A Ready To Use Web-Application Providing a

Personalized Biopsy Schedule for Men With Low-Risk PCa Under Active Surveillance*?*

Anirudh Tomer, MSca,∗, Daan Nieboer, MScb, Monique J. Roobol, PhDc, Anders Bjartell, MD, PhDd, Ewout W. Steyerberg, PhDb,e, Dimitris

Rizopoulos, PhDa, Movember Foundations Global Action Plan Prostate Cancer Active Surveillance (GAP3) consortiumf

*aDepartment of Biostatistics, Erasmus University Medical Center, Rotterdam, the*

*Netherlands*

*bDepartment of Public Health, Erasmus University Medical Center, Rotterdam, the*

*Netherlands*

*cDepartment of Urology, Erasmus University Medical Center, Rotterdam, the Netherlands dDepartment of Urology, Sk˚ane University Hospital, Malm¨o, Sweden eDepartment of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the*

*Netherlands*

*fThe Movember Foundations Global Action Plan Prostate Cancer Active Surveillance (GAP3) consortium members presented in Appendix A*

**Abstract**

Background: Prostate cancer active surveillance (AS) patients undergo repeat biopsies. Active treatment is advised when biopsy Gleason grade group ≥ 2 (*upgrading*). Many patients never experience upgrading, yet undergo biopsies frequently. Personalized biopsy decisions based on upgrading-risk may

*?*Word count abstract (headings excluded): 300; Word count text: 2484

∗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, MD, PhD),

e.w.steyerberg@lumc.nl (Ewout W. Steyerberg, PhD), d.rizopoulos@erasmusmc.nl

(Dimitris Rizopoulos, PhD)

*Preprint submitted to European Urology December 13, 2019*

reduce patient burden.

Objective: Develop a risk prediction model and web-application to assist patients/doctors in personalized biopsy decisions.

Design, Setting, and Participants: Model development: world’s largest AS study PRIAS, 7813 patients, 1134 experienced upgrading; External validation: largest five cohorts of Movember Foundation’s GAP3 database (*>* 20*,*000 patients, 27 centers worldwide); Data: repeat prostate-specific antigen (PSA) and biopsy Gleason grade.

Outcome Measurements, and Statistical Analysis: A Bayesian joint model fitted to the PRIAS dataset. This model was validated in GAP3 cohorts using risk prediction error, calibration, area under ROC (AUC). Model and personalized biopsy schedules based on predicted risks were implemented in a web-application.

Results and Limitations: Cause-specific cumulative upgrading-risk at year five of follow-up: 35% in PRIAS, at most 50% in GAP3 cohorts. 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). Validation: Moderate AUC (0.55–0.75) in PRIAS and GAP3 cohorts. Moderate prediction error (0.1–0.3) in GAP3 cohorts where impact of PSA value and velocity on upgrading-risk was similar to PRIAS, but large (0.3–0.45) otherwise. Recalibration advised for external cohorts.

Conclusions: We successfully developed and validated a model for predicting upgrading-risk, and providing risk-based personalized biopsy decisions, in prostate cancer AS. The model made available via a web-application enables shared decision making of biopsy schedules by comparing fixed and personalized schedules on total biopsies and expected time delay in detecting upgrading.

Patient Summary: Personalized prostate biopsies are a novel alternative to fixed one-size-fits-all schedules. The underlying statistical models are made available through a user-friendly web-application and may help to reduce unnecessary prostate biopsies while maintaining cancer control.

*Keywords:* Active Surveillance, Biopsies, Personalized Medicine, Prostate

Cancer, Shared Decision Making

1. **1. Introduction**
2. Patients with low- and very low-risk screening-detected localized prostate
3. cancer are usually recommended active surveillance (AS) instead of immedi-
4. ate radical treatment [1]. In AS, cancer progression is routinely monitored
5. via prostate-specific antigen (PSA), digital rectal examination, and repeat
6. biopsies. Among these, the strongest indicator of cancer-related outcomes
7. is the biopsy Gleason grade group [2]. When the Gleason grade group in-
8. creases from group 1 (Gleason 3+3) to 2 (Gleason 3+4) or higher, called 9 *upgrading* [3], patients are commonly advised curative treatment [4].
9. In most AS protocols, biopsies are scheduled periodically. Consequently,

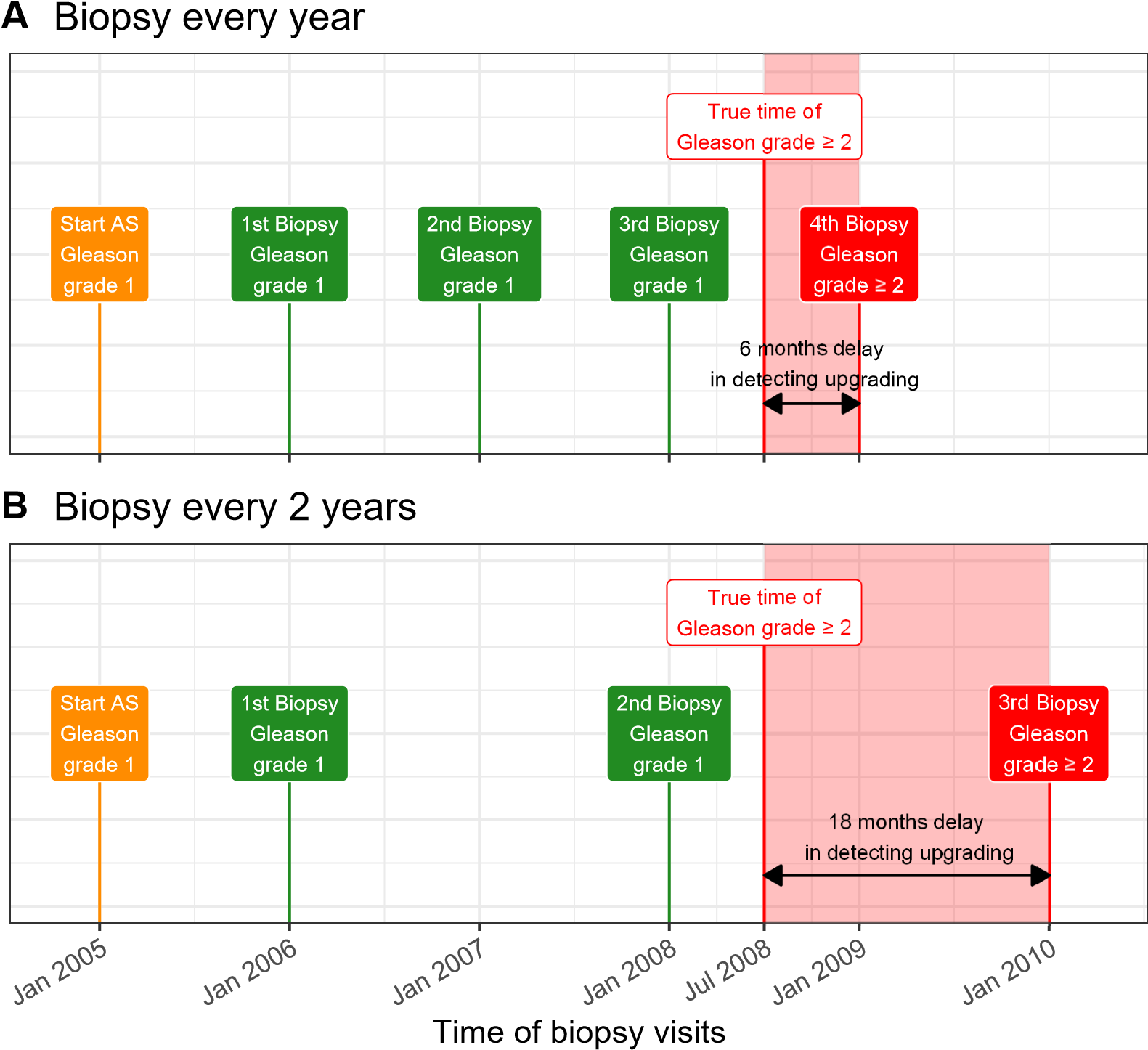


Figure 1: **Trade-off between the number of biopsies and time delay in detecting upgrading (Increase in Gleason grade group from 1 to 2 or higher):** The true

time of upgrading 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.

1. upgrading is always detected with a time delay (Figure 1). For detecting
2. upgrading timely, many AS programs schedule fixed and frequent biopsies
3. (e.g., annually) for all patients [5, 6]. However, this leads to many unnec-
4. essary biopsies in slow/non-progressing patients. Biopsies are invasive, may
5. be painful, and are prone to medical complications such as bleeding and
6. septicemia[7]. Thus, biopsy burden and patient non-compliance to frequent
7. biopsies [8] has raised concerns regarding the optimal biopsy schedule [9, 10].
8. To this end, some cohorts have started using magnetic resonance imaging
9. (MRI) for deciding biopsies. Although, due to currently limited AS data,
10. MRI’s value is not clear. Others have proposed infrequent schedules such
11. as biennial biopsies as an alternative [9, 11]. However, fundamental differ-
12. ences exist in underlying upgrading-risk across cohorts [9]. Thus, biennial
13. biopsies may still lead to five unnecessary biopsies over ten years (current
14. study period of large AS programs) for many slow/non-progressing patients.
15. A promising alternative to fixed and frequent biopsies is personalized biopsy 26 schedules based on the patient-specific upgrading-risk (Figure 2).
16. The first challenge in developing personalized biopsy schedules is consol-
17. idating accumulated patient data (e.g., PSA, previous biopsy results) into
18. estimates for upgrading-risk. Existing calculators for upgrading-risk [12, 13]
19. use only the latest PSA measurement of a patient. In contrast, we intend to
20. utilize all repeated measurements of PSA, previous biopsy results, and base-
21. line characteristics of a patient. To this end, a suitable model is the joint
22. model for time-to-event and longitudinal data [14, 15, 16]. A joint model
23. predicts the upgrading-risk in a personalized manner. A subsequent chal-
24. lenge, however, is translating risks into clinical decisions. For example, a

**A** Should a biopsy be conducted at current visit?

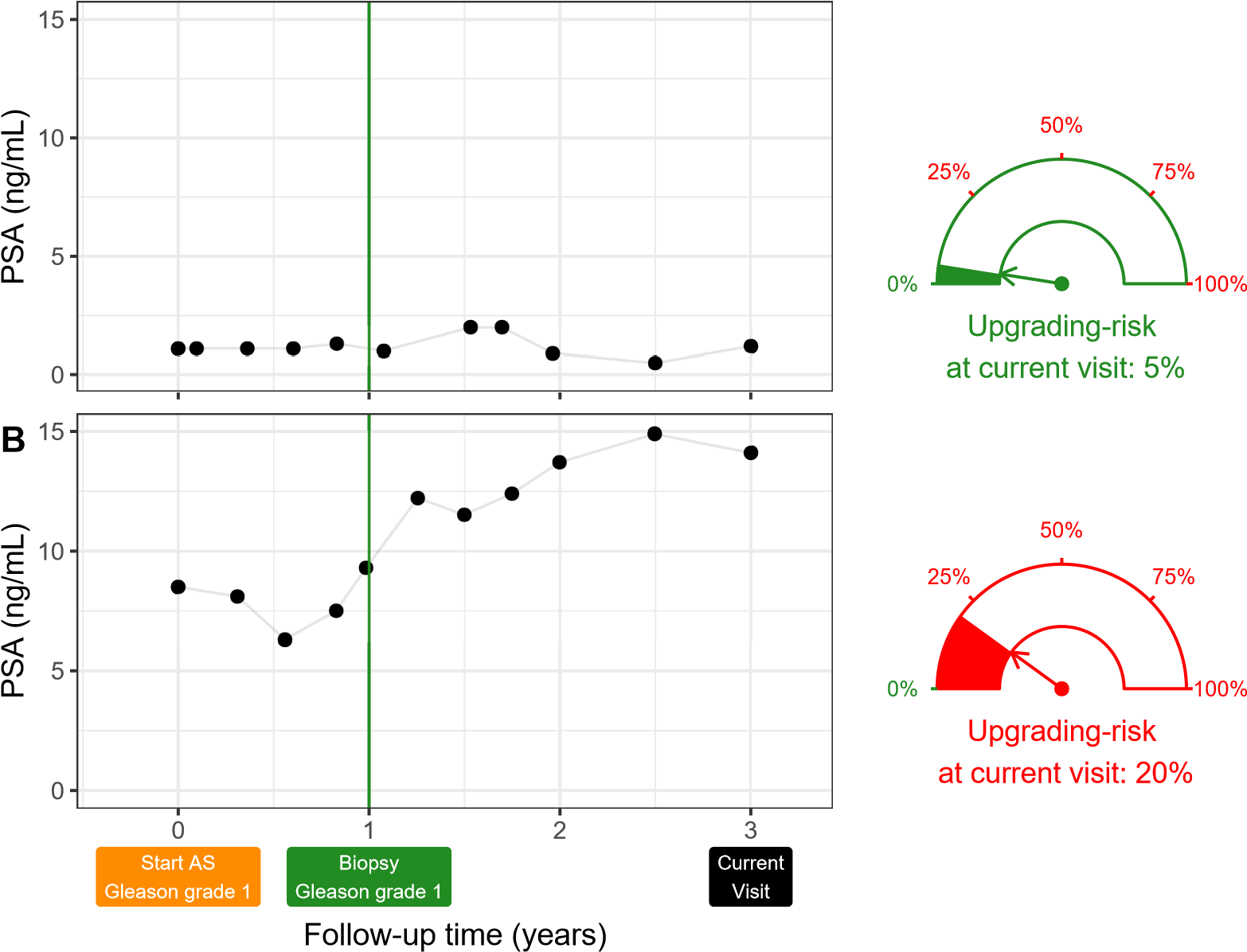


Figure 2: **Motivation for personalized upgrading-risk based decisions of biopsy**: Patient A (**Panel A**) and B (**Panel B**) had their latest biopsy at year one of follow-up (green vertical line). Patient A’s prostate-specific antigen (PSA) profile remained stable until his current visit at year three, whereas patient B’s profile has shown a rise. Consequently, patient B’s upgrading-risk 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.

1. 10% upgrading-risk can be perceived high/low depending upon the patient’s
2. age. Patients may also weigh risks of upgrading with the potential *conse-*
3. *quences* of another biopsy. Two relevant *consequences* of biopsies (Figure 1)
4. are the timing and the total number of biopsies (burden), and the time de-
5. lay in detecting upgrading (smaller is beneficial). The relative importance of
6. these *consequences* can vary between the patients, and also over the follow-up 42 period for the same patient.
7. The goal of this work is to develop a robust, generalizable model that
8. gives reliable **estimates for individual upgrading-risk, and to create personal-**
9. **ized biopsy schedules based on this risk**. To facilitate shared decision making
10. of biopsy schedules, we also aim to provide quantitative estimates of *conse-*
11. *quences* of opting for a personalized versus the standard fixed schedule. For
12. developing our model, we will use the world’s largest AS dataset PRIAS.
13. Subsequently, we want to externally validate our model in the largest five
14. AS cohorts from the Movember Foundation’s GAP3 database [17]. Last, we 51 intend to implement our model and methodology in a web-application.

52 **2. Patients and Methods**

# 53 2.1. Study Cohort

1. For developing a statistical model to predict upgrading-risk, we used the
2. world’s largest AS dataset, Prostate Cancer International Active Surveillance
3. or PRIAS [4] (Table 1). In PRIAS, PSA was measured quarterly for the first
4. two years of follow-up and semiannually thereafter. Biopsies were scheduled
5. at year one, four, seven, and ten of follow-up. Additional yearly biopsies 59 were scheduled when PSA doubling time was between zero and ten years.

Table 1: Summary of the PRIAS dataset. 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](https://www.prias-project.org/)

|  |  |
| --- | --- |
| Characteristic | Value |
| Total centers | *>*100 |
| 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 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) |

1. We selected all 7813 patients who had Gleason grade group 1 at the time
2. of inclusion in PRIAS. Our primary event of interest is an increase in this
3. Gleason grade group upon repeat biopsy, called *upgrading* (1134 patients).
4. Upgrading is a trigger for treatment advice in PRIAS. Although, 2250 pa-
5. tients were provided treatment based on their PSA, or number of biopsy cores
6. with cancer, or anxiety/other reasons. Our reasons for focusing solely on up66 grading are, namely, upgrading is strongly associated with cancer-related 67 outcomes, and other triggers for treatment vary between cohorts [5].
7. For model validation, we selected the largest five cohorts from Movember
8. Foundation’s GAP3 database [17]. These were, namely, the University of
9. Toronto AS (Toronto), Johns Hopkins AS (Hopkins), Memorial Sloan Ket-
10. tering Cancer Center AS (MSKCC), King’s College London AS (KCL), and
11. Michigan Urological Surgery Improvement Collaborative AS (MUSIC). Only 73 patients with a Gleason grade group 1 at the time of inclusion in these cohorts 74 were selected (Supplementary A.2).

# 75 2.2. Statistical Model

1. For developing an upgrading-risk prediction model, the available data in
2. the PRIAS cohort was patient age at inclusion in AS, longitudinally measured
3. PSA, timing of repeat biopsies and corresponding Gleason grades, and ob-
4. served time of upgrading. Analysis of this data required modeling the within-
5. patient correlation for PSA, the association between the Gleason grades and
6. PSA profiles of a patient, and handling missing PSA measurements after a
7. patient experienced upgrading. In such situations, a commonly used model 83 is the joint model for time-to-event and longitudinal data [14, 15, 16].

84 Our joint model consisted of two sub-models. First, a linear mixed sub85 model [18] for longitudinally measured PSA (log-transformed). Second, a

1. relative-risk sub-model (similar to the Cox model) for obtaining the cause-
2. specific upgrading-risk. Patient age was included as a predictor in both sub-
3. models. In the PSA sub-model, we fitted a curve to the PSA measurements
4. (Panel A, Figure 3). From each patient’s fitted PSA profile, we extracted his
5. time-varying PSA velocity (Panel B, Figure 3). This instantaneous velocity
6. is more precise than the widely employed constant PSA velocity [19]. We
7. modeled the impact of PSA on upgrading-risk by including fitted PSA value
8. and velocity as predictors in the relative-risk model. Also, the time of the
9. latest negative biopsy was utilized in the relative-risk sub-model (Panel C,
10. Figure 3). The parameters of the two sub-models were estimated jointly 96 (Supplementary A) using the R package **JMbayes** [20].

# 97 2.3. Model Validation

98 We validated our PRIAS based risk prediction model internally in the 99 PRIAS cohort, and externally using the largest five GAP3 database cohorts

1. (Section 2.1 and Supplementary A.2). We assessed our model’s ability to
2. discriminate between patients who experience/do not experience upgrading
3. via the area under the receiver operating characteristic curve or AUC [21].
4. We employed calibration plots [22, 23] and mean absolute risk prediction
5. error [21] to graphically and quantitatively evaluate our model’s risk predic-
6. tion accuracy. Since AS studies are longitudinal, both AUC and prediction
7. error vary over follow-up (Supplementary B.1). Lastly, to resolve any po-
8. tential model miscalibration in validation cohorts, we aimed to recalibrate
9. our model’s baseline hazard of upgrading (Supplementary A), individually
10. for each cohort.

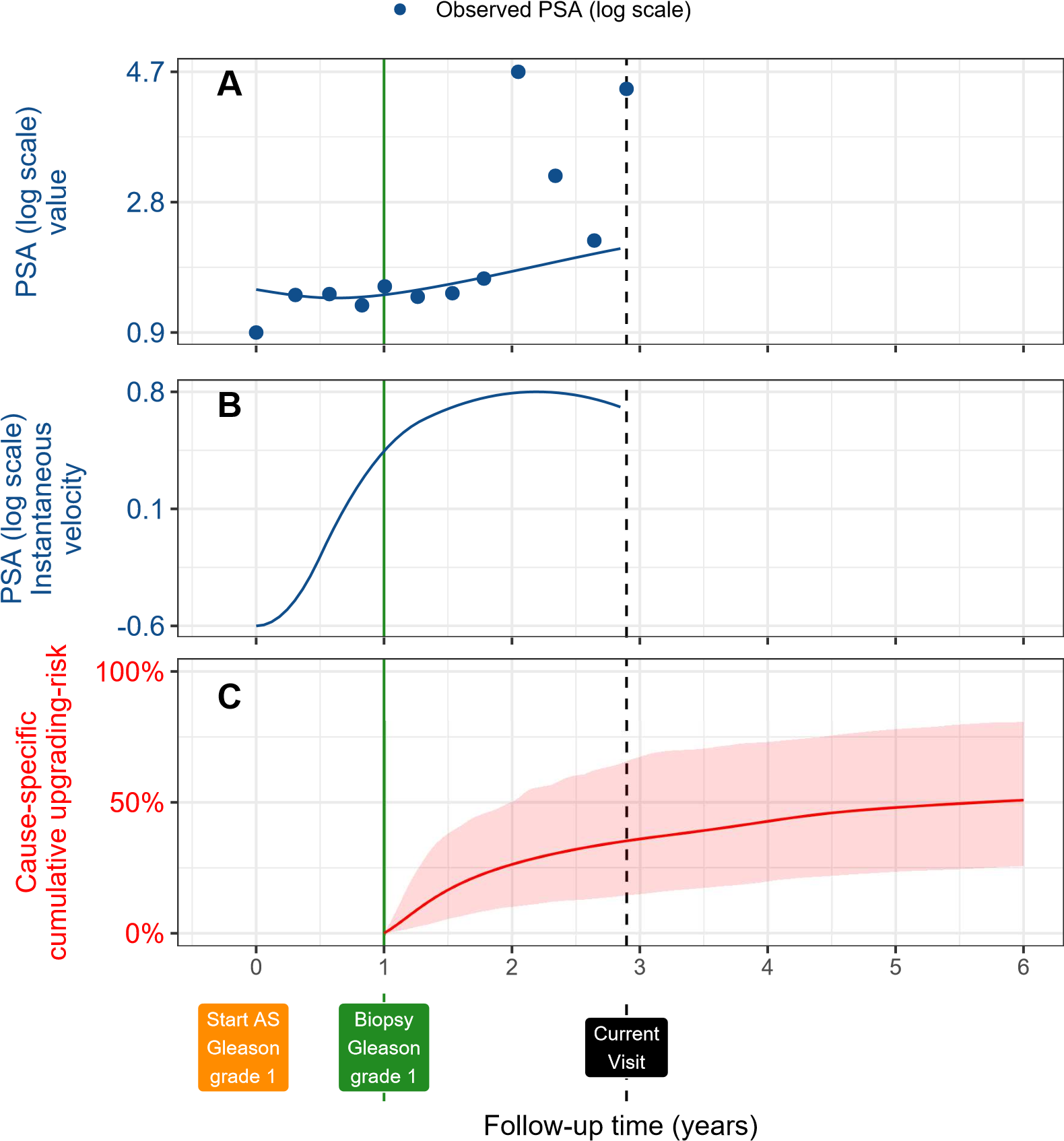


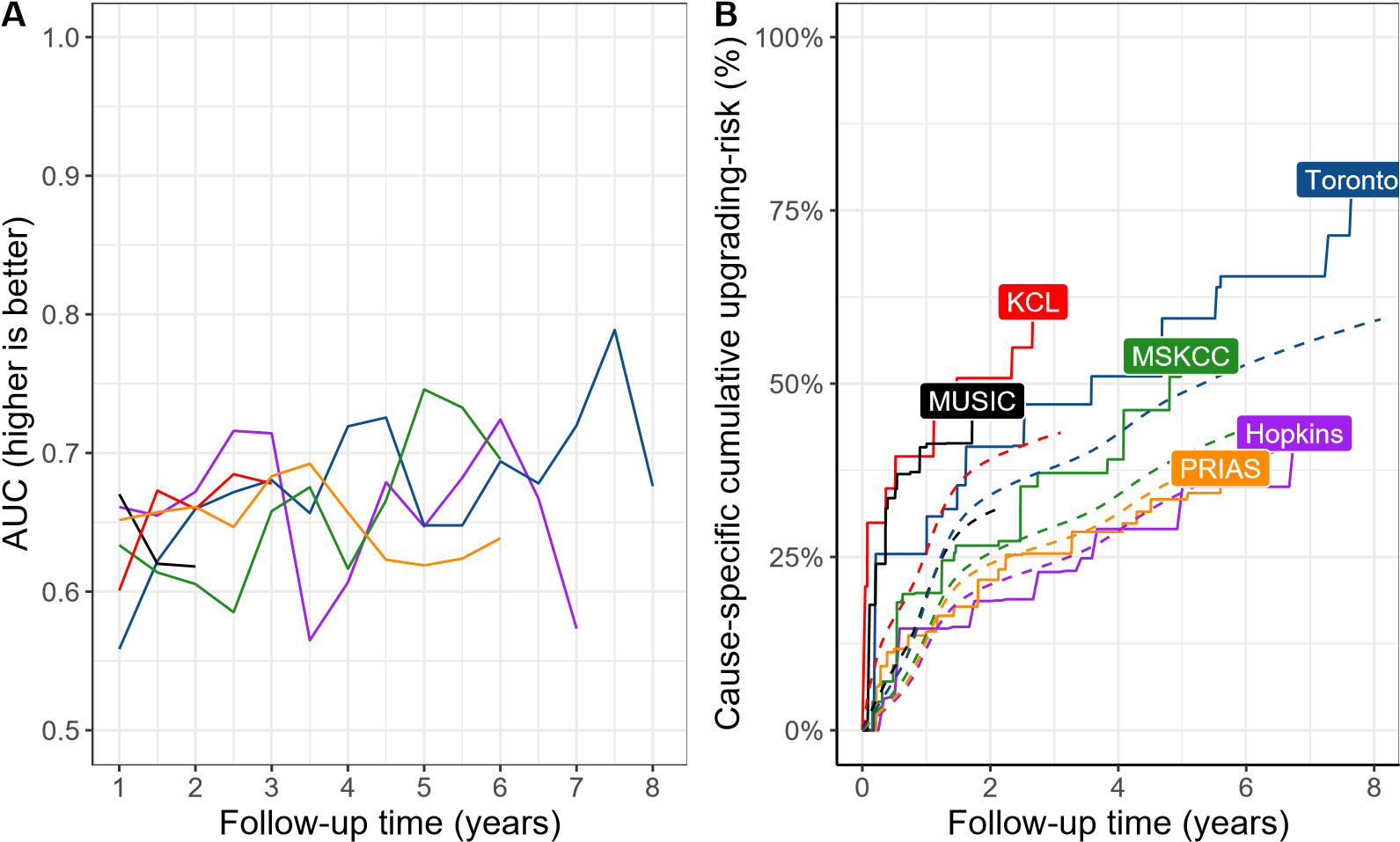
Figure 3: **Illustration of the joint model on a real PRIAS patient**. **Panel A:**

Observed PSA (blue dots) and fitted PSA (solid blue line), log-transformed. **Panel B:** Estimated instantaneous velocity of PSA (log-transformed). **Panel C**: Predicted causespecific 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 1 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 3 denotes the time of current visit.

1. **3. Results**
2. The cause-specific cumulative upgrading-risk at year five of follow-up was
3. 35% in PRIAS, and at most 50% in the five validation cohorts (Panel B,
4. Figure 4). That is, many patients may not require any biopsy in the first 114 five years of AS.
5. In the joint model fitted to the PRIAS dataset, the adjusted hazard ratio
6. of upgrading for an increase in patient **age** from 61 to 71 years (25-th to
7. 75-th percentile) was **1.45** (95%CI: 1.30–1.63). For an increase in fitted **PSA**
8. value (log scale) from 2.36 to 3.07 (25-th to 75-th percentile), the adjusted
9. hazard ratio was **0.99** (95%CI: 0.89–1.11). In contrast to PSA value, **instan-**
10. **taneous PSA velocity** was a stronger predictor of upgrading-risk, because
11. an increase in velocity from -0.09 to 0.31 (25-th to 75-th percentile) had a
12. hazard ratio of **2.47** (95%CI: 1.93–2.99). The impact of PSA value and veloc-
13. ity on upgrading-risk varied between cohorts (Table 6, Supplementary A.2). 124 Detailed results are in Supplementary A.2.
14. The time-varying mean absolute risk prediction error; time-varying AUC;
15. and calibration plot of our model in different validation cohorts are shown in
16. Panel B, Figure 8, Supplementary B; Panel A, Figure 4; and Panel B, Fig-
17. ure 4, respectively. In all cohorts, AUC was moderate (0.55 to 0.75). Mean
18. absolute prediction error was large (0.3 to 0.45) in those cohorts where the
19. impact of PSA value and velocity on upgrading-risk was different from PRIAS
20. (e.g., MUSIC cohort, Table 6, Supplementary A.2), and moderate (0.1 to 0.3)
21. otherwise. To resolve issues in calibration-at-large (Panel B, Figure 4), we
22. recalibrated the baseline hazard of upgrading in all cohorts (Figure 6, Sup-
23. plementary B). We compared risk predictions from the recalibrated models 135 with predictions from separately fitted joint models to each cohort (Figure 7,
24. Supplementary B). The difference in predictions was lowest in Johns Hop-
25. kins cohort (impact of PSA similar to PRIAS, Table 5, Supplementary A.2).
26. Comprehensive validation results are in Supplementary B.

# 139 3.1. Personalized Biopsy Schedules

1. We utilized the fitted joint model to create upgrading-risk based person-
2. alized biopsy schedules. To this end, given a new patient’s accumulated PSA
3. measurements (Panel A, Figure 3) and biopsy results, we first predicted his
4. cause-specific cumulative upgrading-risk at his current as well as future PSA
5. follow-up visits (Panel A, Figure 5). These PSA visits occur every six months
6. in PRIAS. Subsequently, we scheduled personalized biopsies on those future
7. follow-up visits of a patient, where his conditional cumulative upgrading-risk
8. was more than a certain threshold (Supplementary C), for example, 10%
9. risk. We maintained a minimum gap of one year between consecutive biop-
10. sies (PRIAS recommendation). Example personalized schedules based on 5%
11. and 10% risk thresholds are shown in Panel B, Figure 5, and in Figure 9–
12. 11, Supplementary C. Both the risk predictions and resulting personalized
13. schedules were dynamic because they were updated as more follow-up data 153 became available over follow-up (Figure 5, Supplementary B).
14. The choice of the risk threshold in the personalized schedule dictates
15. the timing and the total number of biopsies, and the expected time delay
16. (Figure 1) in detecting upgrading. We estimated the time delay for both
17. personalized and fixed schedules (Panel C in Figure 5 and Figure 9–11, Sup-
18. plementary C). Since we estimated the time delay in a personalized manner as
19. well, patients/doctors can compare personalized schedules based on different



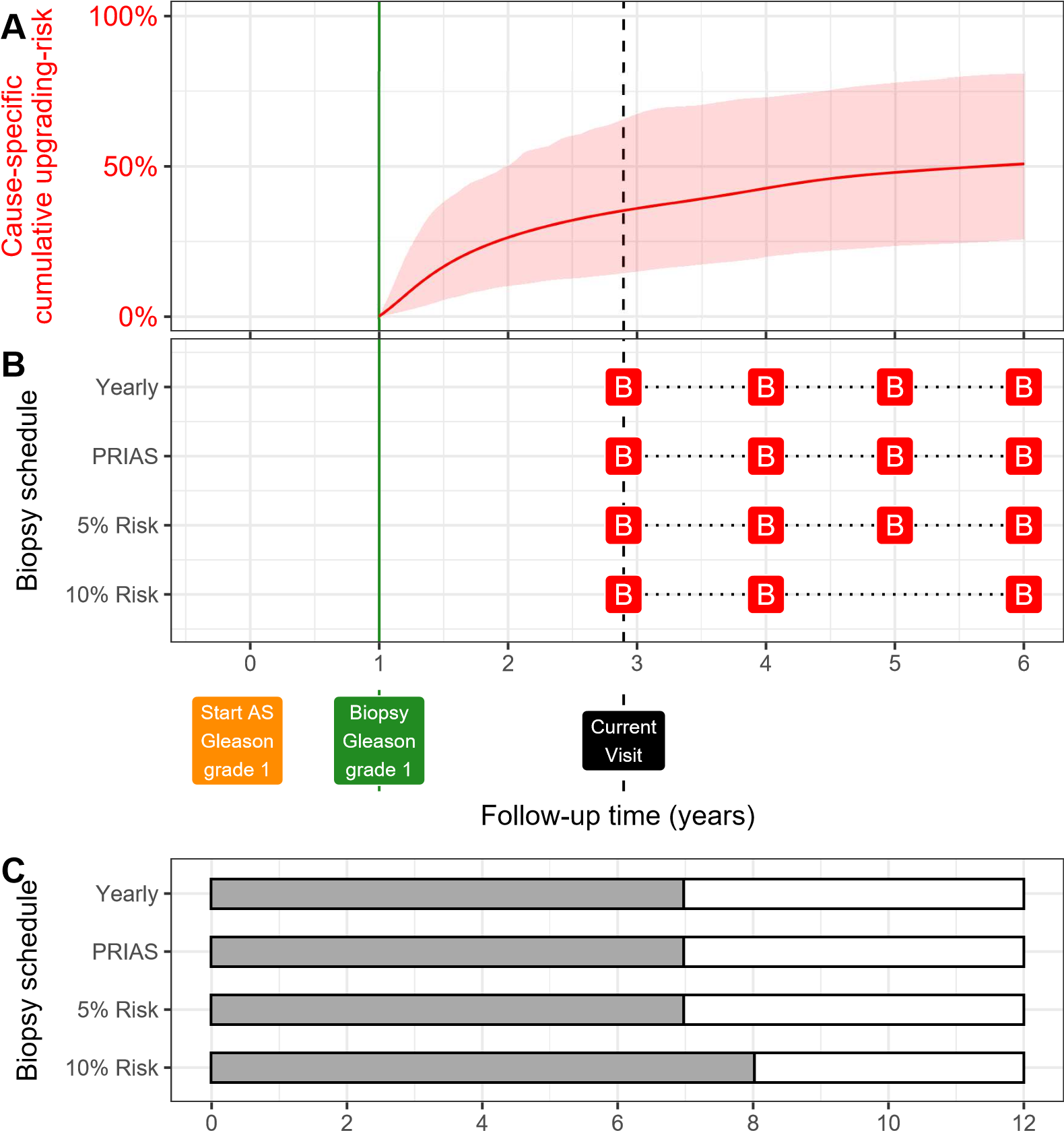
HopkinsMSKCCKCL

TorontoMUSICPRIAS

Figure 4: **Model Validation Results**. **Panel A**: time dependent area under the receiver operating characteristic curve or AUC (measure of discrimination). **Panel B**: calibration-at-large indicates model miscalibration. This is because solid lines depicting the non-parameteric 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. Same plot after recalibration is shown in Figure 6, Supplementary B. 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, *MU-*

*SIC*: Michigan Urological Surgery Improvement Collaborative Active Surveillance.



Expected time delay (months) in detecting upgrading

Figure 5: **Illustration of personalized and fixed schedules of biopsies**. Due to a lack of space, the PSA profile of this patient is shown in Figure 3. **Panel A:** Predicted cumulative upgrading-risk (95% credible interval shaded). **Panel B:** Personalized and fixed schedules of biopsies, with a red ‘B’ indicating a scheduled biopsy. The green vertical line at year 1 denotes the time of the latest negative biopsy. Black dashed line at year 3 denotes the time of the current visit. **Panel C:** Expected time delay in detecting upgrading (months) for different schedules. A compulsory biopsy was scheduled at year six (maximum biopsy scheduling time in PRIAS, Supplementary C) in all schedules for a meaningful comparison between them.

160 risk thresholds, with fixed schedules, before making a choice.

# 161 3.2. Web-Application

1. We implemented our model and personalized schedules in a user-friendly
2. web-application [https://emcbiostatistics.shinyapps.io/prias\_biopsy](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/)\_
3. [recommender/.](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/) Currently, the web-application supports PRIAS and the five
4. validation cohorts. Patient data can be entered manually and in Microsoft
5. Excel format. Predictions for upgrading-risk are available for a currently
6. limited, cohort-specific, follow-up period (Table 7, Supplementary C). The
7. web-application visualizes the timing of biopsies, and expected time delay in
8. detecting upgrading, for personalized schedules based on 5%, 10%, and 15% 170 risk threshold; annual biopsies; biennial biopsies; and PRIAS schedule.
9. **4. Discussion**
10. We successfully developed and externally validated a model for predicting
11. upgrading-risk [3], and providing risk-based personalized biopsy decisions,
12. in prostate cancer AS. Our work has four novel features over earlier risk
13. calculators [15, 25]. First, our model was fitted to the world’s largest AS
14. dataset PRIAS and externally validated in the largest five cohorts of the
15. Movember Foundation’s GAP3 database [17]. Second, the model predicts
16. a patient’s current and future upgrading-risk in a dynamic and personal-
17. ized manner. Third, we use the risks to make a personalized schedule, and
18. also calculate expected time delay in detecting upgrading (less is benefi-
19. cial) if that schedule is followed. Thus, patients/doctors can compare sched-
20. ules before making a choice. Fourth, we implemented our methodology in a 183 user-friendly web-application ([https://emcbiostatistics.shinyapps.io/](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/) 184 [prias\_biopsy\_recommender/)](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/) for PRIAS and validated cohorts.
21. Our PRIAS based model is useful for a large number of patients from
22. the PRIAS and the following validation cohorts: Johns Hopkins AS (Hop-
23. kins), Memorial Sloan Kettering Cancer Center AS, King’s College London
24. AS (KCL), and Michigan Urological Surgery Improvement Collaborative AS
25. (MUSIC). The model had a moderate AUC (0.55–0.75), a measure of dis-
26. crimination, in all validation cohorts. In contrast, the mean absolute risk pre-
27. diction error varied much more between cohorts. It was moderate in cohorts
28. where the effect size for impact of PSA value and velocity on upgrading-risk
29. was similar to that for PRIAS (e.g., Hopkins cohort). Otherwise, as in the
30. case of KCL or MUSIC cohorts, the prediction error was large. Also, in co-
31. horts with longer follow-up periods, prediction error improved over time as
32. more follow-up data became available. Both KCL and MUSIC cohorts cur-
33. rently have a small follow-up period. Hence, we expect that prediction error
34. will improve in the future with more data. Last, we required recalibration
35. of our model’s baseline hazard of upgrading, individually for all validation
36. cohorts.
37. The clinical implications of our work are as follows. First, the cause-
38. specific cumulative upgrading-risk at year five of follow-up was at most 50%
39. in all cohorts (Panel B, Figure 4). That is, many patients may not require any
40. biopsy in the first five years of AS. Given the non-compliance and burden of
41. frequent biopsies [8], the availability of our methodology as a web-application
42. may encourage patients/doctors to consider upgrading-risk based personal-
43. ized schedules instead. An additional advantage of these schedules is that 208 they update as more patient data becomes available over follow-up. Fur-
44. thermore, to assist patients/doctors in choosing between personalized and
45. fixed schedules, the web-application provides a patient-specific estimate of
46. time delay in detecting upgrading, for following both personalized and fixed
47. schedules. We hope that this will objectively address patient apprehensions 213 regarding adverse outcomes in AS.
48. This work has certain limitations. Predictions for upgrading-risk and per-
49. sonalized schedules are available only for a currently limited, cohort-specific,
50. follow-up period (Table 7, Supplementary C). This problem can be mitigated
51. by refitting the model with new follow-up data in the future. It is important
52. to differentiate the instantaneous PSA velocity (predictor for upgrading-risk
53. in our model), from the currently used constant PSA velocity. Unlike the
54. drawbacks suffered by the constant PSA velocity [19], instantaneous PSA ve-
55. locity is more precise. This is because it changes over time and is estimated
56. from the fitted longitudinal PSA profile of a patient. Along with PSA, in
57. some cohorts recently, MRI is also used for deciding biopsies. However, the
58. utility of MRI can only be determined with more follow-up data in the fu225 ture. Subsequently, MRI data can also be added as a predictor in our model.
59. Decisions based on information combined from multiple sources can yield
60. better results than based on MRI or PSA alone. We scheduled biopsies using
61. cause-specific cumulative upgrading-risk. Accounting for competing events,
62. such as treatment based on the number of positive biopsy cores, may lead to
63. improved personalized biopsy decisions. Although, in this work, we did not
64. consider such additional triggers for treatment because, unlike upgrading,
65. they differ between cohorts [5]. Upgrading is susceptible to inter-observer 233 variation too. Models which account for this variation [15, 26] will be in-

234 teresting to investigate further. However, the methodology for personalized 235 scheduling, and for comparison of various schedules need not change.

1. **5. Conclusions**
2. We successfully developed and validated a model for predicting upgrading-
3. risk, and providing risk-based personalized biopsy decisions, in prostate can-
4. cer AS. The model made available via a user-friendly web-application ([https:](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/)
5. [//emcbiostatistics.shinyapps.io/prias\_biopsy\_recommender/)](https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/) enables
6. shared decision making of biopsy schedules by comparing fixed and person-
7. alized schedules on total biopsies and expected time delay in detecting up-
8. grading. Novel biomarkers and MRI data can be added as predictors in the
9. model to improve predictions in the future. Recalibration of the baseline
10. hazard of upgrading-risk is advised before using the model in cohorts other 246 than the PRIAS cohort.
11. **Author Contributions**
12. Anirudh Tomer had full access to all the data in the study and takes
13. responsibility for the integrity of the data and the accuracy of the data anal-
14. ysis.
15. *Study concept and design:* Tomer, Nieboer, Roobol, Bjartell, and Ri-
16. zopoulos
17. *Acquisition of data:* Tomer, Nieboer, and Roobol
18. *Analysis and interpretation of data:* Tomer, Nieboer, and Rizopoulos
19. *Drafting of the manuscript:* Tomer, and Rizopoulos

# 256 Critical revision of the manuscript for important intellectual content: Tomer,

1. Nieboer, Roobol, Bjartell, Steyerberg, and Rizopoulos
2. *Statistical analyses*: Tomer, Nieboer, Steyerberg, and Rizopoulos
3. *Obtaining funding*: Roobol, and Rizopoulos
4. *Administrative, technical or material support*: Nieboer
5. *Supervision:* Roobol, and Rizopoulos

# 262 Other: none

1. **Acknowledgments**
2. The first and last authors would like to acknowledge support by Neder-
3. landse Organisatie voor Wetenschappelijk Onderzoek (the national research
4. council of the Netherlands) VIDI grant nr. 016.146.301, and Erasmus Uni-
5. versity Medical Center funding. Part of this work was carried out on the
6. Dutch national e-infrastructure with the support of SURF Cooperative. The
7. authors also thank the Erasmus University Medical Center’s Cancer Com-
8. putational Biology Center for giving access to their IT-infrastructure and
9. software that was used for the computations and data analysis in this study.
10. This work was supported by the Movember Foundation. The funder did
11. not play any role in the study design, collection, analysis or interpretation of 274 data, or in the drafting of this paper.
12. **Appendix A. Members of The Movember Foundations Global Ac-**
13. **tion Plan Prostate Cancer Active Surveillance (GAP3) consortium**
14. *Principle Investigators:* Bruce Trock (Johns Hopkins University, The
15. James Buchanan Brady Urological Institute, Baltimore, USA), Behfar Ehdaie 279 (Memorial Sloan Kettering Cancer Center, New York, USA), Peter Car-
16. roll (University of California San Francisco, San Francisco, USA), Christo-
17. pher Filson (Emory University School of Medicine, Winship Cancer Insti-
18. tute, Atlanta, USA), Jeri Kim / Christopher Logothetis (MD Anderson
19. Cancer Centre, Houston, USA), Todd Morgan (University of Michigan and
20. Michigan Urological Surgery Improvement Collaborative (MUSIC), Michi-
21. gan, USA), Laurence Klotz (University of Toronto, Sunnybrook Health Sci-
22. ences Centre, Toronto, Ontario, Canada), Tom Pickles (University of British
23. Columbia, BC Cancer Agency, Vancouver, Canada), Eric Hyndman (Uni-
24. versity of Calgary, Southern Alberta Institute of Urology, Calgary, Canada),
25. Caroline Moore (University College London & University College London
26. Hospital Trust, London, UK), Vincent Gnanapragasam (University of Cam-
27. bridge & Cambridge University Hospitals NHS Foundation Trust, Cam-
28. bridge, UK), Mieke Van Hemelrijck (King’s College London, London, UK
29. & Guys and St Thomas NHS Foundation Trust, London, UK), Prokar Das-
30. gupta (Guys and St Thomas NHS Foundation Trust, London, UK), Chris
31. Bangma (Erasmus Medical Center, Rotterdam, The Netherlands/ represen-
32. tative of Prostate cancer Research International Active Surveillance (PRIAS)
33. consortium), Monique Roobol (Erasmus Medical Center, Rotterdam, The
34. Netherlands/ representative of Prostate cancer Research International Active
35. Surveillance (PRIAS) consortium), Arnauld Villers (Lille University Hospi-
36. tal Center, Lille, France), Antti Rannikko (Helsinki University and Helsinki
37. University Hospital, Helsinki, Finland), Riccardo Valdagni (Department of
38. Oncology and Hemato-oncology, Universit degli Studi di Milano, Radia-
39. tion Oncology 1 and Prostate Cancer Program, Fondazione IRCCS Istituto 304 Nazionale dei Tumori, Milan, Italy), Antoinette Perry (University College
40. Dublin, Dublin, Ireland), Jonas Hugosson (Sahlgrenska University Hospital,
41. Gteborg, Sweden), Jose Rubio-Briones (Instituto Valenciano de Oncologa,
42. Valencia, Spain), Anders Bjartell (Skne University Hospital, Malm, Swe-
43. den), Lukas Hefermehl (Kantonsspital Baden, Baden, Switzerland), Lee Lui
44. Shiong (Singapore General Hospital, Singapore, Singapore), Mark Fryden-
45. berg (Monash Health; Monash University, Melbourne, Australia), Yoshiyuki
46. Kakehi / Mikio Sugimoto (Kagawa University Faculty of Medicine, Kagawa,
47. Japan), Byung Ha Chung (Gangnam Severance Hospital, Yonsei University 313 Health System, Seoul, Republic of Korea)
48. *Pathologist:* Theo van der Kwast (Princess Margaret Cancer Centre,
49. Toronto, Canada). Technology Research Partners: Henk Obbink (Royal
50. Philips, Eindhoven, the Netherlands), Wim van der Linden (Royal Philips,
51. Eindhoven, the Netherlands), Tim Hulsen (Royal Philips, Eindhoven, the 318 Netherlands), Cees de Jonge (Royal Philips, Eindhoven, the Netherlands).
52. *Advisory Regional statisticians:* Mike Kattan (Cleveland Clinic, Cleve-
53. land, Ohio, USA), Ji Xinge (Cleveland Clinic, Cleveland, Ohio, USA), Ken321 neth Muir (University of Manchester, Manchester, UK), Artitaya Lophatananon
54. (University of Manchester, Manchester, UK), Michael Fahey (Epworth Health-
55. Care, Melbourne, Australia), Ewout Steyerberg (Erasmus Medical Center,
56. Rotterdam, The Netherlands), Daan Nieboer (Erasmus Medical Center, Rot-
57. terdam, The Netherlands); Liying Zhang (University of Toronto, Sunnybrook
58. Health Sciences Centre, Toronto, Ontario, Canada)

# 327 Executive Regional statisticians: Ewout Steyerberg (Erasmus Medical

1. Center, Rotterdam, The Netherlands), Daan Nieboer (Erasmus Medical Cen-
2. ter, Rotterdam, The Netherlands); Kerri Beckmann (King’s College London,
3. London, UK & Guys and St Thomas NHS Foundation Trust, London, UK),
4. Brian Denton (University of Michigan, Michigan, USA), Andrew Hayen (Uni-
5. versity of Technology Sydney, Australia), Paul Boutros (Ontario Institute of 333 Cancer Research, Toronto, Ontario, Canada).

# 334 Clinical Research Partners IT Experts: Wei Guo (Johns Hopkins Uni-

1. versity, The James Buchanan Brady Urological Institute, Baltimore, USA),
2. Nicole Benfante (Memorial Sloan Kettering Cancer Center, New York, USA),
3. Janet Cowan (University of California San Francisco, San Francisco, USA),
4. Dattatraya Patil (Emory University School of Medicine, Winship Cancer In-
5. stitute, Atlanta, USA), Emily Tolosa (MD Anderson Cancer Centre, Hous-
6. ton, Texas, USA), Tae-Kyung Kim (University of Michigan and Michigan
7. Urological Surgery Improvement Collaborative, Ann Arbor, Michigan, USA),
8. Alexandre Mamedov (University of Toronto, Sunnybrook Health Sciences
9. Centre, Toronto, Ontario, Canada), Vincent LaPointe (University of British
10. Columbia, BC Cancer Agency, Vancouver, Canada), Trafford Crump (Uni-
11. versity of Calgary, Southern Alberta Institute of Urology, Calgary, Canada),
12. Vasilis Stavrinides (University College London & University College Lon-
13. don Hospital Trust, London, UK), Jenna Kimberly-Duffell (University of
14. Cambridge & Cambridge University Hospitals NHS Foundation Trust, Cam-
15. bridge, UK), Aida Santaolalla (King’s College London, London, UK & Guys
16. and St Thomas NHS Foundation Trust, London, UK), Daan Nieboer (Eras-
17. mus Medical Center, Rotterdam, The Netherlands), Jonathan Olivier (Lille
18. University Hospital Center, Lille, France), Tiziana Rancati (Fondazione IR-
19. CCS Istituto Nazionale dei Tumori di Milano, Milan, Italy), Heln Ahlgren 354 (Sahlgrenska University Hospital, Gteborg, Sweden), Juanma Mascars (Insti-
20. tuto Valenciano de Oncologa, Valencia, Spain), Annica Lfgren (Skne Univer-
21. sity Hospital, Malm, Sweden), Kurt Lehmann (Kantonsspital Baden, Baden,
22. Switzerland), Catherine Han Lin (Monash University and Epworth Health-
23. Care, Melbourne, Australia), Hiromi Hirama (Kagawa University, Kagawa,
24. Japan), Kwang Suk Lee (Yonsei University College of Medicine, Gangnam 360 Severance Hospital, Seoul, Korea).
25. *Research Advisory Committee:* Guido Jenster (Erasmus MC, Rotterdam,
26. the Netherlands), Anssi Auvinen (University of Tampere, Tampere, Finland),
27. Anders Bjartell (Skne University Hospital, Malm, Sweden), Masoom Haider
28. (University of Toronto, Toronto, Canada), Kees van Bochove (The Hyve
29. B.V. Utrecht, Utrecht, the Netherlands), Ballentine Carter (Johns Hopkins 366 University, Baltimore, USA until 2018).
30. *Management team:* Sam Gledhill (Movember Foundation, Melbourne,
31. Australia), Mark Buzza / Michelle Kouspou (Movember Foundation, Mel-
32. bourne, Australia), Chris Bangma (Erasmus Medical Center, Rotterdam,
33. The Netherlands), Monique Roobol (Erasmus Medical Center, Rotterdam,
34. The Netherlands), Sophie Bruinsma / Jozien Helleman (Erasmus Medical 372 Center, Rotterdam, The Netherlands).
35. **References**
36. 1. Briganti A, Fossati N, Catto JW, Cornford P, Montorsi F, Mottet N, 375 Wirth M, Van Poppel H. Active surveillance for low-risk prostate cancer:
37. the European Association of Urology position in 2018. *European urology*
38. 2018;74(3):357–68.
39. 2. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA.
40. The 2014 international society of urological pathology (isup) consensus
41. conference on gleason grading of prostatic carcinoma. *The American* 381 *journal of surgical pathology* 2016;40(2):244–52.
42. 3. Bruinsma SM, Roobol MJ, Carroll PR, Klotz L, Pickles T, Moore CM,
43. Gnanapragasam VJ, Villers A, Rannikko A, Valdagni R, et al. Expert
44. consensus document: semantics in active surveillance for men with lo-
45. calized prostate cancerresults of a modified delphi consensus procedure. 386 *Nature reviews urology* 2017;14(5):312.
46. 4. Bul M, Zhu X, Valdagni R, Pickles T, Kakehi Y, Rannikko A, Bjartell A,
47. Van Der Schoot DK, Cornel EB, Conti GN, et al. Active surveillance for
48. low-risk prostate cancer worldwide: the prias study. *European urology* 390 2013;63(4):597–603.
49. 5. Nieboer D, Tomer A, Rizopoulos D, Roobol MJ, Steyerberg EW. Active
50. surveillance: a review of risk-based, dynamic monitoring. *Translational* 393 *andrology and urology* 2018;7(1):106–15.
51. 6. Loeb S, Carter HB, Schwartz M, Fagerlin A, Braithwaite RS, Lepor H.
52. Heterogeneity in active surveillance protocols worldwide. *Reviews in* 396 *urology* 2014;16(4):202–3.
53. 7. Loeb S, Vellekoop A, Ahmed HU, Catto J, Emberton M, Nam R, Rosario
54. DJ, Scattoni V, Lotan Y. Systematic review of complications of prostate
55. biopsy. *European urology* 2013;64(6):876–92.
56. 8. Bokhorst LP, Alberts AR, Rannikko A, Valdagni R, Pickles T, Kakehi Y,
57. Bangma CH, Roobol MJ, PRIAS study group . Compliance rates with
58. the Prostate Cancer Research International Active Surveillance (PRIAS)
59. protocol and disease reclassification in noncompliers. *European Urology* 404 2015;68(5):814–21.
60. 9. Inoue LY, Lin DW, Newcomb LF, Leonardson AS, Ankerst D, Gulati R,
61. Carter HB, Trock BJ, Carroll PR, Cooperberg MR, et al. Comparative
62. analysis of biopsy upgrading in four prostate cancer active surveillance 408 cohorts. *Annals of internal medicine* 2018;168(1):1–9.
63. 10. Bratt O, Carlsson S, Holmberg E, Holmberg L, Johansson E, Josefsson
64. A, Nilsson A, Nyberg M, Robinsson D, Sandberg J, et al. The study of
65. active monitoring in sweden (sams): a randomized study comparing two
66. different follow-up schedules for active surveillance of low-risk prostate 413 cancer. *Scandinavian journal of urology* 2013;47(5):347–55.
67. 11. de Carvalho TM, Heijnsdijk EA, de Koning HJ. Estimating the risks
68. and benefits of active surveillance protocols for prostate cancer: a mi416 crosimulation study. *BJU international* 2017;119(4):560–6.
69. 12. Partin AW, Yoo J, Carter HB, Pearson JD, Chan DW, Epstein JI, Walsh
70. PC. The use of prostate specific antigen, clinical stage and gleason score
71. to predict pathological stage in men with localized prostate cancer. *The* 420 *Journal of urology* 1993;150(1):110–4.
72. 13. Makarov DV, Trock BJ, Humphreys EB, Mangold LA, Walsh PC, Ep-
73. stein JI, Partin AW. Updated nomogram to predict pathologic stage of 423 prostate cancer given prostate-specific antigen level, clinical stage, and

424 biopsy gleason score (partin tables) based on cases from 2000 to 2005. 425 *Urology* 2007;69(6):1095–101.

1. 14. Tomer A, Nieboer D, Roobol MJ, Steyerberg EW, Rizopoulos D. Per-
2. sonalized schedules for surveillance of low-risk prostate cancer patients. 428 *Biometrics* 2019;75(1):153–62.
3. 15. Coley RY, Zeger SL, Mamawala M, Pienta KJ, Carter HB. Prediction
4. of the pathologic gleason score to inform a personalized management 431 program for prostate cancer. *European urology* 2017;72(1):135–41.

432 16. Rizopoulos D. Joint Models for Longitudinal and Time-to-Event Data: 433 With Applications in R. CRC Press; 2012. ISBN 9781439872864.

1. 17. Bruinsma SM, Zhang L, Roobol MJ, Bangma CH, Steyerberg EW,
2. Nieboer D, Van Hemelrijck M, consortium MFGAPPCASG, Trock B,
3. Ehdaie B, et al. The movember foundation’s gap3 cohort: a profile of
4. the largest global prostate cancer active surveillance database to date.
5. *BJU international* 2018;121(5):737–44.
6. 18. Laird NM, Ware JH, et al. Random-effects models for longitudinal data. 440 *Biometrics* 1982;38(4):963–74.
7. 19. Vickers AJ, Savage C, O’Brien MF, Lilja H. Systematic review of pre-
8. treatment prostate-specific antigen velocity and doubling time as pre-
9. dictors for prostate cancer. *Journal of Clinical Oncology* 2009;27(3):398. 444 20. Rizopoulos D. The R package JMbayes for fitting joint models for lon-

445 gitudinal and time-to-event data using MCMC. *Journal of Statistical* 446 *Software* 2016;72(7):1–46.

1. 21. Rizopoulos D, Molenberghs G, Lesaffre EM. Dynamic predictions with
2. time-dependent covariates in survival analysis using joint modeling and 449 landmarking. *Biometrical Journal* 2017;59(6):1261–76.
3. 22. Royston P, Altman DG. External validation of a cox prognostic
4. model: principles and methods. *BMC medical research methodology* 452 2013;13(1):33.
5. 23. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski
6. N, Pencina MJ, Kattan MW. Assessing the performance of prediction
7. models: a framework for some traditional and novel measures. *Epidemi*456 *ology (Cambridge, Mass)* 2010;21(1):128.
8. 24. Turnbull BW. The empirical distribution function with arbitrarily
9. grouped, censored and truncated data. *Journal of the Royal Statisti*459 *cal Society Series B (Methodological)* 1976;38(3):290–5.
10. 25. Ankerst DP, Xia J, Thompson Jr IM, Hoefler J, Newcomb LF, Brooks
11. JD, Carroll PR, Ellis WJ, Gleave ME, Lance RS, et al. Precision
12. medicine in active surveillance for prostate cancer: development of the
13. canary–early detection research network active surveillance biopsy risk 464 calculator. *European urology* 2015;68(6):1083–8.
14. 26. Balasubramanian R, Lagakos SW. Estimation of a failure time distribu-
15. tion based on imperfect diagnostic tests. *Biometrika* 2003;90(1):171–82.