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 | 579 (60) |
| 2 | 223 (23) |
| Missing | 2 (0.2) |
| Maximum cancer involvement, diagnostic biopsy, % | |
| 1–9 | 470 (49) |
| 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

Please cite this article in press as: Coley RY, et al. Prediction of the Pathologic Gleason Score to Inform a Personalized Management Program for Prostate Cancer. Eur Urol (2016), http://dx.doi.org/10.1016/j.eururo.2016.08.005