s GAP3 database [17] version 3.1, covering nearly 73% of the GAP3 patients: the University of Toronto AS (Toronto), Johns Hopkins AS (Hopkins), Memorial Sloan Kettering Cancer Center AS (MSKCC), King’s College London AS (KCL), Michigan Urological Surgery Improvement Collaborative AS (MUSIC), and University of California San Francisco AS (UCSF, version 3.2). Only patients with a Gleason grade group 1 at the time of inclusion in these cohorts were selected. Summary statistics are presented in Supplementary A.2.

*Choice of predictors:.* In our model, we used all repeated PSA measurements, the timing of the previous biopsy and Gleason grade, and age at inclusion in AS. Other predictors such as prostate volume, MRI results can also be important. MRI is utilized already for targeting biopsies, but regarding its use in deciding the time of biopsies, there are arguments both for and against it [? ? ? ]. MRI is still a recent addition in most AS protocols. Consequently, repeated MRI data is very sparsely available in both PRIAS and GAP3 databases to make a stable prediction model. Prostate volume data is also sparsely available, especially in validation cohorts. Based on these reasons, we did not include them in our model.

## 2.2. Statistical Model

Modeling an AS dataset such as PRIAS, posed certain challenges. First, PSA was measured longitudinally, and over follow-up time it did not always increase linearly. Consequently, we expect that PSA measurements of a pa-

Table 1: **Summary of the PRIAS dataset as of April 2019.** The primary event of interest is upgrading, that is, increase in Gleason grade group from group 1 [2] to 2 or higher. IQR: interquartile range, PSA: prostate-specific antigen. Study protocol URL: <https://www.prias-project.org>

Characteristic	Value
Total patients	7813
Upgrading (primary event)	1134
Treatment	2250
Watchful waiting	334
Loss to follow-up	249
Death (unrelated to prostate cancer)	95
Death (related to prostate cancer)	2
Median age at diagnosis (years)	66 (IQR: 61–71)
Median maximum follow-up 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)

104 tient are more similar to each other than of another patient. In other words,  
 105 we need to accommodate the within-patient correlation for PSA. Second,  
 106 PSA was available only until a patient observed upgrading. Thus, we also  
 107 need to model the association between the Gleason grades and PSA profiles of  
 108 a patient, and handle missing PSA measurements after a patient experienced  
 109 upgrading. Third, since the PRIAS biopsy schedule uses PSA, a patient's  
 110 observed time of upgrading was also dependent on their PSA. Thus, the effect  
 111 of PSA on the upgrading-risk need to be adjusted for the effect of PSA on  
 112 the biopsy schedule. Fourth, many patients obtained treatment and watchful  
 113 waiting before observing upgrading. Since we considered events other than  
 114 upgrading as censoring, the model needs to account for patients' reasons for  
 115 treatment or watchful waiting (e.g., age, treatment based on observed data).  
 116 A model that handles these challenges in a statistically sound manner is the  
 117 joint model for time-to-event and longitudinal data [14, 15, 16].

118 Our joint model consisted of two sub-models. Namely, a linear mixed-  
 119 effects sub-model [18] for longitudinally measured PSA (log-transformed),  
 120 and a relative-risk sub-model (similar to the Cox model) for the interval-  
 121 censored time of upgrading. Patient age was used in both sub-models. Re-  
 122 sults and timing of the previous negative biopsies were used only in the risk  
 123 sub-model. To account for PSA fluctuations [? ], we assumed t-distributed  
 124 PSA measurement errors. The correlation between PSA measurements of the  
 125 same patient was established using patient-specific random-effects. We fitted  
 126 a unique curve to the PSA measurements of each patient (Panel A, Figure 3).  
 127 Subsequently, we calculated the mathematical derivative of the patient's fit-  
 128 ted PSA profile (Equation 2, Supplementary A), to obtain his follow-up time

specific instantaneous PSA velocity (Panel B, Figure 3). This instantaneous velocity is a stronger predictor of upgrading than the widely used average PSA velocity [? ]. We modeled the impact of PSA on upgrading-risk by employing fitted PSA value and instantaneous velocity as predictors in the risk sub-model (Panel C, Figure 3). We adjusted the effect of PSA on upgrading-risk for the PSA dependent PRIAS biopsy schedule by estimating parameters using a full likelihood method (proof in Supplementary A). This approach also accommodates watchful waiting and treatment protocols that are also based on patient data. Specifically, the parameters of our two sub-models were estimated jointly under the Bayesian paradigm (Supplementary A) using the R package **JMbayes** [20].

### 2.3. Risk Prediction and Model Validation

Our model provides predictions for upgrading-risk over the entire future follow-up period of a patient (Panel C, Figure 3). However, we recommend using predictions only after year one. This is because most AS programs recommend a confirmatory biopsy at year one, especially to detect patients who may be misdiagnosed as low-grade at inclusion in AS. The risk predictions for a patient are not calculated only once. Rather, as illustrated in Figure 5 of Supplementary B, risk-predictions update over the follow-up, to account for additional patient data (e.g., new biopsy results, PSA measurements) that becomes available. We validated our model internally in the PRIAS cohort, and externally in the largest six GAP3 database cohorts. We employed calibration plots [22, 23] and follow-up *time-dependent* mean absolute risk prediction error or MAPE [21] to graphically and quantitatively evaluate our model’s risk prediction accuracy, respectively. We assessed our

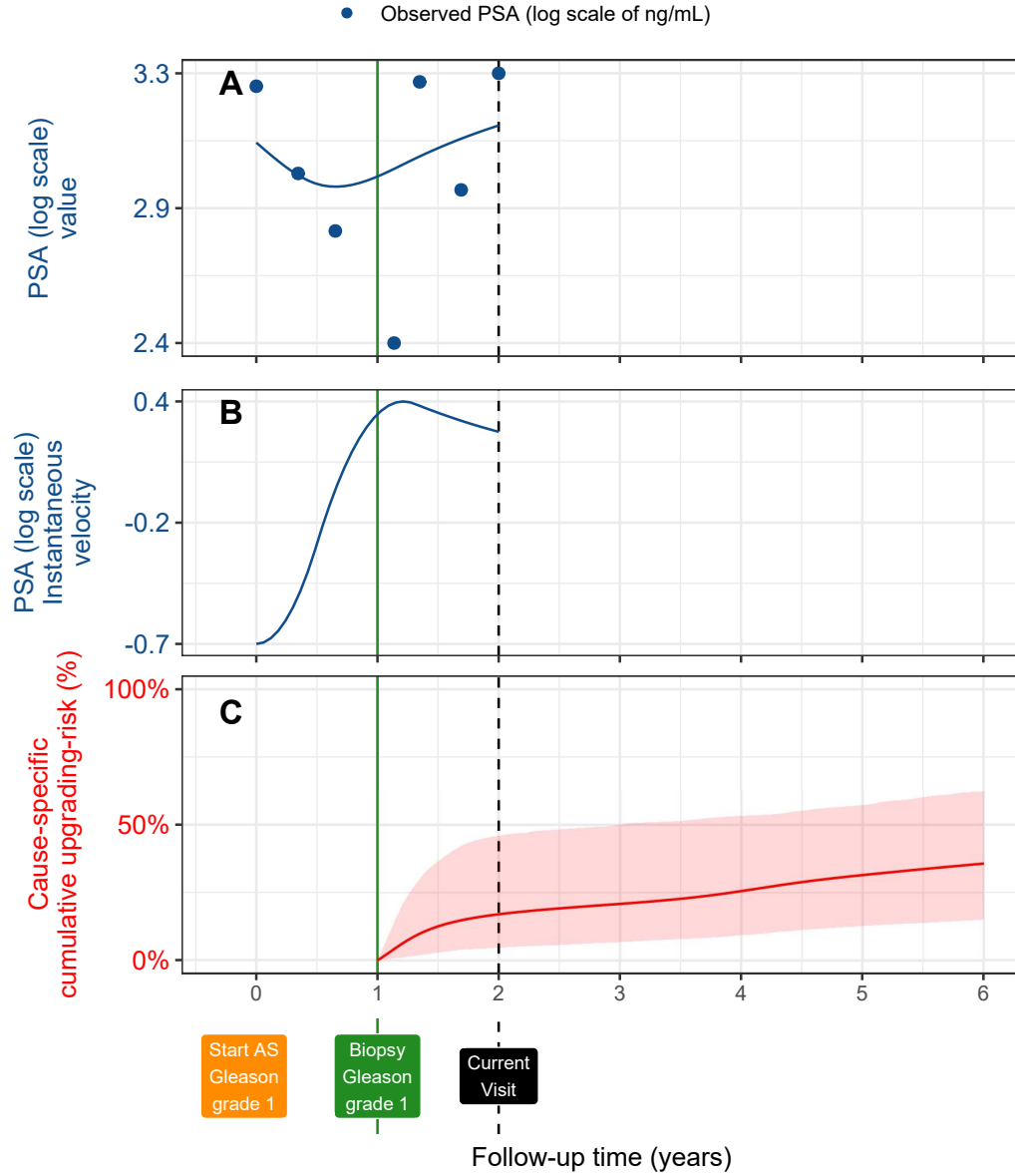


Figure 3: **Illustration of the joint model on a real PRIAS patient.** **Panel A:** Observed PSA (blue dots) and fitted PSA (solid blue line), log-transformed from ng/mL. **Panel B:** Estimated instantaneous velocity of PSA (log-transformed). **Panel C:** Predicted cause-specific cumulative upgrading-risk (95% credible interval shaded). Upgrading is defined as an increase in the Gleason grade group from group 1 [2] to 2 or higher. This upgrading-risk is calculated starting from the time of the latest negative biopsy (vertical green line at year one of follow-up). The joint model estimated it by combining the fitted PSA (log scale) value and instantaneous velocity, and time of the latest negative biopsy. Black dashed line at year two denotes the time of current visit.

154 model’s ability to discriminate between patients who experience/do not ex-  
 155 perience upgrading via the time-dependent area under the receiver operating  
 156 characteristic curve or AUC [21].

157 The aforementioned *time-dependent* AUC and MAPE [21] are temporal  
 158 extensions of their standard versions [23] in a longitudinal setting. Specif-  
 159 ically, at every six months of follow-up, we calculated a unique AUC and  
 160 MAPE for predicting upgrading-risk in the subsequent one year (Supple-  
 161 mentary B.1). For emulating a realistic situation, we calculated the AUC  
 162 and MAPE at each follow-up using only the validation data available un-  
 163 til that follow-up. For example, calculations for AUC and MAPE for the  
 164 time interval year two to year three do not utilize data of patients who pro-  
 165 gressed before year two. Last, to resolve any potential model miscalibration  
 166 in validation cohorts, we aimed to recalibrate our model’s baseline hazard of  
 167 upgrading (Supplementary B.1), individually for each cohort.

### 168 **3. Results**

169 The cause-specific cumulative upgrading-risk at year five of follow-up was  
 170 35% in PRIAS and at most 50% in validation cohorts (Panel B, Figure 4).  
 171 In the fitted PRIAS model, the adjusted hazard ratio (aHR) of upgrading for  
 172 an increase in patient age from 61 to 71 years (25-th to 75-th percentile) was  
 173 1.45 (95%CI: 1.30–1.63). For an increase in fitted PSA value from 2.36 to 3.07  
 174 (25-th to 75-th percentile, log scale), the aHR was 0.99 (95%CI: 0.89–1.11).  
 175 The strongest predictor of upgrading-risk was instantaneous PSA velocity,  
 176 with an increase from -0.09 to 0.31 (25-th to 75-th percentile), giving an  
 177 aHR of 2.47 (95%CI: 1.93–2.99). The aHR for PSA value and velocity was

178 different in each GAP3 cohort (Supplementary Table 8).

179 The time-dependent AUC, calibration plot, and time-dependent MAPE  
 180 of our model are shown in Figure 4, and Supplementary Figure 8. In all co-  
 181 horts, time-dependent AUC was moderate (0.6 to 0.7) over the whole follow-  
 182 up period. Time-dependent MAPE was moderate (0.1 to 0.2) in those cohorts  
 183 where the impact of PSA on upgrading-risk was similar to PRIAS (e.g., Hop-  
 184 kins cohort, Supplementary Table 8), and large (0.2 to 0.3) otherwise. Our  
 185 model was miscalibrated for validation cohorts (Panel B, Figure 4) because  
 186 cohorts had differences in inclusion criteria (e.g., PSA density) and follow-up  
 187 protocols [17] which were not accounted in our model. Consequently, the  
 188 PRIAS based model’s fitted baseline hazard did not correspond to the base-  
 189 line hazard in validation cohorts. To solve this problem, we recalibrated the  
 190 baseline hazard of upgrading in validation cohorts (Supplementary Figure 6).  
 191 We compared risk predictions from the recalibrated models, with predictions  
 192 from separately fitted cohort-specific joint models (Supplementary Figure 7).  
 193 The difference in predictions was lowest in the Johns Hopkins cohort (impact  
 194 of PSA on upgrading-risk similar to PRIAS). Comprehensive results are in  
 195 Supplementary A.3 and Supplementary B.

### 196 3.1. *Personalized Biopsy Schedules*

197 We employed the PRIAS based fitted model to create personalized biopsy  
 198 schedules for real PRIAS patients. Particularly, first using the model and pa-  
 199 tient’s observed data, we predicted his cumulative upgrading-risk (Figure 5)  
 200 on all of his future follow-up visits (biannually in PRIAS). Subsequently,  
 201 we planned biopsies on those future visits where his conditional cumulative  
 202 upgrading-risk was more than a certain threshold (see Supplementary C for

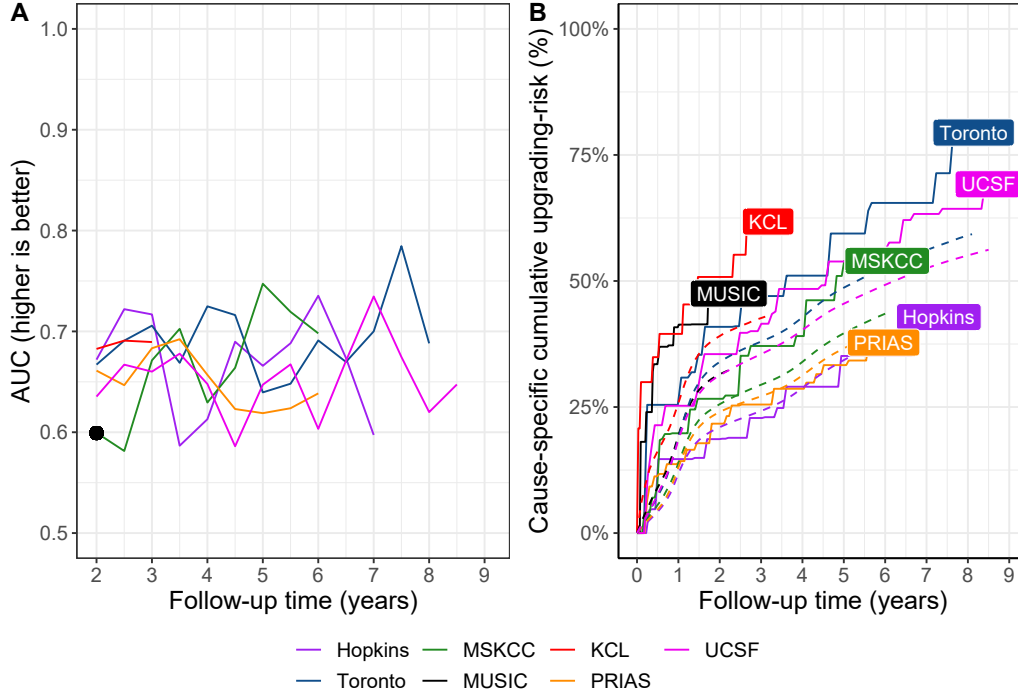


Figure 4: **Model Validation Results.** **Panel A:** time-dependent area under the receiver operating characteristic curve or AUC (measure of discrimination). AUC at year one is not shown because we do not intend to replace the confirmatory biopsy at year one. **Panel B:** calibration-at-large indicates model miscalibration. This is because solid lines depicting the non-parametric estimate of the cause-specific cumulative upgrading-risk [24], and dashed lines showing the average cause-specific cumulative upgrading-risk obtained using the joint model fitted to the PRIAS dataset, are not overlapping. Recalibrating the baseline hazard of upgrading resolved this issue (Supplementary Figure 6). Full names of Cohorts are *PRIAS*: Prostate Cancer International Active Surveillance, *Toronto*: University of Toronto Active Surveillance, *Hopkins*: 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, *UCSF*: University of California San Francisco AS.



203 mathematical details). The choice of this threshold dictates the timing of  
 204 biopsies in a risk-based personalized schedule. For example, personalized  
 205 schedules based on 5% and 10% risk thresholds are shown in Figure 5, and  
 206 in Supplementary Figure 10–12.

207 To facilitate the choice of a risk-threshold, and for comparing the conse-  
 208 quences of opting for a risk-based schedule versus any other schedule (e.g.,  
 209 annual, PRIAS), we predict expected time delay in detecting upgrading for  
 210 following a schedule. We are able to predict this delay for any schedule. For  
 211 example, in Panel C of Figure 5, the annual schedule has the least expected  
 212 delay. In contrast, a personalized schedule based on a 10% risk threshold has  
 213 a slightly larger expected delay, but it also schedules much fewer biopsies.  
 214 An important aspect of this delay is that it is personalized as well. That is,  
 215 even if two different patients are prescribed the same biopsy schedule, their  
 216 expected delays will be different. This is because delay is estimated using all  
 217 available clinical data of the patient (see Supplementary C). While the timing  
 218 and the total number of planned biopsies denote the burden of a schedule, a  
 219 shorter expected time delay in detecting upgrading can be a benefit. These  
 220 two, along with other measures such as a patient’s comorbidities, anxiety,  
 221 etc., can help to make an informed biopsy decision.

### 222 3.2. *Web-Application*

223 We implemented the PRIAS based model, recalibrated models for GAP3  
 224 cohorts, and personalized schedules in a user-friendly web-application [https:](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/)  
 225 [//emcbiostatistics.shinyapps.io/prias\\_biopsy\\_recommender/](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/). This  
 226 application works on both desktop and mobile devices. Users must first  
 227 choose the cohort to which the patient belongs (left panel), and then upload

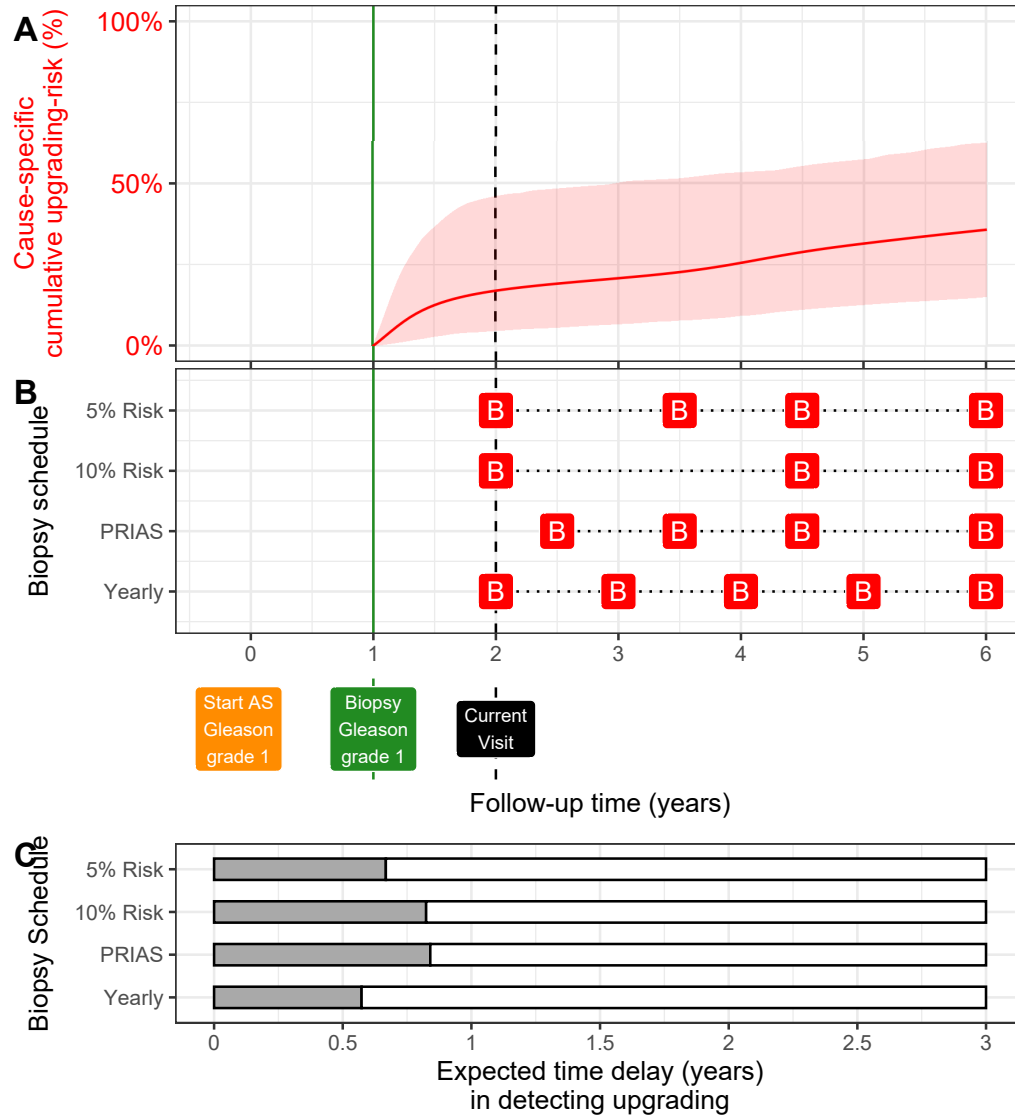


Figure 5: **Illustration of personalized and fixed schedules of biopsies for patient from Figure 3.** **Panel A:** Predicted cumulative upgrading-risk (95% credible interval shaded). **Panel B:** Different biopsy schedules with a red 'B' indicating a future biopsy. Risk: 5% and Risk: 10% are personalized schedules in which a biopsy is planned whenever the conditional cause-specific cumulative upgrading-risk is above 5% or 10% risk, respectively. Smaller risk thresholds lead to more frequently planned biopsies. Green vertical line at year one is the time of the latest negative biopsy. Black dashed line at year two denotes the time of the current visit. **Panel C:** Expected time delay in detecting upgrading (years) if patient progresses before year six. The maximum time delay with limited adverse consequences is three years [11]. A compulsory biopsy was scheduled at year six (maximum biopsy scheduling time in PRIAS, Supplementary C) in all schedules for a meaningful comparison between delays.

228 patient data in Microsoft Excel format. Internally, the web-application loads  
 229 the appropriate validated and recalibrated model for that cohort. The max-  
 230 imum follow-up time up to which predictions can be obtained depends on  
 231 each cohort (Supplementary Table 9). The web-application supports person-  
 232 alized, annual, and PRIAS schedules. For personalized schedules, users can  
 233 control the choice of risk-threshold. The web-application also compares the  
 234 resulting risk-based schedule’s timing of biopsies, and expected time delay  
 235 in detecting upgrading, with annual and PRIAS schedules, to enable sharing  
 236 biopsy decision making.

## 237 4. Discussion

238 We successfully developed and externally validated a statistical model for  
 239 predicting upgrading-risk [3] in prostate cancer AS, and providing risk-based  
 240 personalized biopsy decisions. Our work has four novel features over earlier  
 241 risk calculators [15, 25]. First, our model was fitted to the world’s largest  
 242 AS dataset PRIAS and externally validated in the largest six cohorts of the  
 243 Movember Foundation’s GAP3 database [17]. Second, the model predicts a  
 244 patient’s current and future upgrading-risk in a personalized manner. Third,  
 245 using the predicted risks, we created personalized biopsy schedules. We also  
 246 calculated the expected time delay in detecting upgrading (less is benefi-  
 247 cial) for following any schedule. Thus, patients/doctors can compare sched-  
 248 ules before making a choice. Fourth, we implemented our methodology in a  
 249 user-friendly web-application ([https://emcbiostatistics.shinyapps.io/  
 250 prias\\_biopsy\\_recommender/](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/)) for both PRIAS and validated cohorts.

251 Our model and methods can be useful for numerous patients from PRIAS

252 and the validated GAP3 cohorts (nearly 73% of all GAP3 patients). The  
 253 model utilizes all repeated PSA measurements, results of previous biopsies,  
 254 and baseline characteristics of a patient. We could not include MRI and  
 255 PSA density because of sparsely available data in both PRIAS and GAP3  
 256 databases. But, our model is extendable to include them in the near fu-  
 257 ture. A benefit of our model is that it allows the biopsy schedule, schedule  
 258 of longitudinal measurements, and loss to follow-up in each cohort to de-  
 259 pend on patient age and PSA characteristics. Consequently, in future, when  
 260 MRI data is included in the model, our model will also accommodate biopsy  
 261 schedules dependent on MRI data or MRI schedules dependent on previous  
 262 biopsy results, PSA characteristics, and age of the patient (mathematical  
 263 proof in Appendix A.2). An additional advantage of our model and result-  
 264 ing personalized schedules is that they update as more patient data becomes  
 265 available over follow-up. The current discrimination ability of our model,  
 266 exhibited by the *time-dependent* AUC, was between 0.6 and 0.7 over-follow.  
 267 While this is moderate, it is also so because unlike the standard AUC [23] the  
 268 time-dependent AUC is more conservative as it utilizes only the validation  
 269 data available until the time at which it is calculated. The same holds for the  
 270 time-dependent MAPE (mean absolute prediction error). Although, MAPE  
 271 varied much more between cohorts than AUC. In cohorts where the effect  
 272 size for the impact of PSA value and velocity on upgrading-risk was similar  
 273 to that for PRIAS (e.g., Hopkins cohort), MAPE was moderate. Otherwise,  
 274 MAPE was large (e.g., KCL and MUSIC cohorts). We required recalibration  
 275 of our model’s baseline hazard of upgrading for all validation cohorts.

276 The clinical implications of our work are as follows. First, the cause-

277 specific cumulative upgrading-risk at year five of follow-up was at most 50% in  
 278 all cohorts (Panel B, Figure 4). That is, many patients may not require some  
 279 of the biopsies planned in the first five years of AS. Given the non-compliance  
 280 and burden of frequent biopsies [8], the availability of our methodology as  
 281 a web-application may encourage patients/doctors to consider upgrading-  
 282 risk based personalized schedules instead. Despite the moderate predictive  
 283 performance, we expect the overall impact of our model to be positive. There  
 284 are two reasons for this. First, the risk of adverse outcomes because of  
 285 personalized schedules is quite low because of the low rate of metastases  
 286 and prostate cancer specific mortality in AS patients (Table 1). Second,  
 287 studies [9] have suggested that after the confirmatory biopsy at year one  
 288 of follow-up, biopsies may be done as infrequently as every two to three  
 289 years, with limited adverse consequences. In other words, longer delays in  
 290 detecting upgrading may be acceptable after the first negative biopsy. To  
 291 evaluate the potential harm of personalized schedules, we compared them  
 292 with fixed schedules in a realistic and extensive simulation study [10]. We  
 293 concluded that personalized schedules plan, on average, six fewer biopsies  
 294 compared to annual schedule and two fewer biopsies than the PRIAS schedule  
 295 in slow/non-progressing AS patients, while maintaining almost the same time  
 296 delay in detecting upgrading as PRIAS schedule. Personalized schedules with  
 297 different risk thresholds indeed have different performances across cohorts.  
 298 Thus, to assist patients/doctors in choosing between fixed schedules and  
 299 personalized schedules based on different risk thresholds, the web-application  
 300 provides a patient-specific estimate of the expected time delay in detecting  
 301 upgrading, for both personalized and fixed schedules. We hope that access

302 to these estimates will objectively address patient apprehensions regarding  
 303 adverse outcomes in AS. Last, we note that our web-application should only  
 304 be used to decide biopsies after the compulsory confirmatory biopsy at year  
 305 one of follow-up.

306 This work has certain limitations. Predictions for upgrading-risk and per-  
 307 sonalized schedules are available only for a currently limited, cohort-specific,  
 308 follow-up period (Supplementary Table 9). This problem can be mitigated  
 309 by refitting the model with new follow-up data in the future. Recently, some  
 310 cohorts started utilizing MRI to explore the possibility of targeting visible le-  
 311 sions by biopsy. Presently, the GAP3 database has limited PSA density and  
 312 MRI follow-up data available. Since PSA density is used as an entry criterion  
 313 in some active surveillance studies, including it as a predictor can improve  
 314 the model. In this regard, the current model can be extended to include both  
 315 MRI and PSA density data as predictors when they become available in the  
 316 future. Our model schedules biopsies in a personalized manner, but the pa-  
 317 tient burden can be decreased even more if we also personalize the schedule  
 318 of PSA measurements. A caveat of doing so is that reduction in the number  
 319 of PSA measurements can also lead to an increase in the variance of risk es-  
 320 timates, and also affect the performance of personalized schedules. Although  
 321 we have done a simulation study to conclude that personalized schedules  
 322 may not be any more harmful than PRIAS or annual schedule [? ], with  
 323 an infrequent PSA schedule, these conclusions may not hold. Hence, we do  
 324 not recommend any changes in the schedule of PSA measurements from the  
 325 current protocol of PSA measurements every six months. At the same time,  
 326 personalizing the schedule of both biopsies and PSA measurements together

327 is a research problem we intend to tackle in the near future. We scheduled  
 328 biopsies using cause-specific cumulative upgrading-risk, which ignores com-  
 329 peting events such as treatment based on the number of positive biopsy cores.  
 330 Employing a competing-risk model may lead to improved personalized sched-  
 331 ules. Upgrading is susceptible to inter-observer variation too. Models which  
 332 account for this variation [15, 26] will be interesting to investigate further.  
 333 Even with an enhanced risk prediction model, the methodology for person-  
 334 alized scheduling and calculation of expected time delay (Supplementary C)  
 335 need not change. Last, our web-application only allows uploading patient  
 336 data in Microsoft Excel format. Connecting it with patient databases can  
 337 increase usability.

## 338 5. Conclusions

339 We successfully developed a statistical model and methodology for pre-  
 340 dicting upgrading-risk, and providing risk-based personalized biopsy deci-  
 341 sions, in prostate cancer AS. We externally validated our model, cover-  
 342 ing nearly 73% patients from the Movember Foundations' GAP3 database.  
 343 The model made available via a user-friendly web-application ([https://](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/)  
 344 [emcbiostatistics.shinyapps.io/prias\\_biopsy\\_recommender/](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/)) enables shared  
 345 decision making of biopsy schedules by comparing fixed and personalized  
 346 schedules on total biopsies and expected time delay in detecting upgrading.  
 347 Novel biomarkers and MRI data can be added as predictors in the model to  
 348 improve predictions in the future. Recalibration of baseline upgrading-risk  
 349 is advised for cohorts not validated in this work.

## 350 **Author Contributions**

351 Anirudh Tomer had full access to all the data in the study and takes  
 352 responsibility for the integrity of the data and the accuracy of the data anal-  
 353 ysis.

354 *Study concept and design:* Tomer, Nieboer, Roobol, Bjartell, and Rizopoulos

355 *Acquisition of data:* Tomer, Nieboer, and Roobol

356 *Analysis and interpretation of data:* Tomer, Nieboer, and Rizopoulos

357 *Drafting of the manuscript:* Tomer, and Rizopoulos

358 *Critical revision of the manuscript for important intellectual content:* Tomer,  
 359 Nieboer, Roobol, Bjartell, Steyerberg, and Rizopoulos

360 *Statistical analyses:* Tomer, Nieboer, Steyerberg, and Rizopoulos

361 *Obtaining funding:* Roobol, Steyerberg, and Rizopoulos

362 *Administrative, technical or material support:* Nieboer

363 *Supervision:* Roobol, and Rizopoulos

364 *Other:* none

## 365 **Acknowledgments**

366 We thank Jozien Helleman from the Department of Urology, Erasmus  
 367 University Medical Center, for coordinating the project. The first and last  
 368 authors would like to acknowledge support by Nederlandse Organisatie voor  
 369 Wetenschappelijk Onderzoek (the national research council of the Nether-  
 370 lands) VIDI grant nr. 016.146.301, and Erasmus University Medical Cen-  
 371 ter funding. Part of this work was carried out on the Dutch national e-  
 372 infrastructure with the support of SURF Cooperative. The authors also  
 373 thank the Erasmus University Medical Center's Cancer Computational Bi-



ology Center for giving access to their IT-infrastructure and software that was used for the computations and data analysis in this study. The PRIAS website is funded by the Prostate Cancer Research Foundation, Rotterdam (SWOP). We would like to thank the PRIAS consortium for enabling this research project.

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.

## **Conflicts of Interest**

The authors do not report any conflict of interest, and have nothing to disclose.

## **Appendix A. Members of The Movember Foundation’s Global Action Plan Prostate Cancer Active Surveillance (GAP3) consortium**

*Principle Investigators:* Bruce Trock (Johns Hopkins University, The James Buchanan Brady Urological Institute, Baltimore, USA), Behfar Ehdaie (Memorial Sloan Kettering Cancer Center, New York, USA), Peter Carroll (University of California San Francisco, San Francisco, USA), Christopher Filson (Emory University School of Medicine, Winship Cancer Institute, Atlanta, USA), Jeri Kim / Christopher Logothetis (MD Anderson Cancer Centre, Houston, USA), Todd Morgan (University of Michigan and Michigan Urological Surgery Improvement Collaborative (MUSIC), Michigan, USA), Laurence Klotz (University of Toronto, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada), Tom Pickles (University of British

397 Columbia, BC Cancer Agency, Vancouver, Canada), Eric Hyndman (Uni-  
 398 versity of Calgary, Southern Alberta Institute of Urology, Calgary, Canada),  
 399 Caroline Moore (University College London & University College London  
 400 Hospital Trust, London, UK), Vincent Gnanapragasam (University of Cam-  
 401 bridge & Cambridge University Hospitals NHS Foundation Trust, Cam-  
 402 bridge, UK), Mieke Van Hemelrijck (King's College London, London, UK  
 403 & Guy's and St Thomas' NHS Foundation Trust, London, UK), Prokar  
 404 Dasgupta (Guy's and St Thomas' NHS Foundation Trust, London, UK),  
 405 Chris Bangma (Erasmus Medical Center, Rotterdam, The Netherlands/ rep-  
 406 resentative of Prostate cancer Research International Active Surveillance  
 407 (PRIAS) consortium), Monique Roobol (Erasmus Medical Center, Rotter-  
 408 dam, The Netherlands/ representative of Prostate cancer Research Interna-  
 409 tional Active Surveillance (PRIAS) consortium), The PRIAS study group,  
 410 Arnauld Villers (Lille University Hospital Center, Lille, France), Antti Ran-  
 411 nikko (Helsinki University and Helsinki University Hospital, Helsinki, Fin-  
 412 land), Riccardo Valdagni (Department of Oncology and Hemato-oncology,  
 413 Università degli Studi di Milano, Radiation Oncology 1 and Prostate Cancer  
 414 Program, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy),  
 415 Antoinette Perry (University College Dublin, Dublin, Ireland), Jonas Hugos-  
 416 son (Sahlgrenska University Hospital, Göteborg, Sweden), Jose Rubio-Briones  
 417 (Instituto Valenciano de Oncología, Valencia, Spain), Anders Bjartell (Skåne  
 418 University Hospital, Malmö, Sweden), Lukas Hefermehl (Kantonsspital Baden,  
 419 Baden, Switzerland), Lee Lui Shiong (Singapore General Hospital, Singa-  
 420 pore, Singapore), Mark Frydenberg (Monash Health; Monash University,  
 421 Melbourne, Australia), Yoshiyuki Kakehi / Mikio Sugimoto (Kagawa Uni-

422 versity Faculty of Medicine, Kagawa, Japan), Byung Ha Chung (Gangnam  
423 Severance Hospital, Yonsei University Health System, Seoul, Republic of Ko-  
424 rea)

425 *Pathologist:* Theo van der Kwast (Princess Margaret Cancer Centre,  
426 Toronto, Canada). Technology Research Partners: Henk Obbink (Royal  
427 Philips, Eindhoven, the Netherlands), Wim van der Linden (Royal Philips,  
428 Eindhoven, the Netherlands), Tim Hulsen (Royal Philips, Eindhoven, the  
429 Netherlands), Cees de Jonge (Royal Philips, Eindhoven, the Netherlands).

430 *Advisory Regional statisticians:* Mike Kattan (Cleveland Clinic, Cleve-  
431 land, Ohio, USA), Ji Xinge (Cleveland Clinic, Cleveland, Ohio, USA), Ken-  
432 neth Muir (University of Manchester, Manchester, UK), Artitaya Lophatananon  
433 (University of Manchester, Manchester, UK), Michael Fahey (Epworth Health-  
434 Care, Melbourne, Australia), Ewout Steyerberg (Erasmus Medical Center,  
435 Rotterdam, The Netherlands), Daan Nieboer (Erasmus Medical Center, Rot-  
436 terdam, The Netherlands); Liying Zhang (University of Toronto, Sunnybrook  
437 Health Sciences Centre, Toronto, Ontario, Canada)

438 *Executive Regional statisticians:* Ewout Steyerberg (Erasmus Medical  
439 Center, Rotterdam, The Netherlands), Daan Nieboer (Erasmus Medical Cen-  
440 ter, Rotterdam, The Netherlands); Kerri Beckmann (King's College London,  
441 London, UK & Guy's and St Thomas' NHS Foundation Trust, London, UK),  
442 Brian Denton (University of Michigan, Michigan, USA), Andrew Hayen (Uni-  
443 versity of Technology Sydney, Australia), Paul Boutros (Ontario Institute of  
444 Cancer Research, Toronto, Ontario, Canada).

445 *Clinical Research Partners' IT Experts:* Wei Guo (Johns Hopkins Uni-  
446 versity, The James Buchanan Brady Urological Institute, Baltimore, USA),

447 Nicole Benfante (Memorial Sloan Kettering Cancer Center, New York, USA),  
 448 Janet Cowan (University of California San Francisco, San Francisco, USA),  
 449 Dattatraya Patil (Emory University School of Medicine, Winship Cancer In-  
 450 stitute, Atlanta, USA), Emily Tolosa (MD Anderson Cancer Centre, Hous-  
 451 ton, Texas, USA), Tae-Kyung Kim (University of Michigan and Michigan  
 452 Urological Surgery Improvement Collaborative, Ann Arbor, Michigan, USA),  
 453 Alexandre Mamedov (University of Toronto, Sunnybrook Health Sciences  
 454 Centre, Toronto, Ontario, Canada), Vincent LaPointe (University of British  
 455 Columbia, BC Cancer Agency, Vancouver, Canada), Trafford Crump (Uni-  
 456 versity of Calgary, Southern Alberta Institute of Urology, Calgary, Canada),  
 457 Vasilis Stavriniades (University College London & University College Lon-  
 458 don Hospital Trust, London, UK), Jenna Kimberly-Duffell (University of  
 459 Cambridge & Cambridge University Hospitals NHS Foundation Trust, Cam-  
 460 bridge, UK), Aida Santaolalla (King's College London, London, UK & Guy's  
 461 and St Thomas' NHS Foundation Trust, London, UK), Daan Nieboer (Eras-  
 462 mus Medical Center, Rotterdam, The Netherlands), Jonathan Olivier (Lille  
 463 University Hospital Center, Lille, France), Tiziana Rancati (Fondazione IR-  
 464 CCS Istituto Nazionale dei Tumori di Milano, Milan, Italy), Helén Ahlgren  
 465 (Sahlgrenska University Hospital, Göteborg, Sweden), Juanma Mascarós (In-  
 466 stituto Valenciano de Oncología, Valencia, Spain), Annica Löfgren (Skåne  
 467 University Hospital, Malmö, Sweden), Kurt Lehmann (Kantonsspital Baden,  
 468 Baden, Switzerland), Catherine Han Lin (Monash University and Epworth  
 469 HealthCare, Melbourne, Australia), Hiromi Hiram (Kagawa University, Ka-  
 470 gawa, Japan), Kwang Suk Lee (Yonsei University College of Medicine, Gang-  
 471 nam Severance Hospital, Seoul, Korea).

472 *Research Advisory Committee:* Guido Jenster (Erasmus MC, Rotterdam,  
 473 the Netherlands), Anssi Auvinen (University of Tampere, Tampere, Fin-  
 474 land), Anders Bjartell (Skåne University Hospital, Malmö, Sweden), Ma-  
 475 soom Haider (University of Toronto, Toronto, Canada), Kees van Bochove  
 476 (The Hyve B.V. Utrecht, Utrecht, the Netherlands), Ballentine Carter (Johns  
 477 Hopkins University, Baltimore, USA – until 2018).

478 *Management team:* Sam Gledhill (Movember Foundation, Melbourne,  
 479 Australia), Mark Buzza / Michelle Koussou (Movember Foundation, Mel-  
 480 bourne, Australia), Chris Bangma (Erasmus Medical Center, Rotterdam,  
 481 The Netherlands), Monique Roobol (Erasmus Medical Center, Rotterdam,  
 482 The Netherlands), Sophie Bruinsma / Jozien Helleman (Erasmus Medical  
 483 Center, Rotterdam, The Netherlands).

## 484 **References**

- 485 [1] Briganti A, Fossati N, Catto JW, Cornford P, Montorsi F, Mottet N,  
 486 et al. Active surveillance for low-risk prostate cancer: the European As-  
 487 sociation of Urology position in 2018. *European urology* 2018;74(3):357–  
 488 68.
- 489 [2] Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA.  
 490 The 2014 international society of urological pathology (ISUP) consensus  
 491 conference on Gleason grading of prostatic carcinoma. *The American*  
 492 *journal of surgical pathology* 2016;40(2):244–52.
- 493 [3] Bruinsma SM, Roobol MJ, Carroll PR, Klotz L, Pickles T, Moore CM,  
 494 et al. Expert consensus document: semantics in active surveillance for

- men with localized prostate cancer—results of a modified Delphi consensus procedure. *Nature reviews urology* 2017;14(5):312.
- [4] Bul M, Zhu X, Valdagni R, Pickles T, Kakehi Y, Rannikko A, et al. Active surveillance for low-risk prostate cancer worldwide: the PRIAS study. *European urology* 2013;63(4):597–603.
- [6] Loeb S, Carter HB, Schwartz M, Fagerlin A, Braithwaite RS, Lepor H. Heterogeneity in active surveillance protocols worldwide. *Reviews in urology* 2014;16(4):202–3.
- [7] Loeb S, Vellekoop A, Ahmed HU, Catto J, Emberton M, Nam R, et al. Systematic review of complications of prostate biopsy. *European urology* 2013;64(6):876–92.
- [8] Bokhorst LP, Alberts AR, Rannikko A, Valdagni R, Pickles T, Kakehi Y, et al. Compliance rates with the Prostate Cancer Research International Active Surveillance (PRIAS) protocol and disease reclassification in noncompliers. *European Urology* 2015;68(5):814–21.
- [9] Inoue LY, Lin DW, Newcomb LF, Leonardson AS, Ankerst D, Gulati R, et al. Comparative analysis of biopsy upgrading in four prostate cancer active surveillance cohorts. *Annals of internal medicine* 2018;168(1):1–9.
- [10] Bratt O, Carlsson S, Holmberg E, Holmberg L, Johansson E, Josefsson A, 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.

- 518 [5] Nieboer D, Tomer A, Rizopoulos D, Roobol MJ, Steyerberg EW. Active  
519 surveillance: a review of risk-based, dynamic monitoring. *Translational*  
520 *andrology and urology* 2018;7(1):106–15.
- 521 [] Kasivisvanathan V, Giganti F, Emberton M, Moore CM. Magnetic reso-  
522 nance imaging should be used in the active surveillance of patients with  
523 localised prostate cancer. *European urology* 2020;77(3):318–9.
- 524 [] Chesnut GT, Vertosick EA, Benfante N, Sjoberg DD, Fainberg J, Lee T,  
525 et al. Role of changes in magnetic resonance imaging or clinical stage in  
526 evaluation of disease progression for men with prostate cancer on active  
527 surveillance. *European Urology* 2020;77(4):501–7.
- 528 [11] de Carvalho TM, Heijnsdijk EA, de Koning HJ. Estimating the risks  
529 and benefits of active surveillance protocols for prostate cancer: a mi-  
530 crosimulation study. *BJU international* 2017;119(4):560–6.
- 531 [15] Coley RY, Zeger SL, Mamawala M, Pienta KJ, Carter HB. Prediction  
532 of the pathologic Gleason score to inform a personalized management  
533 program for prostate cancer. *European urology* 2017;72(1):135–41.
- 534 [25] Ankerst DP, Xia J, Thompson Jr IM, Hoeffler J, Newcomb LF, Brooks  
535 JD, et al. Precision medicine in active surveillance for prostate can-  
536 cer: development of the canary–early detection research network active  
537 surveillance biopsy risk calculator. *European urology* 2015;68(6):1083–8.
- 538 [12] Partin AW, Yoo J, Carter HB, Pearson JD, Chan DW, Epstein JI, et al.  
539 The use of prostate specific antigen, clinical stage and Gleason score to

- 540       predict pathological stage in men with localized prostate cancer. The  
541       Journal of urology 1993;150(1):110–4.
- 542 [13] Makarov DV, Trock BJ, Humphreys EB, Mangold LA, Walsh PC, Ep-  
543       stein JJ, et al. Updated nomogram to predict pathologic stage of prostate  
544       cancer given prostate-specific antigen level, clinical stage, and biopsy  
545       Gleason score (Partin tables) based on cases from 2000 to 2005. Urol-  
546       ogy 2007;69(6):1095–101.
- 547 [17] Bruinsma SM, Zhang L, Roobol MJ, Bangma CH, Steyerberg EW,  
548       Nieboer D, et al. The Movember foundation’s GAP3 cohort: a pro-  
549       file of the largest global prostate cancer active surveillance database to  
550       date. BJU international 2018;121(5):737–44.
- 551       □ Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis  
552       M, et al. Eau-estro-siog guidelines on prostate cancer. part 1: screening,  
553       diagnosis, and local treatment with curative intent. European Urology  
554       2017;71(4):618–29.
- 555       □ Schoots IG, Petrides N, Giganti F, Bokhorst LP, Rannikko A, Klotz  
556       L, et al. Magnetic resonance imaging in active surveillance of prostate  
557       cancer: a systematic review. European urology 2015;67(4):627–36.
- 558 [14] Tomer A, Nieboer D, Roobol MJ, Steyerberg EW, Rizopoulos D. Per-  
559       sonalized schedules for surveillance of low-risk prostate cancer patients.  
560       Biometrics 2019;75(1):153–62.
- 561 [16] Rizopoulos D. Joint Models for Longitudinal and Time-to-Event Data:  
562       With Applications in R. CRC Press; 2012. ISBN 9781439872864.



- [18] Laird NM, Ware JH, et al. Random-effects models for longitudinal data. *Biometrics* 1982;38(4):963–74.
- [ ] Nixon RG, Wener MH, Smith KM, Parson RE, Strobel SA, Brawer MK. Biological variation of prostate specific antigen levels in serum: an evaluation of day-to-day physiological fluctuations in a well-defined cohort of 24 patients. *The Journal of urology* 1997;157(6):2183–90.
- [ ] Cooperberg MR, Brooks JD, Faino AV, Newcomb LF, Kearns JT, Carroll PR, et al. Refined analysis of prostate-specific antigen kinetics to predict prostate cancer active surveillance outcomes. *European urology* 2018;74(2):211–7.
- [20] 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.
- [22] Royston P, Altman DG. External validation of a cox prognostic model: principles and methods. *BMC medical research methodology* 2013;13(1):33.
- [23] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology (Cambridge, Mass)* 2010;21(1):128.
- [21] 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.

- 586 [24] Turnbull BW. The empirical distribution function with arbitrarily  
587 grouped, censored and truncated data. *Journal of the Royal Statistical Society Series B (Methodological)* 1976;38(3):290–5.  
588
- 589 [] de Carvalho TM, Heijnsdijk EA, de Koning HJ. When should active  
590 surveillance for prostate cancer stop if no progression is detected? *The Prostate* 2017;77(9):962–9.  
591
- 592 [] Tomer A, Rizopoulos D, Nieboer D, Drost FJ, Roobol MJ, Steyerberg  
593 EW. Personalized decision making for biopsies in prostate cancer active  
594 surveillance programs. *Medical Decision Making* 2019;39(5):499–508.
- 595 [26] Balasubramanian R, Lagakos SW. Estimation of a failure time distribu-  
596 tion based on imperfect diagnostic tests. *Biometrika* 2003;90(1):171–82.