available at www.sciencedirect.com
journal homepage: www.europeanurology.com

European Association of Urology

Prostate Cancer

Prediction of the Pathologic Gleason Score to Inform a Personalized Management Program for Prostate Cancer

R. Yates Coley [a], *Scott L. Zeger* [a], *Mufaddal Mamawala* [b], *Kenneth J. Pienta* [b], *H. Ballentine Carter* [b,*]

[a] *Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA;* [b] *The Johns Hopkins University School of Medicine, The James Brady Urological Institute, Johns Hopkins Hospital, Baltimore, MD, USA*

Article info

Article history:
Accepted August 3, 2016

Associate Editor:
James Catto

Keywords:
Prostate cancer
Active surveillance
Risk prediction
Precision medicine

Abstract

Background: Active surveillance (AS) is an alternative to curative intervention, but overtreatment persists. Imperfect alignment of prostate biopsy and Gleason score after radical prostatectomy (RP) may be a contributing factor.
Objective: To develop a statistical model that predicts the post-RP Gleason score (pathologic Gleason score [PGS]) using clinical observations made in the course of AS.
Design, setting, and participants: Repeated prostate-specific antigen measurements and biopsy Gleason scores from 964 very low-risk patients in the Johns Hopkins Active Surveillance cohort were used in the analysis. PGS observations from 191 patients who underwent RP were also included.
Outcome measurements and statistical analysis: A Bayesian joint model based on accumulated clinical data was used to predict PGS in these categories: 6 (grade group 1), 3 + 4 (grade group 2), 4 + 3 (grade group 3), and 8–10 (grade groups 4 and 5). The area under the receiver operating characteristic curve (AUC) and calibration of predictions was assessed in patients with post-RP Gleason score observations.
Results and limitations: The estimated probability of harboring a PGS >6 was <20% for most patients who had not experienced grade reclassification or elected surgery. Among patients with post-RP Gleason score observations, the AUC for predictions of PGS >6 was 0.74 (95% confidence interval, 0.66–0.81), and the mean absolute error was 0.022.
Conclusions: Although the model requires external validation prior to adoption, PGS predictions can be used in AS to inform decisions regarding follow-up biopsies and remaining on AS. Predictions can be updated as additional data are observed. The joint modeling framework also accommodates novel biomarkers as they are identified and measured on AS patients.
Patient summary: Measurements taken in the course of active surveillance can be used to accurately predict patients' underlying prostate cancer status. Predictions can be communicated to patients via a decision support tool and used to guide clinical decision making and reduce patient anxiety.

* Corresponding author. Department of Urology, Marburg 150, Johns Hopkins School of Medicine, 600 N. Wolfe Street, Baltimore, MD 21218, USA. Tel. +1 410 955 0351; Fax: +1 410 955 0833.
E-mail address: hcarter@jhmi.edu (H.B. Carter).

1. Introduction

Active surveillance (AS) is a recognized management option for reducing overtreatment of favorable risk prostate cancer (PCa) [1,2] that is being increasingly utilized in practice [3]. However, given the prevalence of favorable risk disease diagnosed with screening based on prostate-specific antigen (PSA), substantial overtreatment remains. At Johns

http://dx.doi.org/10.1016/j.eururo.2016.08.005
0302-2838/

Hopkins Hospital, 45% of patients eligible for AS upon diagnosis elect immediate intervention, and 8–9% of patients who initially choose AS later exit in favor of treatment without experiencing disease reclassification [4].

Gleason score (GS) from prostate biopsy is currently the most reliable predictor of cancer-related outcomes in the absence of intervention [5]. Uncertainty regarding the underlying PCa state, that is, the GS at radical prostatectomy (RP) or pathologic Gleason score (PGS), may hinder both physicians and patients from accepting AS as a strategy. Although risk calculators developed for predicting biopsy GS in AS may help reduce the number of biopsies performed [6], there are currently no prediction tools for the actual, rather than biopsied, cancer state for patients participating in AS. Because biopsy GS is subject to measurement error, methods for predicting the true GS can improve clinical decision making.

Previous studies developed nomograms to predict the PGS for patients diagnosed with localized disease [7–9]. These predictions are based on diagnostic biopsy and PSA results and are intended to guide patients choosing between AS and treatment. For patients participating in AS, no comparable tools are available to quantify the accumulated evidence about PGS revealed in repeated PSA and biopsy tests. Development of such a tool requires statistical techniques that incorporate *longitudinal* clinical measurements, and deployment of the tool depends on integration with the electronic health record to allow for more detailed data input than accommodated by a nomogram.

Existing research examining posttreatment outcomes on patients who have left AS in favor of curative intervention is also limited in its relevance because those who choose treatment are not representative of the broader AS population [10]. Patients who choose treatment are more likely to do so based on biopsy findings, PSA kinetics, or personal preference. Prior studies are unable to fully capture the multitude of factors relating both to an individual's underlying cancer state (PGS) and decision to be treated, resulting in unobserved confounding. Therefore, it is necessary to include data on current patients to make accurate predictions.

We hypothesized that the PGS found if a patient underwent RP could be predicted using accumulated data (eg, repeated PSA measurements and biopsy results) from an ongoing AS program. The motivation for this work was twofold: to help patients better understand their risk and to inform decisions regarding follow-up biopsies and remaining on AS.

2. Materials and methods

2.1. Study cohort

The Johns Hopkins Active Surveillance (JHAS) study is an ongoing prospective cohort study of men with very low-risk and low-risk PCa diagnoses. This study is described in detail elsewhere [4]. Our analysis included 964 patients from the JHAS cohort who met the Epstein criteria for very low-risk PCa [11] and had at least two PSA measurements and one postdiagnosis biopsy as of January 1, 2016 (Table 1).

Table 1 – Patient characteristics at diagnosis

Characteristic	JHAS cohort, *n* (%)
Age at diagnosis, yr	
<50	9 (0.9)
50–59	156 (16.2)
60–69	617 (64.0)
70–79	179 (18.6)
>80	3 (0.3)
Year of diagnosis	
Before 2000	73 (7.6)
2000–2004	220 (22.8)
2005–2009	371 (38.5)
2010–2015	300 (31.1)
PSA, ng/ml	
0–2.5	143 (14.8)
2.5–4	166 (17.2)
4–6	435 (45.1)
6–10	188 (19.5)
>10	28 (2.9)
No. of positive cores, diagnostic biopsy	
1	739 (77)
2	223 (23)
Missing	2 (0.2)
Maximum cancer involvement, diagnostic biopsy, %	
1–9	579 (60)
10–19	215 (22)
20–29	165 (17)
Missing	5 (0.5)
Prostate volume	Median 50 (IQR: 37–70)

IQR = interquartile range; JHAS = Johns Hopkins Active Surveillance; PSA = prostate-specific antigen.

Among the patients included in our analysis, 195 patients experienced grade reclassification, 199 patients received RP, and 161 received another curative intervention (primarily radiation therapy) (Fig. 1). Patients who chose treatment in the absence of grade reclassification may have done so based on volume reclassification, PSA kinetics, or anxiety about continued participation in AS. We focus on grade reclassification here because of the association of grade and cancer-specific outcomes [12]. Per JHAS protocol, patients were not recommended for treatment on the basis of PSA kinetics.

As part of the AS regime, PSA was measured every 6–12 mo, and biopsies were performed annually, although some patients chose to delay the procedure. The median number of PSA observations, biopsy assessments of GS, and years of follow-up were 11 (interquartile range [IQR]: 6–16), 4 (IQR: 2–5), and 4.6 (IQR: 2.5–7.9), respectively (Supplementary Table 1). Overall, 85% of patients had an average of at least one PSA measurement per year, and 73% averaged at least one biopsy every 18 mo.

2.2. Statistical methods

The goal of the statistical model was to predict each patient's PGS (Fig. 2), defined as the GS determination that would be made if the entire prostate was surgically removed and analyzed (in four ordered categories: 6, or grade group 1; 3 + 4, or grade group 2; 4 + 3, or grade group 3; and 8–10, or grade groups 4 and 5) [13]. Postsurgery PGS observations available on patients who underwent them

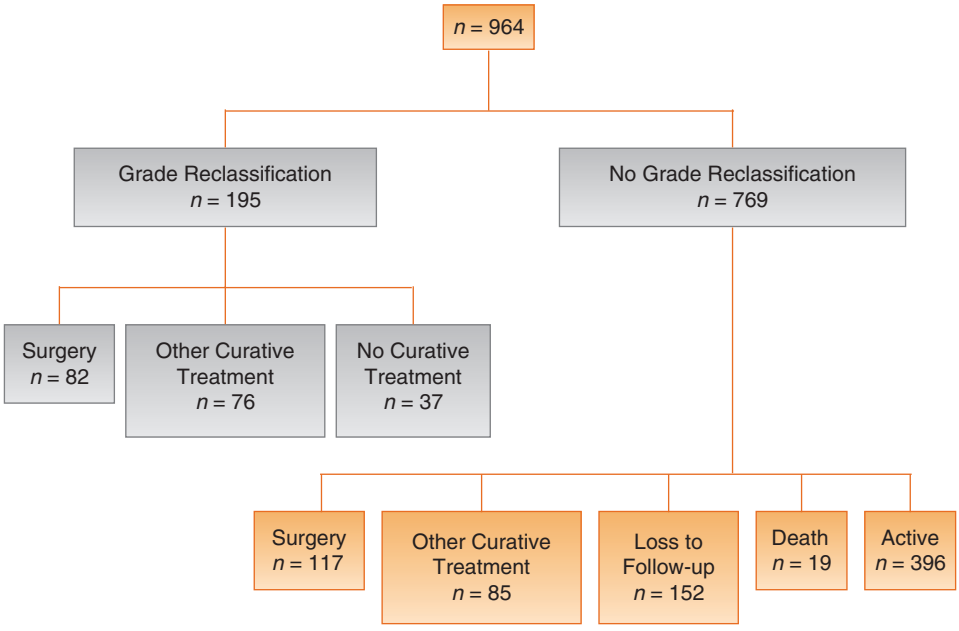

Fig. 1 – Consolidated Standards of Reporting Trials diagram of patient outcomes for the Johns Hopkins Active Surveillance cohort.

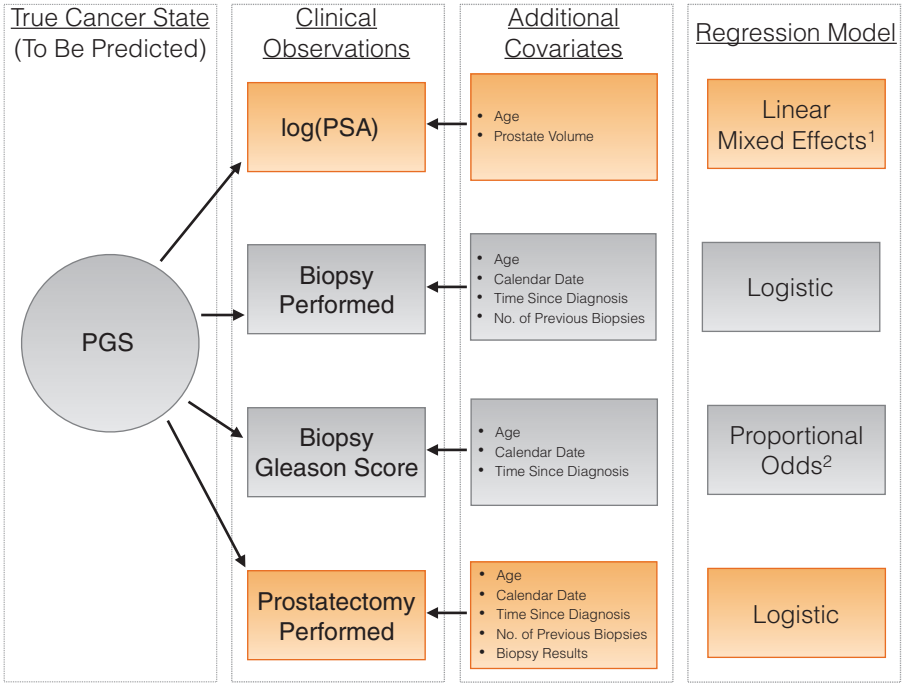

Fig. 2 – Diagram of joint model. Pathologic Gleason score (PGS) (column 1) causes four observed clinical outcomes (column 2): log(PSA), biopsy being performed, the resulting biopsy Gleason score, and prostatectomy being performed. Additional covariates (column 3) may also affect the clinical observations. Dependence of the clinical observations on PGS and additional covariates are described with the specified regression model (column 4). A linear mixed-effects model was used to model a patient's entire prostate-specific antigen history, and distinct regression models were used for each year a patient was eligible to receive a biopsy, to experience grade reclassification, and to elect prostatectomy. For a patient with three biopsies while on surveillance, for example, three proportional odds models were used to model the result of each. Biopsy results included in the logistic regression model for whether surgery was performed in a given year were grade reclassification on a previous biopsy, maximum number of positive cores, and percentage of cancer involvement on all previous biopsies. Variable selection, including the nonlinear transformation of variables, was performed using the Akaike Information Criteria [16,20,21].
PGS = pathologic Gleason score; PSA = prostate-specific antigen.
[1] From Laird and Ware [20].
[2] From Agresti [16].

Table 2 – Cross tabulation of postsurgery pathologic Gleason score and final biopsy

	Final biopsy Gleason score *n* (%)				Total
PGS	6	3 + 4	4 + 3	8–10	
6	75 (67)	19 (40)	1 (4)	2 (25)	97
3 + 4	27 (24)	21 (45)	10 (42)	1 (13)	59
4 + 3	7 (6)	7 (15)	11 (46)	2 (25)	27
8–10	3 (3)	0 (0)	2 (8)	3 (37)	8
Total	112 (100)	47 (100)	24 (100)	8 (100)	191

PGS = pathologic Gleason score.

both for cause (biopsy grade reclassification) and elective RP (not for cause) were included in the analysis. RP cases provided information about biopsy GS misclassification rates (Table 2). Eight patients who underwent RP did not have a postsurgery PGS available.

To predict PGS for patients who did not have surgery, we used all available PSA and biopsy data. To do so requires a joint modeling approach [14]; that is, we did not fit a single multivariate regression model with PGS as the outcome and summary measures of PSA and biopsy results (such as PSA kinetics and number of biopsies without reclassification) as the predictors. This would degrade the information available in the full PSA and biopsy history. Instead, we fit distinct regression models for PSA and biopsy GS with PGS as a predictor in each (Fig. 2); PSA and biopsy models are "joined" by sharing PGS as a predictor. This approach quantifies the effect of the underlying cancer state on PSA and biopsy results.

Using a joint modeling approach also allows us to accommodate missing clinical observations. Because some patients did not receive a biopsy every year (resulting in missing biopsy GS outcomes), we used a logistic regression model for the probability that a patient has a biopsy in a particular year. We also fit a logistic regression model for the probability that a patient elected RP in a given year to account for selection bias of those with PGS observations. To adjust for possible unobserved confounding between the outcome (biopsy or surgery performed) and underlying cancer state, PGS is included as a predictor in both models.

Although the outcome of interest was only observed on patients who underwent RP, data from all patients were used to estimate the prediction model for two reasons. First, patients who chose surgery were a small nonrandom subset of the JHAS patients. Second, clinical data on patients without surgery contained important information about how PSA and biopsy results vary depending on time and patient characteristics. Bayesian statistical methods were used to simultaneously estimate the relationship between the underlying cancer state, clinical outcomes, and covariates, and predict PGS for the 773 patients without postsurgery observations. Specifically, Bayes theorem calculates the probability that a patient's PGS is 6, 3 + 4, 4 + 3, or 8–10 based on the likelihood of observing his clinical data under each possible PGS and the estimated prevalence of each PGS in the JHAS population.

Model estimation was performed with JAGS (https://sourceforge.net/projects/mcmc-jags/files/Manuals/4.x/jags_user_manual.pdf) [15] via the R2JAGS package [16]. Quantile-based point estimates and 95% confidence intervals (CIs) were calculated for all model coefficients. All analyses were conducted using the open source statistical computing software R (R Foundation for Statistical Computing, Vienna, Austria; http://www.R-project.org). Additional analysis details are forthcoming (unpubl. data).

2.3. *Model assessment*

Accuracy of PGS predictions was assessed among patients who underwent RP. So called out-of-sample predictions, that is, predictions that did not rely on the observed PGS, were obtained for patients with postsurgery GS observations by, in turn, withholding each patient's full prostate GS observation to train the prediction model and estimate a probability distribution of that patient's PGS. The accuracy of out-of-sample predictions was then evaluated by examining calibration plots [17] and receiver operating characteristic curves [18] including calculating the area under the curve (AUC) and accompanying bootstrapped intervals. Calibration of predictions was also assessed for outcomes observed on all patients: the choice to receive a biopsy and, if so, the resulting biopsy GS and the choice to undergo RP for each year under observation.

3. Results

The predicted probability of harboring a PGS >6 was <20% for most of the JHAS patients (61%) who had not experienced grade reclassification or elected RP (Fig. 3). As would be expected, our model estimated that having a PGS >6 was positively associated with grade reclassification on biopsy (odds ratio [OR]: 4.2; 95% CI, 2.2–8.2). PGS >6 was also positively associated with the decision to undergo RP (OR: 2.0; 95% CI, 0.98–2.8 for patients without grade reclassification and OR: 14; 95% CI, 3.2–47 for patients with grade reclassification). That is, patients with higher grade disease were more likely to elect surgery, particularly after grade reclassification. The model also estimated higher average intercepts and slopes in patient-level linear regressions of log-PSA for patients with PGS >6 (mean intercept: 1.33 [slope: 1.27–1.38] and 0.24 [slope: 0.19–0.28] for PGS 6; 1.6 [slope: 1.5–1.7] and 0.56 [slope: 0.41–0.58] for PGS >6).

The AUC for prediction of PGS ≥3 + 4 (grade group 2 or above), GS ≥4 + 3 (grade group 3 or above), or GS score ≥8 (grade group 4 or 5) was 0.74 (95% CI, 0.66–0.81), 0.75 (0.66–0.84), and 0.63 (0.50–0.75), respectively (Fig. 4). Calibration plots (Fig. 4) indicate that PGS predictions closely approximate the observed rates of postsurgery PGS (in ranges with sufficient numbers of observations to evaluate calibration). Mean absolute errors of cumulative predictions of PGS ≥3 + 4, ≥4 + 3, and ≥8 were 0.022, 0.014, and 0.003, respectively (Supplementary Table 2). Calibration plots for the choice to receive a biopsy, biopsy results, and electing RP (Supplementary Fig. 1) also indicate model

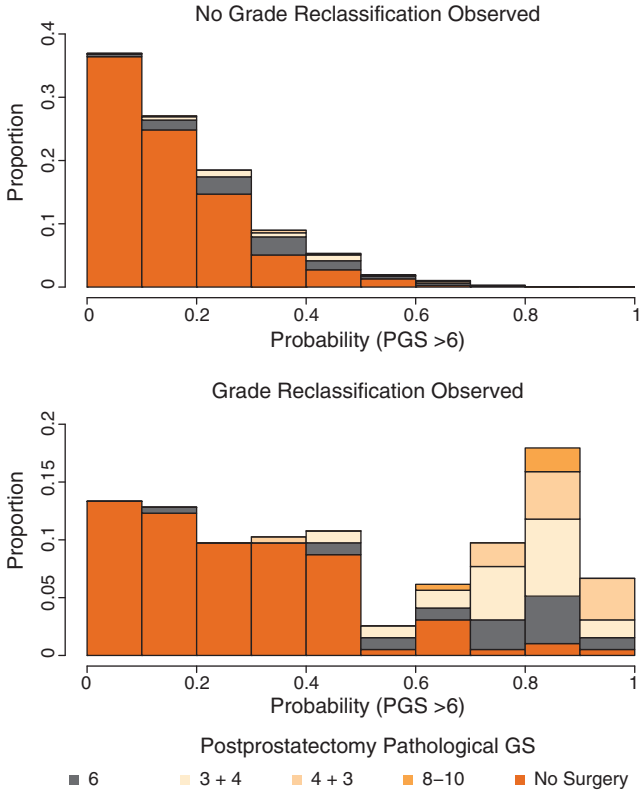

Fig. 3 – Histogram of the predicted probabilities of a pathologic Gleason score (PGS) >6 for Johns Hopkins Active Surveillance patients without (top panel) and with (lower panel) grade reclassification. Colors indicate patients with postsurgery Gleason score observations available and, if so, the observed PGS.
PGS = pathologic Gleason score; RP = radical prostatectomy.

predictions of these outcomes (Supplementary Table 3) were consistent with the observed event rates.

We have designed an example of an interactive web application that displays patient data alongside model predictions of PGS, likely PSA trajectory, and risk of grade reclassification (Supplementary Fig. 2) (https://rycoley.shinyapps.io/dynamic-prostate-surveillance). This app demonstrates how data visualizations can be used throughout a patient's participation in AS to inform clinical decision making. As additional data are observed on a patient, predictions can be updated and precision will improve (Supplementary Fig. 3).

4. Discussion

We have developed a model that can be used in the AS setting to help inform decisions regarding the need for follow-up biopsies and remaining on AS. The model is unique in that it provides an estimated probability of the PGS that would be observed if the prostate was removed and carefully examined, in addition to projections regarding PSA and biopsy results (Supplementary Fig. 3). Prior predictive models have focused on predicting biopsy results [6] rather than the PGS. Models for predicting PGS have been limited to newly diagnosed patients and/or patients who have left AS in favor of treatment [7–9].

Importantly, when our prediction model is integrated into the electronic health record, detailed patient data will be entered automatically—manual entry will not be necessary—and model updates will be ongoing (Supplementary Fig. 3). In clinical practice, visual representations of data and predictions should be accompanied by a written explanation of the model output and implications of predictions, including treatment options and likelihood of cure.

The clinical implications of our findings are as follows. First, men who have a low probability of harboring higher grade cancers could postpone or avoid frequent prostate biopsies during AS. Second, the model predictions can be used during shared decisions that address the safety of embarking and remaining on AS in the setting of favorable disease characteristics for those men who fear an adverse outcome. By quantifying accumulated evidence in biopsy and PSA results, this model is able to provide patients with a more precise prediction of risk than the dichotomous result returned by prostate biopsy (grade reclassification or not). This type of approach could improve patient satisfaction with care and reduce anxiety during AS. The availability of this tool while on AS may also encourage more patients to consider this option.

We were able to construct the current model by using the cumulative results from repeated PSA measurements and prostate biopsies in a prospective longitudinal AS program in which a substantial proportion of men underwent RP informing the "true" prostate state (PGS). These prior measures are important because we have previously shown that future biopsy results are conditional upon prior biopsy results [19]. The model design is unique in that we accounted not just for the effect of PGS on biopsy GS and surgical findings, but also for the association between PGS and whether or not biopsy and surgery observations occurred (Fig. 2). Further, model accuracy improves as additional measurements are made and estimates are updated (Supplementary Fig. 3).

The AUC for prediction of the PGS indicates that our model is comparable with others for predicting the underlying cancer state that used similar clinical measures in men diagnosed with localized disease. For example, Steyerberg et al [8] found that a combination of PSA, prostate volume, biopsy GS, and the percentage of positive cores on biopsy at diagnosis had an AUC of 0.72 (95% CI, 0.66–0.78) for predicting the presence of indolent disease.

Our study has limitations. First, our model was generated from a select group of patients who met our enrollment criteria for AS, and they were not racially diverse. However, unlike models that require application of coefficients estimated in one cohort and then applied to another cohort, our model can easily be refit for different populations that have differing risk profiles. Second, calibration of the model was less certain when limited data were available (GS 8–10) (Fig. 4). Third, our model did not include novel biomarkers or imaging data that could improve predictions of PGS. We would anticipate that predictive accuracy of the proposed model will improve when imaging and other marker results are measured and included in the model. The joint modeling

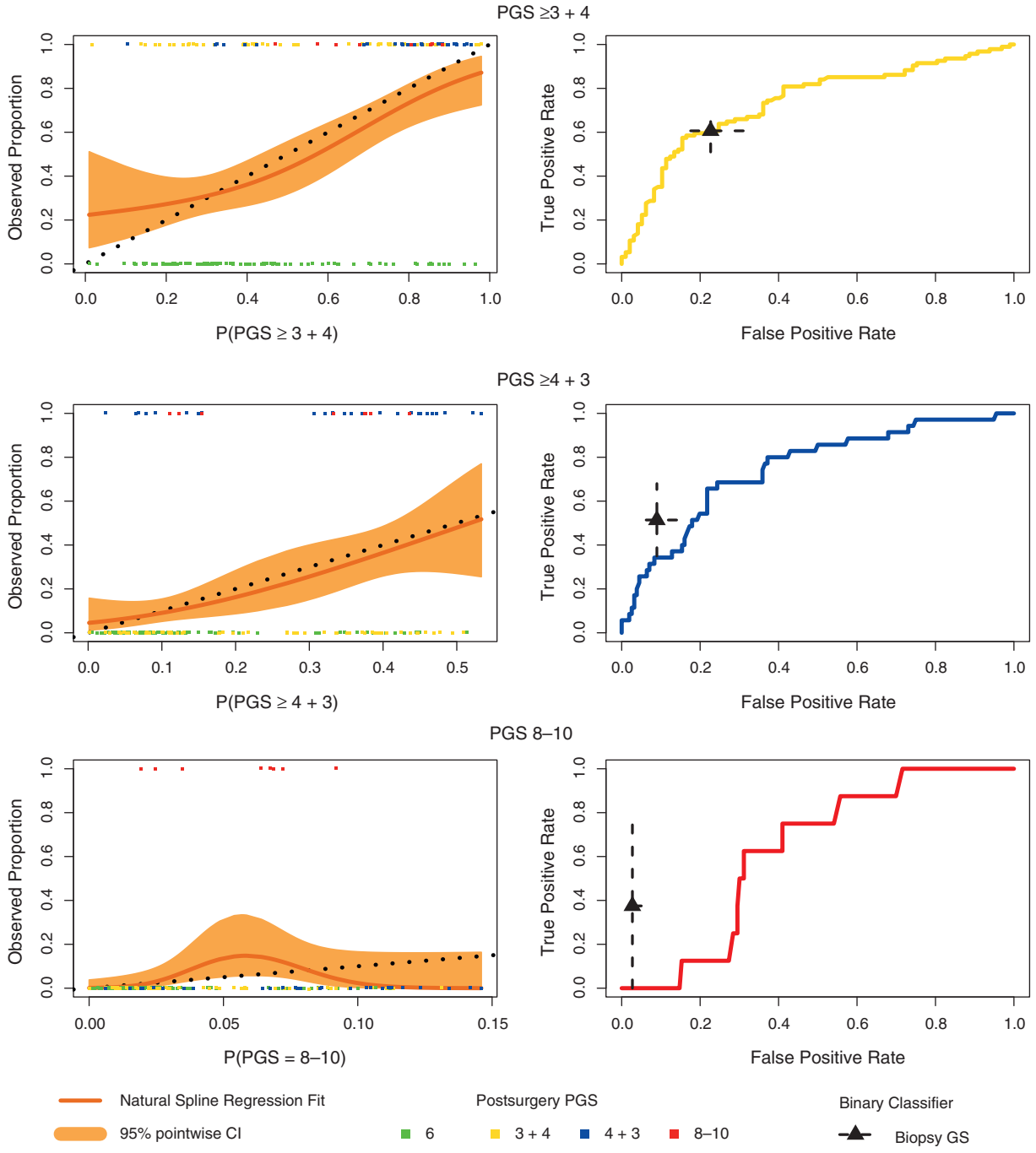

Fig. 4 – Predictive accuracy (P) for pathologic Gleason score (PGS) or "true" Gleason score (GS) in patients with postsurgery GS observations available. Calibration plots (left column) demonstrate that out-of-sample predictions (ie, predictions that did not rely on the observed PGS) of PGS approximate the observed rates of postsurgery GS ≥3 + 4 and ≥4 + 3 (top and middle rows, respectively) in patients who underwent prostatectomy; fitted regression lines closely track the line of equality (dotted line) in both plots in ranges with sufficient data. Specifically, few patients have predicted probabilities of PGS >6 below 20%, so the calibration interval is very wide in this range. Calibration is also more variable for postsurgery GS 8–10 (bottom row) because only eight patients had postsurgery PGS determinations of 8–10, but the confidence band includes the line of equality at all predicted values. Colored squares at the bottom and top of each plot indicate the observed PGS (vertical axis) and predicted PGS (horizontal axis). Receiving operating characteristic curves (right column) show the discriminatory ability of out-of-sample PGS predictions. Classification accuracy is similar for cumulative predictions of PGS ≥3 + 4 and GS ≥4 + 3; the calculated area under the curve (AUC) and 95% bootstrapped confidence intervals (CIs) equal 0.74 (95% CI, 0.66–0.81) and 0.75 (95% CI, 0.66–0.84), respectively. Classification accuracy is considerably lower for GS 8–10 due to a small number of events; AUC equals 0.63 (95% CI, 0.50–0.75). The triangle on each plot indicates the true-positive rate (TPR) and false-positive rate (FPR) of using observed biopsy GS as a predictor of PGS (95% CIs are represented by dashed lines). A biopsy GS ≥3 + 4 has a TPR of 61% and an FPR of 23%. At TPR of 61%, a binary classifier based on PGS predictions from the proposed model has a similar FPR. Biopsy GS >3 + 4 was a less sensitive predictor of full prostate GS, with a TPR of 51% for biopsy GS ≥4 + 3 and 38% for biopsy GS 8–10, but had higher specificity at those TPRs than PGS predictions from the proposed model. AUC = area under the curve; CI = confidence interval; FPR = false-positive rate; GS = Gleason score; P = predictive accuracy; PGS = pathologic Gleason score; TPR = true positive rate.

framework can accommodate new clinical measurements by adding an appropriate regression model for that outcome.

5. Conclusions

We have proposed a statistical model that predicts the PGS for AS patients based on repeated PSA and biopsy measurements and will require external validation prior to adoption. The availability of PGS predictions may increase adoption of AS and improve patients' understanding of their likely cancer outcomes during AS. These predictions may also inform conversations between patients and physicians regarding continued participation in AS and the frequency of follow-up biopsies. Novel markers and imaging measures can easily be included in this type of model to improve predictions in the future.

Author contributions: H. Ballentine Carter 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: Coley, Mamawala, Pienta, Zeger, Carter.
Acquisition of data: Carter, Coley, Mamawala.
Analysis and interpretation of data: Coley, Mamawala, Zeger, Pienta, Carter.
Drafting of the manuscript: Coley, Zeger, Carter, Mamawala.
Critical revision of the manuscript for important intellectual content: Coley, Zeger, Carter.
Statistical analysis: Coley, Zeger, Mamawala.
Obtaining funding: Coley, Zeger, Carter.
Administrative, technical, or material support: Carter, Pienta.
Supervision: Carter, Zeger.
Other (specify): None.

Financial disclosures: H. Ballentine Carter certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/ affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

Funding/Support and role of the sponsor: R. Yates Coley and Scott L. Zeger were funded for this work by the Patrick C Walsh Prostate Cancer Research Fund and a Patient-Centered Outcomes Research Institute (PCORI) Award (ME-1408–20318). The statements presented in this article are solely the responsibility of the authors and do not necessarily represent the views of PCORI, its board of governors, or methodology committee.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eururo.2016.08.005.

References

[1] Dall'Era MA, Albertsen PC, Bangma C, et al. Active surveillance for prostate cancer: a systematic review of the literature. Eur Urol 2012;62:976–83.
[2] Chen RC, Rumble RB, Loblaw DA, et al. Active surveillance for the management of localized prostate cancer (Cancer Care Ontario Guideline): American Society of Clinical Oncology clinical practice guideline endorsement. J Clin Oncol 2016;34:2182–90.
[3] Womble PR, Montie JE, Ye Z, et al. Contemporary use of initial active surveillance among men in Michigan with low-risk prostate cancer. Eur Urol 2015;67:44–50.
[4] Tosoian JJ, Mamawala M, Epstein JI, et al. Intermediate and longer-term outcomes from a prospective active-surveillance program for favorable-risk prostate cancer. J Clin Oncol 2015;33:3379–85.
[5] Popiolek M, Rider JR, Andren O, et al. Natural history of early, localized prostate cancer: a final report from three decades of follow-up. Eur Urol 2013;63:428–35.
[6] Ankerst DP, Xia J, Thompson Jr IM, et al. Precision medicine in active surveillance for prostate cancer: development of the Canary-Early Detection Research Network Active Surveillance Biopsy Risk Calculator. Eur Urol 2015;68:1083–8.
[7] Kattan MW, Eastham JA, Wheeler TM, et al. Counseling men with prostate cancer: a nomogram for predicting the presence of small, moderately differentiated, confined tumors. J Urol 2003;170:1792–7.
[8] Steyerberg EW, Roobol MJ, Kattan MW, van der Kwast TH, de Koning HJ, Schroder FH. Prediction of indolent prostate cancer: validation and updating of a prognostic nomogram. J Urol 2007;177:107–12, discussion 112.
[9] Nakanishi H, Wang X, Ochiai A, et al. A nomogram for predicting low-volume/low-grade prostate cancer: a tool in selecting patients for active surveillance. Cancer 2007;110:2441–7.
[10] Tosoian JJ, Sundi D, Trock BJ, et al. Pathologic outcomes in favorable-risk prostate cancer: comparative analysis of men electing active surveillance and immediate surgery. Eur Urol 2016;69:576–81.
[11] Epstein JI, Walsh PC, Carmichael M, Brendler CB. Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. JAMA 1994;271:368–74.
[12] Eggener SE, Scardino PT, Walsh PC, et al. Predicting 15-year prostate cancer specific mortality after radical prostatectomy. J Urol 2011;185:869–75.
[13] Pierorazio PM, Walsh PC, Partin AW, Epstein JI. Prognostic Gleason grade grouping: data based on the modified Gleason scoring system. BJU Int 2013;111:753–60.
[14] Song X, Davidian M, Tsiatis AA. A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. Biometrics 2002;58:742–53.
[15] Plummer M. JAGS version 4.0.0 user manual. SourceForge Web site. https://sourceforge.net/projects/mcmc-jags/files/Manuals/4.x/jags_user_manual.pdf.
[16] Agresti A. Categorical Data Analysis. ed 3. Hoboken, NJ: Wiley; 2012.
[17] Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38.
[18] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
[19] Alam R, Carter HB, Landis P, Epstein JI, Mamawala M. Conditional probability of reclassification in an active surveillance program for prostate cancer. J Urol 2015;193:1950–5.
[20] Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982;38:963–74.
[21] Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. Selected Papers of Hirotugu Akaike. Springer Series in Statistics: Perspectives in Statistics. New York, NY: Springer; 1998. p. 199–213.