AWARD NUMBER: W81XWH-14-1-0595

TITLE: Biomarkers for Early Detection of Clinically Relevant Prostate Cancer: A Multi-Institutional Validation Trial

PRINCIPAL INVESTIGATOR: Daniel Lin, MD

CONTRACTING ORGANIZATION: Fred Hutchinson Cancer Research Center
Seattle, WA 98109

REPORT DATE: Dec 2018

TYPE OF REPORT: FINAL REPORT

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
Dec 2018	Final	30 Sep 2014 - 29 Sep 2018

4. TITLE AND SUBTITLE
Biomarkers for Early Detection of Clinically Relevant Prostate Cancer:
A Multi-Institutional Validation Trial

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-14-1-0595

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Daniel Lin, MD

E-Mail: DLIN@UW.EDU

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Fred Hutchinson Cancer Research Center
1100 Fairview Ave N J6-300
Seattle, WA 98109-4433

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
For men diagnosed with early stage prostate cancer a critical need exists for molecular assays that accurately distinguish aggressive prostate cancer from those cancers that will not cause harm if left untreated. In this project, we assessed three different panels of established molecular biomarkers for their ability to distinguish aggressive cancers from indolent cancers. We established agreements with three commercial companies to analyze their biomarker platforms in our multi-center, prospectively accrued prostate cancer active surveillance cohort – the Canary Prostate Active Surveillance Study (PASS). We evaluated three biomarker panels in tissue (GHI's OncotypeDX), blood (OPKO's 4K Score), and urine samples (Hologic/GenProbe's PCA3) with well annotated clinical and pathologic data collected as part of PASS. We conducted rigorous statistical evaluation to demonstrate the utility and performance of biomarkers in clinical practice to predict aggressive disease. The accuracy of each biomarker for predicting short- and long-term progression has been characterized with time dependent receiver operating characteristic curves. The successful clinical validation of biomarkers that offer substantially improved predictive and prognostic accuracy should bring extraordinary potential to improve the care of prostate cancer patients.

15. SUBJECT TERMS
Prostate cancer, active surveillance, biomarkers, validation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	99	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

1. INTRODUCTION

Although prostate-specific antigen (PSA) testing and the resulting treatment of prostate cancer (PCa) is likely responsible for some of the 44% decrease in prostate cancer mortality witnessed in the United States since 1992, the detection of low risk tumors has increased. The majority of prostate cancers currently diagnosed are low risk tumors for which there is substantial evidence that the cancer will not cause harm if left untreated. However, enough uncertainty remains in accurately identifying which tumors will not cause harm to a patient that many low risk cancers are still treated, resulting in so-called overtreatment. To reduce this overtreatment, while still diagnosing aggressive high risk tumors early enough that they can be successfully treated, there is a critical need for molecular assays that accurately distinguish more aggressive disease from cancers that will not cause harm. The goal of this project is to perform rigorous clinical validation of established biomarkers in order to improve the accuracy of risk assessment and distinguish aggressive from indolent disease in men with apparently low-risk disease by standard clinical variables. We are evaluating multiple established and analytically validated quantitative molecular biomarkers to predict PCa progression in a multi-center active surveillance cohort with high-quality biospecimens. We aim to unlink the diagnosis of PCa with immediate treatment, thus addressing the overtreatment issue and economic, physical, and emotional burdens of PCa diagnoses. The results have promise to change the standard of care in the treatment of the majority of newly diagnosed PCa with near term impact due to the availability of the biomarkers and execution in an established, prospective cohort of men undergoing AS.

2. KEYWORDS

Prostate cancer; active surveillance; progression; aggressive disease; central pathology review; biomarkers; prediction models; PCA3; TMPRSS2:ERG; kallikreins; 4Kscore; OncotypeDX;

3. ACCOMPLISHMENTS

What were the major goals and objectives of the project?

We hypothesized that biomarkers of disease aggressiveness and prognosis can be interrogated in low risk prostate cancer (PCa) and that these biomarkers will better detect clinically relevant PCa in asymptomatic patients, thus distinguishing aggressive from indolent disease and immediately impacting both the initial choice of therapy and decision-making during AS. The objective of the study was to utilize analytically validated assays that take into account tumor heterogeneity to measure biomarkers in specimens that were collected in a non-invasive manner.

The major goals and milestones of the project, as stated in the scope of work, are outlined below. The activities and accomplishments are described more fully in the text below the outline.

1. Collection of specimens and clinical data. (Coordinated by FHCRC)
   Milestone 1. Completion of a minimum of three years of follow-up with high-quality data and specimen collection. Due: 12/30/2016 **COMPLETED**

2. Analysis of scientific aim 1: Validate a panel of tissue-based biomarkers to determine the presence of or progression to aggressive disease. (Lead site: FHCRC)
   Milestone 2. Execute collaboration agreement with GHI. Due 12/30/2014 **COMPLETED**.
   Milestone 3. Tissue blocks identified for analysis. Due: 12/30/2015 **COMPLETED**
   Milestone 4. Oncotype DX validation complete in PASS cohort. Due 12/30/2016 **COMPLETED**
   Milestone 5. Manuscript submission of Oncotype DX validation. Due 9/30/2017 *in process*

3. Analysis of scientific aim 2: Evaluate a panel of four-kallikrein plasma-based markers to determine the presence of or progression to clinically relevant prostate cancer. (Lead site: FHCRC)
   Milestone 6. Execute collaboration agreement with OPKO. Due 3/30/2015 **COMPLETED**.
   Milestone 7. Plasma samples identified for analysis. Due 12/30/2015 **COMPLETED**
   Milestone 8. OPKO 4KScore validation complete in PASS cohort. Due 9/30/2016 **COMPLETED**
   Milestone 9. Manuscript submission of 4KScore validation. Due 9/30/2017 **COMPLETED**

4. Analysis of scientific Aim 3: Confirm the ability of PCA3 mRNA concentrations in urine, alone or in combination with TMPRSS2:ERG mRNA. (Lead site: FHCRC)
   Milestone 10. Urine specimens identified for analysis. Due 12/30/2014 **COMPLETED**
   Milestone 11. PCA3 and TMPRSS2:ERG validation complete in PASS cohort. Due 12/30/2015 **COMPLETED**
   Milestone 12. Manuscript submission of PCA3 and TMPRSS2:ERG validation. Due 9/30/2017 **COMPLETED**

5. Central pathology review of PASS biopsy and RP slides. (Lead site: CCF)
   Milestone 13. Completion of Central Pathology Review for biopsy-driven endpoints. Due: 12/30/2016 **COMPLETED**

6. Translation of biomarkers into clinical practice. (Lead sites: FHCRC and CCF)

Milestone 14. Construction of integrated model of biomarkers for the prediction of progression in the PASS cohort. Due 9/30/2017 *in process*

Milestone 15. Manuscript submission of integrated model for prediction of progression. Due 9/30/2017 *in process*

What was accomplished under these goals?

Task 1: Collection of specimens and clinical data. *(Coordinated by FHCRC)*

Collection of follow-up data and longitudinal specimens in the PASS cohort is essential to adequately power our funded biomarker analyses. As of 13 Dec 2018, PASS has enrolled 1,811 eligible patients at ten clinical sites. This number includes enrollment from the new PASS site, Emory University, which was added April 2017 outside of the scope of this proposal. We have been highly successful in following participants to obtain outcomes measures, with a median cohort follow-up of over 5.1 years (25th and 75th percentiles: 2.2, 7.6 years). Currently, all of the first 1000 participants enrolled in PASS, which are the subject of this specific research proposal, have at least three years of follow-up. In the past year, we have conducted site visits to three clinical sites (Beth Israel Deaconess Medical Center in Aug 2018, University of California San Francisco in July 2018, University of Michigan in November 2017) to ensure adherence to the protocol. The coordinating center based at the Fred Hutchinson Cancer Center provided data QA and QC.

Task 2: Validate a panel of tissue-based biomarkers to determine the presence of or progression to aggressive disease. *(Lead site: FHCRC)*

The biopsy based 17-gene Genomic Prostate Score (GPS) test has been shown to predict adverse surgical pathology (AP) and recurrence in men diagnosed with low- and intermediate-risk prostate cancer (PCa) who had immediate surgery. However, studies of the predictive value of the GPS test in men initially managed with active surveillance have been limited. We sought to confirm that GPS was associated with time to adverse pathology in men who had surgery after a period of active surveillance. We also evaluated if GPS was associated with time to upgrading at surveillance biopsy. Importantly, we assessed the association of GPS with adverse outcome when adjusted for commonly available clinical variables. We used a retrospective-prospective study design to evaluate GPS at initial diagnosis.

Methods

Fixed Paraffin-Embedded (FPE) tissue blocks from the initial diagnostic biopsies of PASS participants who consented to tissue use at eight sites in PASS (Beth Israel Deaconess Medical Center, Eastern Virginia Medical School, Stanford University, University of British Columbia, University of Michigan, University of Texas Health Sciences Center San Antonio, University of Washington, Veterans Affairs Puget Sound Health Care Systems) were collected and sent to Genomic Health, Inc. (GHI). To obtain the tissue, PASS participants were reconsented for use of the tissue left over from initial diagnosis and tissue blocks collected from pathology departments both at the PASS sites and at

local urology clinics. All blocks were sent to the PASS Central Biospecimen Repository where they underwent QC and were labeled with unique PASS ID numbers prior to sending to GHI.

GHI sectioned all available tissue blocks to make 8 unstained sections. In a few cases, the sectioning was performed at the PASS site's local pathology department and unstained slides were delivered to GHI. The top and bottom section were stained with H&E. Dr. McKenney at Cleveland Clinic reviewed all H&E slides provided to him by GHI and recorded the tumor extent and Gleason Score. For this analysis, H&E stained slides from RPs that occurred after a period of surveillance and recuts of the diagnostic biopsy tissue were centrally reviewed by one urologic pathologist blinded to clinical outcomes and using the 2005 International Society of Urologic Pathology Consensus guidelines. Local pathology data were used for all surveillance biopsies.

GHI manually microdissected the tumor tissue and isolated RNA for Oncotype DX, an analytically validated 17-gene quantitative reverse transcription-polymerase chain reaction assay, resulting in a GPS score scaled from 0-100.

Statistical methods

The primary outcome was time to adverse surgical pathology (AP), defined as Gleason Grade Group (GG) ≥3, any Gleason pattern 5, stage ≥pT3a, or N1 in men who received surgery after a period of surveillance. The secondary outcome was time to upgrading on surveillance biopsy. Follow-up clinical data were collected through February, 2018. The range of possible values for the GPS is 0 to 100. The hazard ratios for continuous GPS are reported per 20-unit increase. Covariates considered in multivariable modeling were: age (continuous or > vs ≤ 65), race (non-white vs white), diagnostic Gleason (3+3 or 3+4), ratio of positive/total biopsy cores), log(PSA), log(prostate size), PSA density (< vs ≥ 0.15 or log2), T-stage, BMI (kg/m2), family history of PCa (yes/no), and year of diagnosis. Tests for proportionality confirmed that the proportional hazards assumptions were valid.

GPS and time to AP

The association between GPS and time to finding AP was assessed in the 101 participants who had RP after a period of AS using both univariable and multivariable parametric accelerated failure time models based on the Weibull distribution. Inverse probability of censoring weighted (IPCW) methodology was used to adjust for possible informative censoring.

GPS and time to biopsy upgrade

The association between GPS at diagnosis and the time to biopsy upgrade was modeled using univariable and multivariable Cox proportional hazards (PH) models. Participants without reclassification were censored at date of last study contact, treatment, or 2 years after their last biopsy, whichever came first. A two-sided p-value <0.05 was considered significant for all analyses, which were performed using SAS or R version 3.3.0.

Results

Among 1041 men using active surveillance and participating in Canary PASS, 634 (61%) had available tissue left over from their diagnostic biopsy (Figure 1). Of these, GPS was obtained for 432 (68%); 7 (1%) did not meet inclusion criteria; 174 (27%) had insufficient tumor tissue; 10 (2%) had Gleason GG ≥3; 11 (2%) had insufficient RNA quality. Of the 432 with GPS, 106 (25%) had a radical

prostatectomy (RP) after a period of surveillance. Radical prostatectomy was performed after biopsy upgrade for 77 men (73%) had RP, while 29 men (27%) had surgery with no biopsy upgrade. After excluding 5 participants with no RP slides available for central pathology review, 101 participants were available for evaluation of the AP endpoint.

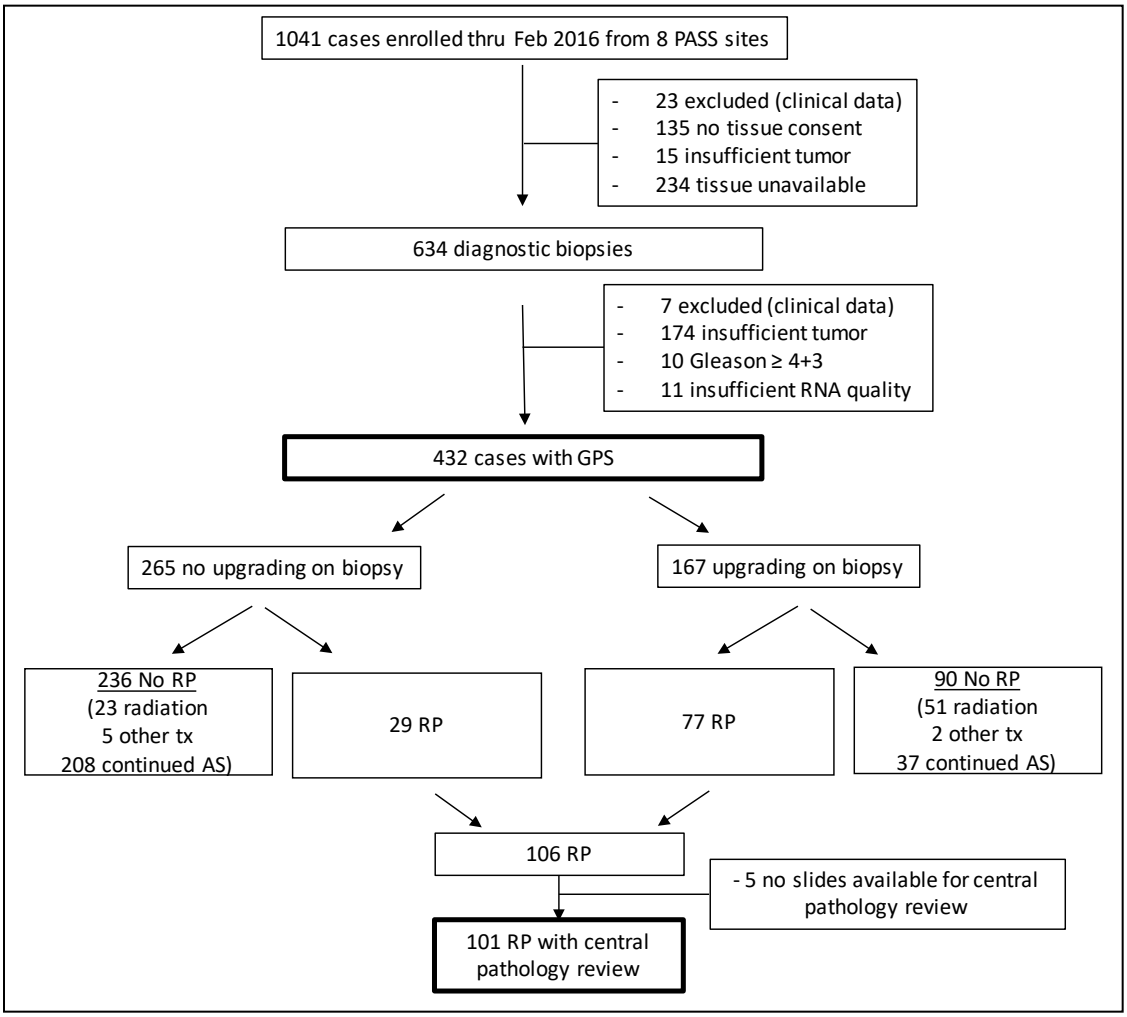


Figure 1. Consort diagram detailing study cohort.

Participant characteristics at diagnosis were similar to the full cohort for the 634 participants with available tissue blocks, and for the 432 evaluable participants (Table 1). In the 432 participants with GPS, median age was 63 years, PSA was 4.8 ng/ml, and PSA density was 0.11 ng/cm3. By local clinical pathology, 395 (91%) were Gleason GG 1; upon central pathology review of the sections used for GPS, 374 (87%) were Gleason GG 1. Median follow-up was 4.6 (IQR: 2.9-6.2) years in participants with no reclassification at biopsy. By local pathology read, which was used for clinical management, 167 (39%) experienced upgrading at a surveillance biopsy, 51 (12%) to Gleason GG ≥ 3. Median time from diagnosis to surgery was 2.1 (IQR: 1.3-4.3) years, and 52 men (51%) had AP at surgery. Median

GPS score in the full cohort (n = 432) was 21 (IQR: 15.4 – 27.3; range 0-67), and in the RP cohort (n = 101) was 20.5 (IQR: 14.6-27.3; range 6-67).

Table 1. Participant characteristics at diagnosis, recorded either as median (IQR) or n (%).

Characteristic	Enrolled thru Feb. 2016 (n=1041)	Available FFPE blocks (n=634)	with GPS (n=432)	RP (n=101)
Age (years)	63 (58, 67)	63 (59, 67)	63 (59, 67)	62 (43, 76)
Race		1		
Asian	25 (2%)	8 (3%)	12 (3%)	5 (5%)
Black	74 (7%)	36 (6%)	24 (6%)	5 (5%)
Other	11 (1%)	8 (1%)	6 (1%)	0
White	931 (89%)	572 (90%)	390 (90%)	91 (90%)
Hispanic ethnicity	42 (4%)	28 (4%)	19 (4%)	3 (3%)
BMI (kg/m^2)	27 (25, 30)	27 (25, 30)	27 (25, 30)	27 (25, 30)
Year of diagnosis	2011 (2009, 2012)	2011 (2009, 2013)	2011 (2009, 2013)	2011 (2009, 2013)
Family history of PCa (first degree)	284 (27%)	165 (26%)	109 (25%)	25 (25%)
T-stage				
T1	929 (89%)	573 (90%)	385 (89%)	91 (90%)
T2a	103 (10%)	56 (9%)	42 (10%)	10 (10%)
T2b	7 (1%)	4 (1%)	4 (1%)	0
T2c	2 (<1%)	1 (<1%)	1 (<1%)	0
Dx biopsy Gleason score (by central review)	N.A.	N.A.		
3+3			374 (87%)	81 (80%)
3+4			58 (13%)	20 (20%)
Dx biopsy Gleason score (by clinical site review)				
≤ 3+3	958 (92%)	584 (92%)	395 (91%)	91 (90%)
3+4	77 (7%)	46 (7%)	36 (8%)	10 (10%)
4+3	6 (1%)	4 (1%)	1 (<1%)	0
% positive biopsy cores	8.3 (8.3, 16.7)	10.0 (8.3, 16.7)	12.5 (8.3, 16.7)	16.7 (8.3, 21.4)
PSA (ng/ml)	5.0 (3.8, 6.5)	4.9 (3.8, 6.6)	4.8 (3.7, 6.5)	4.8 (4.1, 6.1)
Prostate size (cm3)	43 (31, 59)	42 (31, 58)	40 (31, 57)	35 (26, 47)
PSA density (ng/cm3)	0.11 (0.08, 0.16)	0.11 (0.08, 0.16)	0.11 (0.08, 0.15)	0.14 (0.10, 0.19)
NCCN risk group*				
Very low	470 (45%)	310 (49%)	210 (49%)	39 (39%)
Low	411 (39%)	228 (36%)	152 (35%)	48 (48%)
Intermediate	160 (15%)	96 (15%)	70 (16%)	14 (14%)

*****NCCN risk group determined using Gleason score from clinical site review of diagnostic biopsy.

Adverse pathology at RP after a period of surveillance

In univariable analysis of the 101 men who had RP, GPS was not significantly associated with time to AP (HR = 1.70; 95% CI: 1.01, 3.26, p = 0.062). PSA density was the only clinico-pathologic covariate associated with time to AP (Table 2), either as a dichotomous variable (< vs ≥ 0.15 ng/cm3; HR = 0.52; 95% CI: 0.26, 0.95; p = 0.054) or as a continuous variable (HR = 1.78; 95% CI: 1.14, 3.11; p = 0.017). In multivariable models, GPS was associated with time to AP when adjusted for diagnostic Gleason score (Table 3; HR = 1.96; 95% CI: 1.17, 4.28); p = 0.03) or dichotomous PSA density (HR = 1.83; 95%

CI: 1.04, 3.62; p = 0.046). However, it was not associated with time to AP when adjusted for continuous PSA density (HR = 1.61; 95% CI: 0.87, 2.98; p = 0.12), while PSA density was (HR = 1.76; 95% CI: 1.14, 3.24; p = 0.021).

Table 2. Univariable hazard ratios (HRs) for association of variables at diagnosis with time to adverse pathology (AP) in 101 men who had RP after a period of surveillance, and time to biopsy upgrade in 432 men using AS.

	Time to AP (N = 101)		Time to biopsy upgrade (N = 432)	
Variable	**HR (95% CI)**	**p-value**	**HR (95 % CI)**	**p-value**
GPS (per 20 units)	1.70 (1.01,3.26)	0.062	1.02 (0.75, 1.38)	0.93
Age (per year)	1.01 (0.95, 1.05)	0.84	1.00 (0.97, 1.02)	0.7
Age > 65 vs. ≤ 65	1.07 (0.49, 2.16)	0.85	0.85 (0.61, 1.17)	0.31
Nonwhite vs. white race	0.94 (0.26, 2.58)	0.98	1.12 (0.68, 1.85)	0.66
Gleason score 7 vs. 6	0.85 (0.34,1.77)	0.68	0.70 (0.37, 1.34)	0.29
% positive cores	1.02 (0.99, 1.05)	0.29	1.04 (1.02, 1.05)	<0.001
Log PSA	1.65 (0.79, 4.29)	0.21	1.14 (0.88, 1.46)	0.32
Log prostate size	0.61 (0.24,1.45)	0.29	0.44 (0.32, 0.61)	<0.001
PSA density < 0.15 ng/mL2	0.52 (0.26, 0.95)	0.054	0.44 (0.32, 0.61)	<0.001
PSA density (per 0.1 ng/mL^2)	1.69 (1.13, 3.07)	0.025	1.16 (1.08, 1.25)	<0.001
Log_2 PSA density	1.78 (1.14, 3.11)	0.017	1.52 (1.29, 1.79)	<0.001
Clinical stage T2 vs. T1	2.48 (0.92,13.80)	0.14	0.97 (0.59, 1.6)	0.92
BMI (kg/m2)	1.05 (0.96, 1.13)	0.24	1.04 (1.00, 1.07)	0.049
Family history of PCa	0.81 (0.38, 1.59)	0.57	1.12 (0.79, 1.59)	0.52
Diagnosis year (per year)	1.09 (0.98, 1.25)	0.15	1.14 (1.05, 1.24)	0.002

Table 3. Multivariable models for time to AP (n = 101).

Variable	**HR[a] (95% CI)**	**p value**
Model 1		
GPS (per 20 units)	1.96 (1.17, 4.28)	.030
Gleason 7 vs. 6	0.62 (0.24, 1.33)	.26
Model 2		
GPS (per 20 units)	1.83 (1.04, 3.62)	.046
PSA density < 0.15 ng/mL2	0.49 (0.23, 0.90)	.037
Model 3		
GPS (per 20 units)	1.61 (0.87, 2.98)	.12
Log_2 PSA density	1.76 (1.14, 3.24)	.021

[a]Log hazard ratio = regression parameter x Weibull shape parameter. Confidence intervals calculated using the bootstrap quantile method.

Upgrading at surveillance biopsy

In univariable analysis of the 432 men on AS, GPS was not associated with time to upgrade at surveillance biopsy (HR = 1.02; 95% CI: 0.75, 1.38; p = 0.93), while % positive biopsy cores, prostate volume, PSA density, BMI, and year of diagnosis were significantly associated (Table 2). Significant associations with biopsy upgrade did not change in a multivariable model including GPS, % positive biopsy cores, PSA density, and year of diagnosis, either in the cohort of 432 or a subcohort of 395 men diagnosed with GG 1 cancer (adjusted HR (95% CI) = 0.90 (0.66, 1.22), p = 0.48 and 0.96 (0.73, 1.32), p = 0.81, respectively; Table 4). Similar results were observed for a sensitivity analysis using an endpoint of time to high upgrade (to GG ≥ 3) on surveillance biopsy (data not shown).

Table 4. Multivariable analysis for time to biopsy upgrade.

	N = 432, 167 events		N = 395 dx with Gleason 6; 157 events	
Variable	**Hazard Ratio (95% CI)**	**p value**	**Hazard Ratio (95% CI)**	**p value**
GPS (per 20 units)	0.90 (0.66, 1.22)	0.48	0.96 (0.73, 1.32)	0.81
Log_2 PSA density	1.44 (1.21, 1.71)	<.001	1.45 (1.21, 1.72)	<.001
% positive cores	1.03 (1.02, 1.04)	<.001	1.04 (1.02, 1.05)	<.001
Year of diagnosis	1.13 (1.04, 1.23)	0.003	1.11 (1.03, 1.21)	0.010

Conclusions

In a cohort of men on AS, GPS was associated with time to AP when adjusted for diagnostic Gleason grade group or dichotomous PSA density. GPS was not associated with surveillance biopsy Gleason grade upgrading or AP at surgery after adjustment for continuous PSA density, although a trend was seen for AP, suggesting an association may be seen in a larger study.

Dissemination

An abstract describing these results has been accepted to the 2019 GU ASCO meeting, and a manuscript is in preparation.

Task 3: Evaluate a panel of four-kallikrein plasma-based markers to determine the presence of or progression to clinically relevant prostate cancer. *(Lead site: FHCRC)*

In men suspected of having prostate cancer, a panel of four kallikreins (total PSA (tPSA), free PSA (fPSA), intact PSA (iPSA), and human kallikrein 2 (hK2)) combined with age using a mathematical algorithm had been shown to improve the prediction of high-grade cancers compared to the PCPT risk calculator or models using tPSA alone. We collaborated with OPKO to explore utility of prediction models incorporating the pre-defined 4 kallikrein panel algorithm (4Kpanel) to predict the presence of occult high-grade disease in men already diagnosed with Gleason 6 cancer and on active surveillance.

Statistical methods

The objective was to determine whether a model using clinical predictors and kallikrein data collected after diagnosis of Gleason 6 cancer but prior to surveillance biopsy, can predict high-grade cancer in the surveillance biopsy. Sequential surveillance biopsies were considered as two groups: A) the initial biopsy after cancer diagnosis (sometimes called confirmatory biopsy), and B) all subsequent surveillance biopsies. Biopsy data were split 2:1 into training and test sets matched by outcome.

The primary outcome was reclassification from Gleason score 6 to Gleason score ≥7. A value for the 4Kpanel was calculated with tPSA, fPSA, iPSA, hk2 and age using locked down coefficients developed before the study was conducted. This combination of the four kallikreins is the same as in the commercial assay. Additional clinical predictors considered in modeling included age, body mass index (BMI), race (African American or other), digital rectal examination (DRE) results, number of previous biopsies after diagnosis, number of negative biopsies after diagnosis, cores ratio from previous biopsy (ratio of biopsy cores containing cancer to total cores), maximum cores ratio from all previous biopsies, months since diagnosis, prostate volume (prostate size measured closest to time of sampling and imputed within 2-years). Either the 4Kpanel (logit scale) or clinical serum PSA (logarithm

transformed) was used in models. Prediction models were built using data in the training set and clinical performance was assessed with the testing set.

Statistical models were developed to predict reclassification from Gleason 6 cancer to Gleason 7 or greater. The analysis plan was determined before specimens were selected for the study, and included breaking the data/specimens into training and testing cohorts, using a 2/3 to 1/3 split. The models included clinical information and either the 4Kpanel or serum PSA. We used Receiver Operating Characteristic (ROC) curve analyses and area under the curve (AUC) to assess discriminatory capacity and decision curve analysis (DCA) to report clinical net benefit.

Results

Significant predictors for reclassification were 4Kpanel (OR=1.54 [1.31,1.81]) or PSA (OR=2.11 [1.53,2.91]), ≥20% cores positive (OR=2.10 [1.33,3.32]), ≥2 prior negative biopsies (OR=0.19 [0.04,0.85]), prostate volume (OR=0.47 [0.31,0.70]), BMI (OR=1.09 [1.04,1.14]); Table 5. ROC curve analysis comparing 4Kpanel and base models indicated that the 4Kpanel improved accuracy for predicting reclassification (AUC 0.78 versus 0.74) in the first surveillance biopsy (Table 6). Both models performed comparably for prediction of reclassification in subsequent biopsies (AUC=0.75 versus 0.76). In DCA, both models showed higher net benefit compared to biopsy-all and biopsy-none strategies (Figure 2).

Table 5: Summary of fitted models including clinical variables + serum PSA or 4Kpanel in the training set.

Variable	PSA + full clinical model			4K + full clinical model		
	OR	CI	p-value	OR	CI	p-value
Age	1.03	(1.00,1.06)	0.068			
BMI	1.11	(1.06,1.16)	<0.001	1.09	(1.04,1.14)	<0.001
Cores ratio >0.2	2.19	(1.39,3.44)	0.001	2.10	(1.33,3.32)	0.001
Negative biopsies ≥2	0.19	(0.04,0.80)	0.023	0.19	(0.04,0.85)	0.029
Log(prostate volume)	0.31	(0.20,0.48)	<0.001	0.47	(0.31,0.70)	<0.001
Log(PSA)	2.11	(1.53,2.91)	<0.001			
4Kpanel				1.54	(1.31,1.81)	<0.001

Table 6. Results of final regression models for reclassification. AUC (95% CI) of various models for initial surveillance biopsy and subsequent surveillance biopsies. CIs were calculated with bootstrap accounting for correlations among individuals.

Base Model	4K + Clinical Model AUC (95% CI)	PSA + Clinical Model AUC (95% CI)	Difference (95% CI)
Full Clinical Model			
Initial Biopsy	0.783 (0.691,0.871)	0.740 (0.652,0.828)	0.043 (0.003,0.086)
Subsequent Biopsy	0.754 (0.657,0.838)	0.755 (0.653,0.841)	-0.001 (-0.037,0.041)

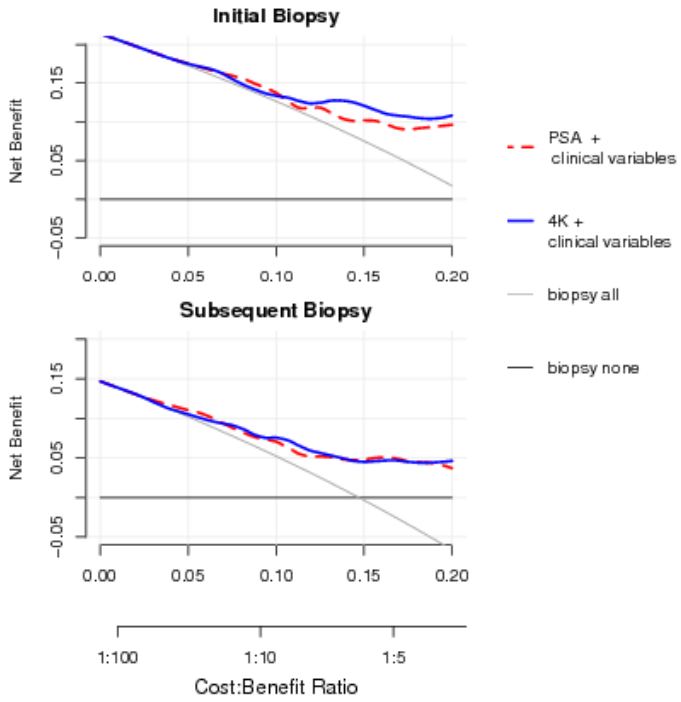


Figure 2. Decision Curve Analysis for full models with serum PSA (dotted red line) or with the 4Kpanel (solid blue line). Strategies for biopsying all men (light grey line) or no men (dark grey line) are also shown. The line with the highest net benefit at any particular threshold probability for biopsy (x-axis) will result in the best clinical results.

Conclusions: The 4Kpanel provided incremental value over routine clinical information in predicting high-grade cancer in the first biopsy after diagnosis. The 4Kpanel did not add predictive value to the base model at subsequent surveillance biopsies.

Dissemination: These results were presented at the 2016 Meeting of the American Urological Association (AUA) and have been published in European Urology (v72, pp448-454.) A reprint of the final publication is included with this report.

Task 4: Confirm the ability of PCA3 mRNA concentrations in urine, alone or in combination with TMPRSS2:ERG mRNA, to predict the presence of or development to clinically relevant prostate cancer. *(Lead site: FHCRC)*

PCA3 and the TMPRSS2:ERG fusion are prostate cancer-specific biomarkers that hold promise for stratifying risk in the setting of AS. Hologic Gen-Probe's assay to quantitate urine PCA3 transcripts in post-digital rectal exam (DRE) urine is FDA-approved for men with a previous negative biopsy, given peer reviewed evidence that it can reduce unnecessary prostate biopsies. We sought to confirm the ability of the PCA3 and TMPRSS2:ERG assays to predict aggressive prostate cancer in the entire PASS cohort. We collaborated with Hologic Gen-Probe to analyze 2,069 urine specimens collected at baseline, 6, 12, and 24 month study visits from 782 PASS participants to evaluate the utility of PCA3 and TMPRSS2:ERG fusion status in an active surveillance setting.

Statistical methods

Reclassification was defined as an increase in primary or secondary Gleason grade at biopsy and/or an increase in the biopsy cores with cancer to total cores collected (cores ratio) to ≥34%. Several models of reclassification were considered, as depicted in the Study Schematic (Figure 3).

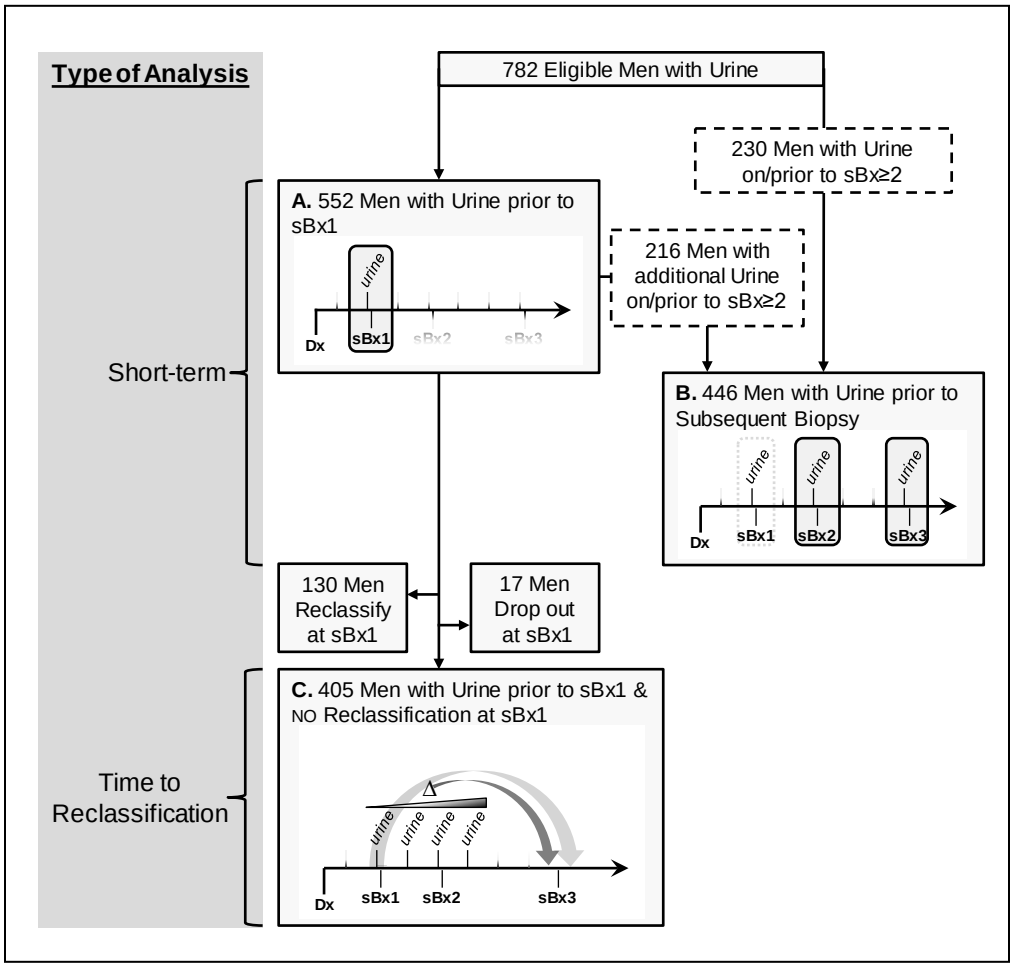


Figure 3. Study design. For short-term prediction, analysis of the association of biomarkers with reclassification in the biopsy immediately following urine collection was performed using logistic regression for A.) 552 men with urine assayed prior to the first surveillance biopsy (sBx1) and B.) 446 men with urine assayed prior to subsequent surveillance biopsies. For longer-term prediction, the association of biomarkers with time to reclassification was examined in C.) 405 men with urine assayed prior to their sBx1 who did not reclassify at that biopsy; the first urine biomarker sample along with the urine biomarker kinetics at each observation time were considered as covariates in a partly conditional Cox model.

Modeling short-term biopsy reclassification (Figure 3, panels A and B)

The association between urine biomarkers collected immediately prior to a biopsy and reclassification at the biopsy was modeled using logistic regression. The analysis was stratified by the first surveillance biopsy (sBx1, sometimes called confirmatory biopsy; Figure 3, Scenario A; n=552) and subsequent surveillance biopsies (Figure 3, Scenario B; n=446).

Modeling time to reclassification with longitudinal biomarkers (Figure 3, panel C)

The association between PCA3 or T2:ERG and the time to biopsy reclassification was modeled using a partly conditional Cox proportional hazards (PH) model. Participants were excluded from this analysis if they reclassified at the sBx1. Each participant had the first urine biomarkers assayed after diagnosis and prior to sBx1 and had up to 3 additional urine samples collected up to 2 years after the first. The first urine biomarker sample along with the urine biomarker kinetics at each observation time were covariates in the partly conditional Cox model (Figure 3, panel C; n=405). Participants without reclassification were censored at date of last study contact, treatment, or 2 years after their last biopsy, whichever came first.

Urine biomarker kinetics were calculated based on a linear mixed effect model (LMEM), in which the natural log of the urine biomarkers was modeled as a linear function of time since diagnosis, with a random intercept indicating the individual-specific ln(urine biomarker) at diagnosis, and a random slope reflecting individual-specific rate of change over time. A PCA3 or T2:ERG kinetics (PCA3k or T2:ERGk) value for each participant based on the first urine sample up to an observation time was then derived based on the best linear unbiased predictor (BLUP) estimator from the LMEM. Intra-class correlation (ICC) was calculated to determine the proportion of total variability in biomarker scores explained by between-participant variability.

Results

Analysis of reclassification at next biopsy: Of the 552 men with urine biomarkers assessed prior to the sBx1 (Figure 3, panel A), 130 (24%) were reclassified at that biopsy. In a logistic regression model adjusted for PSA, cores ratio, and prostate size, PCA3 score was associated with reclassification in the sBx1 (OR = 1.3; 95% CI: 1.0-1.7), and T2:ERG score was not (Table 7). A model in which the endpoint was only Gleason grade reclassification was similar (data not shown).

Table 7. Logistic regression model results for grade and/or tumor volume reclassification in first surveillance biopsy (n = 552).

Variable*	Univariable		Multivariable	
	OR (95% CI)^	p-value^	OR (95% CI)^	p-value^
PSA	1.5 (1.1, 2.0)	0.01	1.8 (1.3, 2.6)	0.001
Dx Cores Ratio	4.0 (2.6, 6.3)	<.001	3.4 (2.1, 5.4)	<.001
Prostate Size	0.3 (0.2, 0.5)	<.001	0.3 (0.1, 0.4)	<.001
PCA3	1.6 (1.2, 1.9)	0.0001	1.3 (1.0, 1.7)	0.02
T2:ERG	1.1 (1.0, 1.2)	0.21	1.0 (0.9, 1.2)	0.52

* The natural log of all variables was used in modeling.

^ Odds ratios, 95% confidence intervals and p-values from logistic regression models.

There was a small change in the AUC of a model for predicting reclassification at sBx1 using clinical variables plus PCA3 versus a model with only clinical variables: 0.753 [95% CI 0.707-0.800] vs 0.743 [0.693-0.791] with and without PCA3 respectively (data not shown). No improvement in AUC was found for a model including T2:ERG. Similar findings were observed in DCA (data not shown). A model with clinical variables and PCA3 showed minimal increase in net benefit relative to a model with only

clinical variables. All models with clinical variables showed an improvement to PCA3 alone, and all models showed an improvement to biopsy-all and biopsy-none strategies.

Neither PCA3 nor T2:ERG was associated with reclassification at subsequent biopsy.

Analysis of time to reclassification for longitudinal biomarkers

There were 405 participants included in the time-to-event analysis who had their first urine sample collected prior to the sBx1 and who did not reclassify at the sBx1 (Figure 1, panel C). With a median follow-up of 3.6 years from the first urine collection, 103 (25%) participants reclassified at any subsequent surveillance biopsy.

The annual percent change in PCA3 estimated by LMEM was 9.8 (95% CI 7.3-12.3, $p<0.001$). As determined by ICC, 85% of the observed variation in PCA3 was explained by between-participant variation, and 15% due to within-participant variation. The annual percent change in T2:ERG was 11.3 (95% CI 5.2-17.8, $p<0.001$), and 68% of the observed variation was explained by between-participant variation and 32% due to within-participant variation. Biomarker kinetics were calculated based on deriving a BLUP estimator from a LMEM. No significant differences in slopes were found between participants with reclassification versus those with no reclassification for either biomarker.

Conclusions

PCA3 but not T2:ERG was associated with cancer reclassification in the first surveillance biopsy, but has negligible improvement over clinical variables alone in ROC or DCA analyses. Neither marker was associated with reclassification in subsequent biopsies.

Dissemination

These results were presented at the Multi-Institutional Prostate Cancer SPORE Program Retreat, 2016. A manuscript describing these results is included with this report and has been accepted for publication at Prostate Cancer and Prostatic Diseases.

Task 5: Central Pathology Review *(Lead Site: CCF)*

The purpose of central pathology review is to standardize endpoints for analyses of biomarkers. With funding from this grant, we have developed a customized pathology review system, in which primary and secondary pathology reviewers can access scanned images and record key data from each slide. All data recorded by the primary and secondary reviewers are reviewed for consistency and if results are discrepant, a consensus review is conducted to resolve. A schematic of this process for biopsies is in Figure 4. All H&E slides from radical prostatectomies (RPs) were collected from clinical sites and reviewed manually at a central location. Review data were recorded and slides returned to the original sites.

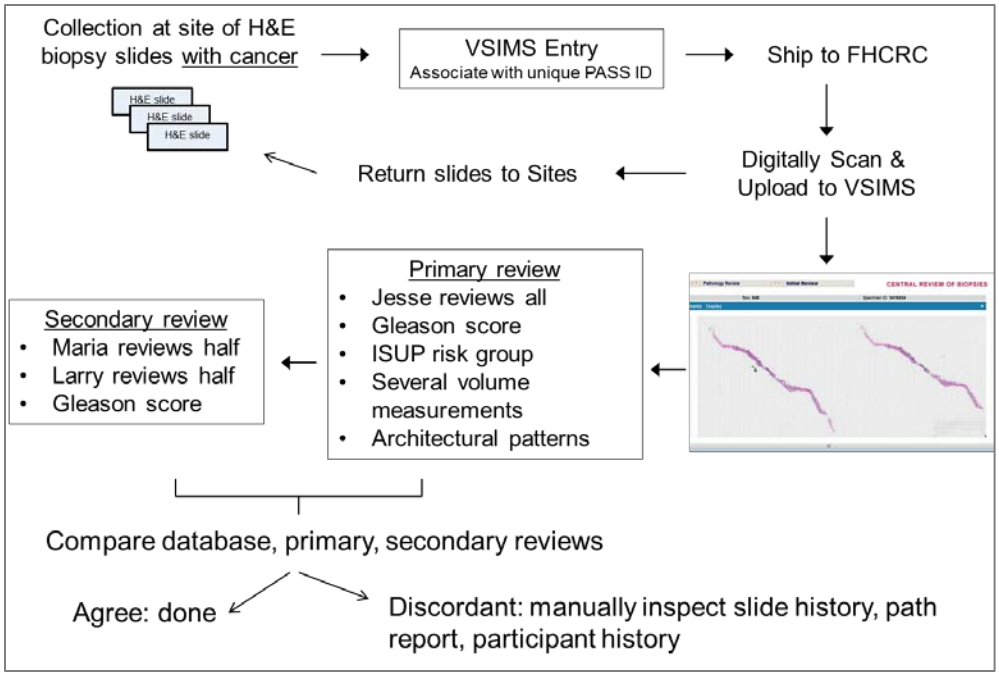


Figure 4. Schematic of central pathology review workflow

All H&E stained slides that contain cancer have been collected for 1,966 of 2,729 diagnostic and surveillance biopsies that occurred at 8 of the PASS sites. These slides have been sent to FHCRC and digitized and are currently being reviewed by the central pathology team. The remaining biopsy slides are not available, mostly because they were at local clinics that could not or would not release slides.

Complete sets of H&E stained slides from 207 RPs that occurred at 8 sites in PASS after a period of surveillance have been collected and reviewed centrally.

In an early analysis of scoring, we evaluated slides from 131 unique diagnostic biopsies, collected from five different PASS study sites. In this small subset, 71% of cases were reviewed concordantly by study pathologists and the original pathologist (**Table 8**). The 29% discordant reviews highlight the need for a centralized review of cases to obtain accurate data, as Gleason is used as an endpoint in many biomarker studies.

Table 2. Analysis of Concordance in Central Pathology Review of 131 PASS Biopsies

Original Path		Total Scenarios				
Gleason	Total Cases	Total Agreement	CR Agreement	Orig & 1° Agree	Orig & 2° Agree	Total Disagreement
3 + 3	120	90 (76)	12 (10)	11 (9)	4 (3)	3 (2)
3 + 4	8	2 (25)	1 (13)	0	5 (63)	0
4 + 3	3	0	3 (67)	0	0	0
TOTAL	**131**	**92 (71)**	**16 (12)**	**11 (9)**	**9 (7)**	**3 (2)**

Task 6. Translation of biomarkers into clinical practice. *(Lead sites: FHCRC and CCF)*

To use any biomarker in clinical practice it is essential that the biomarkers add incremental improvement in prediction of high grade or high volume disease over the clinical variables alone. Until very recently there have been no risk prediction models for use in active surveillance, and the ones that have recently been developed have not fully utilized important variables available in contemporary clinical practice. Thus, as an important component of translating biomarkers into clinical management of active surveillance patients, we have been developing base risk prediction models using commonly available clinical variables (PSA, prostate size, and biopsy information from the diagnostic and surveillance biopsies). These models are used as the basis for evaluation of biomarkers for clinical management, but they may by themselves be exceedingly useful in managing active surveillance patients.

To utilize risk prediction models in clinical management, we are developing user interfaces that not only calculate the risk of finding aggressive cancer (e.g. high grade cancer) at the next or future biopsies, but also show important metrics of how well a given model may be predicting risk and what the clinical consequences are of using a specific risk threshold in decision making. Our first such risk prediction tool uses the "base" model developed to evaluate the 4 Kallikreins (Lin et. al., European Urology, v72, pp448-454, 2017). In this model, a patient's age, body mass index, prostate size, PSA, and information about the ratio of prior biopsy cores containing cancer and of prior biopsies in which no cancer was found are used to predict the risk of finding high grade (Gleason ≥7) cancer in the next biopsy. In addition to the risk of finding high grade cancer, the calculator displays the 95% confidence interval of that risk, how the risk compares to the active surveillance (PASS) population, and where a given patient's value for each variable falls within the population from which the model was developed (**Figure 5**). The clinical consequences of using a specific risk threshold to guide the decision of whether or not to perform a biopsy are shown on a separate tab of the calculator (**Figure 6**). A patient report, or a simpler depiction of the risk information, is also available in the calculator (**Figure 7**). The PASS Risk Calculator can be found at: canarypass.org

Figure 5. First tab of PASS Risk Calculator.

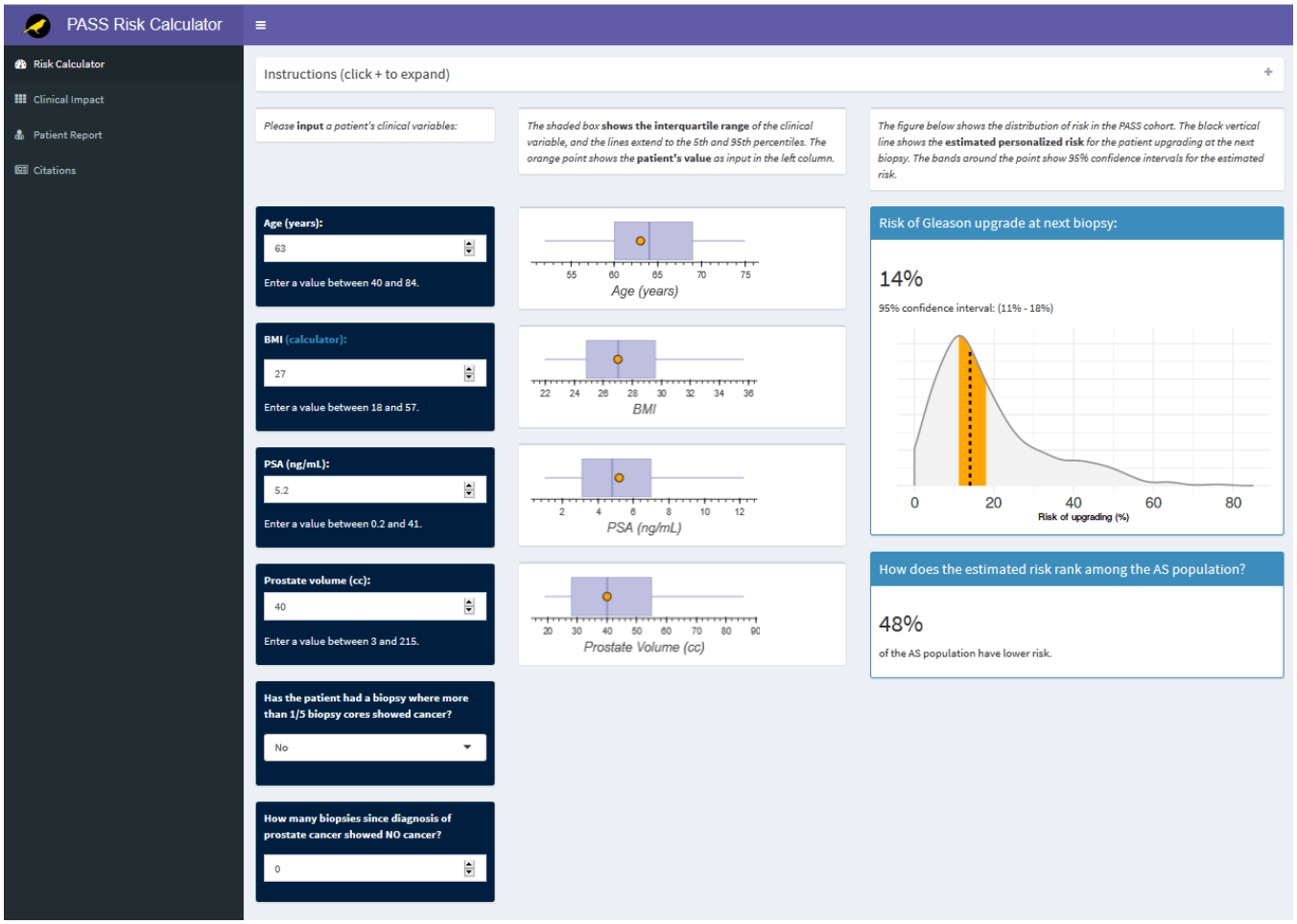


Figure 6. Clinical Impact tab of PASS Risk Calculator.

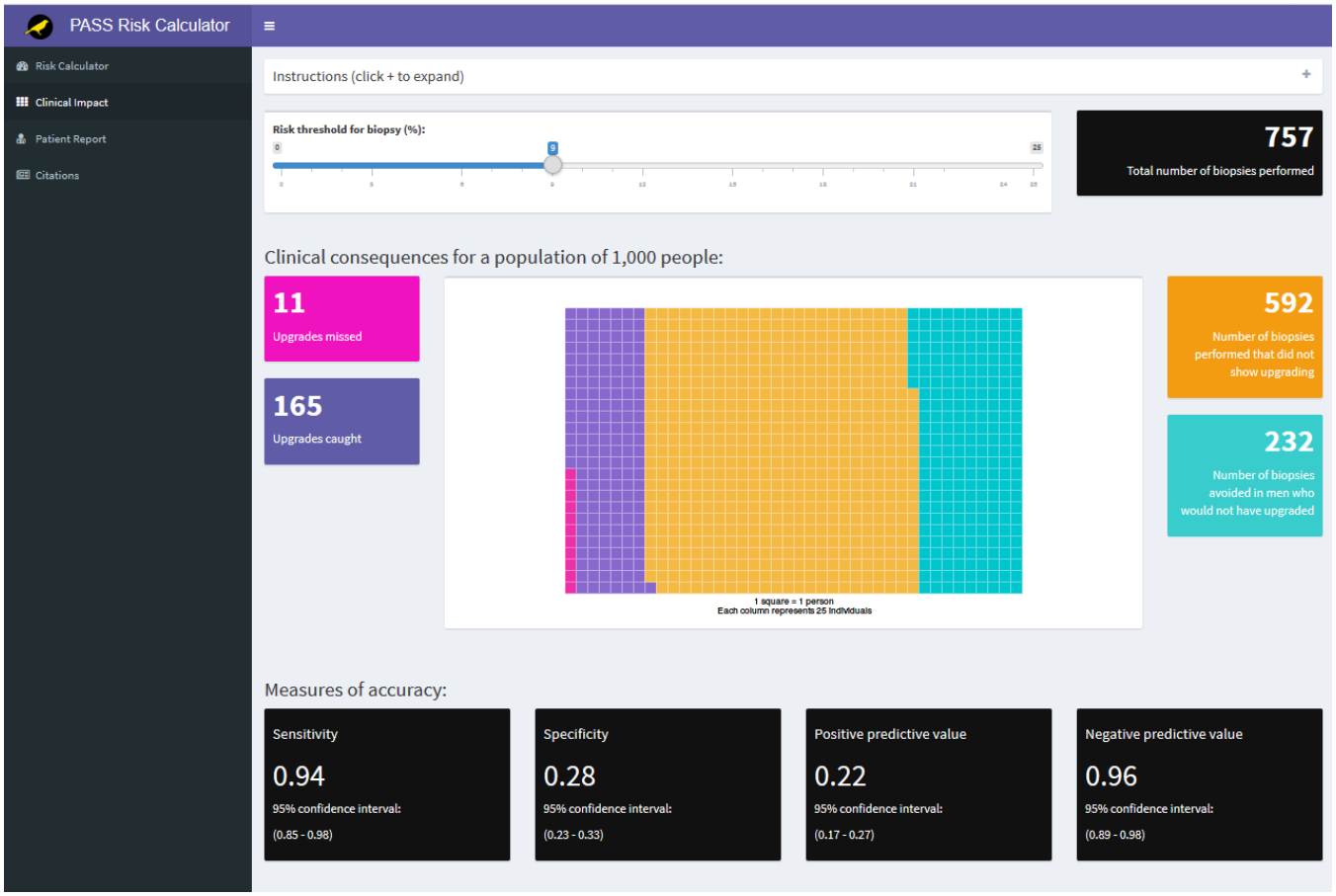
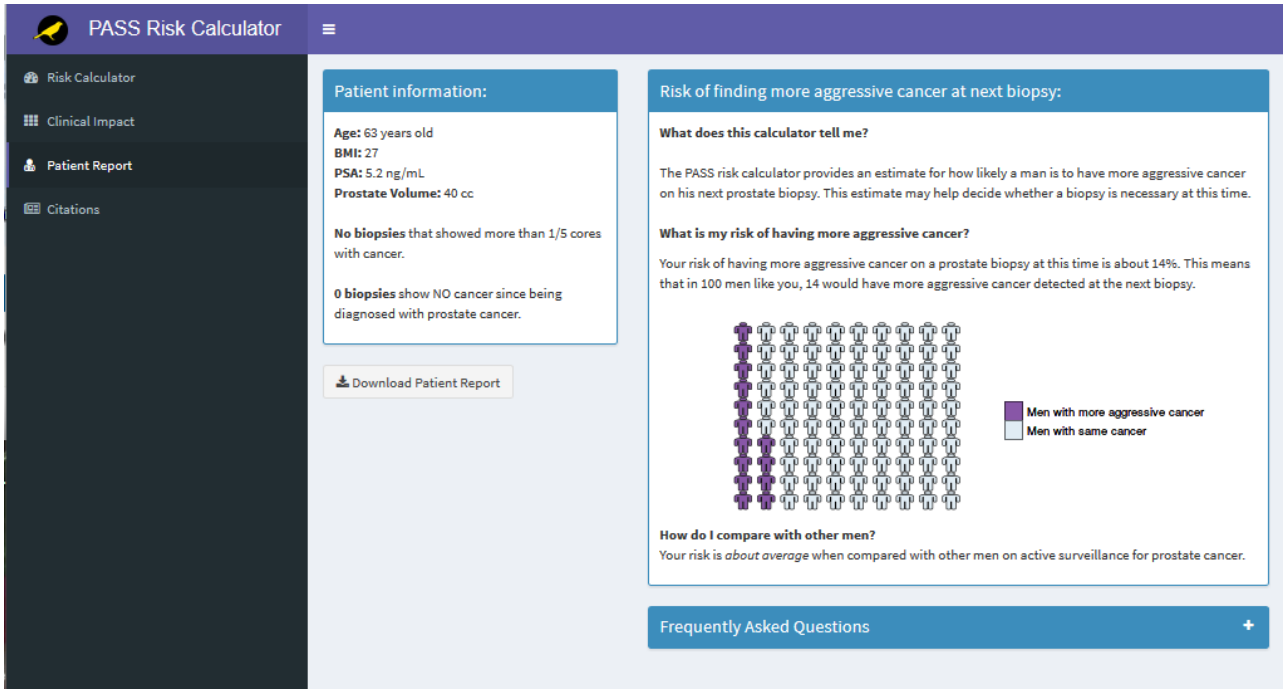


Figure 7. Patient Report of PASS Risk Calculator.



To improve upon our risk prediction models, we have been evaluating individual parameters. We have found that PSA kinetics, when calculated using a linear mixed effects model, is associated with time to biopsy reclassification in a model adjusting for prostate size, time since diagnosis, biopsy parameters, and PSA at diagnosis. The details of this work are described in the attached paper titled "Refined analysis of prostate specific antigen kinetics to predict prostate cancer active surveillance outcomes." (**Cooperberg et al, Eur Urol, 2018**)

We have found that after a diagnosis of low grade prostate cancer, having one or two biopsies in which no cancer are found substantially reduces a man's risk of upgrading at future biopsies. The details of this work are described in the attached manuscript titled "The Role of Surveillance Biopsy with No Cancer as a Prognostic Marker for Reclassification: Results from the Canary Prostate Active Surveillance Study (PASS)." (**Kearns et al, Eur Urol, 2018**)

We also evaluated whether continued use of 5-ARIs during active surveillance is associated with rate of biopsy reclassification, and found that in multivariable analysis, there was no difference in the risk of reclassification between 5-ARI users and never users (HR 0.81, p = 0.31). The details of this work are described in the attached paper titled "Continued Five-Alpha Reductase Inhibitor Use After Prostate Cancer Diagnosis and the Risk of Reclassification and Adverse Pathological Outcomes in the Canary Prostate Active Surveillance Study (PASS)." (**Kearns et al, J Urol, 2018**)

Another aspect of our work designed to move biomarkers into clinical practice is to better address the problem that many men who use active surveillance have biologically indolent tumors, and do not need to follow a surveillance regimen that includes repeat prostate biopsies every 1 to 2 years. We have developed a dynamic risk prediction model that stratifies risk of reclassification in the next 4 years of active surveillance in patients who have not reclassified in the first surveillance biopsy. Men at lowest and highest deciles of this model-based risk faced 6% (95%CI 0-13%) and 53% (40-72%) risks of reclassification within 4 years after the first surveillance biopsy (about 1 year from initial diagnosis). For at least 10% of the men in the cohort, the negative predictive value (NPV) for reclassification was 94% or higher. The model has been validated in the non-PASS UCSF cohort. These results show that a substantial proportion of men with low-risk prostate cancer can safely be followed with a de-intensified active surveillance protocol, which would improve both the tolerability and cost-effectiveness of this management strategy. A manuscript describing these results is in development, and additional work is underway to evaluate if the biomarkers evaluated with support of this grant, especially the 4K panel, improve performance of this dynamic risk prediction model.

An important aspect of our work has been to evaluate the impact of potential ascertainment bias in in our analyses. In active surveillance (AS), disease progression as defined by pathology can only be ascertained at follow-up biopsy. However, clinical factors could also influence the timing of biopsy, resulting in bias that affects conclusions about variables predictive of aggressive disease. To evaluate potential bias, we first defined windows, or date ranges, for on-time (compliant) surveillance biopsies. To determine whether factors were associated with the interval of surveillance biopsy, biopsy timing (on-time, early, or late) was regressed on clinical and prior biopsy variables using multinomial regression models. Probabilities for participants having all on-time biopsies were derived from the models. As an example, we reanalyzed the aforementioned study evaluating the association of PSA kinetics (PSAk) with biopsy reclassification. Model coefficients were compared in the full cohort and

subsets of compliant participants both with and without weighting for the propensity for compliance. Briefly, we found that 78% of surveillance biopsies in PASS were on-time, 11% were early and 11% were late. After adjustment for prostate size, time since diagnosis, cores with cancer and prior negative biopsies, PSAk was associated with biopsy reclassification (HR=1.61 (95% CI: 1.25-2.09), but was also found to be associated with biopsy timing. In multivariable models of PSAk in participants with all biopsies on time, both without and with the propensity adjustment, PSAk was still associated with reclassification [HR=2.09 (95% CI: 1.45-3.02) and HR=2.17 (95% CI: 1.45-2.93), respectively]. Applying propensity weights to account for potential ascertainment bias, we demonstrated that despite its association with biopsy timing, PSA kinetics remains independently associated with reclassification. A manuscript describing these results is in preparation.

Summary

Over 200,000 men are diagnosed with prostate cancer annually, with roughly half having low-risk disease that is unlikely to progress over time. Given persistent uncertainty regarding the accuracy of risk assessment based on biopsy and clinical data available at time of diagnosis, patients and clinicians are hesitant to embrace active surveillance as a management strategy for low-risk disease. Many men with indolent tumors will undergo radical curative treatment immediately, incurring significant costs and, in some cases, morbidity and/or long term side effects.

Our goal in this project was to perform rigorous clinical validation of established biomarkers to improve the accuracy of risk assessment and to allow clinicians to distinguish aggressive from indolent disease in men with apparently low-risk disease. We evaluated three commercially available biomarker panels in this project (GPS, 4Kpanel of the 4Kscore, PCA3) that are used clinically in active surveillance patients but previously had not been thoroughly evaluated for use in active surveillance. The results of our analyses indicate that although these commercially available tests may provide some useful information, they do not greatly improve upon currently available clinical factors to distinguish aggressive from indolent cancer in active surveillance patients. We continue to evaluate clinical data and biomarker results in new combinations to identify a better method for prediction of progression on active surveillance.

What opportunities for training and professional development did the project provide?

Nothing to report. This grant does not provide for training or professional development activities.

How were the results disseminated to communities of interest?

Results are disseminated through presentations at national meetings and through publication in scientific journals. Please see the list of publications and presentations below.

What do you plan to do during the next reporting period to accomplish the goals and objectives?

Nothing to report.

4. IMPACT

What was the impact on the development of the principal discipline(s) of the project?

We anticipated that successful clinical validation of biomarkers would have extraordinary potential to improve the care of PCa patients. Specifically, those men with apparent low-risk tumors that could be confirmed as truly low-risk with greater accuracy could be spared the cost and quality-of-life impact of invasive diagnostic and therapeutic procedures (e.g. surgery, radiation therapy, or even serial biopsies). Conversely, those men with apparent low-risk disease who in fact harbor higher-risk tumors or have the potential to develop lethal disease could be identified earlier, thus avoiding under-treatment. Such a paradigm shift in PCa care would yield near-term changes in the PCa treatment landscape, greatly improving the cost-benefit calculations for population-level PCa screening efforts and reducing the overtreatment of disease. In this project, we did not find strong evidence to support the routine use of the three commercially-available biomarker panels to predict outcomes in AS. Our assessment highlights the need for biomarker development and testing in the active surveillance setting. Although biopsy reclassification is an meaningful endpoint in that it drives treatment decision, our results may also highlight a need for an improved definition of endpoints in active surveillance that are more strongly correlated with development of lethal disease. We are actively pursuing evaluations of biomarkers, alone and in combination, with alternate endpoint definitions and in subsets of patients where they may be more effective.

What was the impact on other disciplines?

We expect that statistical techniques and approaches to risk modeling developed with this funding may be utilized to evaluate biomarker performance in other diseases beyond prostate cancer.

What was the impact on technology transfer?

This project involved evaluation and validation of commercial biomarker panels that have not previously been used in the active surveillance setting. While we did not expect a direct impact on technology transfer, there could be an impact on the commercial use of the molecular diagnostics.

What was the impact on society beyond science and technology?

This project has potential to reduce overutilization of scarce healthcare resources. Commercially available tests are marketed to and routinely used in active surveillance patients but were developed and tested either as diagnostics (4Kscore, PCA3) or in men who were immediately treated for their prostate cancer. This project adds critical information regarding the value of three commercially available clinical tests in the active surveillance setting. None of the three tests evaluated here show clear, strong evidence that substantially improve risk stratification over the stratification that can be achieved with commonly available clinical data. The impact of these findings may be a reduction in spending of limited healthcare resources for tests not predictive of outcome in active surveillance.

5. CHANGES / PROBLEMS

Changes in approach and reasons for change
Nothing to report.

Actual or anticipated problems or delays and actions or plans to resolve them
Nothing to report.

Changes that had a significant impact on expenditures
Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents:

- **Significant changes in use or care of human subjects:** No significant changes in the use or care of human subjects. The Fred Hutchinson Cancer Research Center reviewed the project on 5/17/2018 and approved the study activities through 5/29/2019 under IR file number 8271. A continuing review was submitted to HRPO (Log Number A-18320) and receipt was acknowledged on 06/28/2018.
- **Significant changes in use or care of vertebrate animals:** Nothing to report.
- **Significant changes in use of biohazards and/or select agents:** Nothing to report.

6. PRODUCTS

Publications, Conference Papers, and Presentations

Journal publications

Cooperberg MR, Brooks JD, Faino AV, Newcomb LF, Kearns JT, Carroll PR, Dash A, Etzioni R, Fabrizio MD, Gleave ME, Morgan TM, Nelson PS, Thompson IM, Wagner AA, Lin DW, Zheng Y. **Refined analysis of prostate specific antigen kinetics to predict prostate cancer active surveillance outcomes.** *European Urology.* 2018 Aug;74(2):211-217. PMCID: PMC6263168. Acknowledgement of federal support: Yes.

Kearns JT, Faino AV, Newcomb LF, Brooks JD, Carroll PR, Dash A, Ellis WJ, Fabrizio M, Gleave ME, Morgan TM, Nelson PS, Thompson IM, Wagner AA, Zheng Y, Lin DW. **Role of Surveillance Biopsy with No Cancer as a Prognostic Marker for Reclassification: Results from the Canary Prostate Active Surveillance Study.** *European Urology.* 2018 May;73(5):706-712. PMCID: PMC6064187 Acknowledgement of federal support: Yes.

Kearns JT, Faino AV, Schenk JM, Newcomb LF, Brooks JD, Carroll PR, Dash A, Ellis WJ, Fabrizio M, Gleave ME, Morgan TM, Nelson PS, Thompson IM, Wagner A, Zheng Y, Lin DW. **Continued Five-Alpha Reductase Inhibitor Use After Prostate Cancer Diagnosis and the Risk of Reclassification and Adverse Pathological Outcomes in the Canary Prostate Active Surveillance Study (PASS)**. *Journal of Urology*. In press. Epub 2018 Aug 1.

Lin DW, Newcomb LF, Brown MD, Sjoberg DD, Dong Y, Brooks JD, Carroll PR, Cooperberg M, Dash A, Ellis WJ, Fabrizio M, Gleave ME, Morgan TM, Nelson PS, Thompson IM, Wagner AA, Zheng Y; Canary Prostate Active Surveillance Study Investigators. **Evaluating the four kallikrein panel of the 4Kscore for prediction of high-grade prostate cancer in men in the Canary Prostate Active Surveillance Study (PASS).** *European Urology.* 2017 Sep;72(3):448-454. Acknowledgement of federal support: Yes.

Newcomb LF, Zheng Y, Faino AV, Bianchi-Frias D, Cooperberg MR, Brown MD, Brooks JD, Carroll PR, Dash A, Fabrizio MD, Gleave ME, Liss M, Morgan TM, Thompson IM, Wagner AA, Nelson PS, Lin DW. **Performance of PCA3 and TMPRSS2:ERG urinary biomarkers in prediction of biopsy outcome in the Canary Prostate Active Surveillance Study (PASS).** *Prostate Cancer and Prostatic Diseases*. 2018: In press. Acknowledgement of federal support: Yes.

Books or other non-periodical, one-time publications

Nothing to report.

Other publications, conference papers, and presentations

Cooperberg M, Faino A, Newcomb, L, Carroll, P, Kearns J, Brooks J, Fabrizio M, Gleave M, Morgan T, Dash A, Nelson P, Thompson I, Wagner A, Zheng Y, Lin D. **PD20-01: When can active surveillance (AS) be less active? Prediction of long-term non-reclassification for men with low risk prostate cancer**. Annual Meeting of the American Urological Association; 2018 May 19, San Francisco, CA.

Cooperberg M, Brooks J, Faino A, Zheng Y, Kearns K, Carroll P, Bash A, Fabrizio M, Gleave M, Morgan T, Nelson P, Thompson I, Wagner A, Newcomb L, Lin D. **MP43-11 Refined analysis of prostate specific antigen (PSA) velocity to predict outcomes in active surveillance: Results from the Canary Prostate Active Surveillance Study (PASS).** Annual Meeting of the American Urological Association; 2017 May 13, Boston, MA.

Cooperberg, M, Faino A, Newcomb LF, Carroll P, Kearns JT, Brooks JD, Fabrizio M, Gleave M, Morgan TM, Dash A, Nelson P, Thompson IM, Wagner A, Lin DW, Zheng Y. **Abstract # 140: When can active surveillance be less active? Prediction of long-term non-reclassification for men with low-risk prostate cancer**. 2018 ASCO Genitourinary Cancers Symposium; 2018 Feb 8, San Francisco, CA.

Kearns J, Faino A, Newcomb L, Brooks J, Carroll PR, Dash A, Ellis W, Fabrizio M, Gleave M, Morgan T, Nelson P, Thompson I, Wagner A, Zheng Y, Lin D. **Abstract # 22: The Use of Five-Alpha Reductase Inhibitors and their Association with Reclassification and Pathologic Outcomes in the Canary Prostate Active Surveillance Study (PASS)**. 2017 ASCO Genitourinary Cancers Symposium; 2017 Feb 16, San Francisco, CA.

Kearns J, Faino A, Newcomb L, Brooks J, Carroll PR, Dash A, Ellis W, Fabrizio M, Gleave M, Morgan T, Nelson P, Thompson I, Wagner A, Zheng Y, Lin D. **MP43-10: The role of surveillance biopsy with no cancer as a prognostic marker for reclassification: Results from the Canary Prostate Active Surveillance Study**." Annual Meeting of the American Urological Association; 2017 May 13, San Francisco, CA.

Kearns J, Faino A, Newcomb L, Brooks J, Carroll PR, Dash A, Ellis W, Fabrizio M, Gleave M, Morgan T, Nelson P, Thompson I, Wagner A, Zheng Y, Lin D . **PD55-02: The Use of Five-Alpha Reductase Inhibitors and their Association with Reclassification and Pathologic Outcomes in the Canary Prostate Active Surveillance Study (PASS)**. Annual Meeting of the American Urological Association; 2017 May 15, Boston, MA.

Lin D, Brown M, Newcomb L, Sjoberg D, Brooks J, Carroll P, Dash A, Fabrizio M, Gleave M, Morgan T, Nelson P, Thompson I, Zheng Y. **PD08-02: Evaluating the four kallikrein panel of the 4KScore for prediction of high-grade prostate cancer in men in the Canary Prostate**

Active Surveillance Study (PASS). Annual Meeting of the American Urological Association; 2016 May 6-10, San Diego, CA.

Newcomb L. **Evaluating urinary PCA3 and TMPRSS2:ERG for prediction of adverse biopsy reclassification in men in the Canary Prostate Active Surveillance Study (PASS).** Presentation at the Multi-Institutional Prostate Cancer SPORE Program Retreat, 2016 March 13-15, Fort Lauderdale, FL.

Website(s) or other Internet site(s)

The PASS Risk Calculator can be found at the PASS website: https://canarypass.org

Technologies or techniques

Nothing to report.

Inventions, patent applications, and/or licenses

Nothing to report.

Other Products

We continue to maintain a large biospecimen repository with associated clinical and demographic data, which serves as a rich resource for the scientific community. We will continue to utilize this resource to generate scientific results, validated diagnostics, and prediction models that should make an impact on the clinical management of patients with prostate cancer. Application procedures for external scientists to access the repository are available on our website: https://canarypass.org

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name:	Daniel Lin, MD
Project Role:	Principal Investigator
Researcher Identifier (e.g. ORCID ID):	ORCID: 0000-0002-2135-1534
Nearest person month worked:	1.7 person months
Contribution to Project:	As Principal Investigator, Dr. Lin oversaw the execution of the project, including interactions with industry collaborators and the FDA. He directed overall scientific activities including data collection, interpretation, and manuscript preparation. Dr. Lin took a central role in the analysis of all data from the project, collaborating with the other investigators on manuscript preparations.
Funding Support:	N/A

Name:	Hilary Boyer
Project Role:	Research Scientist
Researcher Identifier (e.g. ORCID ID):	N/A
Nearest person month worked:	5.2 person months
Contribution to Project:	Ms. Boyer worked under the direction of Dr. Newcomb to receive, annotate, and track PASS specimens from the Central Repository. Ms. Boyer was responsible for pulling, tracking, and documenting specimens sent to collaborating sites and coordinating all shipping activities. She also assisted in specimen and clinical data QA and QC, in monitoring study progress, and in preparing reports for study investigators.
Funding Support:	N/A

Name:	Anna Faino, MS
Project Role:	Statistical Research Associate
Researcher Identifier (e.g. ORCID ID):	N/A
Nearest person month worked:	8.3 person months
Contribution to Project:	Ms. Faino worked under the supervision of Dr. Zheng and was responsible for the extensive data analysis involved in this project. She participated in study consultation with project investigators and the data operations group on data and database forms. Under Dr. Zheng's supervision she performed data analyses, data interpretation and manuscript preparation.
Funding Support:	N/A

Name:	Suzanne Kolb, MPH
Project Role:	Project Coordinator
Researcher Identifier (e.g. ORCID ID):	ORCID: 0000-0002-6443-644X
Nearest person month worked:	6.0 person months
Contribution to Project:	Ms. Kolb worked under the direction of Drs. Lin and Newcomb to fulfill daily fiscal and administrative functions of the program. She monitored subaward budgets and provided logistical support. Ms. Kolb worked closely with the PASS Deputy Director to maintain IRB files, material transfer agreements, and other regulatory documents as well as tracking project timelines and deliverables.
Funding Support:	N/A

Name:	Jesse McKenney, MD
Project Role:	Principal Investigator of Partner Award
Researcher Identifier (e.g. ORCID ID):	N/A
Nearest person month worked:	2 person months
Contribution to Project:	Dr. McKenney is the lead pathologist for this project, overseeing all aspects of the central pathology review. He has worked on development of the Centralized Pathology Review system, and leads the group of study pathologists who review all endpoints for PASS participants. He ensures that pathologic review is timely and follows project guidelines.
Funding Support:	Cleveland Clinic Foundation cost sharing

Name:	Lisa Newcomb, PhD
Project Role:	Deputy Director
Researcher Identifier (e.g. ORCID ID):	ORCID: 0000-0003-3505-3754
Nearest person month worked:	6.0 person months
Contribution to Project:	Dr. Newcomb facilitated the day-to-day operations of all aspects of the research, interfacing with the PASS Study to ensure high quality data and specimens. She worked closely with Dr. Lin and all investigators and collaborators in the execution of the project. Dr. Newcomb was responsible for specimen selection, management of the acquisition and distribution of specimens from the biorepository, as well as overseeing regulatory requirements and supervising study staff.
Funding Support:	N/A

Name:	Maria Tretiakova, MD, PhD
Project Role:	Co-investigator, Pathologist
Researcher Identifier (e.g. ORCID ID):	ORCID: 0000-0002-0819-9638
Nearest person month worked:	2.4 person months
Contribution to Project:	Dr. Tretiakova was responsible for reviewing slides of prostate needle biopsies and characterizing the pathologic parameters such as Gleason score and amount of cancer. She also worked with co-investigators at FHCRC and Cleveland Clinic on study design, data analysis, and interpretation.
Funding Support:	N/A

Name:	Lawrence True, MD
Project Role:	Pathologist
Researcher Identifier (e.g. ORCID ID):	ORCID: 0000-0002-8621-9569
Nearest person month worked:	1.2 person month
Contribution to Project:	Dr. True was responsible for reviewing slides of prostate needle biopsies and characterizing the pathologic parameters such as Gleason score and amount of cancer. He worked with co-investigators at FHCRC and Cleveland Clinic on study design, data analysis, and interpretation.
Funding Support:	N/A

Name:	Yingye Zheng, PhD
Project Role:	Co-investigator, Biostatistician
Researcher Identifier (e.g. ORCID ID):	ORCID: 0000-0002-3078-4200
Nearest person month worked:	1.7 person months
Contribution to Project:	Dr. Zheng was responsible for all statistical aspects of this project, including design and analysis. She consulted with investigators on study designs and necessary design modifications if necessary during the course of the study. She ensured that appropriate data items are collected for valid data analyses and QA/QC to be conducted to ensure high quality of clinical and assay data. She also supervised the SRA in data analyses and interpretation of study data.
Funding Support:	N/A

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

N/A, this is a final report.

What other organizations were involved as partners?

Organization Name: University of Washington
Location of Organization: Seattle, WA
Partner's contribution to the project:

Facilities: Staff (Drs. Lin, True, Tretiakova) used facilities provided by the University of Washington for pathology review and office space.

Collaboration: University of Washington personnel provide expertise in pathology (Drs. Tretiakova and True) and study oversight (Dr. Newcomb).

Organization Name: Cleveland Clinic
Location of Organization: Cleveland, OH
Partner's contribution to the project:

Facilities: Dr. McKenney uses facilities provided by the Cleveland Clinic for central pathology review.

Collaboration: Dr. McKenney provides expertise for central pathology review.

Organization Name: Genomic Health, Inc.
Location of Organization: Redwood City, CA
Partner's contribution to the project:

Collaboration: Genomic Health, Inc. performs Prostate OncotypeDx assays free of charge and discussed design of project.

Organization Name: OPKO Diagnostics
Location of Organization: Miami, FL
Partner's contribution to the project:

Collaboration: OPKO Diagnostics performed blood kallikrein assays free of charge and discussed design of project.

Organization Name: Hologic GenProbe
Location of Organization: San Diego, CA
Partner's contribution to the project:

Collaboration: Hologic GenProbe performed the PCA3 and TMPRSS2:ERG urine marker assays free of charge.

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS:
For this project, Dr. Daniel Lin is the initiating PI and Dr. Jesse McKenney is the partnering PI. Drs. Lin and McKenney are independently submitting a duplicate annual project report, with tasks clearly marked with the responsible PI and research site as requested.

QUAD CHARTS: Not applicable.

9. APPENDICES

Appendix 1.
Published paper: “Evaluating the four kallikrein panel of the 4Kscore for prediction of high-grade prostate cancer in men in the Canary Prostate Active Surveillance Study (PASS).”
Lin et al. *European Urology* 2017: 72(3) 448-454.

Appendix 2.
Accepted paper: “Performance of PCA3 and TMPRSS2:ERG urinary biomarkers in prediction of biopsy outcome in the Canary Prostate Active Surveillance Study (PASS)”.
Newcomb et al. *Prostate Cancer and Prostatic Diseases*, Forthcoming 2018.

Appendix 3.
Published paper: “Refined analysis of prostate specific antigen kinetics to predict prostate cancer active surveillance outcomes.”
Cooperberg et al. *European Urology*. 2018 Aug;74(2):211-217.

Appendix 4.
Published paper: “The role of surveillance biopsy with no cancer as a prognostic marker for reclassification: Results from the Canary Prostate Active Surveillance Study (PASS).”
Kearns et al. *European Urology.* 2018 May;73(5):706-712.

Appendix 5.
Published paper: “Continued Five-Alpha Reductase Inhibitor Use After Prostate Cancer Diagnosis and the Risk of Reclassification and Adverse Pathological Outcomes in the Canary Prostate Active Surveillance Study (PASS).”
Kearns et al. *Journal of Urology*, Epub 2018 Aug 1. Advance online publication, doi: 10.1016/j.juro.2018.07.065.

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

European Association of Urology

Prostate Cancer

Evaluating the Four Kallikrein Panel of the 4Kscore for Prediction of High-grade Prostate Cancer in Men in the Canary Prostate Active Surveillance Study

Daniel W. Lin[a,b,c,*], *Lisa F. Newcomb*[a,b], *Marshall D. Brown*[a], *Daniel D. Sjoberg*[d], *Yan Dong*[e], *James D. Brooks*[f], *Peter R. Carroll*[g], *Matthew Cooperberg*[g], *Atreya Dash*[c], *William J. Ellis*[b], *Michael Fabrizio*[h], *Martin E. Gleave*[i], *Todd M. Morgan*[j], *Peter S. Nelson*[a], *Ian M. Thompson*[k], *Andrew A. Wagner*[l], *Yingye Zheng*[a],

for the Canary Prostate Active Surveillance Study Investigators

[a] *Fred Hutchinson Cancer Research Center, Seattle, WA, USA;* [b] *University of Washington, Seattle, WA, USA;* [c] *Veterans Affairs Puget Sound Health Care System, Seattle, WA, USA;* [d] *Memorial Sloan-Kettering Cancer Center, New York, NY, USA;* [e] *OPKO Labs, Nashville, TN, USA;* [f] *Stanford University, Stanford, CA, USA;* [g] *University of California at San Francisco, San Francisco, CA, USA;* [h] *Eastern Virginia Medical School, Norfolk, VA, USA;* [i] *University of British Columbia, Vancouver, Canada;* [j] *University of Michigan, Ann Arbor, MI, USA;* [k] *University of Texas Health Sciences Center at San Antonio, San Antonio, TX, USA;* [l] *Beth Israel Deaconess Medical Center, Boston, MA, USA*

Article info

Article history:
Accepted November 11, 2016

Associate Editor:
Giacomo Novara

Keywords:
Prostatic neoplasms
Prospective studies
Active surveillance
Biomarker
Kallikrein

Abstract

Background: Diagnosis of Gleason 6 prostate cancer can leave uncertainty about the presence of undetected aggressive disease.
Objective: To evaluate the utility of a four kallikrein (4K) panel in predicting the presence of high-grade cancer in men on active surveillance.
Design, setting, and participants: Plasma collected before the first and subsequent surveillance biopsies was assessed for 718 men prospectively enrolled in the multi-institutional Canary PASS trial. Biopsy data were split 2:1 into training and test sets. We developed statistical models that included clinical information and either the 4Kpanel or serum prostate-specific antigen (PSA).
Outcome measurements and statistical analysis: The endpoint was reclassification to Gleason $\geq$7. We used receiver operating characteristic (ROC) curve analyses and area under the curve (AUC) to assess discriminatory capacity, and decision curve analysis (DCA) to report clinical net benefit.
Results and limitations: Significant predictors for reclassification were 4Kpanel (odds ratio [OR] 1.54, 95% confidence interval [CI] 1.31–1.81) or PSA (OR 2.11, 95% CI 1.53–2.91), $\geq$20% cores positive (OR 2.10, 95% CI 1.33–3.32), two or more prior negative biopsies (OR 0.19, 95% CI 0.04–0.85), prostate volume (OR 0.47, 95% CI 0.31–0.70), and body mass index (OR 1.09, 95% CI 1.04–1.14). ROC curve analysis comparing 4K and base models indicated that the 4Kpanel improved accuracy for predicting reclassification (AUC 0.78 vs 0.74) at the first surveillance biopsy. Both models performed comparably for prediction of reclassification at subsequent biopsies (AUC 0.75 vs 0.76). In DCA, both models showed higher net benefit compared to biopsy-all and biopsy-none strategies. Limitations include the single cohort nature of the study and the small numbers; results should be validated in another cohort before clinical use.

* Corresponding author. Department of Urology, University of Washington, 1959 NE Pacific Street, Box 356510, Seattle, WA 98195, USA. Tel. +1 206 2210797; Fax: +1 206 5433272.
E-mail address: dlin@uw.edu (D.W. Lin).

http://dx.doi.org/10.1016/j.eururo.2016.11.017

Conclusions: The 4Kpanel provided incremental value over routine clinical information in predicting high-grade cancer in the first biopsy after diagnosis. The 4Kpanel did not add predictive value to the base model at subsequent surveillance biopsies.
Patient summary: Active surveillance is a management strategy for many low-grade prostate cancers. Repeat biopsies monitor for previously undetected high-grade cancer. We show that a model with clinical variables, including a panel of four kallikreins, indicates the presence of high-grade cancer before a biopsy is performed.

1. Introduction

Active surveillance is a management strategy for low-grade, localized prostate cancer that allows men to delay or be spared the potential morbidities of treatment. Cancers that appear to be low-risk at diagnosis are monitored, typically with serial prostate-specific antigen (PSA) measurements, clinical examinations, and repeat prostate biopsies. Intervention is recommended on evidence of a more aggressive tumor, usually based on changes in biopsy characteristics.

However, fear of occult high-grade cancer, in part because of the known undersampling of systematic prostate biopsies, has tempered widespread adoption of active surveillance. Even with emerging magnetic resonance imaging (MRI)–based biopsy protocols, there remains uncertainty surrounding the presence of more aggressive disease against a background of apparently low-risk cancer. In addition, the optimal surveillance schedule and triggers for intervention have not been established, resulting in substantial variations in the practice of active surveillance. Prostate biopsy can be painful, anxiety-provoking, expensive, and potentially morbid, so avoiding unnecessary surveillance biopsies is attractive. Methods to reduce the number of biopsies in active surveillance regimens, while maximizing the identification of high-grade cancers that may benefit from treatment, would have substantial clinical utility.

A promising approach to determine active surveillance candidacy and surveillance regimens (eg, more intensive vs less intensive biopsy schedules) involves the addition of biomarker panels to prediction models based on known clinical and demographic variables [1]. Among men suspected of having prostate cancer, a panel of four kallikreins (total PSA [tPSA], free PSA [fPSA], intact PSA [iPSA], and human kallikrein 2 [hK2]) combined with age using a mathematical algorithm improves the prediction of high-grade cancers compared to the PCPT risk calculator or models using tPSA alone [2,3]. Here, we explore the utility of prediction models incorporating the predefined four kallikrein panel algorithm (4Kpanel) to predict the presence of occult high-grade disease in men already diagnosed with Gleason 6 cancer and on active surveillance. We use plasma specimens and data from the prospective, multi-institutional Canary Prostate Active Surveillance Study (PASS).

2. Patients and methods

2.1. Study cohort

This study included men from Canary PASS, a multicenter, prospective study enrolling men on active surveillance [4]. Participants in PASS consented to specimen collection as part of the PASS protocol (clinicaltrials.gov NCT00756665), which was approved by institutional review boards at participating sites. The PASS protocol includes monitoring at clinic visits every 6 mo, with the first ≥10-core prostate needle biopsy at 6–12 mo, the second at 24 mo after cancer diagnosis, and subsequent biopsies every 2 yr. Specimens, including EDTA plasma, were collected at study entry and every 6-mo clinic visit, and were stored at −70 °C until use.

In February 2015, 1170 participants were enrolled in PASS at nine sites throughout North America. Of these, 956 participants had an on-study biopsy, of whom 877 had Gleason 3 + 3 disease at study entry, 771 had not used 5α-reductase inhibitors, and EDTA plasma collected before biopsy was available for 753 men. Participants with missing prostate volume or ratio of positive to total biopsy cores were excluded from the modeling ($n = 35$); the remaining 718 men, who had 1111 biopsies, were included in this study.

2.2. Laboratory methods

Blood was collected in K_2EDTA vacutainers, inverted, centrifuged at 1600 × *g*, and frozen at −70 °C within 4 h of collection. Frozen plasma was stored until shipment on dry ice to OPKO Labs (Nashville, TN, USA) for analysis. The analysis laboratory was blinded to all specimen and clinical information. Specimens were thawed immediately before analysis. tPSA, fPSA, iPSA, and hK2 were measured [2].

2.3. Study design and analyses

The objective of the analyses was to determine whether a model using clinical predictors and kallikrein data collected after diagnosis of Gleason 6 cancer, but before surveillance biopsy, can predict high-grade cancer in the surveillance biopsy. Sequential surveillance biopsies were considered as two groups: (1) the initial biopsy after cancer diagnosis (sometimes called confirmatory biopsy) and (2) all subsequent surveillance biopsies. Biopsy data were split 2:1 into training and test sets matched by outcome.

The primary outcome was reclassification from Gleason score 6 to Gleason score ≥7. A value for the 4Kpanel was calculated with tPSA, fPSA, iPSA, hk2, and age using locked down coefficients developed before the study was conducted [3]. This combination of the four kallikreins is the same as in the commercial 4Kscore. However, the commercial 4Kscore is a model containing the 4Kpanel and clinical data available before cancer diagnosis, and is calibrated for a patient before diagnosis. Because we evaluated the kallikreins in a cohort already diagnosed with cancer, we developed a new model that included the 4Kpanel and clinical information available after a diagnosis of cancer, and calibrated to an active surveillance population. Additional clinical predictors considered in modeling included age, body mass index (BMI), race (African American or other), digital rectal examination (DRE) results, number of previous biopsies after diagnosis, number of negative biopsies after diagnosis, core ratio (ratio of biopsy cores containing cancer to total cores) from previous biopsy, maximum core ratio among all previous biopsies, months since diagnosis, and prostate volume (prostate size measured closest to the time of sampling and imputed within 2 yr).

Please cite this article in press as: Lin DW, et al. Evaluating the Four Kallikrein Panel of the 4Kscore for Prediction of High-grade Prostate Cancer in Men in the Canary Prostate Active Surveillance Study. Eur Urol (2016), http://dx.doi.org/10.1016/j.eururo.2016.11.017

Table 1 – Characteristics for 478 participants with kallikreins assayed before the initial surveillance biopsy after diagnosis for combined Gleason score <7 versus ≥7 for the training and test cohorts

Characteristics	Training set			Test set		
	Gleason <7	Gleason ≥7	*p* value	Gleason <7	Gleason ≥7	*p* value
Sample size (*n*)	259	60		125	34	
Age at diagnosis (yr)	63 (58–67)	64 (60–68)	0.109	64 (58–68)	64 (57–67)	0.876
Body mass index (kg/m^2)	27 (25–30)	28 (25–33)	0.116	27 (25–29)	28 (26–31)	0.305
Race						
Non–African American	248 (96)	56 (93)		121 (97)	29 (85)	
African American	11 (4)	4 (7)	0.646	4 (3)	5 (15)	0.522
Time from diagnosis (mo)	12.0 (8.4–14.1)	12.7 (8.6–14.8)	0.237	12.2 (8.8–14.0)	12.6 (10.3–17.6)	0.189
Digital rectal examination						
Normal	238 (92)	55 (92)		118 (94)	30 (88)	
Abnormal	21 (8)	5 (8)	0.971	7 (6)	4 (12)	0.031
Prostate volume (cm^3)	41.0 (30.0–56.5)	35.5 (25.0–50.0)	0.041	40.0 (30.0–51.0)	30.0 (24.0–42.8)	0.006
Positive:total core ratio	0.08 (0.08–0.17)	0.17 (0.08–0.20)	<0.001	0.08 (0.08–0.17)	0.17 (0.17–0.25)	<0.001
Clinical serum PSA (ng/ml)	4.60 (2.91–6.40)	4.81 (4.35–6.42)	0.108	4.56 (3.11–6.24)	5.65 (4.58–7.88)	0.024
4Kpanel (logit)	0.21 (0.08–0.29)	0.32 (0.16–0.44)	<0.001	0.20 (0.07–0.28)	0.36 (0.18–0.53)	<0.001

PSA = prostate-specific antigen.
Data are presented as median (interquartile range) for continuous variables and as n (%) for categorical variables.

Either the 4Kpanel (logit scale) or clinical serum PSA (log-transformed) was used in models. Prediction models were built using data in the training set, and then clinical performance was assessed using the testing set. We followed the principles set forth by the US Food and Drug Administration critical path initiative, using an established biomarker with analytic validity for the intent of clinical validation in the intended use population [7]. Furthermore, we followed reporting recommendations for tumor marker prognostic studies (REMARK) [8] and the Tumor Marker Utility Grading System [9] in reporting the clinical utility of the biomarker panel.

2.3.1. Model building

Data from initial and subsequent biopsy groups were combined for model development. Interaction terms between biopsy group (initial vs subsequent surveillance biopsy) and other variables were evaluated to investigate whether effects may differ for an initial biopsy and a subsequent biopsy. Logistic regression was used to fit the models, with robust variance to account for the correlation among multiple biopsies on the same patient. Forward stepwise model selection procedures were implemented. Variable selection criteria included $p < 0.15$, area under the receiver operating characteristic(ROC) curve(AUC) ≥0.005, or quasi-likelihood under the independence model criterion (QIC) with threshold of zero [5]. Final models were compared to identify variables that were robust to selection procedures. We first identified a full model including clinical predictors and 4Kpanel, and then a base model with serum PSA substituted for the 4Kpanel. In some clinics, prostate volume may not be reliably available, so models without prostate volume were fitted sequentially.

2.3.2. Model validation

Calibration plots were used to gauge the goodness of fit of each model. We used ROC analyses and AUC to assess the discriminatory capacity of a model for separating patients with and without reclassification. Decision curve analysis (DCA) was used to report the clinical net benefit of each model compared to biopsy-all and biopsy-none strategies [6]. The potential clinical impact was illustrated by plotting the number of cancers missed versus the number of biopsies avoided per 1000 individuals. To illustrate the clinical consequence of each model, we report the number of biopsies that could be avoided and the number of Gleason ≥7 cancers that might be missed if a risk-based threshold is applied as a criterion for biopsy. All evaluations were conducted on the initial biopsy and subsequent biopsy groups separately and combined. Confidence intervals (CIs) and significance tests were calculated using the bootstrap resampling procedure to account for within-subject correlations. All analyses were conducted using R version 3.1.1 (www.r-project.org).

3. Results

Of the 718 men in this study, there were 478 participants in the initial biopsy group for whom kallikreins were assayed: 319 in the training set (60 [18.8%] with Gleason ≥7) and 159 in the test set (34 [21.4%] with Gleason ≥7; Table 1). In bivariate analyses, prostate volume, ratio of positive to total cores, and the 4Kpanel were significantly associated with grade reclassification. There were 444 participants (of whom 204 were also in the initial biopsy group) with 633 subsequent surveillance biopsies, 422 in the training set (70 [17%] with Gleason ≥7; Table 2) and 211 in the test set (31 [15%] with Gleason ≥7; Supplementary Table 1). Biopsies in this group ranged from the second to eighth after diagnosis, and most patients had Gleason score 6 or no cancer at their surveillance biopsies, varying slightly across biopsy number.

In the full clinical model (Table 3) including the 4Kpanel, significant predictors for reclassification were BMI (odds ratio [OR] 1.09, 95% CI 1.04–1.14], >20% of cores positive in the prior biopsy (OR 2.10, 95% CI 1.33–3.32), a history of two or more biopsies negative for cancer (OR 0.19, 95% CI 0.04–0.85), prostate volume (per fold increase, OR 0.47, 95% CI 0.31–0.70), and 4Kpanel (OR 1.5, 95% CI 1.31–1.81). In the clinical model with serum PSA replacing the 4Kpanel, PSA was significantly associated with reclassification (per fold increase, OR 2.11, 95% CI 1.53–2.91) and age was not. In models that did not include prostate volume, the effects were similar for covariates left in the model (Supplementary Table 2). Model calibration in the test set showed predicted probabilities of reclassification closely matching the empirical rates (Supplementary Fig. 1).

Please cite this article in press as: Lin DW, et al. Evaluating the Four Kallikrein Panel of the 4Kscore for Prediction of High-grade Prostate Cancer in Men in the Canary Prostate Active Surveillance Study. Eur Urol (2016), http://dx.doi.org/10.1016/j.eururo.2016.11.017

Table 2 – Biopsy characteristics at each sequential surveillance biopsy after diagnosis for 558 participants in the training set

Parameter	Initial biopsy	Subsequent surveillance biopsies						
	First	Second	Third	Fourth	Fifth	Sixth	Seventh	Eighth
Biopsies (*n*)	319	246	108	34	20	10	3	1
CR for previous biopsy[a]								
Median (IQR)	0.08 (0.08)	0.07 (0.17)	0.08 (0.17)	0.06 (0.12)	0.06 (0.12)	0 (0.07)	0.11 (0.06)	0 (0)
Missing, *n* (%)	0	5 (2)	5 (5)	0	0	0	0	0
Median MCR[b] (IQR)	0.08 (0.08)	0.11 (0.08)	0.13 (0.15)	0.17 (0.13)	0.10 (0.17)	0.14 (0.15)	0.17 (0.08)	0.17 (0.00)
Negative biopsies[c], *n* (%)								
0	319 (100)	145 (59)	44 (41)	10 (29)	4 (20)	1 (10)	1 (33)	0
1	0	101 (41)	38 (35)	13 (38)	6 (30)	3 (30)	2 (67)	0
2	0	0	26 (24)	6 (18)	3 (15)	1 (10)	0	1 (100)
3	0	0	0	5 (15)	2 (10)	3 (30)	0	0
4	0	0	0	0	5 (25)	2 (20)	0	0
Median PV, cm^3 (IQR)	41.0 (26.5)	38.0 (27.0)	41.0 (27.0)	48.5 (25.0)	59.5 (36.5)	43.5 (27.8)	41.0 (19.5)	97.0 (0.0)
Biopsy GS, *n* (%)								
Negative	107 (34)	95 (39)	38 (35)	11 (32)	8 (40)	6 (60)	2 (67)	0
6	152 (48)	108 (44)	48 (45)	21 (62)	10 (50)	3 (30)	1 (33)	1 (100)
7	58 (18)	42 (17)	21 (19)	2 (6)	2 (10)	1 (10)	0	0
8	1 (0)	1 (0)	1 (1)	0	0	0	0	0
9	1 (0)	0	0	0	0	0	0	0

CR = core ratio; IQR = interquartile range; MCR = maximum CR; PV = prostate volume; GS = Gleason score.
[a] CR is defined as the number of biopsy cores containing cancer divided by the total number of biopsy cores in the previous biopsy.
[b] MCR among all previous biopsies.
[c] Number of surveillance biopsies in which no cancer was found.

Table 3 – Summary of fitted models including clinical variables + serum PSA or 4Kpanel in the training set

Variable	PSA + full clinical model		4K + full clinical model	
	OR (95% CI)	*p* value	OR (95% CI)	*p* value
Age	1.03 (1.00–1.06)	0.068		
Body mass index	1.11 (1.06–1.16)	<0.001	1.09 (1.04–1.14)	<0.001
Positive ore ratio >0.2	2.19 (1.39–3.44)	0.001	2.10 (1.33–3.32)	0.001
Negative biopsies ≥2	0.19 (0.04–0.80)	0.023	0.19 (0.04–0.85)	0.029
Log(prostate volume)	0.31 (0.20–0.48)	<0.001	0.47 (0.31–0.70)	<0.001
Log(PSA)	2.11 (1.53–2.91)	<0.001		
4Kpanel			1.54 (1.31–1.81)	<0.001

PSA = prostate-specific antigen; OR = odds ratio; CI = confidence interval.

ROC curve analysis (Table 4, Supplementary Fig. 2) comparing the full model with the 4Kpanel and the full clinical model with serum PSA indicated that the 4Kpanel significantly improved the accuracy for predicting reclassification (AUC 0.78 vs 0.74) in the initial surveillance biopsy, with a significant incremental value in AUC of 0.04 (95% CI 0.003–0.09). In a model without prostate volume, the incremental value in AUC was 0.07 (95% CI 0.02–0.11). The 4Kpanel did not improve prediction of reclassification in subsequent biopsies relative to PSA (AUC 0.75 vs 0.76).

Similar findings were observed in DCA. Compared to a clinical model with serum PSA, the model with 4Kpanel showed a higher net benefit for the initial surveillance biopsy, but there was no benefit for subsequent biopsies. All models showed substantial gain in net benefit compared with the biopsy-all and biopsy-none strategies across

Table 4 – Results of final regression models for reclassification

Base model	Area under the curve (95% confidence interval)		
	4K + clinical model	PSA + clinical model	Difference
Full clinical model			
Initial biopsy	0.783 (0.691–0.871)	0.740 (0.652–0.828)	0.043 (0.003–0.086)
Subsequent biopsy	0.754 (0.657–0.838)	0.755 (0.653–0.841)	−0.001 (−0.037–0.041)
Clinical model without prostate volume			
Initial biopsy	0.748 (0.654–0.840)	0.678 (0.579–0.774)	0.069 (0.016–0.114)
Subsequent biopsy	0.738 (0.633–0.825)	0.718 (0.611–0.810)	0.02 (−0.023–0.07)

PSA = prostate-specific antigen.
Confidence intervals were calculated with bootstrap accounting for correlations among individuals.

Please cite this article in press as: Lin DW, et al. Evaluating the Four Kallikrein Panel of the 4Kscore for Prediction of High-grade Prostate Cancer in Men in the Canary Prostate Active Surveillance Study. Eur Urol (2016), http://dx.doi.org/10.1016/j.eururo.2016.11.017

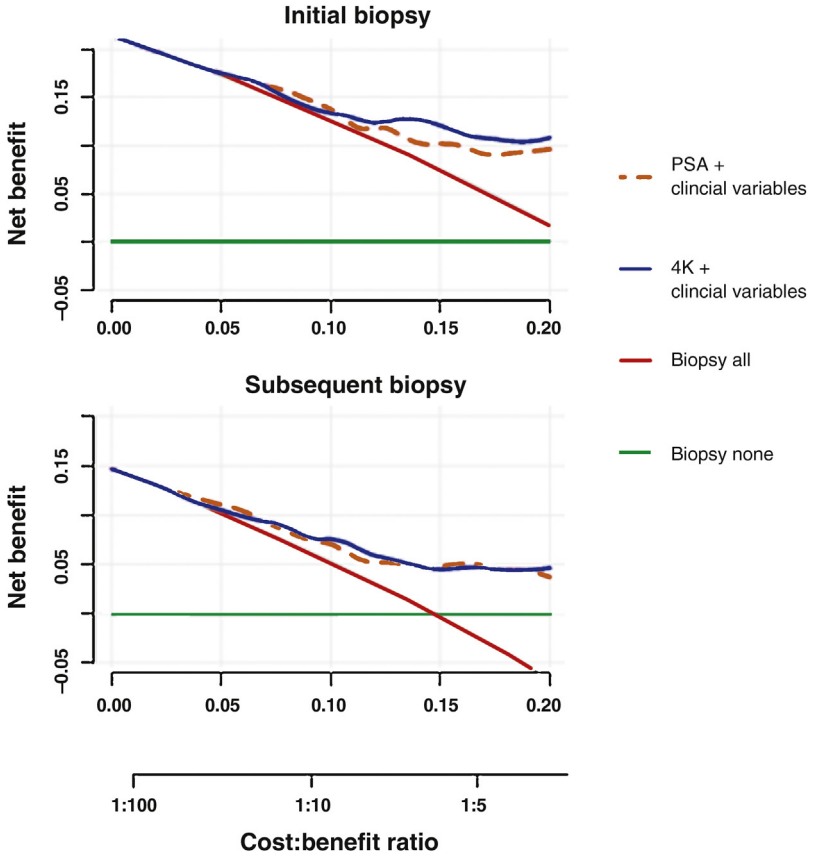


Fig. 1 – Decision curve analysis for full models with serum Prostate-specific antigen (PSA) or with the 4Kpanel. Strategies for biopsying all men (biopsy all) or no men (biopsy none) are also shown. The line with the highest net benefit at any particular threshold probability for biopsy (*x*-axis) will yield the best clinical results.

a range of plausible cost and benefit ratios (Fig. 1 and Supplementary Fig. 3).

The clinical consequences, or the number of biopsies and the number of high-grade cancers that could be avoided or delayed per 1000 patients, were illustrated based on prediction models with the 4Kpanel or PSA (Table 5). For example, using a model with the 4Kpanel and a clinical rule of only performing an initial surveillance biopsy in patients whose risk of high-grade cancer exceeded 10%, 252 biopsies would be avoided, 19 of which would contain high-grade cancer as defined by any pattern 4 disease, and zero biopsies with primary Gleason 4. Comparing the two models at the same numbers of biopsies avoided (Supplementary Fig. 4) shows that the 4K model appears to miss fewer higher-grade cancers while avoiding the same number of initial biopsies.

4. Discussion

In this study using a prospectively enrolled multi-institutional cohort of men on active surveillance, we show that addition of a panel of four kallikrein markers to a model that includes clinical information can significantly improve prediction of the outcome in the first surveillance biopsy. Both models performed comparably for prediction of reclassification in subsequent biopsies. Importantly, in DCA both models showed a higher net benefit compared to biopsy-all and biopsy-none strategies. Lastly, we showed that the 4Kpanel added to currently available clinical metrics and how the results impact clinical management.

There is a growing body of evidence that true Gleason 6 prostate cancer is indolent and will not cause harm if left untreated [10–12]. This knowledge is balanced by the known undersampling in prostate needle biopsies, and while some have advocated that select Gleason 3 + 4 cancers may undergo surveillance, level 1 clinical trial data and treatment guidelines generally recommend treatment of higher-grade cancers, including Gleason 3 + 4 disease [13,14]. Our efforts focus on developing tools for use after diagnosis of Gleason 6 prostate cancer to provide a higher degree of certainty that no occult high-grade cancer was missed at diagnosis. More accurate tools would not only support the practice of active surveillance but could also promote less intensive monitoring regimens.

A panel of four kallikreins, when combined in a mathematical algorithm, improves the prediction of newly diagnosed high-grade (Gleason $\geq$7) cancer [3]. This panel of markers also improved long-term prediction of metastatic disease among men with PSA $\geq$2 in a Swedish cohort [15]. In this study, we asked whether the same panel of markers [3] improved the prediction of high-grade disease in surveillance biopsies of men already diagnosed with Gleason

Please cite this article in press as: Lin DW, et al. Evaluating the Four Kallikrein Panel of the 4Kscore for Prediction of High-grade Prostate Cancer in Men in the Canary Prostate Active Surveillance Study. Eur Urol (2016), http://dx.doi.org/10.1016/j.eururo.2016.11.017

Table 5 – Clinical consequences showing the number of biopsies that could be avoided for initial surveillance biopsy or subsequent surveillance biopsy

HGC probability	Biopsies		High-grade cancers		Primary Gleason 4 cancers	
	Performed	Avoided	Found	Missed	Found	Missed
Initial surveillance biopsy						
Biopsy all	1000	0	214	0	44	0
Initial biopsy: risk by clinical variables + PSA						
>5%	943 (896–970)	57 (30–104)	214 (157–284)	0 (0–24)	44 (21–88)	0 (0–24)
>10%	761 (689–821)	239 (179–311)	201 (146–270)	13 (3–45)	44 (21–88)	0 (0–24)
>15%	509 (432–586)	491 (414–568)	164 (114–229)	50 (26–96)	38 (17–80)	6 (1–35)
Initial biopsy: risk by clinical variables + 4K						
>5%	956 (912–979)	44 (21–88)	214 (157–284)	0 (0–24)	44 (21–88)	0 (0–24)
>10%	748 (676–809)	252 (191–324)	195 (141–263)	19 (6–54)	44 (21–88)	0 (0–24)
>15%	522 (445–598)	478 (402–555)	182 (130–250)	31 (14–71)	44 (21–88)	0 (0–24)
Subsequent surveillance biopsies						
Biopsy all	1000	0	147	0	47	0
Risk by clinical variables + PSA						
>5%	844 (789–886)	156 (114–211)	147 (105–201)	0 (0–18)	47 (26–85)	0 (0–18)
>10%	692 (627–750)	308 (250–373)	133 (93–185)	14 (5–41)	43 (23–79)	5 (1–26)
>15%	445 (380–513)	555 (487–620)	109 (74–158)	38 (19–73)	43 (23–79)	5 (1–26)
Risk by clinical variables + 4K						
>5%	848 (794–890)	152 (110–206)	142 (101–196)	5 (1–26)	47 (26–85)	0 (0–18)
>10%	654 (588–715)	346 (285–412)	133 (93–185)	14 (5–41)	47 (26–85)	0 (0–18)
>15%	408 (344–475)	592 (525–656)	100 (66–147)	47 (26–85)	38 (19–73)	9 (3–34)

HGC = high-grade cancer.
Results are presented as the number (95% confidence interval) per 1000 men.

6 cancer. We found that when the kallikreins were assessed before the initial surveillance biopsy (sometimes called the confirmatory biopsy), the 4Kpanel provided incremental benefit for prediction of high-grade cancer (Gleason ≥7) over the clinical factors that are available at diagnosis. Specifically, depending on the choice from the various cutpoints that are based on the risk of high-grade disease, a substantial number of biopsies could be avoided while minimizing the number of missed high-grade cancers, few of which had primary pattern 4. The 4Kpanel was not of value over PSA for the prediction of reclassification in subsequent biopsies after the first surveillance biopsy. We found that the impact of other biopsy information, primarily volume of core involvement in previous biopsies and the number of previous negative biopsies, carries such a statistical weight in modeling that the impact of the 4Kpanel is minimized. For example, if a patient had low-volume disease at the initial surveillance biopsy or had subsequent negative biopsies after the initial diagnosis, then these factors were highly protective against biopsy reclassification at subsequent biopsy. It should be noted that our analysis of these subsequent biopsies used the 4Kpanel from the plasma sample that was closest to the subsequent biopsy, not necessarily the plasma sample from study entry, which could be months or years earlier than the subsequent biopsy.

We included serum PSA and prostate volume separately in our models instead of calculating PSA density, as we find a better model fit when the variables enter the model independently. While transurethral ultrasound prostate volume measurements may suffer from imprecision [16], statistical models that included prostate volume appeared to provide slightly improved predictive performance (AUC for all groups 0.77 with volume vs 0.75 without volume). Furthermore, prostate volume is a strong predictor of finding higher-grade cancers, with larger prostates being protective, as previously reported [17].

This study has limitations that merit mention. First, the model was developed and tested in the same cohort and with relatively limited numbers that resulted in wide confidence intervals and minor differences between the training and test sets. The results should clearly be validated in other cohort before clinical application. However, we expect that our results will be similar to those found in a community setting, as PASS is a multicenter center study that represents a broad spectrum of men utilizing active surveillance. Similarly, as PASS is primarily a Caucasian cohort, the findings of this study may not be generalizable to African American patients. Another limitation is that the serum PSA measurements used were obtained as part of standard clinical care, and the local site assays may differ from the one used with the 4Kpanel. Thus, the comparative modeling using PSA versus 4Kpanel may have slightly different tPSA values, with caution suggested for comparisons between the models. Lastly, as the use of imaging such as multiparametric MRI (mpMRI) is increasing, we do not have MRI data for most of our participants and recognize the potential value of future studies incorporating results from mpMRI and biomarkers in active surveillance.

5. Conclusions

The 4Kpanel was significantly associated with reclassification at the first surveillance biopsy, providing incremental

Please cite this article in press as: Lin DW, et al. Evaluating the Four Kallikrein Panel of the 4Kscore for Prediction of High-grade Prostate Cancer in Men in the Canary Prostate Active Surveillance Study. Eur Urol (2016), http://dx.doi.org/10.1016/j.eururo.2016.11.017

value over routine clinical information, and the 4K model performed significantly better than the base model in this group. The 4Kpanel did not add predictive value to a PSA clinical model for biopsy decision-making for men at subsequent surveillance biopsies. This work aims to provide clinical validation of a biomarker that will help determine those men who have or will develop aggressive prostate cancer, allowing for the accurate determination of those men who may avoid or delay the burden of immediate treatment safely, while concurrently identifying men who may optimally benefit from early treatment.

Author contributions: Daniel W. Lin 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: Lin, Newcomb, Sjoberg, Dong, Zheng.
Acquisition of data: Lin, Newcomb, Brooks, Carroll, Cooperberg, Dash, Ellis, Fabrizio, Gleave, Morgan, Nelson, Thompson, Zheng.
Analysis and interpretation of data: Lin, Newcomb, Brown, Sjoberg, Dong, Nelson, Thompson, Zheng.
Drafting of the manuscript: Lin, Newcomb, Brown, Sjoberg, Dong, Nelson, Thompson, Zheng.
Critical revision of the manuscript for important intellectual content: Lin, Newcomb, Brown, Sjoberg, Dong, Brooks, Carroll, Cooperberg, Dash, Ellis, Fabrizio, Gleave, Morgan, Nelson, Thompson, Zheng.
Statistical analysis: Brown, Sjoberg, Dong, Zheng.
Obtaining funding: Lin, Newcomb, Nelson, Zheng.
Administrative, technical, or material support: Lin, Nelson.
Supervision: Newcomb, Lin, Nelson.
Other: None.

Financial disclosures: Daniel W. Lin 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: Yan Dong is an employee of OPKO. The remaining authors have nothing to disclose.

Funding/Support and role of the sponsor: This work was supported by the Department of Defense (PC130355), the Canary Foundation, and the Institute for Prostate Cancer Research. The sponsors played a role in the design and conduct of the study, and collection, management, and analysis of the data.

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.2016.11.017.

References

[1] 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.
[2] Parekh DJ, Punnen S, Sjoberg DD, et al. A multi-institutional prospective trial in the USA confirms that the 4Kscore accurately identifies men with high-grade prostate cancer. Eur Urol 2015;68:464–70.
[3] Bryant RJ, Sjoberg DD, Vickers AJ, et al. Predicting high-grade cancer at ten-core prostate biopsy using four kallikrein markers measured in blood in the ProtecT study. J Natl Cancer Inst 2015;107(7).
[4] Newcomb LF, Thompson Jr IM, Boyer HD, et al. Outcomes of active surveillance for clinically localized prostate cancer in the prospective, multi-institutional Canary PASS cohort. J Urol 2016;195:313–20.
[5] Pan W. Akaike's information criterion in generalized estimating equations. Biometrics 2001;57:120–5.
[6] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.
[7] US Food and Drug Administration. Innovation or stagnation: challenge and opportunity on the critical path to new medical products. Washington, DC: US Department of Health and Human Services; 2004.
[8] McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst 2005;97:1180–4.
[9] Hayes DF, Bast RC, Desch CE, et al. Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. J Natl Cancer Inst 1996;88:1456–66.
[10] Ross HM, Kryvenko ON, Cowan JE, Simko JP, Wheeler TM, Epstein JI. Do adenocarcinomas of the prostate with Gleason score (GS) $\leq$6 have the potential to metastasize to lymph nodes? Am J Surg Pathol 2012;36:1346–52.
[11] Eggener SE, Badani K, Barocas DA, et al. Gleason 6 prostate cancer: translating biology into population health. J Urol 2015;194:626–34.
[12] Carter HB, Partin AW, Walsh PC, et al. Gleason score 6 adenocarcinoma: should it be labeled as cancer? J Clin Oncol 2012;30:4294–6.
[13] Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. N Engl J Med 2012;367:203–13.
[14] Mohler JL, Kantoff PW, Armstrong AJ, et al. Prostate cancer, version 1.2014. J Natl Compr Canc Netw 2013;11:1471–9.
[15] Stattin P, Vickers AJ, Sjoberg DD, et al. Improving the specificity of screening for lethal prostate cancer using prostate-specific antigen and a panel of kallikrein markers: a nested case-control study. Eur Urol 2015;68:207–13.
[16] Rodriguez Jr E, Skarecky D, Narula N, Ahlering TE. Prostate volume estimation using the ellipsoid formula consistently underestimates actual gland size. J Urol 2008;179:501–3.
[17] Roobol MJ, van Vugt HA, Loeb S, et al. Prediction of prostate cancer risk: the role of prostate volume and digital rectal examination in the ERSPC risk calculators. Eur Urol 2012;61:577–83.

Prostate Cancer and Prostatic Diseases
Accepted for publication 12 Dec 2018
Manuscript # PCAN-18-0427R

Performance of PCA3 and TMPRSS2:ERG urinary biomarkers in prediction of biopsy outcome in the Canary Prostate Active Surveillance Study (PASS)

Lisa F. Newcomb PhD[1,2], Yingye Zheng PhD[3], Anna V. Faino MS[3], Daniella Bianchi-Frias PhD[4], Matthew R. Cooperberg MD MPH[5,6], Marshall D. Brown MS[3], James D. Brooks MD[7], Atreya Dash MD[8], Michael D. Fabrizio MD[9], Martin E. Gleave MD[10], Michael Liss MD[11], Todd M. Morgan MD[12], Ian M. Thompson MD[13], Andrew A. Wagner MD[14], Peter R. Carroll MD[5], Peter S. Nelson MD[4*], and Daniel W. Lin MD[1,2*]

Author affiliations:

[1] Cancer Prevention Program, Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle WA
[2] Department of Urology, University of Washington, Seattle WA
[3] Biostatistics Program, Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle WA
[4] Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle WA
[5] Department of Urology, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco CA
[6] Department of Epidemiology & Biostatistics, University of California, San Francisco CA
[7] Department of Urology, Stanford University, Stanford CA
[8] Department of Urology, VA Puget Sound Health Care System, Seattle, WA
[9] Department of Urology, Eastern Virginia Medical School, Virginia Beach VA
[10] Department of Urologic Sciences, University of British Columbia, Vancouver BC
[11] Department of Urology, University of Texas Health Sciences Center, San Antonio TX
[12] Department of Urology, University of Michigan, Ann Arbor MI
[13] CHRISTUS Medical Center Hospital, San Antonio TX
[14] Division of Urology, Beth Israel Deaconess Medical Center, Boston MA

Running title: PCA3, TMPRSS2:ERG in prostate cancer surveillance

Keywords: PCA3, TMPRSS2:ERG, biomarkers, prostate cancer, active surveillance

Financial Support: Department of Defense Prostate Cancer Research Program (Grants W81XWH1110489 and W81XWH1410595); National Institutes of Health R01 CA181605; Canary Foundation, Institute for Prostate Cancer Research

*Correspondence:
Daniel W. Lin, MD
Department of Urology
University of Washington
1959 NE Pacific Street, Box 356510
Seattle, WA 98195
dlin@uw.edu
phone: 206-221-0797

fax: 206-543-3272

Peter Nelson, MD
Division of Human Biology
Fred Hutchinson Cancer Research Center
1100 Fairview Avenue
Seattle, WA 98109
pnelson@fredhutch.org
phone: 206-667-3377
fax: 206-667-2917

Word count: 2998
Tables and Figures: 5 total (3 tables and 2 figures)

The authors declare no potential conflicts of interest.

ABSTRACT

Background:
For men on active surveillance for prostate cancer, biomarkers may improve prediction of reclassification to higher grade or volume cancer. This study examined the association of urinary PCA3 and TMPRSS2:ERG (T2:ERG) with biopsy-based reclassification.

Methods:
Urine was collected at baseline, 6, 12, and 24 months in the multi-institutional Canary Prostate Active Surveillance Study (PASS), and PCA3 and T2:ERG levels were quantitated. Reclassification was an increase in Gleason score or ratio of biopsy cores with cancer to ≥34%. The association of biomarker scores, adjusted for common clinical variables, with short-term and long-term reclassification was evaluated. Discriminatory capacity of models with clinical variables alone or with biomarkers was assessed using receiver operating characteristic (ROC) curves and decision curve analysis (DCA).

Results:
782 men contributed 2,069 urine specimens. After adjusting for PSA, prostate size and ratio of biopsy cores with cancer, PCA3 but not T2:ERG was associated with short-term reclassification at the first surveillance biopsy (OR=1.3; 95% CI 1.0-1.7, p=0.02). The addition of PCA3 to a model with clinical variables improved area under the curve from 0.743 to 0.753 and increased net benefit minimally. After adjusting for clinical variables, neither marker, nor marker kinetics, was associated with time to reclassification in subsequent biopsies.

Conclusions:
PCA3 but not T2:ERG was associated with cancer reclassification in the first surveillance biopsy, but has negligible improvement over clinical variables alone in ROC or DCA analyses. Neither marker was associated with reclassification in subsequent biopsies.

INTRODUCTION

Active surveillance (AS) is a management strategy for many clinically localized prostate cancers that allows men to delay or be spared the potential morbidities of treatment.[1] Cancers that appear relatively low-risk at diagnosis are monitored, typically with regular clinical exams, serial prostate-specific antigen (PSA) measurements, and repeat prostate biopsies. However, uncertainty about the possibility of occult aggressive cancer is one factor tempering widespread adoption of AS.[2,3] Furthermore, optimal surveillance schedules and triggers for intervention have not yet been established,[1] resulting in substantial variation in the practice of AS.[4] Biomarkers that are collected non-invasively and that improve the prediction of aggressive behavior could improve the utilization of AS and inform decisions about the intensity of surveillance regimens.

During AS, treatment is usually recommended when higher grade or volume disease is found by biopsy. Biomarkers that detect the presence of occult high grade or high volume disease or that predict future reclassification to high grade or volume cancer could have substantial clinical utility. Importantly, biomarkers should incrementally improve upon models that are based on commonly available clinical variables.

In the present study, we evaluated PCA3 and TMPRSS2:ERG (T2:ERG) mRNA, which are two prostate-specific biomarkers present in urine,[5,6] for their potential utility in AS. Clinical grade assays have been developed for both of these biomarkers,[6,7] and both have demonstrated ability to improve on PSA and clinical factors in detection of prostate cancer.[6,8] Here, we evaluated whether PCA3 and T2:ERG assayed at multiple time-points during AS associated with reclassification to high grade or high volume cancer. We assessed if these biomarkers were

associated with reclassification at the next surveillance biopsy. We also evaluated if PCA3 and T2:ERG collected early in or serially during AS associated with time to reclassification in later surveillance biopsies.

MATERIALS AND METHODS

Study Population

The multi-center Canary Prostate Active Surveillance Study (PASS) enrolls men diagnosed with clinically localized prostate cancer who chose to use AS to manage their cancer. Men provide informed consent under institutional review board supervision at nine centers (clinicaltrials.gov NCT00756665). Under the PASS protocol, PSA is measured every 3 months, clinic visits occur every 6 months, and ultrasound-guided biopsies are performed 6-12 months and 24 months after diagnosis, and then every 2 years. Specimens, including post-DRE urine, are collected at study entry and every 6 months. Follow-up data was collected September 2008 to February 2017. Men were included in the analysis if they had Gleason score ≤3+4 and <34% ratio of biopsy cores with cancer to total cores collected prior to urine collection. Men were excluded if they enrolled in PASS >5 years after diagnosis or if they had no on-study biopsy for endpoint determination.

Biomarker Collection

Urine samples were collected, assays performed, and biomarker scores calculated as described previously.[9] For this analysis, all urine specimens from study entry, 6, 12, and 24 month visits that were available in July 2014 were assayed.

Statistical Models for Reclassification

Reclassification was defined as an increase in primary or secondary Gleason grade at biopsy and/or an increase in the biopsy cores with cancer to total cores collected (cores ratio) to ≥34%.[10] Several models of reclassification were considered, as depicted in the Study Schematic (Figure 1).

Modeling short-term biopsy reclassification (Figure 1, panels A and B)

The association between urine biomarkers collected immediately prior to a biopsy and reclassification at the biopsy was modeled using logistic regression. The analysis was stratified by the first surveillance biopsy (sBx1, sometimes called confirmatory biopsy; Figure 1, Scenario A; n=552) and subsequent surveillance biopsies (Figure 1, Scenario B; n=446).

Modeling time to reclassification with longitudinal biomarkers (Figure 1, panel C)

The association between PCA3 or T2:ERG and the time to biopsy reclassification was modeled using a partly conditional Cox proportional hazards (PH) model.[11] Participants were excluded from this analysis if they reclassified at the sBx1. Each participant had the first urine biomarkers assayed after diagnosis and prior to sBx1 and had up to 3 additional urine samples collected up to 2 years after the first. The first urine biomarker sample along with the urine biomarker kinetics at each observation time were covariates in the partly conditional Cox model (Figure 1, panel C; n=405). Participants without reclassification were censored at date of last study contact, treatment, or 2 years after their last biopsy, whichever came first.

Urine biomarker kinetics were calculated based on a linear mixed effect model (LMEM), in which the natural log of the urine biomarkers was modeled as a linear function of time since diagnosis, with a random intercept indicating the individual-specific ln(urine biomarker) at diagnosis, and a random slope reflecting individual-specific rate of change over time. A PCA3 or

T2:ERG kinetics (PCA3k or T2:ERGk) value for each participant based on the first urine sample up to an observation time was then derived based on the best linear unbiased predictor (BLUP) estimator from the LMEM.[11] Intra-class correlation (ICC) was calculated to determine the proportion of total variability in biomarker scores explained by between-participant variability.

For statistical models, the natural log (ln) of the urine biomarkers was calculated as ln(PCA3) and, due to possible values of 0 in T2:ERG, ln(T2:ERG + 1). We also adjusted for, as appropriate, most recent ln(PSA), ln[time since diagnosis (in years)], ln(prostate size), ln[maximum ratio of positive to total biopsy cores (cores ratio)], diagnostic Gleason score, number of prior biopsies, number of prior negative biopsies, cT-stage (T1a-c versus T2a-c), BMI (obese, overweight or normal), race (Caucasian, African American or other), ethnicity (Hispanic versus non-Hispanic/other), age at diagnosis, and family history of PCa. Insignificant clinical variables were backwards eliminated based on a p-value cutoff of 0.05. Robust variance estimators were used to account for multiple biopsies or urine specimens within participant, where appropriate (panels B and C).

Evaluation of clinical performance

If either PCA3 or T2:ERG was significant in a model, model performance was compared between a clinical model containing clinical and biopsy variables, and a model with the urine biomarkers added. Model performance was assessed with receiver operating characteristic (ROC) curves, area under the curve (AUC). Confidence intervals for model performance assessments were obtained by calculating 1000 bootstrap samples. Decision curve analysis (DCA) was used to report the clinical net benefit of each model compared to biopsy-all and

biopsy-none strategies.[12] All analyses were performed with SAS version 9.4 and R version 3.4.1. Code is available upon request.

RESULTS

There were 782 participants included with a median follow-up of 4.9 years (IQR: 3.5, 6.5) among censored participants. Urine specimens were collected at six to twelve month increments such that 627 participants contributed at least 2 specimens, 448 contributed at least 3 specimens, and 209 contributed 4 specimens. There were 552 participants who had urine collected prior to the sBx1 (Figure 1). Median age was 63, PSA was 4.8, prostate size was 40 cc, 94% were initially diagnosed with Gleason 3+3 cancer, and the median ratio of cores containing cancer to total biopsy cores (cores ratio) was 8.3% (Table 1). Results of each analysis depicted in Figure 1 are described below.

Analysis of reclassification at next biopsy

Of the 552 men with urine biomarkers assessed prior to the sBx1 (Figure 1, panel A), 130 (24%) were reclassified at that biopsy. In a logistic regression model adjusted for PSA, cores ratio, and prostate size, PCA3 score was associated with reclassification in the sBx1 (OR = 1.3; 95% CI: 1.0-1.7), and T2:ERG score was not (Table 2). A model in which the endpoint was only Gleason grade reclassification was similar (data not shown).

There was a small change in the AUC of a model for predicting reclassification at sBx1 using clinical variables plus PCA3 versus a model with only clinical variables: 0.753 [95% CI 0.707-0.800] vs 0.743 [0.693-0.791] with and without PCA3 respectively (Figure 2). No improvement

in AUC was found for a model including T2:ERG (Figure 2). Similar findings were observed in DCA (Figure 2). A model with clinical variables and PCA3 showed minimal increase in net benefit relative to a model with only clinical variables. All models with clinical variables showed an improvement to PCA3 alone, and all models showed an improvement to biopsy-all and biopsy-none strategies.

In the 446 men with urine biomarkers assessed prior to subsequent surveillance biopsies (Figure 1, panel B), 85 (19%) reclassified at the biopsy immediately following biomarker assessment (Supplemental Table 1). In a logistic regression model adjusted for clinical variables, neither PCA3 nor T2:ERG were associated with reclassification (OR = 1.01; 95% CI: 0.77, 1.32, p = 0.96, and OR = 1.12; 95% CI: 1.00, 1.27, p = 0.06, respectively; Supplemental Table 2).

Analysis of time to reclassification for longitudinal biomarkers

There were 405 participants included in the time-to-event analysis who had their first urine sample collected prior to the sBx1 and who did not reclassify at the sBx1 (Figure 1, panel C). With a median follow-up of 3.6 years from the first urine collection, 103 (25%) participants reclassified at any subsequent surveillance biopsy (Supplemental Table 3).

The annual percent change in PCA3 estimated by LMEM was 9.8 (95% CI 7.3-12.3, p<0.001). As determined by ICC, 85% of the observed variation in PCA3 was explained by between-participant variation, and 15% due to within-participant variation. The annual percent change in T2:ERG was 11.3 (95% CI 5.2-17.8, p<0.001), and 68% of the observed variation was explained by between-participant variation and 32% due to within-participant variation. Biomarker kinetics were calculated based on deriving a BLUP estimator from a LMEM.[11] No significant

differences in slopes were found between participants with reclassification versus those with no reclassification for either biomarker (Supplemental Figure 1).

In a Cox PH model adjusted for time since diagnosis, BMI, prostate size, cores ratio, biopsies since diagnosis (0 vs 1+), negative biopsies since diagnosis (0 vs 1+), and PSA, no significant association was found between baseline PCA3 or T2:ERG and reclassification (HR for PCA3 = 1.16; 95% CI 0.86 – 1.57, p = 0.33, and HR for T2:ERG = 0.92; 95% CI = 0.75 – 1.12, p = 0.40) or PCA3 or T2:ERG kinetics and reclassification (HR for 0.10 increase in PCA3k = 0.96; 95% CI 0.44 - 2.09, p = 0.92, and HR for 0.10 increase in T2:ERGk = 1.56; 95% CI 0.73 - 3.34, p = 0.26) (Table 3).

DISCUSSION

Non-invasive assays for diagnosing or monitoring prostate cancer have the potential to aid treatment decisions and improve clinical outcomes. In this context, biomarkers assayed in urine represent an attractive approach and several urine-based assays have been developed. The Progensa PCA3 assay is a commercially available, analytically-validated diagnostic test that has been FDA approved to inform biopsy decision making in men with no known cancer and a previous negative biopsy. PCA3 is a prostate-specific non-coding mRNA and has been shown in many studies to improve predictive accuracy for cancer on initial biopsy,[7,8,13,14] and to be correlated with more aggressive cancer at prostatectomy.[15,16] At the time we initiated this work, the T2:ERG assay had been analytically validated[6] and was being developed as a commercial assay. Thus, we hypothesized that both biomarker assays could improve management of patients using AS.

In this report from a multi-center contemporary AS cohort, we evaluated the association between urinary PCA3 and T2:ERG and biopsy reclassification using urine collected at multiple times during surveillance. The study was designed according to PRoBE criteria,[17] and analyses were tailored to help inform varying decisions made during AS. After adjusting for clinical and biopsy variables, we observed a significant association of PCA3 with reclassification at the sBx1, but only a modest improvement in AUC was found between a model with clinical variables only and a model with clinical variables plus PCA3. Similarly, in decision curve analysis minimal improvement in net benefit was observed when PCA3 was added to models. We also found no association between either baseline PCA3 or T2:ERG and time to reclassification, and no association between changes in the biomarker scores over time and time to reclassification.

Our motivation for performing this study was that biomarkers that improve discrimination of indolent cancers and more aggressive tumors will not only support the practice of AS but also promote less intensive biopsy regimes. However, to optimize patient management and clinical utility, biomarkers should improve upon existing information. Multimodal risk assessment approaches that combine several sources of information into a risk score have been developed, such as the Prostate Cancer Prevention Trial (PCPT) Risk Calculator, which is used prior to a diagnosis of cancer to predict the risk of finding high grade cancer in the next biopsy,[18] and PCA3 has been shown to improve the net benefit of the PCPT Risk Calculator.[19] Several commercial biomarker panels employ such a multimodal strategy for predicting the risk of finding high grade cancer prior to a diagnosis of cancer, including the 4Kscore[20] and more recently the Select score.[21] Because men who are candidates for AS have already been diagnosed with cancer and have available more clinical information than prior to diagnosis, risk models that include information from the index biopsy have been developed for the AS

setting.[22,23] These models utilize commonly available clinical variables and provide utility in risk prediction while on AS.

Studies evaluating the use of PCA3 in AS have been limited and sample sizes have been small.[24,25] In the current study, which is the largest study to date of PCA3 in men using AS, after adjustment for clinical variables available after cancer diagnosis, we found a significant association of PCA3 with reclassification at the sBx1 (adjusted OR = 1.3, p = 0.02), but not for subsequent biopsies (adjusted OR = 1.01, p = 0.96). Although we found no association between T2:ERG and biopsy reclassification, some studies have suggested improved performance when PCA3 and T2:ERG are used in combination[26] or combined into a MiPS score for the initial diagnosis of PCa.[27] We thus combined PCA3 and T2:ERG into a MIPS score, but found little or no improvement over PCA3 alone (data not shown).

In ROC curve analysis of models for predicting reclassification at sBx1, addition of PCA3 to a model that contained PSA, prostate size, and ratio of positive to total biopsy cores improved the AUC of the model very minimally. However, in a clinical setting, predictions are not being made over the full range of sensitivity and specificity. The likely use for a biomarker measured after a diagnosis of low-risk prostate cancer would be to “rule out” men who are at very low risk of harboring high grade or volume cancer, allowing them to delay having a biopsy. Thus we examined specificity at 95% sensitivity. The addition of PCA3 to a clinical model increased specificity at 95% sensitivity (0.156 vs 0.095), but the difference was not significant. Similarly, a slight improvement was seen in the net benefit of proceeding to biopsy for models including PCA3 at about 15% risk threshold, but at other risk thresholds addition of PCA3 to clinical variables made no improvement in the benefit of a decision to proceed to biopsy.

We also evaluated whether the urinary biomarkers, alone or in combination, improved upon clinical variables for prediction of time to future reclassification. Although PCA3 alone was associated with time to reclassification, when incorporated into a multivariable model neither PCA3 or T2:ERG were.

Using the same model, we evaluated if changes in PCA3 or T2:ERG scores measured over time (biomarker kinetics) were associated with reclassification. We used samples collected prior to the sBx1, and at 6 month intervals up to 2 years, and employed an analytic strategy that allowed the models to account for each individual biomarker measurement while utilizing information from the general trend across all participants and accommodating for random variability in the biomarkers. Although PCA3 increased over time, and both PCA3 score and T2:ERG score were higher in men who reclassified than those who did not, there were no significant differences in the slopes of either biomarker over time for event versus non-event participants (Supplemental Figure 1), and we found no association between biomarker kinetics and reclassification. Our results are consistent with the one other longitudinal study of PCA3 in a smaller, more uniform risk cohort, suggesting that longitudinal PCA3 measurements do not add value over a single PCA3 measurement.[25]

We found a surprisingly large amount of variability in the longitudinal samples, in particular for T2:ERG. To assess the variation statistically, we used ICC, which can be interpreted as the proportion of total variability explained by between-participant variability. The variation in longitudinal measurements of PCA3 is very similar to that of PSA[11] : for both PCA3 and PSA, 15% of the variability can be attributed to random variability. However, 32% of the variability in T2:ERG was attributed to noise.

This study has limitations. Our design does not allow for us to address if the observed variability is due to biology, specimen collection methods, or assay performance. Second, although this is the largest study of urine biomarkers in AS published, the sample size is somewhat modest. However, while we expect that a larger sample size may reduce confidence intervals, we do not expect that it would result in different conclusions. Another limitation may be the reliability of our endpoint. Biopsy reclassification is imperfect in that it may reflect minimal changes in the tumor that may have little clinical importance. Nonetheless, our definition is consistent with those used in most AS cohorts, and importantly, drives treatment decisions in contemporary clinical practice. We did evaluate biomarker scores in only the 94% of men who were diagnosed with 3+3 disease and found no difference in results. Similarly, we evaluated biomarker scores in the 31 men who reclassified to 4+3 or higher at sBx1, and found no difference from the scores in the men who reclassified to 3+4 (data not shown).

In conclusion, we found that PCA3 was associated with reclassification at sBx1 in a multivariable model, but PCA3, or PCA3 and T2:ERG together, demonstrated minimal improvement to the clinical utility of a multivariable model. Neither PCA3 or T2:ERG was associated with time to reclassification at subsequent biopsies. Overall, we found that these markers add little or no improvement over clinical variables in predicting biopsy reclassification during AS.

ACKNOWLEDGEMENTS

The authors would like to thank the large and dedicated PASS team, including the coordinators and coordinating center staff. We also thank Scott Tomlins and Jack Groskopf for helpful comments and the research team at Hologic GenProbe for running the biomarker assays. Importantly, the authors also thank all of the men who have participated in PASS.

REFERENCES

1. Chen RC, Rumble RB, Loblaw DA, Finelli A, Ehdaie B, Cooperberg MR, 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(18):2182-2190.

2. Ganz PA, Barry JM, Burke W, Col NF, Corso PS, Dodson E, et al. National Institutes of Health State-of-the-Science Conference: Role of Active Surveillance in the Management of Men With Localized Prostate Cancer. *Ann Intern Med.* 2012;156(8):591-595.

3. Cooperberg MR, Carroll PR. Trends in Management for Patients With Localized Prostate Cancer, 1990-2013. *JAMA.* 2015;314(1):80-82.

4. Luckenbaugh AN, Auffenberg GB, Hawken SR, Dhir A, Linsell S, Kaul S, et al. Variation in Guideline Concordant Active Surveillance Followup in Diverse Urology Practices. *J Urol.* 2017;197(3 Pt 1):621-626.

5. Bussemakers MJ, van Bokhoven A, Verhaegh GW, Smit FP, Karthaus HF, Schalken JA, et al. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res.* 1999;59(23):5975-5979.

6. Tomlins SA, Aubin SM, Siddiqui J, Lonigro RJ, Sefton-Miller L, Miick S, et al. Urine TMPRSS2:ERG Fusion Transcript Stratifies Prostate Cancer Risk in Men with Elevated Serum PSA. *Sci Transl Med.* 2011;3(94):94ra72.

7. Groskopf J, Aubin SM, Deras IL, Blase A, Bodrug S, Clark C, et al. APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis of prostate cancer. *Clin Chem.* 2006;52(6):1089-1095.

8. Wei JT, Feng Z, Partin AW, Brown E, Thompson I, Sokoll L, et al. Can urinary PCA3 supplement PSA in the early detection of prostate cancer? *J Clin Oncol.* 2014;32(36):4066-4072.

9. Lin DW, Newcomb LF, Brown EC, Brooks JD, Carroll PR, Feng Z, et al. Urinary TMPRSS2:ERG and PCA3 in an Active Surveillance Cohort: Results from a Baseline Analysis in the Canary Prostate Active Surveillance Study. *Clin Cancer Res.* 2013;19(9):2442-2450.

10. Newcomb LF, Thompson IM, Jr., Boyer HD, Brooks JD, Carroll PR, Cooperberg MR, et al. Outcomes of Active Surveillance for Clinically Localized Prostate Cancer in the Prospective, Multi-Institutional Canary PASS Cohort. *J Urol.* 2016;195(2):313-320.

11. Cooperberg MR, Brooks JD, Faino AV, Newcomb LF, Kearns JT, Carroll PR, et al. Refined Analysis of Prostate-specific Antigen Kinetics to Predict Prostate Cancer Active Surveillance Outcomes. *Eur Urol.* 2018;74(2):211-217.

12. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* 2006;26(6):565-574.

13. Chevli KK, Duff M, Walter P, Yu C, Capuder B, Elshafei A, et al. Urinary PCA3 as a predictor of prostate cancer in a cohort of 3,073 men undergoing initial prostate biopsy. *J Urol.* 2014;191(6):1743-1748.

14. Deras IL, Aubin SM, Blase A, Day JR, Koo S, Partin AW, et al. PCA3: a molecular urine assay for predicting prostate biopsy outcome. *J Urol.* 2008;179(4):1587-1592.

15. Whitman EJ, Groskopf J, Ali A, Chen Y, Blase A, Furusato B, et al. PCA3 score before radical prostatectomy predicts extracapsular extension and tumor volume. *J Urol.* 2008;180(5):1975-1978; discussion 1978-1979.

16. Nakanishi H, Groskopf J, Fritsche HA, Bhadkamkar V, Blase A, Kumar SV, et al. PCA3 molecular urine assay correlates with prostate cancer tumor volume: implication in selecting candidates for active surveillance. *J Urol.* 2008;179(5):1804-1809; discussion 1809-1810.

17. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *J Natl Cancer Inst.* 2008;100(20):1432-1438.

18. Ankerst DP, Hoefler J, Bock S, Goodman PJ, Vickers A, Hernandez J, et al. Prostate Cancer Prevention Trial risk calculator 2.0 for the prediction of low- vs high-grade prostate cancer. *Urology.* 2014;83(6):1362-1367.

19. Ankerst DP, Goros M, Tomlins SA, Patil D, Feng Z, Wei JT, et al. Incorporation of Urinary Prostate Cancer Antigen 3 and TMPRSS2:ERG into Prostate Cancer Prevention Trial Risk Calculator. *Eur Urol Focus.* 2018.

20. Parekh DJ, Punnen S, Sjoberg DD, Asroff SW, Bailen JL, Cochran JS, et al. A multi-institutional prospective trial in the USA confirms that the 4Kscore accurately identifies men with high-grade prostate cancer. *Eur Urol.* 2015;68(3):464-470.

21. Van Neste L, Hendriks RJ, Dijkstra S, Trooskens G, Cornel EB, Jannink SA, et al. Detection of High-grade Prostate Cancer Using a Urinary Molecular Biomarker-Based Risk Score. *Eur Urol.* 2016;70(5):740-748.

22. Ankerst DP, Xia J, Thompson IM, Jr., Hoefler J, Newcomb LF, Brooks JD, 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(6):1083-1088.

23. Lin DW, Newcomb LF, Brown MD, Sjoberg DD, Dong Y, Brooks JD, et al. Evaluating the Four Kallikrein Panel of the 4Kscore for Prediction of High-grade Prostate Cancer in Men in the Canary Prostate Active Surveillance Study. *Eur Urol.* 2017;72(3):448-454.

24. Tosoian JJ, Loeb S, Kettermann A, Landis P, Elliot DJ, Epstein JI, et al. Accuracy of PCA3 measurement in predicting short-term biopsy progression in an active surveillance program. *J Urol.* 2010;183(2):534-538.

25. Tosoian JJ, Patel HD, Mamawala M, Landis P, Wolf S, Elliott DJ, et al. Longitudinal assessment of urinary PCA3 for predicting prostate cancer grade reclassification in favorable-risk men during active surveillance. *Prostate Cancer Prostatic Dis.* 2017;20(3):339-342.

26. Sanda MG, Feng Z, Howard DH, Tomlins SA, Sokoll LJ, Chan DW, et al. Association Between Combined TMPRSS2:ERG and PCA3 RNA Urinary Testing and Detection of Aggressive Prostate Cancer. *JAMA Oncol.* 2017;3(8):1085-1093.

27. Tomlins SA, Day JR, Lonigro RJ, Hovelson DH, Siddiqui J, Kunju LP, et al. Urine TMPRSS2:ERG Plus PCA3 for Individualized Prostate Cancer Risk Assessment. *Eur Urol.* 2016;70(1):45-53.

Figure Legends

Figure 1. Study design. For short-term prediction, analysis of the association of biomarkers with reclassification in the biopsy immediately following urine collection was performed using logistic regression for A.) 552 men with urine assayed prior to the first surveillance biopsy (sBx1) and B.) 446 men with urine assayed prior to subsequent surveillance biopsies. For longer-term prediction, the association of biomarkers with time to reclassification was examined in C.) 405 men with urine assayed prior to their sBx1 who did not reclassify at that biopsy; the first urine biomarker sample along with the urine biomarker kinetics at each observation time were considered as covariates in a partly conditional Cox model.

Figure 2. Comparison of model performance from logistic regression model with grade and/or tumor volume reclassification at the first surveillance biopsy (sBx1). A) Receiver operating characteristic curves; dotted line corresponds to 95% sensitivity, B) .Decision curve analysis. Strategies for biopsying all men (grey) and no men (dark green) are also shown. The line with the highest net benefit at any particular risk threshold (x-axis) will yield the best clinical results.

Fig. 1

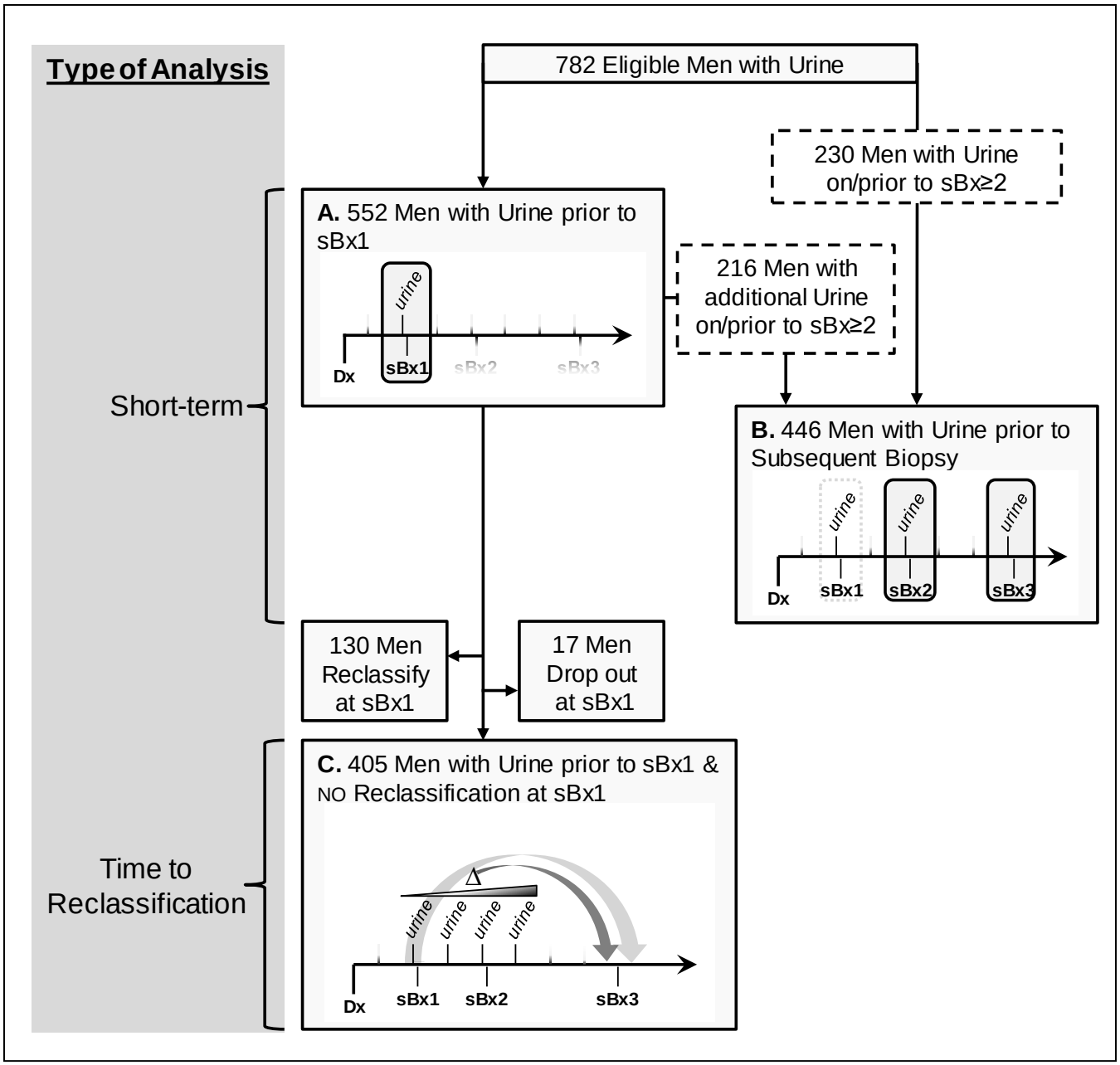
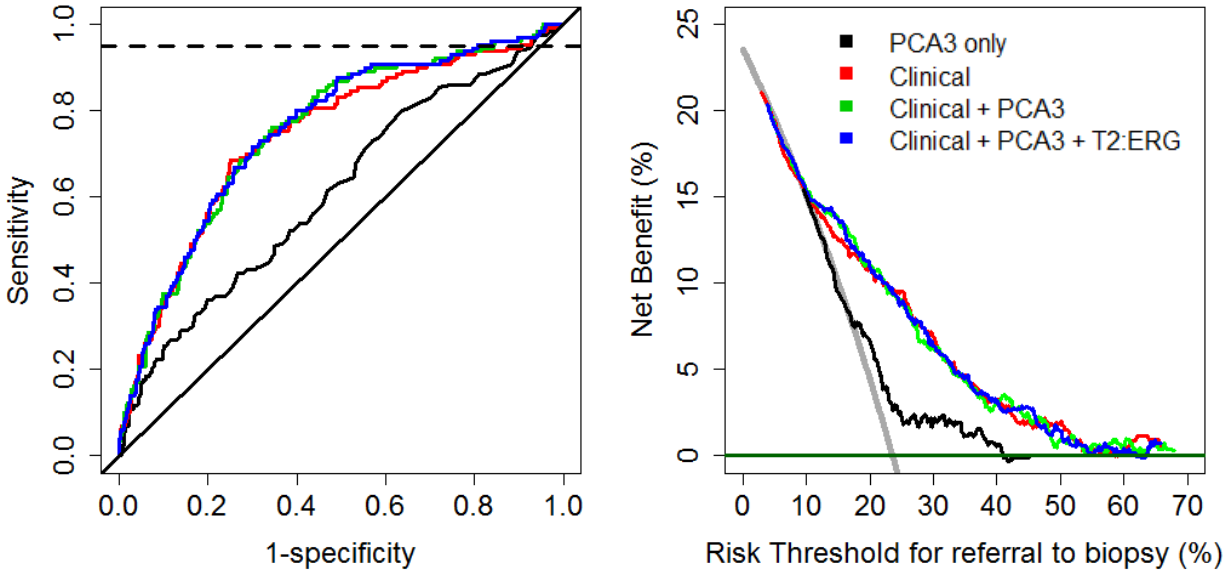

Type of Analysis
782 Eligible Men with Urine
230 Men with Urine on/prior to sBx≥2
A. 552 Men with Urine prior to sBx1
urine
Dx
sBx1
sBx2
sBx3
216 Men with additional Urine on/prior to sBx≥2
B. 446 Men with Urine prior to Subsequent Biopsy
urine
urine
urine
Dx
sBx1
sBx2
sBx3
Short-term
130 Men Reclassify at sBx1
17 Men Drop out at sBx1
C. 405 Men with Urine prior to sBx1 & NO Reclassification at sBx1
Δ
urine
urine
urine
urine
Dx
sBx1
sBx2
sBx3
Time to Reclassification

Fig. 2



Model	AUC (95% CI)^	Difference in AUC (95% CI)^
PCA3 only	0.611 (0.553, 0.668)	-0.132 (-0.212, -0.055)
Clinical Variables Alone	0.743 (0.693, 0.791)	reference
Clinical + PCA3	0.753 (0.707, 0.800)	0.010 (0.0004, 0.030)
Clinical + T2:ERG	0.745 (0.696, 0.798)	0.002 (-0.002, 0.017)
Clinical + PCA3 + T2:ERG	0.754 (0.710, 0.803)	0.011 (0.001, 0.033)

^ 95% confidence interval calculated using 1000 bootstrap samples. Difference calculated as AUC or specificity of clinical variables + biomarker model minus clinical variables only model.

Table 1. Description of participants with urine collected prior to the first surveillance biopsy (sBx1), distributed by outcome at sBx1.

Variable	All Participants n=552	Reclassifiers, n=130	Non-Reclassifiers, n=422
	Median [IQR]	Median [IQR]	Median [IQR]
Age at Dx	63 [58, 67]	63 [58, 69]	63 [58, 67]
Race, n(%)			
Caucasian American	500 (91)	115 (88)	385 (91)
African American	27 (5)	10 (8)	17 (4)
Other	25 (5)	5 (4)	20 (5)
Dx PSA	4.8 [3.8, 6.4]	5.1 [4.3, 6.3]	4.7 [3.6, 6.4]
Prostate size	40 [30, 55]	35 [25, 48]	42 [32, 60]
cT-stage, n(%)			
T1a-T1c	508 (92)	118 (91)	390 (92)
T2a-T2c	44 (8)	12 (9)	32 (8)
Dx Gleason			
3+3	521 (94)	120 (92)	401 (95)
3+4	31 (6)	10 (8)	21 (5)
Dx cores ratio	8.3 [8.3, 16.7]	16.7 [8.3, 24.5]	8.3 [8.3, 16.7]
BMI, n(%)			
Normal	131 (24)	32 (25)	99 (23)
Overweight	283 (51)	63 (48)	220 (52)
Obese	138 (25)	35 (27)	103 (24)
PCA3	32 [18, 61]	39.5 [24, 89]	30 [16, 57]
T2:ERG	14 [2, 57]	27 [1, 82]	13 [2, 53]

Table 2. Logistic regression model results for grade and/or tumor volume reclassification in first surveillance biopsy (n = 552).

Variable*	Univariable		Multivariable	
	OR (95% CI)^	p-value^	OR (95% CI)^	p-value^
PSA	1.5 (1.1, 2.0)	0.01	1.8 (1.3, 2.6)	0.001
Dx Cores Ratio	4.0 (2.6, 6.3)	<.001	3.4 (2.1, 5.4)	<.001
Prostate Size	0.3 (0.2, 0.5)	<.001	0.3 (0.1, 0.4)	<.001
PCA3	1.6 (1.2, 1.9)	0.0001	1.3 (1.0, 1.7)	0.02
T2:ERG	1.1 (1.0, 1.2)	0.21	1.0 (0.9, 1.2)	0.52

* The natural log of all variables was used in modeling.

^ Odds ratios, 95% confidence intervals and p-values from logistic regression models.

Table 3. Cox proportional hazards model results for grade and/or tumor volume reclassification using longitudinally collected samples (405 participants, 103 (25%) with event). PCA3k and T2:ERGk refer to the biomarker kinetics, respectively.

Variable	Univariable[Ŧ]		Multivariable	
	HR (95% CI)^	p-value^	HR (95% CI)^	p-value^
Time since Dx*	1.09 (0.95, 1.25)	0.22	1.01 (0.83, 1.23)	0.94
BMI				
Obese vs Normal	1.28 (0.74, 2.22)	0.38	1.80 (1.04, 3.12)	0.04
Overweight vs Normal	0.92 (0.55, 1.52)	0.74	1.21 (0.74, 1.99)	0.44
Prostate Size*	0.49 (0.31, 0.77)	0.002	0.29 (0.17, 0.50)	<.001
Max Prior Cores Ratio (10% increase)	1.68 (1.32, 2.13)	<.001	1.30 (1.00, 1.69)	0.05
Prior Biopsy (1+ vs 0)	1.33 (1.05, 1.68)	0.02	2.09 (1.38, 3.14)	<.001
Prior No Cancer Biopsy (1+ vs 0)	0.47 (0.30, 0.71)	<.001	0.38 (0.23, 0.62)	<.001
PSA*	1.66 (1.28, 2.17)	<.001	2.18 (1.59, 2.99)	<.001
PCA3*	1.57 (1.16, 2.12)	0.003	1.16 (0.86, 1.57)	0.33
PCA3k (0.10 unit increase)	1.62 (0.67, 3.91)	0.28	0.96 (0.44, 2.09)	0.92
T2:ERG*	0.98 (0.79, 1.21)	0.84	0.92 (0.75, 1.12)	0.40
T2:ERGk (0.10 unit increase)	1.40 (0.66, 2.95)	0.38	1.56 (0.73, 3.34)	0.26

* The natural log of all variables was used in modeling.

Ŧ PCA3 and PCA3k entered together; T2:ERG and T2:ERGk entered together.

^ Hazard ratios, 95% confidence intervals and p-values from Cox proportional hazards models.

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

European Association of Urology

Prostate Cancer

Refined Analysis of Prostate-specific Antigen Kinetics to Predict Prostate Cancer Active Surveillance Outcomes

Matthew R. Cooperberg[a,b,*], *James D. Brooks*[c], *Anna V. Faino*[d], *Lisa F. Newcomb*[e,f], *James T. Kearns*[f], *Peter R. Carroll*[a], *Atreya Dash*[f], *Ruth Etzioni*[d], *Michael D. Fabrizio*[g], *Martin E. Gleave*[h], *Todd M. Morgan*[i], *Peter S. Nelson*[j], *Ian M. Thompson*[k], *Andrew A. Wagner*[l], *Daniel W. Lin*[e,f], *Yingye Zheng*[d]

[a] *Department of Urology, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA, USA;* [b] *Department of Epidemiology & Biostatistics, University of California, San Francisco, CA, USA;* [c] *Department of Urology, Stanford University, Stanford, CA, USA;* [d] *Biostatistics Program, Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA;* [e] *Cancer Prevention Program, Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA;* [f] *Department of Urology, University of Washington, Seattle, WA, USA;* [g] *Department of Urology, Eastern Virginia Medical School, Virginia Beach, VA, USA;* [h] *Department of Urologic Sciences, University of British Columbia, Vancouver, BC, Canada;* [i] *Department of Urology, University of Michigan, Ann Arbor, MI, USA;* [j] *Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA, USA;* [k] *CHRISTUS Medical Center Hospital, San Antonio, TX, USA;* [l] *Division of Urology, Beth Israel Deaconess Medical Center, Boston, MA, USA*

Article info

Article history:
Accepted January 19, 2018

Associate Editor:
Giacomo Novara

Keywords:
Prostate-specific antigen
Kinetics
Prostate cancer
Active surveillance
Outcomes

Abstract

Background: For men on active surveillance for prostate cancer, utility of prostate-specific antigen (PSA) kinetics (PSAk) in predicting pathologic reclassification remains controversial.
Objective: To develop prediction methods for utilizing serial PSA and evaluate frequency of collection.
Design, setting, and participants: Data were collected from men enrolled in the multicenter Canary Prostate Active Surveillance Study, for whom PSA data were measured and biopsies performed on prespecified schedules. We developed a PSAk parameter based on a linear mixed-effect model (LMEM) that accounted for serial PSA levels.
Outcome measurements and statistical analysis: The association of diagnostic PSA and/or PSAk with time to reclassification (increase in cancer grade and/or volume) was evaluated using multivariable Cox proportional hazards models.
Results and limitations: A total of 851 men met the study criteria; 255 (30%) had a reclassification event within 5 yr. Median follow-up was 3.7 yr. After adjusting for prostate size, time since diagnosis, biopsy parameters, and diagnostic PSA, PSAk was a significant predictor of reclassification (hazard ratio for each 0.10 increase in PSAk = 1.6 [95% confidence interval 1.2–2.1, $p < 0.001$]). The PSAk model improved stratification of risk prediction for the top and bottom deciles of risk over a model without PSAk. Model performance was essentially identical using PSA data measured every 6 mo to those measured every 3 mo. The major limitation is the reliability of reclassification as an end point, although it drives most treatment decisions.
Conclusions: PSAk calculated using an LMEM statistically significantly predicts biopsy reclassification. Models that use repeat PSA measurements outperform a model incorporating only diagnostic PSA. Model performance is similar using PSA assessed every 3 or 6 mo. If validated, these results should inform optimal incorporation of PSA trends into active surveillance protocols and risk calculators.
Patient summary: In this report, we looked at whether repeat prostate-specific antigen (PSA) measurements, or PSA kinetics, improve prediction of biopsy outcomes in men using active surveillance to manage localized prostate cancer. We found that in a large multicenter active surveillance cohort, PSA kinetics improves the prediction of surveillance biopsy outcome.

* Corresponding author. UCSF Helen Diller Family Comprehensive Cancer Center, Box 1695, 550 16th Street, San Francisco, CA 94143-1695, USA. Tel. +1 415 8853660; Fax: +1 415 8857443.
E-mail address: matthew.cooperberg@ucsf.edu (M.R. Cooperberg).

https://doi.org/10.1016/j.eururo.2018.01.017
0302-2838/

1. Introduction

Given the prolonged natural history and indolent behavior of most low-risk prostate cancers [1], active surveillance (AS) has been developed as an alternative to immediate treatment. Surveillance is now recognized as a preferred strategy for low-risk disease [2], and is offered to a large and growing proportion of men, both in the USA [3,4] and internationally [5]. While substantial variation persists in terms of eligibility criteria for surveillance, follow-up intervals, and triggers for intervention, all AS protocols are based principally on repeated prostate-specific antigen (PSA) measurements and periodic rebiopsy [2].

However, it remains unclear how to collect and interpret serial PSA data optimally in the AS setting. In most centers, PSA is collected quarterly, with the goal of identifying men with a rapid PSA rise, which may signify aggressive disease. However, studies to date have not shown analyses of PSA kinetics to be informative in most cases. In multiple cohorts, PSA kinetics consistently failed to predict reclassification based on biopsy parameters (ie, increase in biopsy Gleason grade and/or tumor volume) [6–8]. In the prospective, multicenter Canary Prostate Active Surveillance Study (PASS), PSA doubling time (PSADT) of <36 mo was originally a criterion for progression, but since consistently few men met this threshold it was dropped from the protocol [9].

Limiting factors in most AS cohorts reporting outcomes are the relatively short duration of follow-up and limited longitudinal PSA data. As the PASS cohort has matured with longer follow-up, additional PSA measurements, and more reclassification events, we have an opportunity to determine the extent to which PSA kinetics might facilitate improved decision making for men on surveillance for low-risk prostate cancer. We also aimed to determine whether quarterly PSA measurements are necessary for accurate assessment of PSA kinetics or whether semiannual measurement would be sufficient.

2. Patients and methods

The Canary PASS is a multicenter, prospective cohort study enrolling men on AS at nine North American centers. Men eligible for AS provide informed consent under institutional review board supervision (clinicaltrials.gov NCT00756665). In PASS, PSA is measured every 3 mo, clinic visits occur every 6 mo, and ultrasound-guided biopsies are performed 6–12 mo after diagnosis, 24 mo after diagnosis, and then every 2 yr. Other tests, including magnetic resonance imaging, are performed at the clinicians' discretion; however, as enrollment started in 2008, the majority of men did not undergo these procedures. For the current study, participants were enrolled before February 2016 and had diagnostic Gleason grade $\leq 3+4$ and <34% of biopsy cores involved with cancer, no history of 5α-reductase inhibitor (5ARI) use, and at least one PSA and one biopsy following diagnosis. The primary outcome was tumor reclassification, defined as an increase in primary or secondary Gleason grade, or an increase in tumor volume to ≥34% of total biopsy cores involved. Tumor risk at diagnosis was summarized using the validated Cancer of the Prostate Risk Assessment (CAPRA) score [10].

2.1. Statistical analysis

PSA may be measured irregularly during AS and is characterized by within-individual random variation, which may attenuate associations between PSA kinetics and clinical outcomes. To study longitudinal PSA measurements as predictors of reclassification while accommodating these complicating factors, a two-stage procedure was used [11,12]. Through this process, we derived a novel PSA kinetic parameter (designated PSAk), which we treated like a biomarker, and our approach conformed to the REMARK criteria for novel biomarkers [13].

First, we calculated PSAk using a linear mixed-effect model (LMEM), in which the natural logarithm of PSA (ln[PSA]) was modeled as a linear function of time since diagnosis, with a random intercept indicating the individual-specific ln(PSA) at diagnosis and a random slope reflecting the individual-specific rate of change over time. PSAk for each participant based on all his PSA measurements from diagnosis to a specific observation time was derived using the best linear unbiased predictor (BLUP) estimator from the LMEM (see the Supplementary material, Methods). Intraclass correlation (ICC) was calculated to assess how much of the variability in PSA was explained by between-participant variance compared with total variance. A high ICC indicates strong correlations among PSA measurements from the same individual.

Two other approaches for calculating PSA kinetics were considered: a linear regression model using all the PSA measurements from diagnosis to an observation time (simple PSAk [PSAkS]), and a slope change using two PSA measurements closest to and including the observation time (restricted simple PSAk [PSAkRS]). Models were adjusted for prostate size.

Second, Cox proportional hazards (PH) models were used to determine the risk of future reclassification as a function of covariates at each observation time. The outcome was defined as time from each PSA measurement to reclassification or censoring. Participants were censored at treatment, last study contact, or 2 yr after biopsy; the latter criterion was included to control for patients who do not undergo ongoing serial biopsies, and therefore may accrue long-term follow-up but do not have the possibility of meeting the reclassification outcome. Individual-specific PSAk at each measurement time estimated from stage 1 was the key covariate. Other covariates considered were the following: age, ln(prostate size), ln(observation time since diagnosis), diagnostic Gleason (3 + 3 or 3 + 4), percent of positive biopsy cores, number of biopsies since diagnosis (0, 1, 2, 3, or 4+), negative biopsy since diagnosis, recent biopsy result (cancer vs no cancer), and ln(diagnostic PSA). Tests for proportionality confirmed that the PH assumptions were valid.

Hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated with robust variance estimates to account for correlations from multiple observations from the same individual. Model fit was compared using the Akaike information criterion (AIC); a smaller AIC indicated better goodness of fit. Nonsignificant variables were backward eliminated using a p value cutoff of 0.05.

To address whether our results were biased by an increase or a decrease in PSAk that influenced the decision to undergo or delay a biopsy, several steps were taken. Timing of each biopsy was defined as "on time," "early," or "late" based on the PASS protocol. Multinomial regression analyses were used to determine whether biopsy timing was associated with PSAk. Three different sensitivity analyses were performed: compliant participants only (all biopsies needed to be compliant to the protocol), compliant biopsies only (only data preceding on-time biopsies were included), and adjusted event or censor time (early and late biopsies were adjusted by a randomly selected time within the "on-time" window). Further details are provided in the Supplementary material.

To assess the performance of the multivariable model incorporating PSAk, the Cox PH model was used to calculate individual risk of having a

Please cite this article in press as: Cooperberg MR, et al. Refined Analysis of Prostate-specific Antigen Kinetics to Predict Prostate Cancer Active Surveillance Outcomes. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.017

reclassification event at 4 yr from 1 yr after diagnosis for each participant, using PSA data from diagnosis to 1 yr. Time-dependent receiver operating characteristic (ROC) curves and areas under the curve (AUCs) were used to quantify the performance of models. Bootstrapping methods were used to obtain 95% CIs for AUCs. ROC curves and AUCs accounted for censoring prior to 4 yr, and compared model-based individual risk with each participant's event or censor outcome. To evaluate the usefulness of PSAk for risk stratification, we categorized the model-based risk as follows: lowest 10%, middle, and highest 10% risk. We compared reclassification-free probabilities among these risk groups using a Kaplan–Meier (KM) analysis. Risk groups were generated using models with and without PSAk.

We ran an additional analysis comparing modeling performance for PSA measured every 6 versus every 3 mo since diagnosis, and assessed whether semiannual and quarterly PSA measurement yielded similar predictions. As an additional sensitivity analysis, we also analyzed PSAk based on the end point of grade reclassification only. A two-sided p value of <0.05 was considered significant for analyses, which were performed using R version 3.3.0.

3. Results

Of 1278 men in PASS, 107 did not yet have a surveillance biopsy, 193 had a history of 5ARI use, 70 did not meet study risk criteria, and 57 had missing data. Thus, 851 (67%) men were included in this analysis. Median (interquartile range [IQR]) follow-up among censored participants was 3.7 (2.4–5.2) yr. Among all participants, 291 (34%) were reclassified by an increase in biopsy Gleason grade or tumor volume to $\geq$34% of total biopsy cores with cancer, of whom 210 (25%) and 255 (30%) were reclassified within 3 and 5 yr of diagnosis, respectively. Of the 291 men, 247 (85%) were reclassified based on Gleason grade and only 44 by the extent of biopsy involvement only. Only 46 men (8%) were censored based on treatment in the absence of progression. The median participant age was 62 yr. Six percent were African Americans and 4% belonged to other non-Caucasian races. Eighty percent of biopsies were per protocol (on time), 10% early, and 9% late. Table 1 summarizes the clinical characteristics of the cohort. Reclassified participants had similar clinical risk to censored participants, as assessed by the CAPRA score ($p = 0.95$); however, compared with censored participants, reclassified participants had smaller prostates, a higher PSA density (PSAD), and a higher proportion of diagnostic biopsy cores involved (all $p < 0.001$).

The annual percent change in PSA estimated by the LMEM was 4.3 (95% CI 3.4–5.2, $p < 0.001$). As determined by the ICC, 85% of the observed variation in PSAk was explained by between-participant variation and 15% by within-participant variation. By LMEM estimation, PSA increased 8.1% annually (95% CI 6.0–10.3, $p < 0.001$) for reclassified participants and 0.8% (95% CI –0.7 to 2.4, $p = 0.33$) for censored participants (Fig. 1).

The median PSAk at 1 yr from diagnosis was 0.04. In a Cox PH model including both ln(diagnostic PSA) and PSAk, PSAk was independently associated with reclassification, with an HR of 1.5 (95% CI 1.1–1.9) for each 0.1 unit increase. The HR for ln(diagnostic PSA) at diagnosis was 1.3 (95% CI 1.0–1.6). In a multivariable model adjusted for prostate size, time since diagnosis, percent of biopsy cores involved in the most recent biopsy, and any negative biopsy after diagnosis, the HRs for ln(diagnostic PSA) and PSAk remained significant: 1.7 (95% CI 1.3–2.2) and 1.6 (95% CI 1.2–2.1), respectively.

In a secondary analysis modeling PSAk calculated using three different methods, a PSAkRS (see methods) was not associated with time to reclassification ($p = 0.10$). The association between PSAkS from linear regression and time to reclassification was less significant ($p = 0.002$, compared with $p = 0.0006$ for PSAk) and not as strong (HR = 1.02, 95% CI 1.01–1.03 for each IQR increase) as PSAk based on the BLUP estimator from the LMEM (HR = 1.25, 95% CI 1.10–1.42 for each IQR increase). The model with PSAk had the highest goodness of fit with respect to AIC. In a model that contained all three PSAk measurements, PSAk was statistically significant, while PSAkS and PSAkRS were not (Table 2). Thus, the simple methods of calculating PSAk were not considered further. No meaningful differences in parameter estimates or statistical significance were observed in sensitivity analyses that minimized potential ascertainment bias (see the Supplementary material for more details).

The AUC for the full multivariable model including PSAk in predicting 4-yr reclassification outcomes from a measurement time of 1 yr after diagnosis was 0.80 (95% CI 0.75–0.85). As illustrated in Figure 2, when subgroups of low,

Table 1 – Participant characteristics

	All participants ($n = 851$)	Reclassified participants ($n = 291$)	Censored participants ($n = 560$)
Time to event/censor (yr), median (IQR)	3.0 (1.7–4.8)	2.0 (1.1–3.2)	3.7 (2.4–5.2)
No. of PSA values, median (IQR)	8 (4–13)	5 (3–9)	9 (5–14)
Dx PSA, median (IQR)	4.8 (3.6–6.3)	4.9 (3.9–6.3)	4.7 (3.5–6.3)
Prostate size, median (IQR)	40 (30–54)	35 (27–46)	44 (32–58)
Dx PSA density, median (IQR)	0.11 (0.08–0.16)	0.14 (0.10–0.18)	0.10 (0.07–0.14)
Dx core percentage, median (IQR) [a]	8 (8, 17)	17 (8, 17)	8 (8, 17)
Dx age, median (IQR)	62 (57–67)	63 (58–67)	62 (57–67)
Dx CAPRA score, n (%) [a]			
0	30 (4)	9 (3)	21 (4)
1	510 (64)	177 (64)	333 (64)
2	206 (26)	73 (26)	133 (25)
3+	56 (7)	19 (7)	37 (7)

CAPRA = Cancer of the Prostate Risk Assessment; Dx = at diagnosis; IQR = interquartile range; PSA = prostate-specific antigen.
[a] Of the participants, 28 and 49 are missing cores percentage and CAPRA score at diagnosis, respectively.

Please cite this article in press as: Cooperberg MR, et al. Refined Analysis of Prostate-specific Antigen Kinetics to Predict Prostate Cancer Active Surveillance Outcomes. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.017

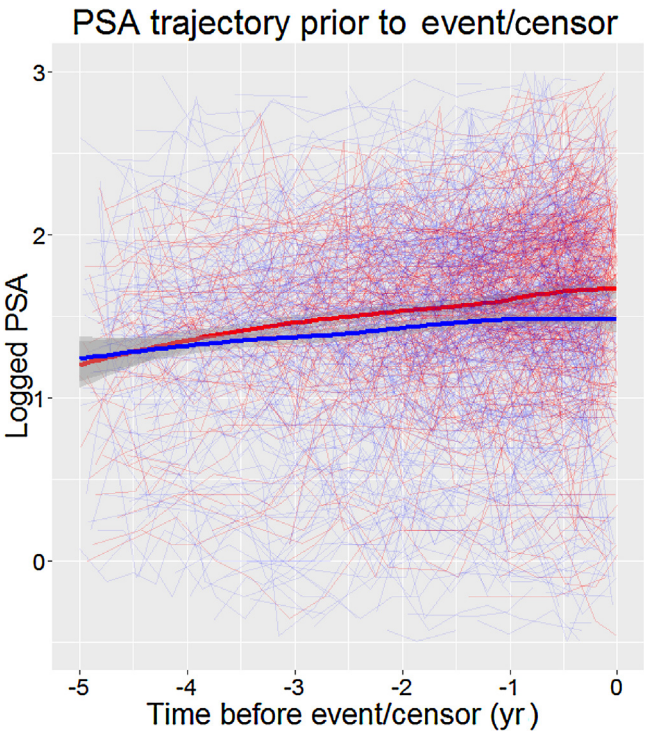


Fig. 1 – PSA trajectory prior to reclassification or censoring. In this spaghetti plot, individual ln(PSA) trajectories are plotted in red for reclassified participants and in blue for censored participants, omitting PSA data within 2 yr prior to censor date to look at long-term nonevents. Smoothed trend lines were added using LOESS. A separate LMEM analysis found a slope of 8.1%/yr for reclassified participants and 0.8% for censored participants (for interaction between reclassification group and PSA change: $p < 0.001$). LMEM = linear mixed-effect model; PSA = prostate-specific antigen.

middle, and high risk for reclassification were identified based on models with and without PSAk, the inclusion of PSAk was better able to distinguish between extreme subgroups of individuals (10% of each cohort) with low and high event rates in the years after the prediction. The reclassification-free probability based on the KM estimator in the low-risk group at 4 yr after the 1-yr measurement time was 1.00 (95% CI 1.00–1.00) with PSAk and 0.94 (95% CI 0.87–1.00) without PSAK in the model. In contrast, the reclassification-free probability in the high-risk group at 2 yr after the 1-yr measurement time was 0.34 (95% CI 0.22–0.53) with PSAk versus 0.41 (95% CI 0.28–0.60) without PSAk. The analysis based on grade reclassification only yielded very similar results (HR for each 0.1 unit increase in PSAk 1.62, 95% CI 1.22–2.17).

Calculating PSAk based on semiannual rather than quarterly PSA measurements yielded tightly correlated results: $r = 0.95$, $p < 0.001$ (Fig. 3). Recalculating the multivariable Cox PH model described above using only semiannual PSA measurements yielded substantially similar results, with the new HRs for ln(PSA) and PSAk being 1.6 (95% CI 1.3–2.1) and 1.9 (95% CI 1.3–2.6), respectively. The AUCs for 3- and 6-mo models are similar (Supplementary Fig. 2).

4. Discussion

PSA kinetics have long been studied as an indicator of prostate cancer prognosis, at decision points ranging from whether a man should undergo initial prostate biopsy [14] to early identification of advanced disease progression [15]. The utility of PSA kinetics in the pretreatment setting has been difficult to establish for a number of reasons, including a close correlation with static PSA at diagnosis [16], relatively short follow-up and limited longitudinal data in most series, PSA "noise" from noncancer sources, and the myriad published definitions of PSA kinetics such as velocity, doubling time, and other measures of growth [17]. We found stronger strength of association and better prediction calibration when PSAk was calculated based on an LMEM that accounted for both the general trend of increasing PSA over time in the cohort and individual-specific trajectories, while discounting the random noise in the PSA measurements.

For men on AS for low-risk prostate cancer, a rapidly rising PSA would intuitively seem to predict aggressive disease and adverse outcomes. In fact, in the Toronto cohort,

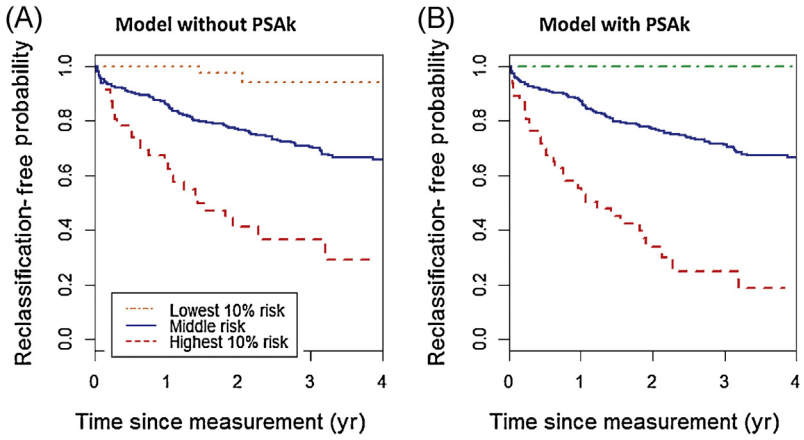
Table 2 – Comparing simple PSAk (PSAkS), restricted simple PSAk (PSAkRS), and PSAk in Cox PH models ($n = 841$) [a]

Variable	PSAkS model		PSAkRS model		PSAk model		Model containing all PSAk measurements	
	HR (95% CI)	*p* value	HR (95% CI)	*p* value	HR (95% CI)	*p* value	HR (95% CI)	*p* value
Dx PSA	1.84 (1.41, 2.39)	<0.001	1.80 (1.39, 2.34)	<0.001	1.88 (1.43, 2.46)	<0.001	1.90 (1.44, 2.49)	<0.001
PSAkS (IQR increase)	1.02 (1.01, 1.03)	0.002					1.01 (1.00, 1.02)	0.12
PSAkRS (IQR increase)			1.01 (0.99, 1.02)	0.10			1.01 (0.99,1.02)	0.3
PSAk (IQR increase)					1.25 (1.10, 1.42)	<0.001	1.24 (1.09, 1.41)	0.001
AIC	26 388		26 403		26 287		26 286	

AIC = Akaike information criterion; CI = confidence interval; Dx = at diagnosis; HR = hazard ratio; IQR = interquartile range; PH = proportional hazards; PSA = prostate-specific antigen; PSAk = prostate-specific antigen kinetics.

[a] Participants were required to have nonmissing PSAkS, PSAkRS, and PSAk to be considered in the model comparison. All models were adjusted for prostate size. Note that IQR increase was equivalent to 0.23 for PSAkS, 0.94 for PSAkRS, and 0.05 for PSAk.

Please cite this article in press as: Cooperberg MR, et al. Refined Analysis of Prostate-specific Antigen Kinetics to Predict Prostate Cancer Active Surveillance Outcomes. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.017



Model-based risk group [a]	Reclassification-free probability at 4 yr after the 1-yr measurement time (95% CI)	
	Model without PSAk	Model with PSAk
Lowest 10% risk	0.94 (0.87, 1.00)	1.00 (1.00, 1.00)
Middle risk	0.66 (0.60, 0.72)	0.66 (0.61, 0.72)
Highest 10% risk	0.41 (0.28, 0.60) [b]	0.34 (0.22, 0.53) [b]

Fig. 2 – Predicting reclassification outcomes. Kaplan–Meier plots showing reclassification-free probabilities at 4 yr after the 1-yr measurement time, using data up to 1 yr after diagnosis. (A) Model-based risk categories from Cox PH model adjusted for PSA at diagnosis, prostate size, time since diagnosis, most recent percent of biopsy cores involved, and history of any negative biopsy—but not PSAk. (B) Similar analysis adjusted for the same variables in addition to PSAk. The PSAk model improved stratification of risk prediction for the top and bottom deciles of risk over a model without PSAk. CI = confidence interval; PH = proportional hazards; PSA = prostate-specific antigen; PSAk = prostate-specific antigen kinetics. [a] Model-based risk is calculated at a measurement time of 1 yr after diagnosis using all data available up to the measurement time. [b] Reclassification-free probability at 2 yr after the 1-yr measurement time due to small numbers.

one of the earliest AS cohorts in North America, PSADT of <2 yr was initially the primary trigger for intervention. This threshold was found to be inadequately sensitive and was extended to <3 yr. However, while the definition of rapid PSADT in this cohort predicts outcomes after definitive treatment for men initially on AS, PSADT alone was found to be nonspecific and is now considered in the context of other indicators, particularly grade reclassification [18]. In other large cohort studies, PSA kinetics have not been proved useful with relatively short-term follow-up. In the international, multicenter Prostate Cancer Research International Active Surveillance study, PSA kinetics (PSADT <3 yr) was not predictive of pathologic reclassification [8]. In the University of California, San Francisco, cohort, PSADT of <3 yr was associated with an increased risk of reclassification—but only one man in the first 241 enrolled met this threshold [7]. In the Johns Hopkins cohort, both PSADT and PSA velocity, calculated as PSA multiplied by the slope of a linear regression of log(PSA), were poor predictors of reclassification, with AUCs of 0.59 and 0.61, respectively [6]. In this cohort, however, enrollment criteria for AS are very restrictive, yielding a narrow dynamic range in terms of progression risk in which to evaluate PSA kinetics.

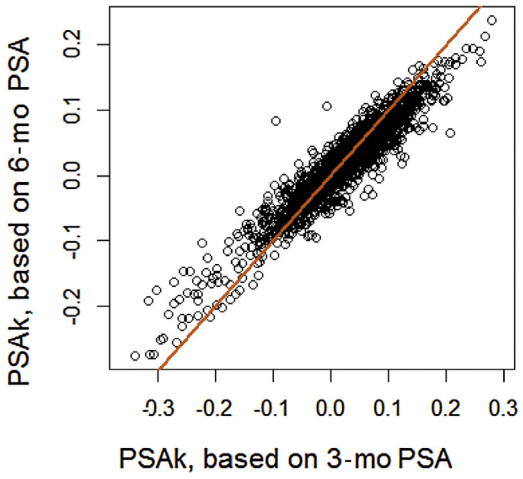


Fig. 3 – Quarterly versus semiannual PSA measurement. Correlation between PSAk calculations based on every 3- versus every 6-mo PSA measurements is illustrated. Pearson correlation (r) 0.95, $p < 0.001$. PSA = prostate-specific antigen; PSAk = prostate-specific antigen kinetics.

In this study, we analyzed PSA kinetics in a multicenter cohort with PSA data collected at protocol-mandated intervals, relatively long follow-up, and centralized analysis. We employed an analytic strategy that allowed the models to account for prior PSA history at each individual PSA measurement in an individual participant's trajectory, while borrowing information from the general trend across all participants and accommodating for random variability in PSA. In a plot of individual PSA trajectories, a higher overall slope was found for those who were reclassified versus those who were not (Fig. 1). Moreover, the addition of PSAk to a rich multivariable model improved the performance of the model, suggesting that PSAk may be considered an additional biomarker for outcomes on AS and is predictive independent of the absolute PSA level. In general, this

Please cite this article in press as: Cooperberg MR, et al. Refined Analysis of Prostate-specific Antigen Kinetics to Predict Prostate Cancer Active Surveillance Outcomes. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.017

finding suggests that collecting PSA measurements over time to provide an updated outcome is clinically useful, and our approach of calculating PSAk provides a summary that effectively reflects the changes of PSA over time. On the contrary, given essentially identical results analyzing every 3- versus every 6-mo PSA data, we suggest that in most cases PSA may not need to be measured any more often than semiannually, with the important qualification that this finding remains to be validated in other cohorts.

The imaging and molecular tests available to supplement standard clinical data to guide decision making for men with low-risk disease are proliferating rapidly, and the potential clinical utility of PSAk should be considered in this context. We have adhered to the REMARK criteria for biomarker reporting [13] to as great an extent as possible. In particular, we stress that all PSA and outcome data have been collected and reported prospectively throughout the duration of PASS, and all analyses conducted centrally. Although several biomarkers, as well as multiparametric magnetic resonance imaging, are currently marketed for decision making with respect to AS [19], so far none has been validated in a prospective AS cohort. Moreover, PSAk has the advantage of requiring neither any additional biomaterial nor any incremental cost.

A few caveats should be noted. PSA levels are reported directly from the Canary PASS clinical sites, and reflect different laboratories. We, therefore, cannot control for interassay variability in PSA levels. However, men are instructed to use the same laboratory consistently for their PSA measurements, and we expect that assay variability would introduce a bias toward a null result rather than a false-positive result. We examined the performance of PSAk above and below the PSAD threshold of 0.15 to better understand the performance of PSAk relative to PSAD. Our finding of differential performance at high and low PSAD is intriguing—perhaps reflecting better information from PSA trends in the absence of substantial benign prostatic hyperplasia —and merits further examination. As changes in PSA may affect decisions regarding biopsy performance (and therefore the opportunity to identify reclassification), a risk of ascertainment bias exists. However, compliance with biopsy schedules in PASS is generally excellent (80% of biopsies on time), and the sensitivity analysis excluding men with noncompliant biopsies did not change the results. The BLUP methodology does not lend itself to simple calculation at the point of care and requires a robust background of PSA data. We plan to incorporate PSAk, together with other parameters predictive of AS outcomes, in a web-based, multivariable risk calculator that will be presented in a future publication.

Perhaps the most important limitation is the reliability of our end point. The principal question was the ability of PSAk to predict biopsy reclassification. We acknowledge that reclassification itself is an imperfect end point, as it may reflect initial undersampling [20], variation in the interpretation of different pathologists [21], and/or minimal changes in the tumor, which have little clinical importance. However, our reclassification definition is consistent with those used by most other AS cohorts, and these changes frequently drive treatment decision making in contemporary practice. Therefore, while perhaps not biologically optimal, we believe that our findings are quite relevant for current clinical management and can in fact improve AS care.

5. Conclusions

We found that a sophisticated mathematical approach to measuring PSAk, as reflected in the novel PSAk parameter, can improve prediction of outcomes for men on surveillance for prostate cancer and that PSA may need to be measured no more often than semiannually. Obviously, PSA should never be interpreted in a vacuum, and we did not identify a PSAk threshold that should always indicate treatment. These results, which must be validated in other surveillance cohorts, suggest that PSAk or similar assessments of kinetics should be considered in future multivariable models of AS outcomes.

Author contributions: Matthew R. Cooperberg 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: Cooperberg, Brooks, Lin, Zheng.
Acquisition of data: Newcomb, Carroll, Dash, Fabrizio, Gleave, Thompson, Wagner, Lin.
Analysis and interpretation of data: Faino, Zheng, Cooperberg, Brooks, Lin, Newcomb, Kearns.
Drafting of the manuscript: Cooperberg, Brooks, Lin, Zheng, Newcomb, Faino.
Critical revision of the manuscript for important intellectual content: Cooperberg, Brooks, Faino, Newcomb, Kearns, Carroll, Dash, Etzioni, Fabrizio, Gleave, Nelson, Thompson, Wagner, Lin, Zheng.
Statistical analysis: Faino, Zheng.
Obtaining funding: Cooperberg, Lin, Nelson.
Administrative, technical, or material support: None.
Supervision: None.
Other: None.

Financial disclosures: Matthew R. Cooperberg 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: Department of Defense Prostate Cancer Research Program (grants W81XWH1110489 and W81XWH1410595); Canary Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.eururo.2018.01.017.

References

[1] Ip S, Dahabreh IJ, Chung M, et al. An evidence review of active surveillance in men with localized prostate cancer. Evid Rep Technol Assess (Full Rep) 2011;204:1–341.

[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] Cooperberg MR, Carroll PR. Trends in management for patients with localized prostate cancer, 1990-2013. JAMA 2015;314:80–2.
[5] Loeb S, Folkvaljon Y, Curnyn C, Robinson D, Bratt O, Stattin P. Uptake of active surveillance for very-low-risk prostate cancer in Sweden. JAMA Oncol 2017;3:1393–8.
[6] Ross AE, Loeb S, Landis P, et al. Prostate-specific antigen kinetics during follow-up are an unreliable trigger for intervention in a prostate cancer surveillance program. J Clin Oncol 2010;28:2810–6.
[7] Whitson JM, Porten SP, Hilton JF, et al. The relationship between prostate specific antigen change and biopsy progression in patients on active surveillance for prostate cancer. J Urol 2011;185:1656–60.
[8] Bokhorst LP, Valdagni R, Rannikko A, et al. A decade of active surveillance in the PRIAS study: an update and evaluation of the criteria used to recommend a switch to active treatment. Eur Urol 2016;70:954–60.
[9] Newcomb LF, Thompson IM, Boyer HD, et al. Outcomes of active surveillance for clinically localized prostate cancer in the prospective, multi-institutional Canary PASS cohort. J Urol 2016;195:313–20.
[10] Cooperberg MR, Pasta DJ, Elkin EP, et al. The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. J Urol 2005;173:1938–42.
[11] Zheng Y, Heagerty PJ. Semiparametric estimation of time-dependent ROC curves for longitudinal marker data. Biostatistics 2004;5:615–32.
[12] Maziarz M, Heagerty P, Cai T, Zheng Y. On longitudinal prediction with time-to-event outcome: comparison of modeling options. Biometrics 2017;73:83–93.
[13] Mcshane L, Altman D, Sauerbrei W, Taube S, Gion M, Clark G. Reporting recommendations for tumor marker prognostic studies. J Clin Oncol 2005;23:9067.
[14] Carter HB, Ferrucci L, Kettermann A, et al. Detection of life-threatening prostate cancer with prostate-specific antigen velocity during a window of curability. J Natl Cancer Inst 2006;98:1521–7.
[15] Scher HI, Halabi S, Tannock I, et al. Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: recommendations of the Prostate Cancer Clinical Trials Working Group. J Clin Oncol 2008;26:1148–59.
[16] Vickers AJ, Till C, Tangen CM, Lilja H, Thompson IM. An empirical evaluation of guidelines on prostate-specific antigen velocity in prostate cancer detection. J Natl Cancer Inst 2011;103:462–9.
[17] den Bergh van RCN, Roemeling S, Roobol MJ, Wolters T, Schröder FH, Bangma CH. Prostate-specific antigen kinetics in clinical decision-making during active surveillance for early prostate cancer—a review. Eur Urol 2008;54:505–16.
[18] Klotz L, Vesprini D, Sethukavalan P, et al. Long-term follow-up of a large active surveillance cohort of patients with prostate cancer. J Clin Oncol 2015;33:272–6.
[19] Leapman MS, Nguyen HG, Cooperberg MR. Clinical utility of biomarkers in localized prostate cancer. Curr Oncol Rep 2016;18:30.
[20] Porten SP, Whitson JM, Cowan JE, et al. Changes in prostate cancer grade on serial biopsy in men undergoing active surveillance. J Clin Oncol 2011;29:2795–800.
[21] McKenney JK, Simko J, Bonham M, et al. The potential impact of reproducibility of Gleason grading in men with early stage prostate cancer managed by active surveillance: a multi-institutional study. J Urol 2011;186:465–9.

Please cite this article in press as: Cooperberg MR, et al. Refined Analysis of Prostate-specific Antigen Kinetics to Predict Prostate Cancer Active Surveillance Outcomes. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.017

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

European Association of Urology

Platinum Priority – Prostate Cancer
Editorial by XXX on pp. x–y of this issue

Role of Surveillance Biopsy with No Cancer as a Prognostic Marker for Reclassification: Results from the Canary Prostate Active Surveillance Study

James T. Kearns [a,*], *Anna V. Faino* [b], *Lisa F. Newcomb* [a,b], *James D. Brooks* [c], *Peter R. Carroll* [d], *Atreya Dash* [a], *William J. Ellis* [a], *Michael Fabrizio* [e], *Martin E. Gleave* [f], *Todd M. Morgan* [g], *Peter S. Nelson* [b], *Ian M. Thompson* [h], *Andrew A. Wagner* [i], *Yingye Zheng* [b], *Daniel W. Lin* [a]

[a] *Department of Urology, University of Washington, Seattle, WA, USA;* [b] *Fred Hutchinson Cancer Research Center, Seattle, WA, USA;* [c] *Stanford University, Stanford, CA, USA;* [d] *University of California, San Francisco, CA, USA;* [e] *Eastern Virginia Medical School, VA, USA;* [f] *University of British Columbia, Vancouver, BC, Canada;* [g] *University of Michigan, Ann Arbor, MI, USA;* [h] *University of Texas Health Sciences Center at San Antonio, TX, USA;* [i] *Beth Israel Deaconess Medical Center, Boston, MA, USA*

Article info

Article history:
Accepted January 17, 2018

Associate Editor:
Giacomo Novara

Keywords:
Active surveillance
Prostate biopsy
Prostate cancer

Abstract

Background: Many patients who are on active surveillance (AS) for prostate cancer will have surveillance prostate needle biopsies (PNBs) without any cancer evident.
Objective: To define the association between negative surveillance PNBs and risk of reclassification on AS.
Design, setting, and participants: All men were enrolled in the Canary Prostate Active Surveillance Study (PASS) between 2008 and 2016. Men were included if they had Gleason $\leq 3+4$ prostate cancer and $<34\%$ core involvement ratio at diagnosis. Men were prescribed surveillance PNBs at 12 and 24 mo after diagnosis and then every 24 mo.
Outcome measurements and statistical analysis: Reclassification was defined as an increase in Gleason grade and/or an increase in the ratio of biopsy cores to cancer to $\geq 34\%$. PNB outcomes were defined as follows: (1) no cancer on biopsy, (2) cancer without reclassification, or (3) reclassification. Kaplan–Meier and Cox proportional hazard models were performed to assess the risk of reclassification.
Results and limitations: A total of 657 men met inclusion criteria. On first surveillance PNB, 214 (32%) had no cancer, 282 (43%) had cancer but no reclassification, and 161 (25%) reclassified. Among those who did not reclassify, 313 had a second PNB. On second PNB, 120 (38%) had no cancer, 139 (44%) had cancer but no reclassification, and 54 (17%) reclassified. In a multivariable analysis, significant predictors of decreased future reclassification after the first PNB were no cancer on PNB (hazard ratio [HR] = 0.50, $p = 0.008$), lower serum prostate-specific antigen, larger prostate size, and lower body mass index. A finding of no cancer on the second PNB was also associated with significantly decreased future reclassification in a multivariable analysis (HR = 0.15, $p = 0.003$), regardless of the first PNB result. The major limitation of this study is a relatively small number of patients with long-term follow-up.
Conclusions: Men who have a surveillance PNB with no evidence of cancer are significantly less likely to reclassify on AS in the PASS cohort. These findings have implications for tailoring AS protocols.
Patient summary: Men on active surveillance for prostate cancer who have a biopsy showing no cancer are at a decreased risk of having worse disease in the future. This may have an impact on how frequently biopsies are required to be performed in the future.

* Corresponding author. Department of Urology, University of Washington School of Medicine, 1959 NE Pacific, Box 356510, Seattle, WA 98195, USA. Tel. +1 864 650 5255; Fax: +1 206 543 3272.
E-mail address: jkearns1@uw.edu (J.T. Kearns).

https://doi.org/10.1016/j.eururo.2018.01.016
0302-2838/

1. Introduction

Active surveillance (AS) for prostate cancer is an increasingly popular management strategy for Gleason 3 + 3 and low-volume 3 + 4 prostate cancer [1]. Patients are generally assessed by periodic serum prostate-specific antigen (PSA) testing, digital rectal examination, and prostate biopsy. Despite increasing use, an optimal AS protocol that defines precise timing of these assessments has not yet been established or defined by practice guidelines. In published series, biopsies are performed as frequently as annually [2] to every 3–4 yr [3]. Furthermore, within a given protocol, there has been no formal strategy for tailoring biopsy frequency based on a patient's individualized risk.

Prostate biopsies yield a wealth of information about an individual's cancer, but many men find them to be unpleasant, the biopsies are costly [4], and there is an approximately 5% risk of infection following biopsy [5]. Furthermore, published AS series report that although the majority of surveillance biopsies find no change in the Gleason grade, 21–50% [6] of surveillance biopsies have no cancer found on the biopsy specimens, suggesting a low cancer volume. Given these considerations, it is a common clinical scenario for an AS patient who has one or more surveillance biopsies with the finding of no cancer to question the need for further biopsy.

In this context, we examined the predictive value of no cancer on surveillance biopsy for future pathological reclassification after a diagnosis of very-low– and low-risk prostate cancer in the large, multicenter Canary Prostate Active Surveillance Study (PASS). We assessed the significance of biopsy results in the first and second biopsies after the initial diagnosis and performed modeling to take into account variables that contribute to risk of reclassification.

2. Patients and methods

2.1. Patient population

PASS is a multi-institutional prostate cancer AS cohort study in North America [7]. All patients were enrolled in PASS and approved by institutional review boards at all participating sites (clinicaltrials.gov NCT000756665). Under the PASS protocol, PSA is measured every 3 mo, clinic visits occur every 6 mo, and ultrasound-guided biopsies are performed first between 6 and 12 mo after diagnosis, second at 24 mo after diagnosis, and then every 2 yr. In addition, the PASS protocol allows for off-protocol, "for-cause" biopsies. Eighty percent of biopsies were per protocol (on time), with 20% occurring either earlier or later than the protocol schedule. At least 10-core templates were required, with the median (interquartile range [IQR]) number of total biopsy cores collected being 12 (12, 14). Other tests, including magnetic resonance imaging (MRI), may be performed at the clinicians' discretion, but as the study started enrollment in 2008, the majority of men have not undergone these procedures. Patients were included in the current analysis if they were enrolled as of February 2016, had Gleason $\leq$3 + 4 prostate cancer, had <34% ratio of biopsy cores containing cancer to total biopsy cores (core ratio) at diagnosis, and had their first surveillance biopsy after the initial diagnosis of prostate cancer (aka, confirmatory biopsy) within 2 yr of diagnosis and while enrolled in PASS.

2.2. Outcomes and statistical methods

The primary outcome was time to reclassification from either the first or the second surveillance biopsy. Reclassification was defined as an increase in primary or secondary Gleason grade at biopsy and/or an increase in the core ratio to $\geq$34%. All pathology outcomes were determined by uropathologists at each site. Sensitivity analyses were also performed for participants diagnosed with Gleason 3 + 3 only or for grade-only reclassification. Patients without reclassification were censored on the date of last study contact, treatment, or 2 yr after their last biopsy, whichever came first.

Patients were stratified by the outcome of their first or second surveillance biopsy as follows: (1) no evidence of cancer on biopsy, (2) evidence of cancer on biopsy without reclassification, or (3) reclassification. Kaplan–Meier curves were plotted to examine how reclassification-free probability varied with surveillance biopsy outcome over the follow-up period. Log-rank tests were used to compare differences in reclassification-free probabilities.

Associations between previous surveillance biopsy result (no cancer vs cancer without reclassification) and time to future reclassification were modeled using Cox proportional hazard models. In order to assess whether the first surveillance biopsy result was associated with future reclassification, we considered a time since first surveillance biopsy model, where the association of interest was the result of the first surveillance biopsy. In order to assess whether the aggregate effect of the first and second surveillance biopsy results was associated with future reclassification, we considered a time since second surveillance biopsy model, where the two associations of interest were the results of the first and second surveillance biopsies, respectively. Owing to our hypotheses of interest, previous surveillance biopsy result(s) remained in the two models regardless of statistical significance. In addition, the following covariates were considered: natural log-transformed PSA closest and prior to surveillance biopsy, maximum core ratio from either diagnostic biopsy or surveillance biopsy, natural log-transformed diagnostic PSA, body mass index (BMI), natural log-transformed prostate volume, age at diagnosis, clinical T stage (T1 vs T2), diagnostic Gleason (3 + 4 or 3 + 3), and race (Caucasian vs others). Study site was accounted for by stratifying the baseline hazard. In order to account for potential collinearity among the variables, insignificant covariates were backward eliminated based on a p value cutoff of 0.05.

To address whether our results were biased by a negative biopsy influencing the decision to undergo or delay a biopsy, several steps were taken. The timing of each biopsy was defined as "on time," "early," or "late" based on the PASS protocol. Multinomial regression analyses were used to determine if biopsy timing was associated with prior biopsy result. A sensitivity analysis was performed on a subset of participants with all biopsies compliant to the protocol. Further details are in the Supplementary material. Analyses were performed with SAS version 9.4 and R version 3.3.0.

3. Results

Six hundred fifty-seven men were included in this analysis. Overall median follow-up from diagnosis for participants without a reclassification event was 2.9 yr (IQR 1.8–4.7). All participants received a first surveillance biopsy, which occurred at a median of 1.0 yr after diagnosis (IQR 0.7–1.2 yr). The outcomes of the first surveillance biopsy were as follows: 214 (32%) with no cancer on this biopsy, 282 (43%) with cancer on biopsy but no reclassification, and 161 (25%) with reclassification (Fig. 1). Of the 496 men who did not reclassify, 313 had a second biopsy at a median of 2.3 yr

Please cite this article in press as: Kearns JT, et al. Role of Surveillance Biopsy with No Cancer as a Prognostic Marker for Reclassification: Results from the Canary Prostate Active Surveillance Study. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.016

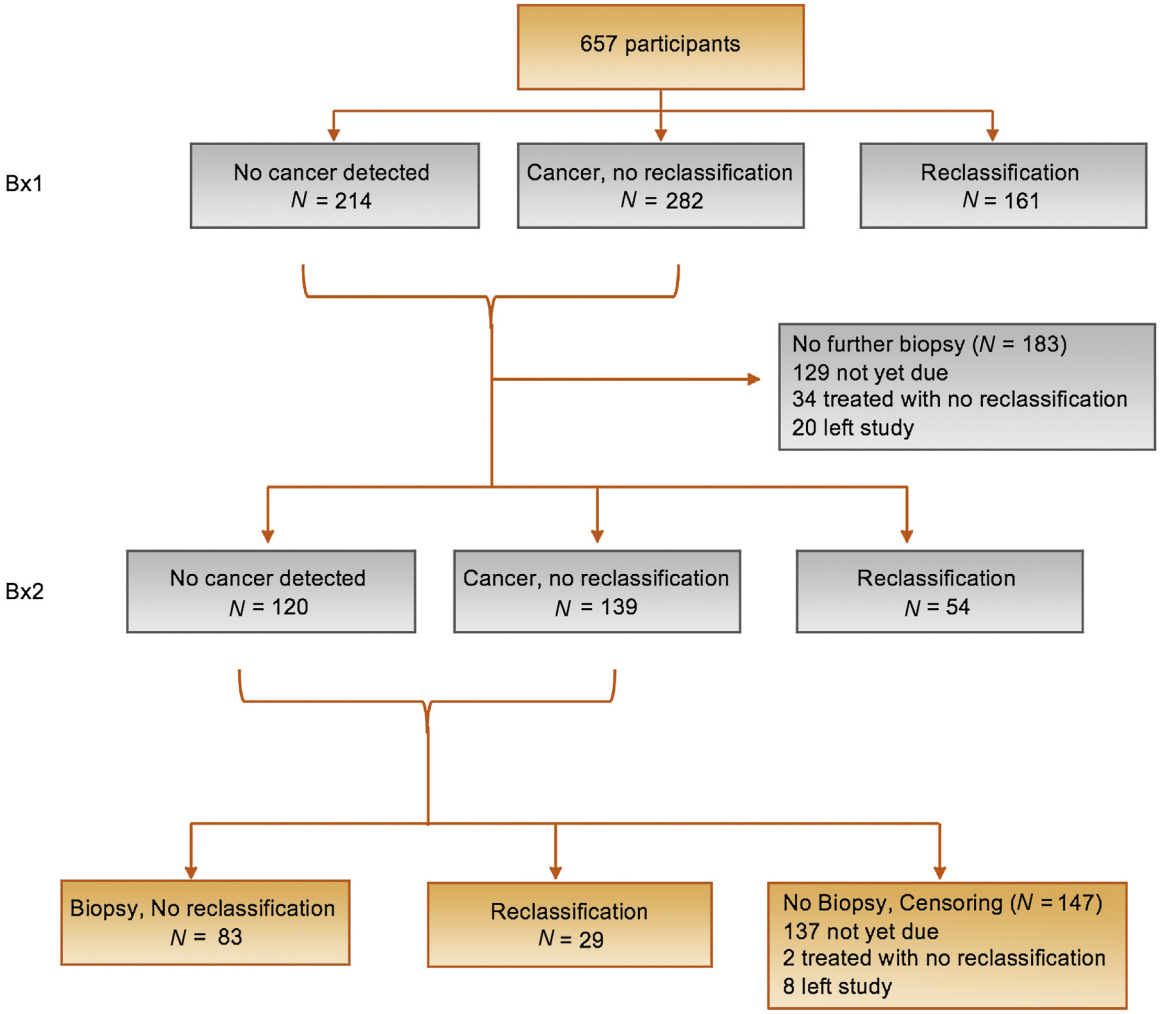


Fig. 1 – Consort diagram of patients receiving surveillance biopsy and biopsy outcomes. Bx1 = first surveillance biopsy; Bx2 = second surveillance biopsy.

from diagnosis (IQR 2.0–3.0 yr). Among these 313 men, 120 (38%) had no cancer on this biopsy, 139 (45%) had some cancer but no reclassification, and 54 (17%) had a reclassification event at second biopsy (Fig. 1).

The mean age of the cohort was 63 yr, median PSA was 4.9 ng/ml, median prostate volume was 42 cc, 94% were diagnosed with Gleason 3 + 3, and the median core ratio was 8% (which corresponds to 1/12 biopsy cores with cancer; Table 1). When stratified by the outcome of the first surveillance biopsy, the groups were similar with respect to racial makeup, age, clinical stage, family history of prostate cancer, and BMI. There were statistically significant differences across groups for prostate volume, serum PSA level, PSA density, diagnostic Gleason grade, and diagnostic core ratio positive for prostate cancer (Table 1). The results for patients who underwent a second surveillance biopsy are similar and are given in Supplementary Table 1.

Kaplan–Meier analysis of reclassification stratified by outcome of the first surveillance biopsy is shown in Figure 2. There was a statistically significant difference in time to reclassification in men whose first biopsy had no evidence of cancer versus men having evidence of cancer without reclassification ($p < 0.001$). Similarly, there was a statistically significant difference in time to reclassification based on the outcome of the second biopsy ($p < 0.001$), as shown in Figure 3. When patients who had two surveillance biopsies without reclassification were stratified by outcome of both first and second surveillance biopsies, the reclassification-free probability was similar for patients whose second surveillance biopsy showed no cancer, regardless of the result of the first biopsy (Supplementary Fig. 1).

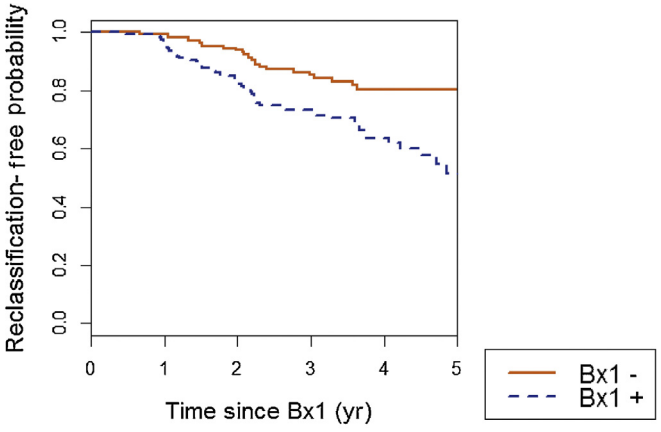
A first surveillance prostate biopsy negative for any cancer versus positive for cancer without reclassification was associated with less risk of reclassification in future biopsies (hazard ratio [HR] = 0.44, $p < 0.001$). After adjusting for serum PSA, prostate volume, and BMI, no cancer on initial surveillance biopsy was still significantly protective against reclassification (HR 0.50, $p = 0.008$; Table 2). Finding no cancer in the second surveillance biopsy was also significantly protective against reclassification in both unadjusted (HR 0.12, $p < 0.001$) and adjusted (HR 0.18, $p = 0.01$) analyses (Table 3).

Please cite this article in press as: Kearns JT, et al. Role of Surveillance Biopsy with No Cancer as a Prognostic Marker for Reclassification: Results from the Canary Prostate Active Surveillance Study. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.016

Table 1 – Patient characteristics based on results of first surveillance biopsy

	No cancer 1st surveillance biopsy	Cancer without reclassification 1st surveillance biopsy	Reclassification on 1st surveillance biopsy	*p* value [a]
N	214	282	161	
Race, *n* (%)				0.12
Caucasian American	187 (87)	258 (91)	143 (89)	
African American	16 (7)	9 (3)	13 (8)	
Other	11 (5)	15 (5)	5 (3)	
Prostate volume (cc), median (IQR)	46 (34–64)	43 (32–56)	36 (27–48)	<0.001
Age (yr), mean (SD)	62 (7)	63 (7)	63 (7)	0.22
PSA (ng/ml), median (IQR)	5.1 (3.7–6.6)	4.7 (3.7–6.1)	5.3 (4.4–6.6)	0.02
PSA density, median (IQR)	0.10 (0.07–0.14)	0.11 (0.08–0.15)	0.15 (0.11–0.21)	<0.001
Clinical stage, *n* (%)				0.37
T1a-T1c	197 (92)	249 (88)	146 (91)	
T2a-T2c	17 (8)	33 (12)	15 (9)	
Diagnostic Gleason score, *n* (%)				0.03
3 + 3	208 (97)	259 (92)	148 (92)	
3 + 4	6 (3)	23 (8)	13 (8)	
Diagnostic core ratio, median (IQR) [b]	8 (8–14)	13 (8–17)	17 (8–18)	<0.001
Family history of prostate cancer, *n* (%) [b]	55 (27)	79 (29)	42 (27)	0.89
BMI, mean (SD)	28.2 (4.3)	27.6 (4.0)	28.4 (5.0)	0.08

ANOVA = analysis of variance; BMI = body mass index; IQR = interquartile range; PSA = prostate-specific antigen; SD = standard deviation.
[a] *p* values comparing biopsy outcomes from the first surveillance biopsy (no cancer, cancer without reclassification, or reclassification), from chi-square test for categorical variables and from ANOVA for continuous variables. For prostate volume, PSA, PSA density, core ratio, and *p* value from Kruskal–Wallis test.
[b] Core ratio missing for 38 participants and family history of prostate cancer missing for 21 participants.



Log-rank test $p < 0.001$

Number at risk:

	Time since Bx1 (yr)					
	0	**1**	**2**	**3**	**4**	**5**
Bx1–	214	176	132	90	40	18
Bx1+	282	196	132	82	41	12

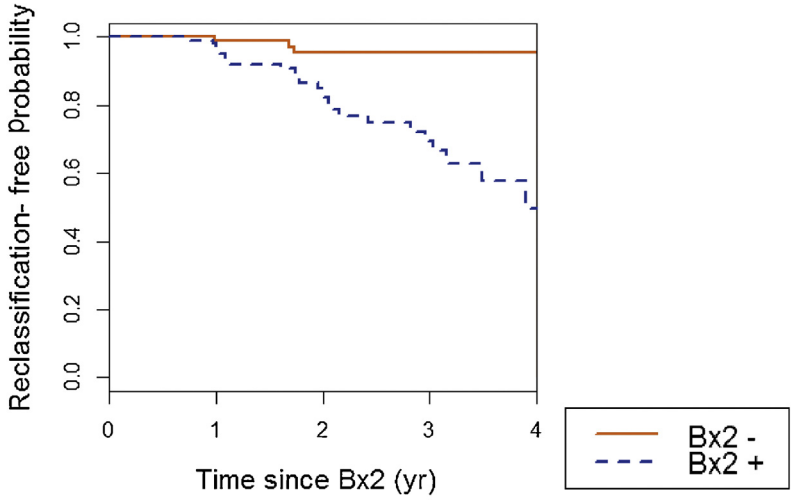
Fig. 2 – Time to grade and/or tumor volume reclassification by first surveillance biopsy outcome. Bx1 = first surveillance biopsy; Bx1– = no cancer detected on first surveillance biopsy; Bx1+ = cancer but no reclassification detected on first surveillance biopsy.

All results were similar when sensitivity analysis was performed for grade-only reclassification or for the subset of participants diagnosed with Gleason 3 + 3 cancer, and can be found in the Supplementary material. Prior biopsy result was not found to be associated with biopsy timing in an adjusted analysis. Similar significance was observed in a sensitivity analysis that minimized potential ascertainment bias (see the Supplementary material for more details).

4. Discussion

Our present study examined the risk of pathological reclassification in AS patients who have no cancer on first or second surveillance biopsy. In both Kaplan–Meier and multivariable-adjusted Cox proportional hazard analyses, no cancer on surveillance biopsy was prognostic against future reclassification. When there was no detectable

Please cite this article in press as: Kearns JT, et al. Role of Surveillance Biopsy with No Cancer as a Prognostic Marker for Reclassification: Results from the Canary Prostate Active Surveillance Study. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.016



Log-rank test $p < 0.001$

Number at risk:

	Time since Bx2 (yr)				
	0	**1**	**2**	**3**	**4**
Bx2–	120	88	55	26	10
Bx2+	139	98	57	26	5

Fig. 3 – Time to grade and/or tumor volume reclassification by second surveillance biopsy outcome. Bx2: second surveillance biopsy; Bx2–: no cancer detected on second surveillance biopsy; Bx2+:cancer but no reclassification detected on second surveillance biopsy.

Table 2 – Time to grade and/or tumor volume reclassification, from time of first surveillance biopsy (*n* = 494 [a], 85 with event)

Variable	Univariable		Multivariable	
	HR (95% CI) [b]	*p* value [b]	HR (95% CI) [b]	*p* value [b]
No cancer on first surveillance biopsy (vs cancer without reclassification)	0.44 (0.27, 0.71)	<0.001	0.50 (0.30, 0.83)	0.008
Ln (PSA on/prior to first surveillance biopsy)	1.93 (1.32, 2.83)	<0.001	2.74 (1.83, 4.10)	<0.001
Ln (prostate volume, cc)	0.38 (0.22, 0.69)	0.001	0.19 (0.10, 0.37)	<0.001
BMI	1.03 (0.98, 1.09)	0.28	1.07 (1.01, 1.13)	0.02

BMI = body mass index; CI = confidence interval; HR = hazard ratio; PSA = prostate-specific antigen.
[a] Two participants were missing core ratio data and were not included in the modeling.
[b] 95% confidence intervals and *p* values from Cox proportional hazards models.

Table 3 – Time to grade and/or tumor volume reclassification, from the time of second surveillance biopsy (*n* = 259, 29 with event)

Variable	Univariable		Multivariable	
	HR (95% CI) [a]	*p* value [a]	HR (95% CI) [a]	*p* value [a]
No cancer on second surveillance biopsy (vs cancer without reclassification)	0.12 (0.03, 0.39)	<0.001	0.18 (0.05, 0.66)	0.01
No cancer on first surveillance biopsy (vs cancer without reclassification)	0.35 (0.15, 0.82)	0.02	0.53 (0.20, 1.41)	0.20
Ln (PSA on/prior to second surveillance biopsy)	4.66 (2.22, 9.78)	<0.001	6.10 (2.62, 14.17)	<0.001
Ln (prostate volume, cc)	0.45 (0.16, 1.26)	0.13	0.18 (0.05, 0.64)	0.008

CI = confidence interval; HR = hazard ratio; PSA = prostate-specific antigen.
^ 95% confidence intervals and p-values from Cox proportional hazards models.

cancer in the first surveillance biopsy, the risk of future reclassification was decreased by 50%, and if no cancer was seen on second surveillance biopsy, then there was an 82% decreased risk of future reclassification.

We also found that patients with no cancer on first surveillance biopsy were more likely to have no cancer on the second surveillance biopsy when compared with those who had a first surveillance biopsy with cancer but no reclassification. This is consistent with previous work suggesting that no cancer found on initial surveillance biopsy is protective against future reclassification [8–11] and work suggesting that negative biopsy prior to diagnosis

Please cite this article in press as: Kearns JT, et al. Role of Surveillance Biopsy with No Cancer as a Prognostic Marker for Reclassification: Results from the Canary Prostate Active Surveillance Study. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.016

is associated with lower adverse pathological outcomes at radical prostatectomy [12]. Importantly, it also appears that continued presence of cancer on subsequent surveillance biopsy results in a significantly higher risk of pathological reclassification. Within 5 yr of diagnosis, ~3–5% of patients with no cancer on surveillance biopsies reclassify compared with ~20–30% of those who have some cancer on subsequent biopsies. These findings indicate that even in men who do not initially reclassify, there is a persistent risk of pathological reclassification and thus a need for continued surveillance. Decreasing risk of reclassification with increasing biopsy number was seen in this cohort, with 25% of men reclassifying on first biopsy and 17% reclassifying on second biopsy. This is consistent with our previously reported data and other AS cohorts that demonstrate decreasing rates of reclassification over time [3,7,13–15].

One of the major goals of evaluating factors that predict reclassification of prostate cancer on AS is to use all available data in the best possible manner to decrease the number of prostate biopsies required without sacrificing the detection of potentially lethal prostate cancer. Laviana et al [4] found that the economic cost of AS increases steadily with time, surpassing the cost of brachytherapy within 9 yr and nearly equaling that of robotic-assisted laparoscopic prostatectomy by 12 yr. These costs were driven chiefly by serial prostate biopsy. In addition to the financial cost of biopsies, there are biopsy-related morbidities, most notably an approximately 5% risk of infection [5]. However, as seen in the ProtecT trial, a strategy of "active monitoring" that relies solely upon large increases in serum PSA levels to trigger prostate biopsy may be an inadequate paradigm, with a 2.6-time increased risk of clinical progression [16]. One or more mandatory surveillance biopsies are likely necessary to better risk stratify patients before making decisions regarding future biopsy frequency. Using a finding that is prognostic against reclassification, such as surveillance biopsy without cancer, to decrease biopsy frequency may decrease patient discomfort, cost, and risk of infection while maintaining detection of significant disease.

In order to best use available clinical information, it is worth noting that the risk of reclassification associated with a given variable changes depending on what has transpired with the patient during his course of surveillance. Previously published nomograms for reclassification while on AS [17,18] do not adjust their covariates over the course of AS, despite patients having different risk profiles as they undergo biopsies without reclassification. We found that no cancer on second surveillance biopsy was much more prognostic against reclassification than no cancer on the first surveillance biopsy (HR 0.18 vs 0.50). This finding is consistent with previous reported outcomes where fewer men reclassify on AS over time [10]. Given that clinical variables may confer different risks at different time points, models and risk assessment tools should account for these varying risks.

The major strengths of our study include the fact that it is a multicenter, prospectively designed study with quality control of all clinical data collected. All participants were recommended the same biopsy schedule (6–12 mo after diagnosis, 24 mo after diagnosis, and then every 2 yr), regardless of whether or not they had detectable disease on surveillance biopsies. Overall, 80% of biopsies were per protocol (on time), and finding no cancer in the first surveillance biopsy was not associated with delayed subsequent biopsies. The inclusion at diagnosis of both Gleason 3 + 3 and 3 + 4 disease makes the results more generalizable to community AS protocols. In addition, the use of pathological reclassification as the end point does not rely upon patient factors such as tolerance for risk or anxiety that may sway treatment decisions. The study is limited by the lack of a centralized pathological review, lack of information for all patients regarding MRI use in the surveillance of these men, and relatively small numbers of patients with long-term follow-up. These limitations are mitigated by the fact that an early central pathology review indicates ~80% concordance with local pathology scoring, and most patients in PASS have not had prostatic MRI. Additionally, MRI is still not considered the standard of care in AS according to National Comprehensive Cancer Network guidelines [19]. Inclusion of more patients over time with similar risk profiles would be expected to tighten the confidence intervals rather than significantly change hazard ratios. In addition, our study would benefit from validation by an external AS cohort.

5. Conclusions

No detectable cancer in a biopsy during AS was prognostic for a decreased risk of pathological reclassification. The clinical impact of no cancer on surveillance biopsy becomes stronger on subsequent biopsy, suggesting that the risk of reclassification changes with time. Men with Gleason 3 + 3 prostate cancer and two initial surveillance biopsies with no detectable cancer may not warrant annual or semiannual biopsy, and may perhaps lengthen the biopsy interval to several years, similar to other published protocols [3]. Further work with models should include the concept of varying risk by taking into account real-time variables along the course of AS in order to individualize biopsy intervals and patient assessments. Portions of this work were presented as a moderated poster at the AUA Annual Meeting, May 2017.

Author contributions: James T. Kearns 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: Kearns, Faino, Newcomb, Zheng, Lin.
Acquisition of data: Brooks, Carroll, Dash, Ellis, Fabrizio, Gleave, Morgan, Nelson, Thompson, Wagner, Lin.
Analysis and interpretation of data: Kearns, Faino, Newcomb, Zheng, Lin.
Drafting of the manuscript: Kearns, Faino, Newcomb, Zheng, Lin.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Kearns, Faino, Newcomb, Zheng, Lin.
Obtaining funding: None.
Administrative, technical, or material support: None.

Supervision: None.
Other: None.

Financial disclosures: James T. Kearns 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: Canary Foundation, Department of Defense (PC130355), and Institute for Prostate Cancer Research.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.eururo.2018.01.016.

References

[1] Cooperberg MR, Carroll PR. Trends in management for patients with localized prostate cancer, 1990-2013. JAMA 2015;314:80–2.
[2] 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.
[3] Klotz L, Vesprini D, Sethukavalan P, et al. Long-term follow-up of a large active surveillance cohort of patients with prostate cancer. J Clin Oncol 2015;33:272–7.
[4] Laviana AA, Ilg AM, Veruttipong D, et al. Utilizing time-driven activity-based costing to understand the short- and long-term costs of treating localized, low-risk prostate cancer. Cancer 2016;122:447–55.
[5] Wagenlehner FME, van Oostrum E, Tenke P, et al. Infective complications after prostate biopsy: outcome of the Global Prevalence Study of Infections in Urology (GPIU) 2010 and 2011, a prospective multinational multicentre prostate biopsy study. Eur Urol 2013;63:521–7.
[6] 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.
[7] Newcomb LF, Thompson IM, Boyer HD, et al. Outcomes of active surveillance for clinically localized prostate cancer in the prospective, multi-institutional Canary PASS cohort. J Urol 2016;195:313–20.
[8] Eggener SE, Mueller A, Berglund RK, et al. A multi-institutional evaluation of active surveillance for low-risk prostate cancer. J Urol 2009;181:1635–41.
[9] Tseng KS, Landis P, Epstein JI, Trock BJ, Carter HB. Risk stratification of men choosing surveillance for low risk prostate cancer. J Urol 2010;183:1779–85.
[10] Porten SP, Whitson JM, Cowan JE, et al. Changes in prostate cancer grade on serial biopsy in men undergoing active surveillance. J Clin Oncol 2011;29:2795–800.
[11] Cary KC, Cowan JE, Sanford M, et al. Predictors of pathologic progression on biopsy among men on active surveillance for localized prostate cancer: the value of the pattern of surveillance biopsies. Eur Urol 2014;66:337–42.
[12] ElShafei A, Nyame Y, Kara O, et al. More favorable pathological outcomes in men with low risk prostate cancer diagnosed on repeat versus initial transrectal ultrasound guided prostate biopsy. J Urol 2016;195:1767–72.
[13] Tosoian JJ, Trock BJ, Landis P, et al. Active surveillance program for prostate cancer: an update of the Johns Hopkins experience. J Clin Oncol 2011;29:2185–90.
[14] Bul M, Zhu X, Valdagni R, et al. Active surveillance for low-risk prostate cancer worldwide: the PRIAS study. Eur Urol 2013;63:597–603.
[15] Dall'Era MA, Konety BR, Cowan JE, Shinohara K, Stauf F, Cooperberg MR, et al. Active surveillance for the management of prostate cancer in a contemporary cohort. Cancer 2008;112:2664–70.
[16] Hamdy FC, Donovan JL, Lane JA et al. 10-Year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer. N Engl J Med. In press. doi:10.1056/NEJMoa1606220.
[17] 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.
[18] Mamawala MM, Rao K, Landis P, et al. Risk prediction tool for grade re-classification in men with favourable-risk prostate cancer on active surveillance. BJU Int 2017;120:25–31.
[19] National Comprehensive Cancer Network. Prostate cancer (version 2.2017). https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf.

Please cite this article in press as: Kearns JT, et al. Role of Surveillance Biopsy with No Cancer as a Prognostic Marker for Reclassification: Results from the Canary Prostate Active Surveillance Study. Eur Urol (2018), https://doi.org/10.1016/j.eururo.2018.01.016

Author's Accepted Manuscript

Continued Five-Alpha Reductase Inhibitor Use After Prostate Cancer Diagnosis and the Risk of Reclassification and Adverse Pathological Outcomes in the Canary Prostate Active Surveillance Study (PASS)

James T. Kearns , Anna V. Faino , Jeanette M. Schenk , Lisa F. Newcomb , James D. Brooks , Peter R. Carroll , Atreya Dash , William J. Ellis , Michael Fabrizio , Martin E. Gleave , Todd M. Morgan , Peter S. Nelson , Ian M. Thompson , Andrew Wagner , Yingye Zheng , Daniel W. Lin

PII: S0022-5347(18)43635-X
DOI: 10.1016/j.juro.2018.07.065
Reference: JURO 15752

To appear in: *The Journal of Urology*
Accepted Date: 25 July 2018

Please cite this article as: Kearns JT, Faino AV, Schenk JM, Newcomb LF, Brooks JD, Carroll PR, Dash A, Ellis WJ, Fabrizio M, Gleave ME, Morgan TM, Nelson PS, Thompson IM, Wagner A, Zheng Y, Lin DW, Continued Five-Alpha Reductase Inhibitor Use After Prostate Cancer Diagnosis and the Risk of Reclassification and Adverse Pathological Outcomes in the Canary Prostate Active Surveillance Study (PASS), *The Journal of Urology*® (2018), doi: 10.1016/j.juro.2018.07.065.

Continued Five-Alpha Reductase Inhibitor Use After Prostate Cancer Diagnosis and the Risk of Reclassification and Adverse Pathological Outcomes in the Canary Prostate Active Surveillance Study (PASS)

James T Kearns[1], Anna V Faino[2], Jeanette M Schenk[2], Lisa F Newcomb[1,2], James D Brooks[3], Peter R Carroll[4], Atreya Dash[1], William J Ellis[1], Michael Fabrizio[5], Martin E Gleave[6], Todd M Morgan[7], Peter S Nelson[2], Ian M Thompson[8], Andrew Wagner[9], Yingye Zheng[2], Daniel W Lin[1,2]

[1]Department of Urology, University of Washington
[2]Fred Hutchinson Cancer Research Center
[3]Stanford University
[4]University of California San Francisco
[5]Eastern Virginia Medical School
[6]University of British Columbia
[7]University of Michigan
[8]University of Texas Health Sciences Center at San Antonio
[9]Beth Israel Deaconess Medical Center

Correspondence:
James T Kearns
Department of Urology
University of Washington School of Medicine
1959 NE Pacific, Box 356510
Seattle, WA 98195
Phone: (864) 650-5255 Fax: 206-543-3272
Email: jkearns1@uw.edu

Keywords: prostate cancer; 5-alpha reductase inhibitor; active surveillance
Word count (abstract): 246
Word count (text only): 2374

Abstract

Purpose

Outcomes of patients who enroll in active surveillance (AS) programs for prostate cancer (PCa) while taking five-alpha reductase inhibitors (5-ARIs) have not been well defined. We sought to determine the association of 5-ARI use with risk of reclassification in the Canary Prostate Active Surveillance Study (PASS).

Materials and Methods

Participants in the multicenter PASS were enrolled between 2008-2016. Inclusion criteria were current or never 5-ARI user, Gleason ≤ 3+4 PCa at diagnosis, < 34% core involvement ratio at diagnosis, and ≥ 1 surveillance biopsy. 1009 men (107 5-ARI users and 902 never users) were included. Reclassification was defined as increase in Gleason score and/or increase in ratio of biopsy cores positive for cancer to ≥ 34%. Adverse pathology at prostatectomy was defined as Gleason ≥ 4+3 and/or non organ-confined disease (pT3 or N1).

Results

On multivariable analysis, there was no difference in the risk of reclassification between 5-ARI users and never 5-ARI users (HR 0.81, p = 0.31). 5-ARI users were less likely to undergo radical prostatectomy (RP) (8% vs 18%, p=0.01) or any definitive treatment (19% vs 24%, p=0.04). Among participants who underwent RP (n=167), there was no suggestion of a difference in the rate of adverse pathology between 5-ARI users and non-users at prostatectomy.

Conclusions

Continued 5-ARI use after initial diagnosis of prostate cancer was not associated with risk of reclassification on AS for men in the PASS cohort.

Introduction

Five-alpha reductase inhibitors (5-ARIs) are widely used to treat benign prostatic hyperplasia (BPH). Since they inhibit the conversion of testosterone to the more potent dihydrotestosterone, large randomized clinical trials have evaluated the efficacy of 5-ARIs for the primary prevention of prostate cancer (PCa). In these trials, both finasteride[1] and dutasteride[2] were associated with decreased incidence of low grade PCa but slightly increased incidence of high grade PCa compared to placebo. These findings led to an FDA safety advisory regarding the risk of developing high grade PCa while taking 5-ARIs.[3]

However, many men continue to take 5-ARIs given their effectiveness in the treatment of BPH. Evidence also suggests that men on a 5-ARI for BPH who are also being annually screened for PCa with digital rectal examination and serum prostate specific antigen (PSA) undergo fewer biopsies, but the biopsies more frequently show PCa, with a similar Gleason score distribution.[4] As active surveillance (AS) for PCa becomes more popular and recommended[5,6] in the management of low- and very low-risk PCa,[7] more of these men will likely choose AS as their initial management strategy.

Previous work evaluating the use of 5-ARI therapy *after* enrollment in AS on the effect of pathological reclassification have yielded conflicting results.[8,9] It is still unclear whether these agents alter tumor biology to decrease pathological disease progression or if they lead to decreased treatment that is independent of effects on pathological disease progression. Furthermore, the effect of 5-ARIs in men using AS to manage their cancer and who intitated 5-ARI use prior to diagnosis of their cancer is not known. The goal of this study was to evaluate whether continuing 5-ARIs after a diagnosis of PCa is associated with adverse outcomes on AS. Specifically, we assessed whether 5-ARI therapy was

associated with risk of pathological reclassification on surveillance biopsy, and adverse pathology (Gleason grade ≥ 4+3 and/or non-organ confined disease) on radical prostatectomy.

Methods:

Patient Population

Data are from the multicenter Canary Prostate Active Surveillance Study (PASS), approved by the institutional review boards at all participating sites (clinicaltrials.gov NCT000756665).[10] Under the PASS protocol, serum PSA is recommended every 3 months, and clinic visits occur every 6 months, and ultrasound-guided biopsies are prescribed between 6 and 12 months after diagnosis, 24 months after diagnosis and then every 24 months thereafter. At least 10-core study biopsy regimens were required, and 91% of regimens were 12-core or more (median [interquartile range] for both 5-ARI users and non users = 12 [12-12]). Other tests, including magnetic resonance imaging (MRI), may be performed at the clinicians' discretion, but as the study started enrollment in 2008, the majority of men have not undergone MRI. Data on Gleason score, clinical stage, cores ratio and corresponding PSA for diagnostic and follow-up biopsies are extracted from medical records. Participants were asked to report current and 10-year history of 5-ARI use at study enrollment, and current use was assessed at each follow-up visit. Men who indicated use prior to diagnosis and current use at all follow up visits were defined as 5-ARI users in this study. Men included in this analysis were enrolled in PASS by February 2016, had a prostate cancer diagnosis within 5 years of enrollment, had Gleason ≤ 3+4 cancer and < 34% ratio of biopsy cores containing cancer to total biopsy cores (cores ratio) at diagnosis, and had at least one surveillance biopsy after diagnostic biopsy (1069 participants). We excluded participants who were former 5-ARI users (had a history of 5-ARI use prior to diagnosis but discontinued use; n=20), who initiated 5-ARI use after diagnosis (n=36), or who had unknown 5-ARI use at diagnosis (n=4), resulting in 1009 participants remaining for analysis.

Outcomes

The primary outcome for these analyses was time to reclassification while on AS. Reclassification was defined as either: 1) increase in primary or secondary Gleason grade at biopsy only or 2) a composite of increase in Gleason grade and/or an increase in the ratio of biopsy cores with cancer to total cores (cores ratio) to ≥ 34% . Participants without reclassification were censored at date of last study contact, treatment, or 2 years after their last biopsy, whichever came first. A total of 13 deaths occurred in this study population, none due to prostate cancer. Among the subset of men who underwent radical prostatectomy (RP), we also examined whether 5-ARI use was associated with risk of adverse pathology, defined as Gleason grade ≥ 4+3 and/or non organ-confined disease (pT3 or N1).

Statistical Methods

Descriptive statistics were used to characterize the study sample. Differences between 5-ARI users and non-users were evaluated using either t-test or Wilcoxon sign rank test for continuous variables and chi-square or Fisher's test for categorical variables.

All time-dependent analyses were based on the time between PCa diagnosis and either incident reclassification or censoring event. Kaplan Meier curves were plotted in order to examine how reclassification-free probability varied by 5-ARI status. Cox proportional hazards models were used to estimate the unadjusted and covariate-adjusted hazards ratios for the association between 5-ARI use and risk of reclassification. Covariate adjusted models considered the following variables: diagnostic PSA (natural log-transformed, continuous), body mass index (BMI, continuous), prostate volume (natural log-transformed, continuous), age at diagnosis (continuous), self-reported BPH (yes, no), diagnostic T stage (T1a-c, T2a-c), diagnostic Gleason (3+3, 3+4), family history of PCa (yes, no), and diagnostic cores ratio

(continuous). The final adjusted model included BMI, cores ratio, PSA and prostate volume. PSA and prostate volume were modeled as separate variables instead of as the composite variable PSA density. Participants missing cores ratio (n=61) dropped out of multivariable models. The baseline hazard in all Cox proportional hazards models was stratified by study site. Non-significant variables were backwards eliminated using a p-value cutoff of 0.05.

Sensitivity analyses were performed among the subset of men with Gleason 3+3 PCa. Exploratory analyses compared the rates of adverse pathologic outcomes between 5-ARI users and non-users among the subset of men (n=167) who underwent radical prostatectomy (RP). To address whether or not our results were biased by an affect of 5-ARI use on biopsy timing, biopsies were defined as "on-time," "early" or "late" based on the PASS protocol. Multinomial regression was used to determine if biopsy timing was associated with 5-ARI use. All analyses were performed using SAS version 9.4 and R version 3.3.0.

Results:

A total of 1009 men were included in this analysis with a median follow up of 3.6 (IQR 2.2-5.4) years among censored participants. Demographic data are shown in Table 1. There were 107 men on a 5-ARI at diagnosis, and 902 who had never used a 5-ARI. Men in the 5-ARI group were more likely to have a BPH diagnosis (77 vs 28%, $p < 0.001$), had larger prostate volume (median 51 g vs 40 g, $p < 0.001$), and were older (65 vs 62 years, $p < 0.001$). Men who were on a 5-ARI were less likely to undergo RP (8 vs 18%, $p = 0.01$) or any curative treatment (19 vs 28%, $p = 0.04$). Men in the two groups were statistically similar in terms of racial background, serum PSA level, PSA density, clinical stage, Gleason score, and diagnostic positive cores ratio.

Overall, there was no significant difference in the time to grade and/or volume reclassification (p = 0.10) or grade only reclassification (p = 0.30) between 5-ARI users and non-users (Figure 1). A sensitivity analysis limited to men who entered the study with 3+3 PCa also did not reveal any association between 5-ARI use and time to reclassification (data not shown). In an unadjusted Cox proportional hazards model (Table 2), continued use of 5-ARIs while on AS was associated with a decreased risk of reclassification (HR 0.63, 95% CI 0.43-0.94); however, after adjustment for diagnostic PSA, BMI, prostate size and diagnostic cores ratio, 5-ARI use was not associated with risk of reclassification (HR 0.81, 95% CI 0.55-1.21). 5-ARI use did not significantly affect biosy timing; compared to on-time biopsies, the odds of an early or late biopsy were 1.09 (95%CI: 0.63 to 1.89; p=0.77) and 1.20 (0.70 to 2.08; p=0.51), respectively.

In an exploratory analysis among the 167 participants who underwent RP, 158 men had never used a 5-ARI, while 9 men used 5-ARIs (Table 3). There was no suggestion of a difference in the rate of adverse pathology based on either grade ≥4+3 (p=0.99) or grade ≥4+3 and/or non-organ confined disease (p=0.73). Furthermore, among men who underwent RP, there were no cases of Gleason 8+ disease among 5-ARI users, compared to 12 cases (8%) among non-users (data not shown).

Discussion

To our knowledge, this is the first study to evaluate the relationship between continued 5-ARI use after the initial diagnosis of PCa and the risk of pathological reclassification during subequent AS. We found that continued 5-ARI use after cancer diagnosis did not appear to be associated with higher risk of reclassification of PCa on a subsequent biopsy. Although 5-ARI use was associated with decreased pathological reclassification on unadjusted analysis, when controlling for diagnostic PSA, BMI, prostate size, and ratio of cores positive for PCa to total cores sampled on prostate biopsy, continued 5-ARI use

did not significantly protect against grade and/or volume reclassification. Furthermore, there was no evidence that men who proceeded to prostatectomy while on 5-ARI had worse pathological outcomes than men who did not use 5-ARI.

Several studies have evaluated the effects of 5-ARI initiation after diagnosis on reclassification during AS; however, results have been inconsistent. The clinical benefit of 5-ARI use after diagnosis was first evaluated by Finelli et al., who reported a significantly lower rate of pathological progression (defined as Gleason score >6 or ≥3 cores involved or >50% of core involvement) among men who starting using 5-ARIs after diagnosis.[11,12] More recently, in a review of medical records from AS patients at one academic institution, Dai et al. reported no overall difference in the risk of reclassification (defined as increase in Gleason score or the predominant Gleason pattern) between men who started using 5-ARIs within 12 months of diagnosis and who never used 5-ARIs.[13] In a retrospective analysis of 587 men enrolled in an AS cohort Ross et al. found that initiation of 5-ARI in 47 men was not associated with risk of reclassification (defined as any Gleason ≥ 4, ≥ 3 cores involved with cancer, or > 50% of any core involved with cancer).[8] In contrast, REDEEM, a randomized controlled trial of dutasteride versus placebo among men on AS, reported that men in the dutasteride arm had a significantly lower risk of progression than men in the placebo arm (HR 0.62, 95% CI 0.43-0.89, $p < 0.001$).[9] However the definition of progression used in the REDEEM trial includes definitive treatment (radical prostatectomy, brachytherapy and hormonal treatment) as well as pathological reclassification. Given the differential rate of definitive treatment in the placebo arm (12.3%), compared to the 5-ARI arm (7.5%), the primary results reported for this trial are likely biased by the inclusion of treatment as an endpoint. Moreover, in stratified analyses there was no difference in the risk of pathologic progression between the 5-ARI and placebo arms ($p=0.079$).[9]

Given previous findings of increased risk of high-grade PCa in 5-ARI users in the PCPT[1] and REDUCE[2] trials, there was potential concern for risk of adverse pathology in an AS population. However, consistent with post hoc analyses of both trials,[14,15] we found no evidence to suggest that the occurrence of high grade (Gleason ≥ 4+3) cancer among men who proceeded to RP (p = 0.99) differed between 5-ARI users and non-users. The rate of adverse pathology (Gleason ≥ 4+3 or pT3 or pN1) was also similar between 5-ARI users and non-users (p = 0.73). However, this analysis was limited by the small number of 5-ARI users who elected RP and the potential bias in reasons that men elect RP. While 5-ARI use does not appear to be associated with time to PCa reclassification in the PASS cohort, 5-ARI users were less likely to elect definitive treatment than non users (p=0.04). Avoiding definitive treatment and its associated morbidities[16] may have value to many men who choose AS. While this study does not evaluate the reasons why men avoided radical prostatectomy while using 5-ARIs, the lower treatment rate may be related to the well described phenomenon of decreased PSA rise while on 5-ARIs.[17]

Although studies evaluating the associations of 5-ARI use with risk of progression are presumably interested in biologic effects, interpretation of their results are complicated by the complex relationship between use of these drugs and factors that influence the outcome. For example, 5-ARI use is associated with an approximate 50% decrease in PSA over the first year of use and a continued decline thereafter[17], and since higher PSA is associated with adverse reclassification, 5-ARI use could be expected to decrease the risk of reclassification. However, 5-ARIs are known to decrease prostate size[18,19], and we have shown that smaller prostates are associated with a higher risk of reclassification.[10] Thus, these competing influences may substantially affect the timing of biopsy or the ability to detect reclassification to higher grade cancer. However, in PASS, the use of protocol-directed PSA tests and biopsies at pre-specified

time-points allows a similar opportunity to detect progression, which helps minimize the potential for bias. Indeed, we found no evidence that 5-ARI use affected the timing of biopsies in PASS.

Major strengths of our study include the fact that is a multicenter, prospectively designed study with extensive collection and quality control of clinical data. In addition, the inclusion of Gleason 3+3 and 3+4 disease at diagnosis in the Canary PASS cohort makes the results of this study more generalizable to community AS protocols. This study is not without limitations. First, 5-ARI use was determined through self-report at study entry with discrete response options for duration of use, which could result in inaccurate assessment of the duration of 5-ARI use. In addition, complete data on overall duration or duration prior to diagnosis were not available. Second, because few participants reported discontinuing 5-ARI use after diagnosis, we could not examine associations of discontinued use with progression. In addition, data on type of 5-ARI used were not available. Finally, the number of men who used 5-ARI in the RP cohort was small, making it difficult to draw definitive conclusions regarding pathologic outcomes.

Conclusions

Continued 5-ARI use in men diagnosed with PCa does not appear to affect risk of pathological reclassification while on active surveillance in the PASS cohort. Our data suggest that men on 5-ARIs who have radical prostatectomy after a period of AS do not have increased incidence of high-grade PCa, and 5-ARI users undergo definitive treatment at lower rates than non-users.

Acknowledgements

This work was supported by the Canary Foundation, NIH grants R01CA181605 and P50CA097186, the Institute for Prostate Cancer Research, and the Department of Defense Prostate Cancer Research Program grant W81XWH1410595.

References

1. Thompson IM, Goodman PJ, Tangen CM, et al: The Influence of Finasteride on the Development of Prostate Cancer. N. Engl. J. Med. 2003; **349**: 215–224.

2. Andriole GL, Bostwick DG, Brawley OW, et al: Effect of Dutasteride on the Risk of Prostate Cancer. N. Engl. J. Med. 2010; **362**: 1192–1202.

3. Research C for DE and: Drug Safety and Availability - FDA Drug Safety Communication: 5-alpha reductase inhibitors (5-ARIs) may increase the risk of a more serious form of prostate cancer. Available at: http://www.fda.gov/drugs/drugsafety/ucm258314.htm, accessed January 12, 2017.

4. Roehrborn CG, Andriole GL, Wilson TH, et al: Effect of Dutasteride on Prostate Biopsy Rates and the Diagnosis of Prostate Cancer in Men with Lower Urinary Tract Symptoms and Enlarged Prostates in the Combination of Avodart and Tamsulosin Trial. Eur. Urol. 2011; **59**: 244–249.

5. Sanda MG, Chen RC, Crispino T, et al: Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. Risk 2017; **6**: 27.

6. Anon: Prostate Cancer (Version 2.2017. Natl. Compr. Cancer Netw. Available at: https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf, accessed June 26, 2017.

7. Cooperberg MR and Carroll PR: Trends in Management for Patients With Localized Prostate Cancer, 1990-2013. JAMA 2015; **314**: 80–82.

8. Ross AE, Feng Z, Pierorazio PM, et al: Effect of treatment with 5-α reductase inhibitors on progression in monitored men with favourable-risk prostate cancer. BJU Int. 2012; **110**: 651–657.

9. Fleshner NE, Lucia MS, Egerdie B, et al: Dutasteride in localised prostate cancer management: the REDEEM randomised, double-blind, placebo-controlled trial. The Lancet 2012; **379**: 1103–1111.

10. Newcomb LF, Thompson IM, Boyer HD, et al: Outcomes of Active Surveillance for Clinically Localized Prostate Cancer in the Prospective, Multi-Institutional Canary PASS Cohort. J. Urol. 2016; **195**: 313–320.

11. Finelli A, Trottier G, Lawrentschuk N, et al: Impact of 5α-Reductase Inhibitors on Men Followed by Active Surveillance for Prostate Cancer. Eur. Urol. 2011; **59**: 509–514.

12. Wong L-M, Fleshner N and Finelli A: Impact of 5-Alpha Reductase Inhibitors on Men Followed by Active Surveillance for Prostate Cancer: A Time-dependent Covariate Reanalysis. Eur. Urol. 2013; **64**: 343.

13. Dai C, Ganesan V, Zabell J, et al: Impact of 5α-Reductase Inhibitors on Disease Reclassification among Men on Active Surveillance for Localized Prostate Cancer with Favorable Features. J. Urol. 2017. Available at: http://linkinghub.elsevier.com/retrieve/pii/S002253471777259X, accessed November 16, 2017.

14. Lucia MS, Epstein JI, Goodman PJ, et al: Finasteride and high-grade prostate cancer in the Prostate Cancer Prevention Trial. J. Natl. Cancer Inst. 2007; **99**: 1375–1383.

15. Andriole GL, Bostwick DG, Gomella LG, et al: Modeling and analysis of Gleason score 8-10 prostate cancers in the REDUCE study. Urology 2014; **84**: 393–399.

16. Chou R, Croswell JM, Dana T, et al: Screening for Prostate Cancer: A Review of the Evidence for the U.S. Preventive Services Task Force. Ann. Intern. Med. 2011; **155**: 762.

17. Andriole GL, Marberger M and Roehrborn CG: Clinical Usefulness of Serum Prostate Specific Antigen for the Detection of Prostate Cancer is Preserved in Men Receiving the Dual 5α-Reductase Inhibitor Dutasteride. J. Urol. 2006; **175**: 1657–1662.

18. Moore E, Bracken B, Bremner W, et al: Proscar: five-year experience. Eur. Urol. 1995; **28**: 304–309.

19. Roehrborn CG, Marks LS, Fenter T, et al: Efficacy and safety of dutasteride in the four-year treatment of men with benign prostatic hyperplasia. Urology 2004; **63**: 709–715.

Table and Figure Titles and Legends

Figure 1. Kaplan-Meier analysis demonstrating time to pathological reclassification based on 5-ARI use for a) any increase in Gleason grade and/or tumor volume to ≥34% ratio of positive core, and b) increase in Gleason grade only.

Table 1. Demographic and clinical information

	No 5-ARI	5-ARI	P-value^
N	902	107	
Reclassified, n (%)			
Gleason grade and/or volume	327 (36)	32 (30)	
Gleason grade only	284 (31)	31 (29)	
Race, n (%)			0.16
Caucasian American	804 (89)	102 (95)	
African American	59 (7)	3 (3)	
other	39 (4)	2 (2)	
BPH, n (%)	254 (28)	82 (77)	< 0.001
Prostate volume (cc), median [IQR]	40 [30 - 54]	51 [34 - 67]	< 0.001
Age, mean (SD)	62 (7)	65 (7)	< 0.001
PSA, median [IQR]	4.8 [3.6 - 6.3]	5.0 [3.6 - 7.0]	0.25
PSA density, median [IQR]	0.11 [0.08 - 0.16]	0.10 [0.07 - 0.15]	0.06
Clinical stage, n (%)			0.51
T1a-T1c	803 (89)	93 (87)	
T2a-T2c	99 (11)	14 (13)	
Gleason score, n (%)			0.20
3+3	847 (94)	97 (91)	
3+4	55 (6)	10 (9)	
Cores ratio, median [IQR]	8 [8 - 17]	8 [8 - 17]	0.74
Family history of PCa, n (%)	259 (29)	21 (21)	0.12
BMI, mean (SD)	27.9 (4.3)	27.0 (4.0)	0.05
IPSS, median [IQR]	6 [3 - 11]	9 [6 - 12]	< 0.001
Any treatment, n (%)*	254 (28)	20 (19)	0.04
Radical prostatectomy, n (%)	160 (18)	9 (8)	0.01

* subsequent radiation or prostatectomy

^ P-value from either t-test or Wilcoxon sign rank test for continuous variables, and from Chi-Square test or Fisher's test for categorical variables

5-ARI: 5-alpha reductase inhibitor; BPH: benign prostatic hyperplasia; PSA: prostate specific antigen; PCa: prostate cancer; BMI: body mass index; IPSS: international prostate symptom score

Table 2. Unadjusted and adjusted time to event model for grade and/or volume reclassification

Variable	Unadjusted		Adjusted	
	HR (95% CI)^	**P-value^**	**HR (95% CI)^**	**P-value^**
5-ARI use	0.63 (0.43, 0.94)	0.02	0.81 (0.55, 1.21)	0.31
Log(PSA)*	1.41 (1.17, 1.71)	0.0003	1.75 (1.44, 2.13)	<.0001
BMI	1.03 (1.00, 1.06)	0.02	1.04 (1.02, 1.07)	0.001
Log(prostate volume)	0.52 (0.41, 0.66)	<.0001	0.45 (0.35, 0.57)	<.0001
Cores ratio*	1.06 (1.04, 1.07)	<.0001	1.05 (1.03, 1.06)	<.0001

* at diagnosis

^ Hazard ratios, 95% confidence intervals and p-values from Cox proportional hazard models

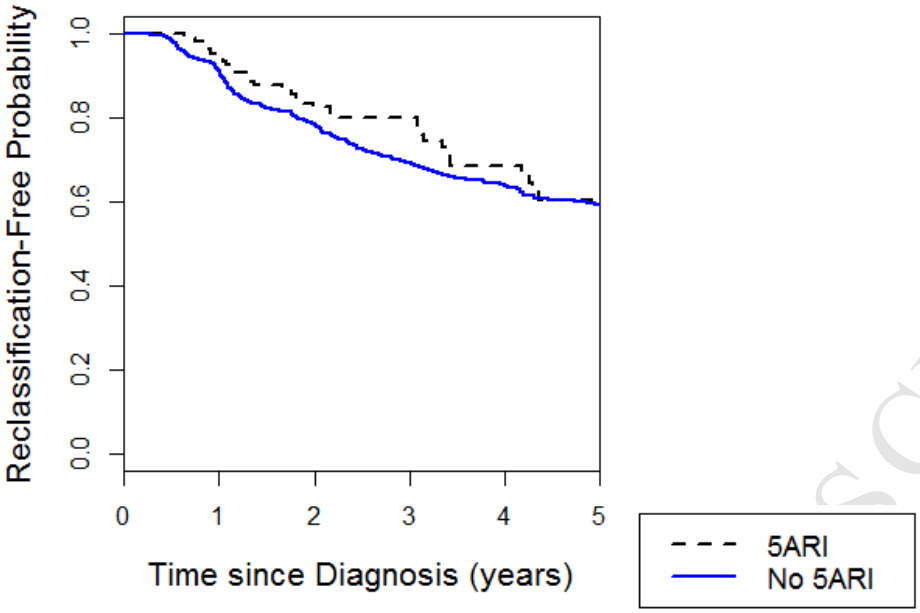
5-ARI: 5-alpha reductase inhibitor; PSA: prostate specific antigen; BMI: body mass index

Table 3. Adverse pathology outcomes at radical prostatectomy

Outcome, n(%)	**No 5-ARI n=158**	**5-ARI n=9**	**p-value^**
Gleason grade ≥ 4+3 only	39 (25)	2 (22)	0.99
Gleason grade ≥ 4+3 and/or non organ-confined disease	59 (37)	4 (44)	0.73

^ P-value from Fisher's exact test
5-ARI: 5-alpha reductase inhibitor

Figure 1a

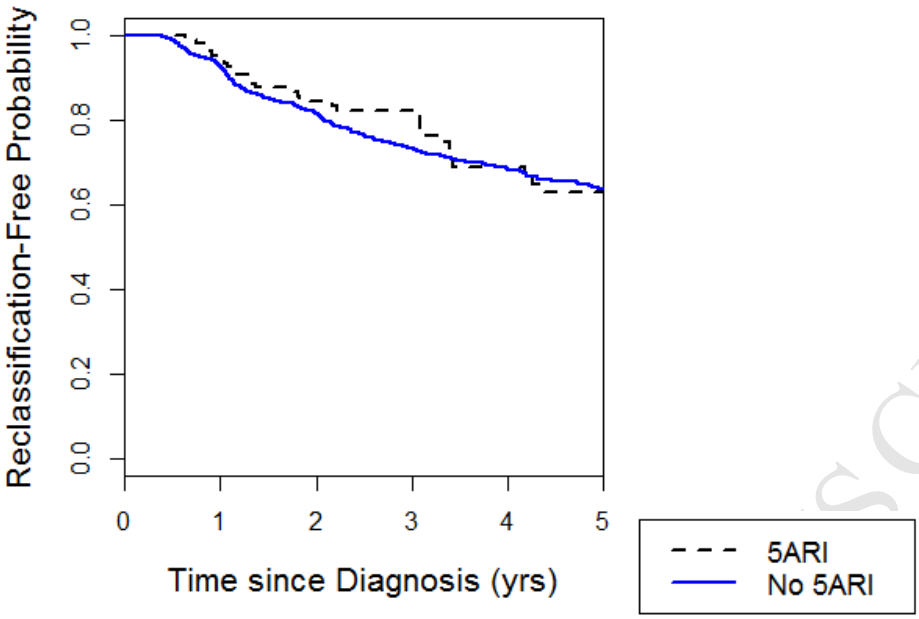


Log-rank test p = 0.10

Number at risk:

	Time since Diagnosis (years)				
	1	**2**	**3**	**4**	**5**
5-ARI user	100	76	59	41	24
5-ARI non user	784	604	422	297	182

Figure 1b



Log-rank test p = 0.30

Number at risk:

	Time since Diagnosis (years)				
	1	**2**	**3**	**4**	**5**
5-ARI user	100	78	60	41	25
5-ARI non user	792	616	434	305	186

Abbreviation	Definition
AS	Active surveillance
5-ARI	5-alpha reductase inhibitor
PASS	Canary Prostate Active Surveillance Study
RP	Radical prostatectomy
PCa	Prostate cancer
BPH	Benign prostatic hyperplasia
PSA	Prostate specific antigen
MRI	Magnetic resonance imaging
BMI	Body mass index