A Ready To Use Web-Application Providing a Personalized Biopsy Schedule for Men With Low-Risk PCa Under Active Surveillance*

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

[a] *Department of Biostatistics, Erasmus University Medical Center, Rotterdam, the Netherlands*
[b] *Department of Public Health, Erasmus University Medical Center, Rotterdam, the Netherlands*
[c] *Department of Urology, Erasmus University Medical Center, Rotterdam, the Netherlands*
[d] *Department of Urology, Skåne University Hospital, Malmö, Sweden*
[e] *Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands*
[f] *The 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 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: 300; Word Count Text: 2500

*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)

reduce patient burden.

`Objective`: Develop 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`: Bayesian joint model fitted to PRIAS dataset. Model validated in GAP3 cohorts using risk prediction error, calibration, area under ROC (AUC). Model and personalized biopsy schedules based on predicted risks implemented in a web-application.

`Results and Limitations`: Upgrading rate at year five of follow-up: 35% in PRIAS, at most 50% in GAP3 cohorts. PSA velocity 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 with upgrading rate 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. Introduction

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

In most AS protocols, biopsies are conducted periodically. Consequently, upgrading is always detected with a time delay (Figure 1). For detecting

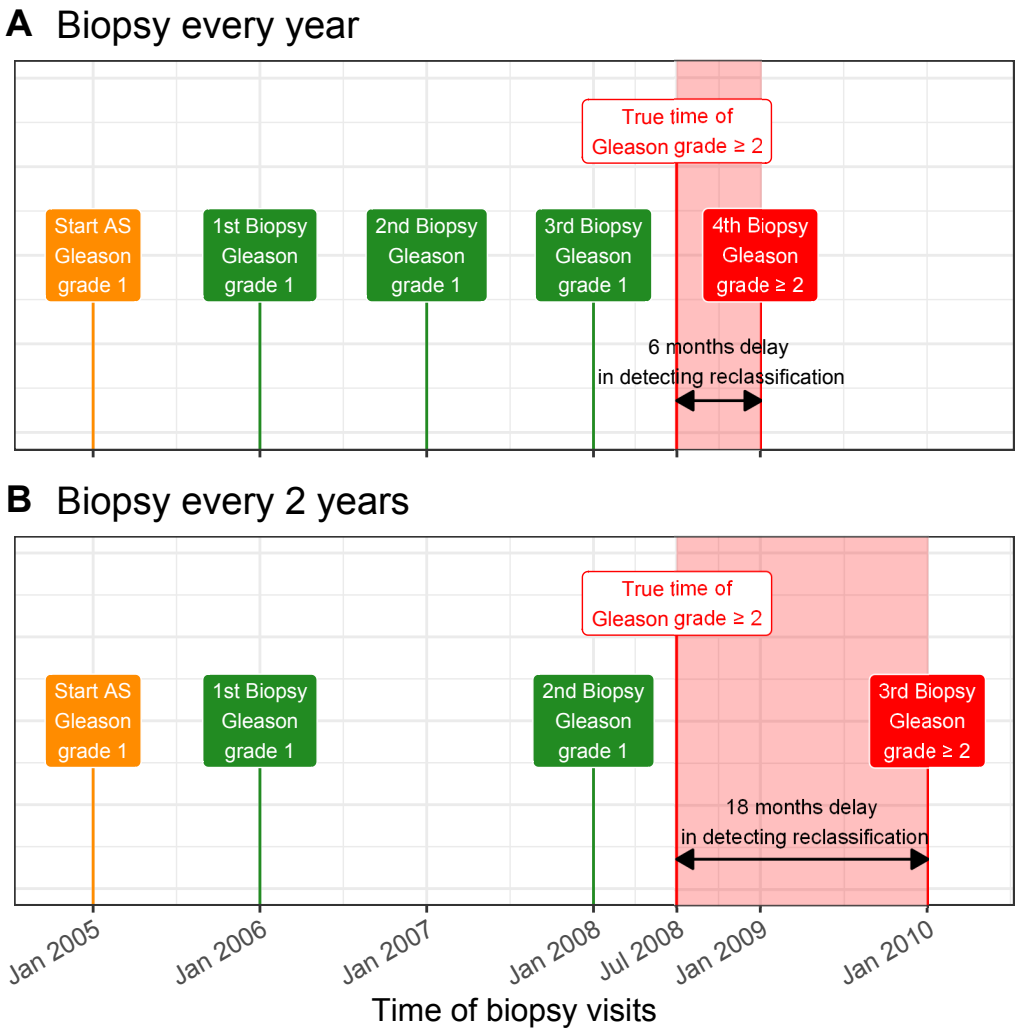

Figure 1: **Trade-off between the number of biopsies and time delay in detecting upgrading (Increase in Gleason grade group from 1 to 2 or higher):** The true time of upgrading for the patient in this figure is July 2008. When biopsies are scheduled annually (**Panel A**), upgrading is detected in January 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**.

upgrading timely, many AS programs schedule fixed and frequent biopsies (e.g., annually) for all patients [5, 6]. However, this leads to many unnecessary biopsies in slow/non-progressing patients. Biopsies are invasive, may be painful, and are prone to medical complications such as bleeding and septicemia[7]. Thus, biopsy burden and patient non-compliance to frequent biopsies [8] has raised concerns regarding the optimal biopsy schedule [9, 10]. To this end, infrequent schedules such as biennial biopsies have been proposed as an alternative [9, 11]. Although, biennial biopsies may still lead to five unnecessary biopsies over ten years (current study period of large AS programs) for slow/non-progressing patients. A promising alternative to fixed and frequent biopsies is personalized biopsy schedules based on the patient-specific risk of upgrading (Figure 2).

The first challenge in developing personalized biopsy schedules is consolidating accumulated patient data (e.g., PSA, previous biopsy results) into risk estimates for upgrading. Existing calculators for risk of upgrading [12, 13] use only the latest PSA measurement of a patient. In contrast, we intend to utilize all repeated measurements of PSA, previous biopsy results, and baseline characteristics of a patient. To this end, a suitable model is the joint model for time-to-event and longitudinal data [14, 15, 16]. A joint model predicts risk of upgrading in a personalized manner. A subsequent challenge, however, is translating risks into clinical decisions. For example, a 10% risk of upgrading can be perceived high/low depending upon the patient age. Patients may also weigh risks of upgrading with the potential *consequences* of another biopsy. Two relevant *consequences* of biopsies (Figure 1) are the timing and total number of biopsies (burden), and the time delay in detecting

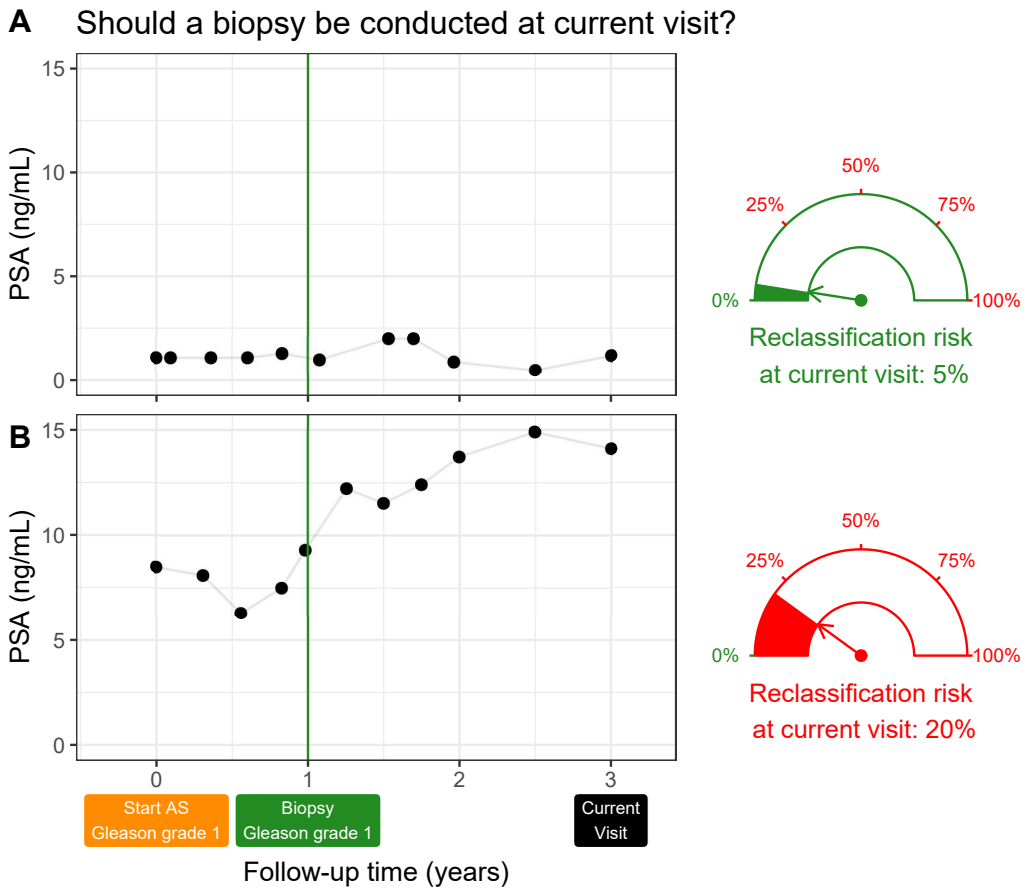

Figure 2: **Motivation for personalized risk-based decisions of biopsy**: Patient A (**Panel A**) and B (**Panel B**) had their latest biopsy at year one of follow-up (green vertical line). Patient A's prostate-specific antigen (PSA) profile remained stable until his current visit at year three, whereas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

upgrading (smaller is beneficial). The relative importance of these *consequences* can vary between the patients, and also over the follow-up period for the same patient.

The goal of this work is to develop a robust, generalizable model that gives reliable estimates for individualized risk of upgrading, and to create personalized biopsy schedules based on these risks. To further assist patients and doctors in shared decision making of a biopsy schedule, we also aim to provide quantitative estimates of *consequences* of opting for a personalized biopsy schedule versus the standard fixed protocol. For developing our model we will use the worlds largest AS dataset PRIAS. Subsequently, we will externally validate our model in the largest five AS cohorts from the Movember Foundation's GAP3 database [17]. Last, we intend to implement our model and methodology in a web-application.

2. Patients and Methods

2.1. Study Cohort

For developing a statistical model to power our web-application, we used the Prostate Cancer International Active Surveillance (PRIAS) database. It is an ongoing (December 2006 – to date) prospective cohort study of men with low- and very-low risk prostate cancer diagnoses [4]. More than 100 medical centers from 17 countries contributed to PRIAS, using a common protocol (`https://www.prias-project.org`). Upon inclusion in PRIAS, PSA was measured quarterly for the first two years of follow-up and semiannually thereafter. Biopsies were scheduled at year one, four, seven, and ten of follow-up. 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 [2] at the time of inclusion in PRIAS (Table 1). Our primary event of interest is increase in this Gleason grade group upon repeat biopsy, called *upgrading* (1134 patients). Upgrading is a trigger for treatment advice in PRIAS. Although, 2250 patients were provided treatment on the basis of their PSA, or number of biopsy cores with cancer, or anxiety/other reasons. Our reasons for focusing solely on upgrading are, namely, upgrading is strongly associated with cancer-related outcomes, and other triggers for treatment vary between cohorts.

2.2. Statistical Model

To create personalized biopsy schedules based on patient-specific risk of upgrading, we aimed to develop a risk prediction model. Available data was patient age at inclusion in AS, longitudinally measured PSA, timing of repeat biopsies and corresponding Gleason grades, and observed time of upgrading. Analysis of this data required modeling the within-patient correlation for PSA, association between the Gleason grades and PSA profiles of a patient, and handling missing PSA measurements after a patient experienced upgrading. In such situations, a commonly used model is the joint model for time-to-event and longitudinal data [14, 15, 16].

Our joint model consisted of two sub-models. First, a linear mixed model [18] for longitudinally measured PSA (log-transformed). Second, a relative-risk model (similar to Cox model) for obtaining the risk of upgrading. In the model for PSA, we fitted a curve to PSA measurements (Panel A, Figure 3). From each patient's fitted PSA profile, we extracted the instan-

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

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 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)

taneous PSA velocity. This velocity varies over time (Panel B, Figure 3). Consequently, it is more precise than the currently used constant PSA velocity assumption [19]. We connected the two sub-models by using the fitted PSA and instantaneous velocity as predictors in the sub-model for risk of upgrading (Panel C, Figure 3). Patient age was included in both sub-models. The parameters of the two sub-models were estimated jointly (Supplementary A) using the R package **JMbayes** [20].

2.3. Risk of Upgrading Based Personalized Biopsies

The key component in personalized schedules is the cumulative-risk of upgrading. Given a patient's accumulated PSA measurements and biopsy results, our joint model predicted the cumulative-risk of upgrading at his current as well as future visit times (Panel C, Figure 3). This cumulative-risk is updated with more patient data over follow-up (Figure 5, Supplementary B).

In PRIAS, patient PSA was measured every six months. If during a PSA visit, a patient's predicted cumulative-risk of upgrading was more than a certain threshold (e.g., 10%), we scheduled an immediate biopsy. We scheduled future biopsies too because our model predicts patient's cumulative-risk at his future follow-up visits as well. We achieved this by repeatedly applying the same risk threshold rule at each future follow-up visit (Supplementary C). We maintained a minimum gap of one year between consecutive biopsies (PRIAS recommendation). Example personalized schedules based on 5% and 10% risk thresholds are shown in Panel B, Figure 4. Due to the currently limited follow-up period of PRIAS, we were able to schedule biopsies during the first six years of follow-up only (Table 12, Supplementary C).

The choice of the risk threshold in the personalized schedule dictates the

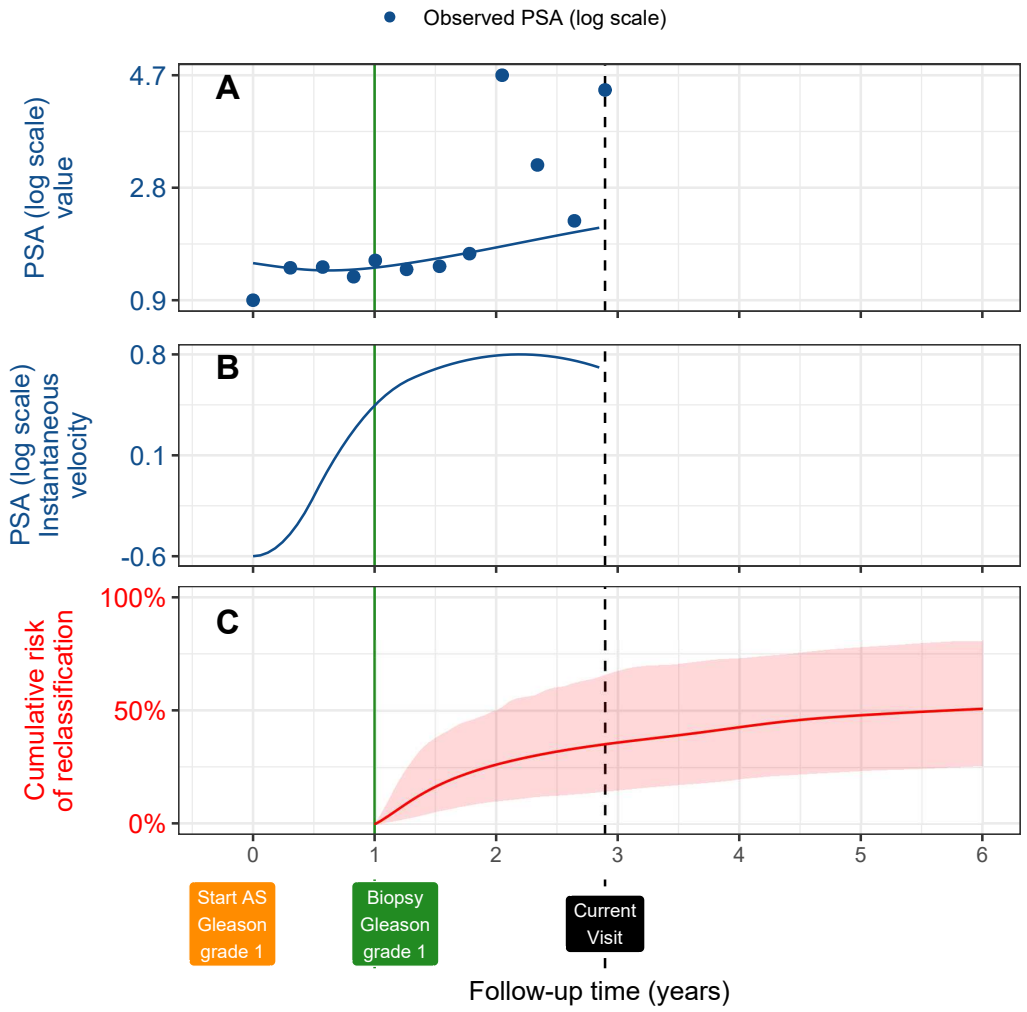

Figure 3: **Illustration of the joint model on a real PRIAS patient**. **Panel A:** Observed PSA (blue dots) and fitted PSA (solid blue line), log-transformed. **Panel B:** Estimated instantaneous velocity of PSA (log-transformed). **Panel C**: Predicted cumulative-risk of upgrading (95% credible interval shaded). Upgrading is defined as an increase in Gleason grade group from group 1 to 2 or higher. This risk of upgrading is available 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 value and fitted instantaneous velocity of PSA (log scale), and time of the latest negative biopsy. Black dashed line at year 3 denotes the time of current visit.

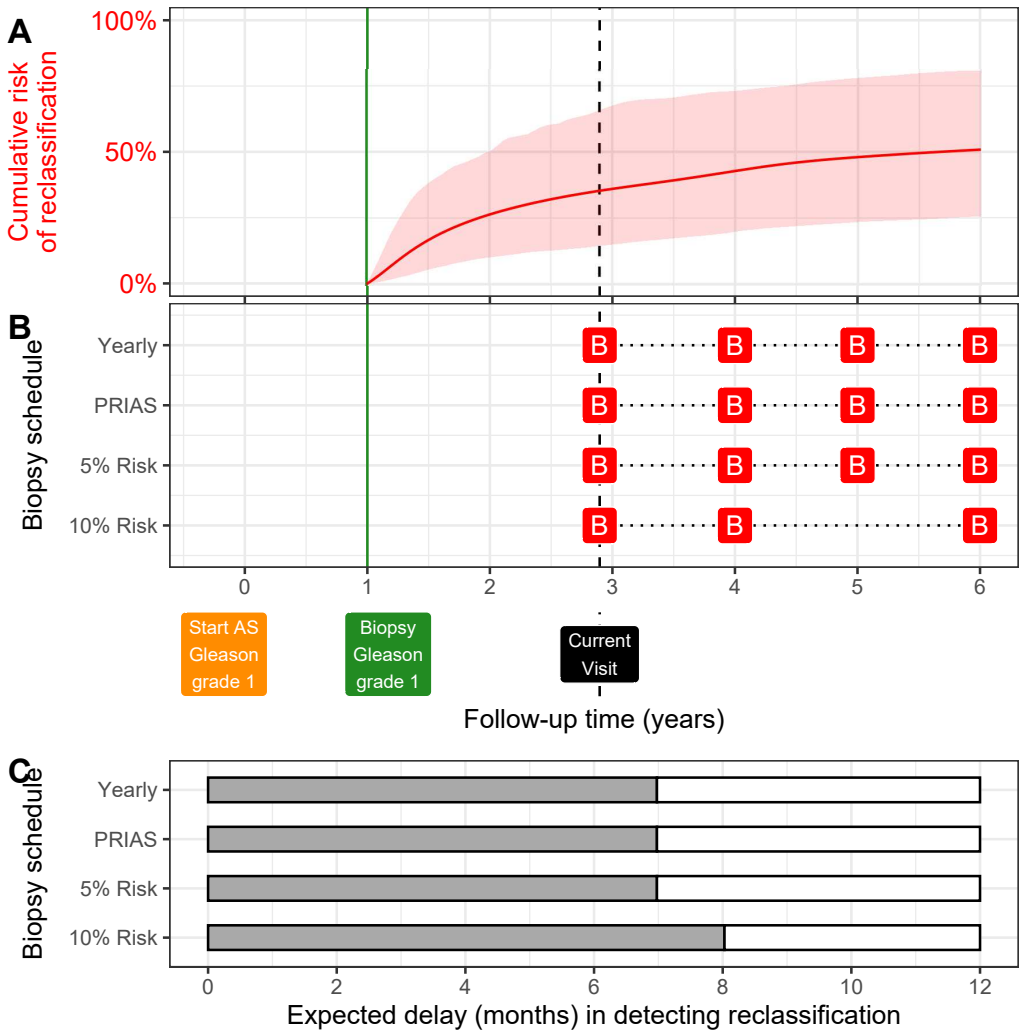

Figure 4: **Illustration of personalized and fixed schedules of biopsies**. The PSA profile of this patient is shown in Figure 3. **Panel A:** Predicted cumulative-risk of upgrading (95% credible interval shaded). **Panel B:** Personalized and fixed schedules of biopsies, with a red 'B' indicating a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

consequences of following that schedule. *Consequences* are the timing and the total number of biopsies, and the expected time delay in detecting upgrading. Our model estimated *consequences* in a personalized manner (Panel B,C in Figure 4, and Figure 9–11 in Supplementary C) given any schedule of biopsies. Thus, patients can compare personalized schedules based on different risk thresholds, with fixed schedules, before making a choice.

2.4. Model Validation

We validated the PRIAS based model internally using the PRIAS cohort, and externally using the largest five GAP3 database [17] cohorts. These were, namely, University of Toronto AS (Toronto), Johns Hopkins AS (Hopkins), Memorial Sloan Kettering Cancer Center AS (MSKCC), King's College London AS (KCL), and Michigan Urological Surgery Improvement Collaborative AS (MUSIC). We assessed our model's ability to discriminate between patients who experience/do not experience upgrading, via the area under the receiver operating characteristic curve or AUC [21]. We employed calibration plots [22, 23] and mean absolute prediction error [21] to graphically and quantitatively evaluate the prediction accuracy of our model. Due to the longitudinal nature of AS studies, the AUC and prediction error varies over follow-up (Supplementary B.1). Lastly, to resolve any potential model miscalibration in external GAP3 cohorts we aimed to recalibrate our model's baseline hazard of upgrading, individually for each GAP3 cohort (Supplementary B.1).

IMPORTATN...make this a subsection and move web-application to results Finally we aim to transform our model into a user friendly web application including visualization of different biopsy risk thresholds and the

consequence (i.e dealy time???)

2.5. Web-Application

We implemented our methodology in a web-application `https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/`. It utilizes the joint model fitted to the PRIAS dataset. Currently, the web-application supports PRIAS and the five external cohorts in which we validated our model. Patient data can be entered manually or can be uploaded in Microsoft Excel format. Predictions for risk of upgrading are shown for a currently limited, cohort-specific, follow-up period (Table 12, Supplementary C). The web-application allows comparison of the *consequences* of following these schedules: personalized schedules based on 5%, 10%, and 15% risk threshold; annual biopsies; biennial biopsies; and PRIAS schedule.

3. Results

The rate of upgrading at year five of follow-up was 35% in PRIAS, and at most 50% in the five validation GAP3 cohorts (Panel B, Figure 5). That is, many patients do not require any biopsy in the first five years of AS.

In the fitted joint model, when patient age increased from 61 to 71 years (25-th to 75-th percentile), the adjusted hazard ratio of upgrading was 1.45 (95%CI: 1.30–1.63). When fitted PSA value (log scale) increased from 2.36 to 3.07 (25-th to 75-th percentile), the adjusted hazard ratio was 0.99 (95%CI: 0.89–1.11). When estimated instantaneous PSA (log scale) velocity increased from -0.09 to 0.31 (25-th to 75-th percentile), the adjusted hazard ratio was 2.47 (95%CI: 1.93–2.99). Hence, instantaneous velocity of

PSA was a stronger predictor of upgrading than PSA value. Detailed parameter estimates are in Supplementary A.2.

The time-varying mean absolute prediction error, time-varying AUC, and calibration plot of our model in different cohorts are shown in Panel B, Figure 8, Supplementary B; Panel A, Figure 5; and Panel B, Figure 5, respectively. The AUC was moderate (0.55 to 0.75) in all cohorts. Mean absolute prediction error was large (0.3 to 0.45) in cohorts with rate of upgrading different from PRIAS, and moderate (0.1 to 0.3) otherwise. Our model required recalibration of baseline hazard of upgrading in all cohorts (Figure 6, Supplementary B). Although, calibration was fine in Johns Hopkins cohort, whose rate of upgrading was similar to PRIAS (Panel B, Figure 5). The risk predictions from the recalibrated models were as good as risk predictions from joint models fitted separately to each cohort (Figure 7, Supplementary B). Comprehensive validation results are in Supplementary B.

Various personalized and fixed biopsy schedules for a demonstration patient in Figure 4 show that a personalized schedule based on 10% risk threshold leads to one less biopsy than other schedules. At the same time, the corresponding time delay in detection of upgrading is expected to be only one month more than other 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. Additional demonstrations are in Figure 9–11, Supplementary C.

4. Discussion

First mention overall result and performance

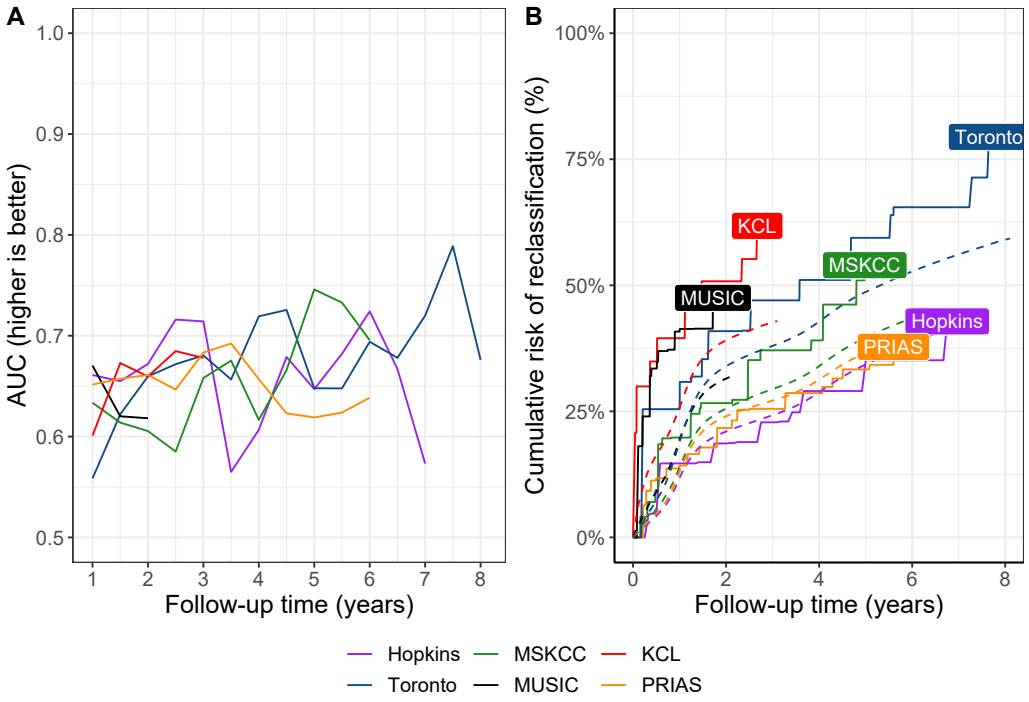

Figure 5: **Model Validation Results**. **Panel A**: time dependent area under the receiver operating characteristic curve or AUC (measure of discrimination). **Panel B**: calibration-at-large [22, 23], with solid lines depicting the non-parameteric estimate of the cumulative-risk of upgrading [24], and dashed lines showing the average cumulative-risk of upgrading obtained using the joint model fitted to the PRIAS dataset. 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.

We developed a web-application for assisting patients/doctors in making biopsy decisions during prostate cancer active surveillance (AS). Our web-application provides the patient's current and future risks of upgrading (increase in Gleason grade [2] from grade 1 to 2 or higher), and personalized biopsy schedules based on this risk. Our work has four novel features over earlier risk calculators [15, 25]. First, for personalized biopsy schedules, we developed a statistical model using the world's largest AS dataset PRIAS. Second, for following any biopsy schedule, fixed or personalized, our model predicts the corresponding time delay in detection of upgrading (less is beneficial). Thus, patients/doctors can compare schedules before making a choice. Third, we externally validated our model in the largest five GAP3 database [17] AS cohorts. Fourth, we implemented our methodology in a web-application (`https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/`) for PRIAS and validated GAP3 cohorts.

Currently, biopsies are decided either according to fixed schedules (e.g., annual biopsies) or utilize PSA. Both approaches have drawbacks [19, 8]. In particular, PSA has not been exploited fully and correctly. For example, using observed PSA is incorrect because it has measurement error. Other approaches utilize only the latest PSA, and/or when they utilize all PSA data, they assume constant PSA velocity. In contrast, our model employs all PSA measurements to build a patient-specific profile of PSA. This profile is allowed to increase/decrease non-linearly over time (non-constant PSA velocity). Subsequently, the model consolidates the PSA profile, previous biopsy results, and baseline characteristics of a patient, into a single personalized risk of upgrading. This risk also gets updated as more patient data becomes

available over follow-up. Due to currently limited magnetic resonance imaging (MRI) data, we could not incorporate it into our model. However, MRI data can be added as a predictor in our model in the future. Decisions based on information combined from multiple sources can yield better results than based on MRI or PSA alone.

MRI is used a lot today and more and more before the first biopsy and a lot in AS patients to avoid re-biopsy. I would suggest to write that it is used in revised protocols and that the value of it is still to be determined when more follow-up data has been collected.

Our model is useful for a large number of patients from PRIAS (model development), and the largest five GAP3 database AS cohorts (model external validation). These are the University of Toronto AS, Johns Hopkins AS, Memorial Sloan Kettering Cancer Center AS, King's College London AS, and Michigan Urological Surgery Improvement Collaborative AS. During validation, we required recalibration of our model's baseline hazard of upgrading, individually for all validation cohorts. Our model's prediction error was moderate in cohorts with rate of upgrading similar to PRIAS, and large otherwise. Both prediction error and AUC can be improved with newer biomarkers or MRI data in the future.

Our work has important clinical implications. The rate of upgrading after five years of follow-up was at most 50% in all cohorts (Figure 5). That is, a large number of patients do not require any biopsy during the first five years of follow-up. Given the non-compliance and burden of frequent biopsies [8], the availability of our methodology as a web-application may encourage patients/doctors to consider personalized schedules instead. To

assist them in this decision making, the web-application provides an estimate of time delay in detection of upgrading for both personalized and fixed schedules, in a personalized manner. We hope this will objectively address patient apprehensions regarding adverse outcomes in AS.

This work has certain limitations. Due to currently limited follow-up period of PRIAS and GAP3 cohorts, the proposed model is valid only for a restricted period (Table 12, Supplementary C). This problem can be mitigated by refitting the model with new follow-up data in the future. While we focused only on upgrading, the number of positive biopsy cores can also be used to trigger treatment. We did not consider such additional criteria because they differ between cohorts [5], whereas upgrading is used widely. Upgrading is susceptible to inter-observer variation too. Models which account for this variation [15, 26] will be interesting to investigate further. However, the methodology for personalized scheduling, and for comparison of various schedules need not change.

5. Conclusions

Prediction tool for risk of upgrading presented in a web-based application (`https://emcbiostatistics.shinyapps.io/prias_biopsy_recommender/`), for assisting patients/doctors in making biopsy decisions during prostate cancer AS. Our web-application provides the patient's current and future risks of upgrading, and personalized biopsy schedules based on this risk. Our web-application enables shared decision making of biopsy schedule by comparing fixed and personalized schedules on the total biopsies and expected time delay in detection of upgrading.

Currently supported cohorts are the worlds largest AS cohort PRIAS (model development), and the largest five GAP3 database cohorts (model external validation). Risk prediction accuracy in validation cohorts was better only if they had the rate of reclassification similar to PRIAS

Instead of mentionre currently supported cohorts say: Not in conclusion, more a general remarks that calibration to a specific setting is adviasble or something like that.

Author Contributions

Anirudh Tomer had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

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

Acquisition of data: Tomer, Nieboer, and Roobol

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

Drafting of the manuscript: Tomer, and Rizopoulos

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

Statistical analyses: Tomer, Nieboer, Steyerberg, and Rizopoulos

Obtaining funding: Roobol, and Rizopoulos

Administrative, technical or material support: Nieboer

Supervision: Roobol, and Rizopoulos

Other: none

Acknowledgments

The first and last authors would like to acknowledge support by Nederlandse Organisatie voor Wetenschappelijk Onderzoek (the national research council of the Netherlands) VIDI grant nr. 016.146.301, and Erasmus University Medical Center funding. Part of this work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. The authors also thank the Erasmus University Medical Center's Cancer Computational Biology Center for giving access to their IT-infrastructure and software that was used for the computations and data analysis in this study.

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.

Appendix A. Members of The Movember Foundations 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

Columbia, BC Cancer Agency, Vancouver, Canada), Eric Hyndman (University of Calgary, Southern Alberta Institute of Urology, Calgary, Canada), Caroline Moore (University College London & University College London Hospital Trust, London, UK), Vincent Gnanapragasam (University of Cambridge & Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK), Mieke Van Hemelrijck (King's College London, London, UK & Guys and St Thomas NHS Foundation Trust, London, UK), Prokar Dasgupta (Guys and St Thomas NHS Foundation Trust, London, UK), Chris Bangma (Erasmus Medical Center, Rotterdam, The Netherlands/ representative of Prostate cancer Research International Active Surveillance (PRIAS) consortium), Monique Roobol (Erasmus Medical Center, Rotterdam, The Netherlands/ representative of Prostate cancer Research International Active Surveillance (PRIAS) consortium), Arnauld Villers (Lille University Hospital Center, Lille, France), Antti Rannikko (Helsinki University and Helsinki University Hospital, Helsinki, Finland), Riccardo Valdagni (Department of Oncology and Hemato-oncology, Universit degli Studi di Milano, Radiation Oncology 1 and Prostate Cancer Program, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy), Antoinette Perry (University College Dublin, Dublin, Ireland), Jonas Hugosson (Sahlgrenska University Hospital, Gteborg, Sweden), Jose Rubio-Briones (Instituto Valenciano de Oncologa, Valencia, Spain), Anders Bjartell (Skne University Hospital, Malm, Sweden), Lukas Hefermehl (Kantonsspital Baden, Baden, Switzerland), Lee Lui Shiong (Singapore General Hospital, Singapore, Singapore), Mark Frydenberg (Monash Health; Monash University, Melbourne, Australia), Yoshiyuki Kakehi / Mikio Sugimoto (Kagawa University Faculty of Medicine, Kagawa,

Japan), Byung Ha Chung (Gangnam Severance Hospital, Yonsei University Health System, Seoul, Republic of Korea)

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

Advisory Regional statisticians: Mike Kattan (Cleveland Clinic, Cleveland, Ohio, USA), Ji Xinge (Cleveland Clinic, Cleveland, Ohio, USA), Kenneth Muir (University of Manchester, Manchester, UK), Artitaya Lophatananon (University of Manchester, Manchester, UK), Michael Fahey (Epworth HealthCare, Melbourne, Australia), Ewout Steyerberg (Erasmus Medical Center, Rotterdam, The Netherlands), Daan Nieboer (Erasmus Medical Center, Rotterdam, The Netherlands); Liying Zhang (University of Toronto, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada)

Executive Regional statisticians: Ewout Steyerberg (Erasmus Medical Center, Rotterdam, The Netherlands), Daan Nieboer (Erasmus Medical Center, Rotterdam, The Netherlands); Kerri Beckmann (King's College London, London, UK & Guys and St Thomas NHS Foundation Trust, London, UK), Brian Denton (University of Michigan, Michigan, USA), Andrew Hayen (University of Technology Sydney, Australia), Paul Boutros (Ontario Institute of Cancer Research, Toronto, Ontario, Canada).

Clinical Research Partners IT Experts: Wei Guo (Johns Hopkins University, The James Buchanan Brady Urological Institute, Baltimore, USA), Nicole Benfante (Memorial Sloan Kettering Cancer Center, New York, USA),

Janet Cowan (University of California San Francisco, San Francisco, USA), Dattatraya Patil (Emory University School of Medicine, Winship Cancer Institute, Atlanta, USA), Emily Tolosa (MD Anderson Cancer Centre, Houston, Texas, USA), Tae-Kyung Kim (University of Michigan and Michigan Urological Surgery Improvement Collaborative, Ann Arbor, Michigan, USA), Alexandre Mamedov (University of Toronto, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada), Vincent LaPointe (University of British Columbia, BC Cancer Agency, Vancouver, Canada), Trafford Crump (University of Calgary, Southern Alberta Institute of Urology, Calgary, Canada), Vasilis Stavrinides (University College London & University College London Hospital Trust, London, UK), Jenna Kimberly-Duffell (University of Cambridge & Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK), Aida Santaolalla (King's College London, London, UK & Guys and St Thomas NHS Foundation Trust, London, UK), Daan Nieboer (Erasmus Medical Center, Rotterdam, The Netherlands), Jonathan Olivier (Lille University Hospital Center, Lille, France), Tiziana Rancati (Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Milan, Italy), Heln Ahlgren (Sahlgrenska University Hospital, Gteborg, Sweden), Juanma Mascars (Instituto Valenciano de Oncologa, Valencia, Spain), Annica Lfgren (Skne University Hospital, Malm, Sweden), Kurt Lehmann (Kantonsspital Baden, Baden, Switzerland), Catherine Han Lin (Monash University and Epworth HealthCare, Melbourne, Australia), Hiromi Hirama (Kagawa University, Kagawa, Japan), Kwang Suk Lee (Yonsei University College of Medicine, Gangnam Severance Hospital, Seoul, Korea).

Research Advisory Committee: Guido Jenster (Erasmus MC, Rotterdam,

the Netherlands), Anssi Auvinen (University of Tampere, Tampere, Finland), Anders Bjartell (Skne University Hospital, Malm, Sweden), Masoom Haider (University of Toronto, Toronto, Canada), Kees van Bochove (The Hyve B.V. Utrecht, Utrecht, the Netherlands), Ballentine Carter (Johns Hopkins University, Baltimore, USA until 2018).

Management team: Sam Gledhill (Movember Foundation, Melbourne, Australia), Mark Buzza / Michelle Kouspou (Movember Foundation, Melbourne, Australia), Chris Bangma (Erasmus Medical Center, Rotterdam, The Netherlands), Monique Roobol (Erasmus Medical Center, Rotterdam, The Netherlands), Sophie Bruinsma / Jozien Helleman (Erasmus Medical Center, Rotterdam, The Netherlands).

References

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

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

3. Bruinsma SM, Roobol MJ, Carroll PR, Klotz L, Pickles T, Moore CM, Gnanapragasam VJ, Villers A, Rannikko A, Valdagni R, et al. Expert consensus document: semantics in active surveillance for men with lo-

calized prostate cancerresults 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, Bjartell A, Van Der Schoot DK, Cornel EB, Conti GN, et al. Active surveillance for low-risk prostate cancer worldwide: the prias study. *European urology* 2013;63(4):597–603.

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

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, Rosario DJ, Scattoni V, Lotan Y. 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, Bangma CH, Roobol MJ, PRIAS study group . 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, Carter HB, Trock BJ, Carroll PR, Cooperberg MR, et al. Comparative

analysis of biopsy upgrading in four prostate cancer active surveillance cohorts. *Annals of internal medicine* 2018;168(1):1–9.

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

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

12. Partin AW, Yoo J, Carter HB, Pearson JD, Chan DW, Epstein JI, Walsh PC. The use of prostate specific antigen, clinical stage and gleason score to predict pathological stage in men with localized prostate cancer. *The Journal of urology* 1993;150(1):110–4.

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

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

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

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

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

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

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

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.

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.

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, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology (Cambridge, Mass)* 2010;21(1):128.

24. Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society Series B (Methodological)* 1976;38(3):290–5.

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

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