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

Urine *TMPRSS2:ERG* Plus PCA3 for Individualized Prostate Cancer Risk Assessment

Scott A. Tomlins [a,b,c,*], *John R. Day* [d], *Robert J. Lonigro* [a,c], *Daniel H. Hovelson* [e], *Javed Siddiqui* [a], *L. Priya Kunju* [a], *Rodney L. Dunn* [b], *Sarah Meyer* [d], *Petrea Hodge* [d], *Jack Groskopf* [d], *John T. Wei* [b], *Arul M. Chinnaiyan* [a,b,c,e,f,**]

[a] *Michigan Center for Translational Pathology, Department of Pathology, University of Michigan Medical School, Ann Arbor, MI, USA;* [b] *Department of Urology, University of Michigan Medical School, Ann Arbor, MI, USA;* [c] *Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, MI, USA;* [d] *Hologic/Gen-Probe Inc., San Diego, CA, USA;* [e] *Department of Computational Medicine and Bioinformatics, University of Michigan Medical School, Ann Arbor, MI, USA;* [f] *Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, MI, USA*

Article info

Article history:
Accepted April 29, 2015

Associate Editor:
Giacomo Novara

Keywords:
Prostate cancer
Urine biomarkers
Early detection
Gene fusions
PCA3

Abstract

Background: *TMPRSS2:ERG* (T2:ERG) and prostate cancer antigen 3 (PCA3) are the most advanced urine-based prostate cancer (PCa) early detection biomarkers.
Objective: Validate logistic regression models, termed Mi-Prostate Score (MiPS), that incorporate serum prostate-specific antigen (PSA; or the multivariate Prostate Cancer Prevention Trial risk calculator version 1.0 [PCPTrc]) and urine T2:ERG and PCA3 scores for predicting PCa and high-grade PCa on biopsy.
Design, setting, and participants: T2:ERG and PCA3 scores were generated using clinical-grade transcription-mediated amplification assays. Pretrained MiPS models were applied to a validation cohort of whole urine samples prospectively collected after digital rectal examination from 1244 men presenting for biopsy.
Outcome measurements and statistical analysis: Area under the curve (AUC) was used to compare the performance of serum PSA (or the PCPTrc) alone and MiPS models. Decision curve analysis (DCA) was used to assess clinical benefit.
Results and limitations: Among informative validation cohort samples (n = 1225 [98%], 80% from patients presenting for initial biopsy), models incorporating T2:ERG had significantly greater AUC than PSA (or PCPTrc) for predicting PCa (PSA: 0.693 vs 0.585; PCPTrc: 0.718 vs 0.639; both $p < 0.001$) or high-grade (Gleason score >6) PCa on biopsy (PSA: 0.729 vs 0.651, $p < 0.001$; PCPTrc: 0.754 vs 0.707, $p = 0.006$). MiPS models incorporating T2:ERG score had significantly greater AUC (all $p < 0.001$) than models incorporating only PCA3 plus PSA (or PCPTrc or high-grade cancer PCPTrc [PCPThg]). DCA demonstrated net benefit of the MiPS_PCPTrc (or MiPS_PCPThg) model compared with the PCPTrc (or PCPThg) across relevant threshold probabilities.
Conclusions: Incorporating urine T2:ERG and PCA3 scores improves the performance of serum PSA (or PCPTrc) for predicting PCa and high-grade PCa on biopsy.
Patient summary: Incorporation of two prostate cancer (PCa)-specific biomarkers (*TMPRSS2:ERG* and *PCA3*) measured in the urine improved on serum prostate-specific antigen (or a multivariate risk calculator) for predicting the presence of PCa and high-grade PCa on biopsy. A combined test, Mi-Prostate Score, uses models validated in this study and is clinically available to provide individualized risk estimates.

* Corresponding author. University of Michigan Medical School, 1524 BSRB, 109 Zina Pitcher Place, Ann Arbor, MI 48109-2200, USA. Tel. +1 734 764 1549.
** Corresponding author. University of Michigan Medical School, 1400 E. Medical Center Drive, 5316 CCGC, Ann Arbor, MI 48109-0602, USA. Tel. +1 734 615 4062.
E-mail addresses: arul@umich.edu (A.M. Chinnaiyan), tomlinss@umich.edu (S.A. Tomlins).

http://dx.doi.org/10.1016/j.eururo.2015.04.039
0302-2838/

1. Introduction

Approximately 1 million men undergo prostate biopsy each year in the United States, most for elevated serum prostate-specific antigen (PSA or KLK3). Serum PSA's lack of prostate cancer (PCa) specificity, the unclear benefits of PSA screening for reducing PCa deaths, and the harms of overdiagnosing indolent disease have called PSA screening into question [1–3]. Although aggressive PCa-specific biomarkers may eventually replace serum PSA, at present, methods to individualize management of elevated PSA are needed. Such approaches include multivariate risk models, such as the Prostate Cancer Prevention Trial risk calculator (PCPTrc), which includes serum PSA and clinical factors [4–6]. Likewise, multiple PSA derivatives and other related kallikreins have been advanced as early detection biomarkers, including free PSA and [−2]proPSA (both of which are incorporated, with total PSA, in the Prostate Health Index [PHI]), with free PSA and PHI approved by the US Food and Drug Administration (FDA) for PCa risk estimation in men with serum PSA of 4–10 ng/ml [7–9]. Similarly, a panel of free and total PSA, single-chain intact PSA, and a related kallikrein (KLK2) outperforms serum PSA alone for predicting PCa on biopsy, and a test incorporating these kallikreins along with clinical parameters (4Kscore) is available [7,9].

An alternative to using tissue-specific biomarkers, such as serum PSA and other kallikreins, for predicting the presence of PCa is to utilize PCa-specific biomarkers. Prostate cancer antigen 3 (PCA3; a noncoding RNA) and *TMPRSS2:ERG* (T2:ERG) gene fusions are the most advanced PCa-specific early detection biomarkers [10–13]. In tissues, both biomarkers show markedly improved PCa specificity compared with PSA or derivatives or related kallikreins [12,14,15]. In addition, both PCA3 and T2:ERG transcripts are detectable and quantifiable in urine collected after digital rectal examination (DRE) [10–13]. The Progensa PCA3 test (Hologic Inc, Bedford, MA, USA), which reports a quantitative PCA3 score using a transcription-mediated amplification (TMA) assay, has been extensively studied as a urine-based PCa biomarker [11–13] and is FDA approved for estimating PCa risk following a negative biopsy.

Previously, we reported the development and application of a clinical-grade TMA assay for quantifying T2:ERG messenger RNA (mRNA), which generates a T2:ERG score by normalizing urine T2:ERG mRNA to urine PSA mRNA (to control for prostate cell and mRNA abundance) [16]. This assay is based on the same technology as the Progensa PCA3 test and can be performed on the same post-DRE whole urine sample. Previously, we applied initial T2:ERG TMA assay versions to post-DRE whole urine from 1312 men presenting for biopsy or prostatectomy at multiple centers [16]. More recently, we and others have evaluated the performance of a final clinical-grade T2:ERG TMA assay [17–20]. In this study, we evaluated pretrained multivariate regression models combining urine T2:ERG and/or PCA3 scores with serum PSA (or the PCPTrc) in a large independent validation cohort to develop methods for individualized PCa risk estimates.

2. Materials and methods

2.1. Patients

The regression models were developed in a training cohort and validated in an independent cohort. For the training cohort, post-DRE urine was prospectively collected from 733 patients presenting for diagnostic prostate biopsy at three US academic institutions and assessed for urine T2:ERG and PCA3 scores at the University of Michigan Health System (training cohort), predominantly as part of an Early Detection Research Network (EDRN) biopsy cohort [21], using standardized protocols. For the validation cohort, prospectively collected post-DRE urine samples were obtained from 1244 men presenting for diagnostic biopsy at seven community clinics throughout the United States. In both cohorts, men with prior treatment for PCa or surgical treatment of the prostate within 6 mo of urine collection (or previous biopsy within 6 wk) were excluded. In the validation cohort, men with a history of PCa were excluded. A flow diagram of all specimens from the training and validation cohorts is shown in Figure 1. Additional cohort information is available in the Supplement and Supplementary Table 1.

2.2. Urine T2:ERG and PCA3 score generation

Urine processing and T2:ERG and PCA3 score determination on the same post-DRE whole urine specimen using TMA assays were performed as described [16,17,19]. Details of the T2:ERG and PCA3 (Progensa) assays are provided in the Supplement and Supplementary Table 2.

2.3. Statistical analysis

Statistical analyses were performed using R version 2.10.1 (R Foundation for Statistical Computing, http://www.R-project.org) or MedCalc, version 12.4.0.0. Two-tailed tests were used, and p values <0.05 were considered statistically significant.

The ability of combinations of various biomarker (serum PSA, PCPTrc [version 1.0], high-grade cancer PCPTrc [PCPThg; version 1.0], and urine T2:ERG and PCA3 scores) to predict cancer (vs no cancer) or high-grade cancer (Gleason score >6 vs 6 plus no cancer) on biopsy was assessed by multivariable logistic regression models. Diagnostic potential was quantified using the area under the receiver operating characteristic curve (AUC). Decision curve analysis (DCA) was performed using the DCA R package, as described [22]. All models and additional details, including AUC comparison and DCA methodology, are described in the Supplement and Supplementary Table 3.

3. Results

3.1. Development of logistic regression models incorporating urine T2:ERG and PCA3 scores

The multivariable logistic regression models evaluated in this study were developed using a 733-specimen training cohort. Of the 711 samples (97%) that were informative for both urine T2:ERG and PCA3 scores (sufficient urine PSA [10 000 copies per milliliter] to ensure adequate prostatic-derived RNA), 689 samples were collected as part of an EDRN protocol [21] and were from men without PCa presenting for biopsy. The remaining 22 informative samples were collected in the same manner from men with PCa (on active surveillance) presenting for rebiopsy. Models were trained on all 711 informative samples in the

Please cite this article in press as: Tomlins SA, et al. Urine *TMPRSS2:ERG* Plus PCA3 for Individualized Prostate Cancer Risk Assessment. Eur Urol (2015), http://dx.doi.org/10.1016/j.eururo.2015.04.039

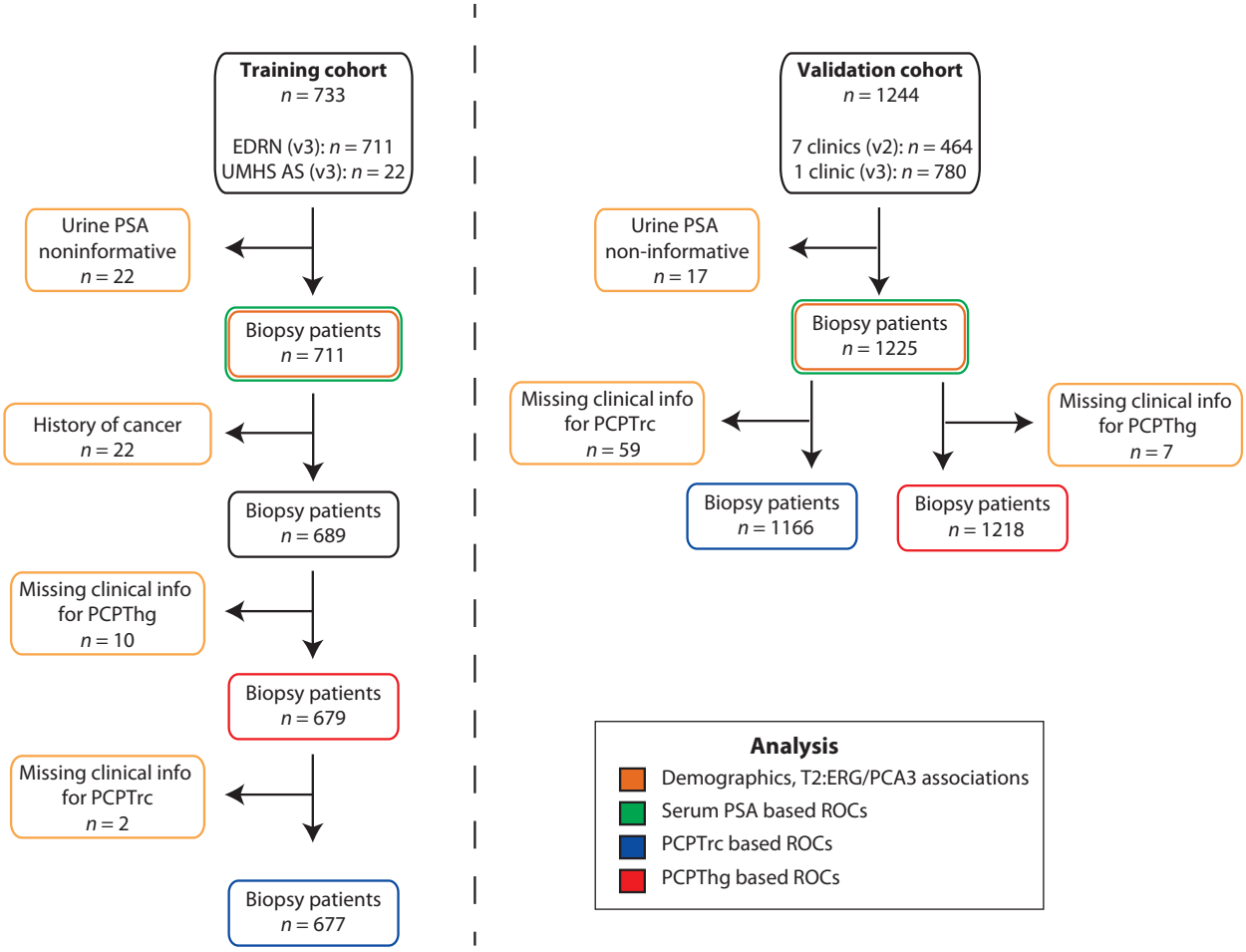

Fig. 1 – Flow diagram of specimens in the training and validation cohort. Specimen cohorts for all urine samples assessed for *TMPRSS2:ERG* (T2:ERG) and PCA3 are shown. Samples excluded from various analyses (indicated by legend and described in the text) are indicated in orange. Samples assessed using version 2 or version 3 (final assay) T2:ERG transcription-mediated amplification assays are indicated.
AS = active surveillance; EDRN = Early Detection Research Network; PCPTrc = Prostate Cancer Prevention Trial risk calculator; PCPThg = PCPT high grade cancer risk calculator; PSA = prostate-specific antigen; ROC = receiver operating characteristic; T2:ERG = *TMPRSS2:ERG*;UMHS = University of Michigan Health System; v3 = version 3.

training cohort (98% of men underwent ≥12 core biopsy) and incorporated serum PSA (or PCPTrc), urine T2:ERG score, and/or PCA3 score for predicting PCa presence on biopsy. Similar models were also trained using serum PSA (or PCPThg) for predicting high-grade (Gleason score >6) cancer (Supplementary Table 3).

These pretrained models were then evaluated in an independent cohort consisting of 1244 prospectively collected, post-DRE urine samples from seven US community clinics assessed for urine T2:ERG and PCA3 scores at Gen-Probe (San Diego, CA, USA; validation cohort). Clinicopathologic characteristics of the 1225 of 1244 (98%) validation cohort specimens informative for both urine T2:ERG and PCA3 scores are given in Table 1. Of note, in the validation cohort, no patient had a previous history of PCa, 80% were from initial biopsy, 73% were white, and 99% underwent ≥12 core biopsy. Moreover, 42% and 18% of patients were diagnosed with cancer and high-grade cancer, respectively, on biopsy. A flow diagram of all samples from training and validation cohorts is shown in Figure 1. Associations with clinicopathologic parameters are shown in Table 1 and described in the Supplement.

3.2. *Incorporating urine T2:ERG and PCA3 with serum prostate-specific antigen or the PCPT risk calculators for predicting cancer or high-grade cancer on biopsy*

We next assessed the ability of the trained models incorporating urine T2:ERG and PCA3, either alone or in combination, to improve on serum PSA or the PCPTrc for predicting the presence of cancer on biopsy in the validation cohort using AUC comparisons (Table 2). Among 1225 informative validation cohort patients, AUCs for PSA, PSA plus T2:ERG score, PSA plus PCA3 score, and Mi-Prostate Score (MiPS) were 0.585, 0.693, 0.726, and 0.751, respectively. Among 1166 informative validation cohort patients with calculable PCPTrc scores, AUCs for PCPTrc, PCPTrc plus T2:ERG, PCPTrc plus PCA3, and PCPTrc plus T2:ERG plus PCA3 (MiPS_PCPT) were 0.639, 0.718, 0.739, and 0.762, respectively (Table 2). Calibration

Please cite this article in press as: Tomlins SA, et al. Urine *TMPRSS2:ERG* Plus PCA3 for Individualized Prostate Cancer Risk Assessment. Eur Urol (2015), http://dx.doi.org/10.1016/j.eururo.2015.04.039

Table 1 – Associations of urine T2:ERG score, urine PCA3 score, and the Mi-Prostate Score risk model for predicting cancer on biopsy and clinicopathologic parameters in evaluable validation cohort patients (*n* = 1225)

Parameter	Patients, *n*	Median, IQR	T2:ERG score, r_s	*p* value	PCA3 score, r_s	*p* value	MiPS model risk, r_s	*p* value
Age, yr	1225	64 (58–70)	0.10	*0.0004*	0.32	*< 0.0001*	0.32	*<0.0001*
Serum PSA	1225	4.7 (3.3–6.5)	0.07	*0.01*	0.11	*< 0.0001*	0.33	*<0.0001*
Ultrasound volume, ml	1181	48 (35–68)	−0.02	0.55	−0.08	*0.006*	0.00	0.90
PSAD	1181	0.091 (0.057–0.140)	0.08	*0.008*	0.17	*< 0.0001*	0.32	*<0.0001*
PCPTrc risk, %	1166	40 (33–51)	0.09	*0.002*	0.16	*< 0.0001*	0.32	*<0.0001*
PCPTrc high-grade risk, %	1218	11 (6–18)	0.09	*0.001*	0.25	*< 0.0001*	0.40	*<0.0001*
Bx cores, positive, no.	518	3 (1–5)	0.24	*<0.0001*	0.24	*< 0.0001*	0.34	*<0.0001*
Bx cores, positive, %	518	25 (8–42)	0.23	*<0.0001*	0.23	*< 0.0001*	0.34	*<0.0001*
Greatest involvement of a single bx core, %	201	38 (17–64)	0.27	*0.0001*	0.03	0.67	0.25	*0.0003*

Parameter	Total patients, *n*	Patients, *n* (%)	T2:ERG score, median (IQR)	*p* value	PCA3 score, median (IQR)	*p* value	MiPS model risk, median (IQR)	*p* value
Diagnosis	1225							
Noncancer		707 (58)	2 (0–14)	*<0.0001*	15 (8–36)	*<0.0001*	30 (17–50)	*<0.0001*
Cancer		518 (42)	15 (1–71)		40 (19–84)		60 (38–76)	
Diagnosis	707							
Atypia and/or HGPIN		269 (38)	6 (0.5–25)	*<0.0001*	26 (12–57)	*<0.0001*	43 (26–63)	*<0.0001*
Other benign		438 (62)	1 (0–7)		12 (6–26)		24 (14–39)	
Race	1222							
White		890 (73)	5 (0.1–35)	0.07	26 (11–58)	*0.002*	42 (23–67)	*0.01*
Not white		332 (27)	4 (0–25)		21 (8–50)		38 (19–62)	
Family History	1170							
Negative		940 (80)	4 (0.1–33)	0.44	23 (10–55)	0.23	39 (21–65)	0.10
Positive		230 (20)	6 (0.3–28)		28 (11–61)		46 (24–67)	
DRE	1223							
Normal		936 (77)	5 (0.2–31)	0.40	24 (10–54)	0.22	42 (23–66)	0.17
Abnormal		287 (23)	5 (0–35)		27 (10–58)		37 (20–65)	
Previous biopsy	1223							
No		977 (80)	5 (0.1–33)	0.25	24 (10–57)	0.51	40 (21–66)	0.21
Yes		246 (20)	4 (0–31)		24 (12–53)		43 (25–65)	
Bx cores (#)	773							
<12		11 (1)	6 (0.5–70)	0.96	39 (17–68)	0.98	62 (50–79)	0.28
12		687 (89)	7 (0.3–48)		32 (15–73)		50 (30–71)	
>12		75 (10)	6 (0.2–94)		32 (16–76)		44 (32–75)	
Bx Gleason score	518							
6		294 (57)	11 (0.7–52)	*0.008*	34 (16–75)	*0.0006*	53 (32–72)	*<0.0001*
>6		224 (43)	23 (2–113)		50 (23–94)		67 (46–81)	
Clinical stage	517							
T1		382 (74)	16 (1–68)	0.81	37 (19–78)	0.17	59 (39–76)	0.92
>T1		135 (26)	12 (0.4–95)		47 (20–95)		59 (35–77)	
Bx signficance[a]	195							
Insignificant		32 (16)	5 (0.3–29)	0.24	34 (15–55)	*0.008*	44 (28–66)	*0.007*
Significant		163 (84)	15 (0.4–70)		50 (23–96)		63 (42–78)	
Bx significance (pathology)[b]	201							
Insignificant		65 (32)	5 (0.3–29)	*0.004*	36 (16–71)	*0.002*	45 (28–64)	*<0.0001*
Significant		136 (68)	22 (1–97)		55 (28–103)		68 (45–82)	

Bx = biopsy; bx = previous biopsy; DRE = digital rectal examination; HGPIN = high-grade prostatic intraepithelial neoplasia; IQR = interquartile range; MiPS = Mi-Prostate Score, serum PSA plus T2:ERG plus PCA3; PCPTrc = Prostate Cancer Prevention Trial risk calculator; PSA = prostate-specific antigen; T2:ERG = *TMPRSS2:ERG*.

For all evaluable patients in the validation cohort, the number of patients with data for each parameter, the median (and interquartile range), and correlations (Spearman's rho [r_s]) and *p* values with urine T2:ERG score, urine PCA3 score, and the urine T2:ERG score plus urine PCA3 score plus serum PSA (MiPS) logistic regression model (logit) for predicting the presence of prostate cancer on biopsy are given in the upper panel. In the lower panel, the total number of patients with data for each parameter, the number (and percentage) of patients within each categorical parameter, and the median T2:ERG score, PCA3 score, and MiPS model (logit*100) and *p* value for each categorical comparison (Mann-Whitney or Kruskal-Walis test) are given.

[a] Any clinical stage higher than T1c, PSA density (serum PSA/prostate volume on ultrasound) ≥0.15 ng/ml/ml, Gleason score >6, ≥25% cores positive, or >50% greatest single core involvement as significant.

[b] Any Gleason score >6, ≥25% cores positive, or >50% greatest single core involvement as significant.

plots demonstrate that MiPS_PCPT predicted and observed cancer risks were similar in this cohort, with the MiPS_PCPT model showing enhanced risk stratification compared with the PCPT model (Fig. 2A; Supplementary Fig. 1a).

We then assessed the ability of the trained models incorporating urine T2:ERG and PCA3, either alone or in combination, to improve on serum PSA or PCPThg for predicting the presence of high-grade cancer (Gleason score >6) on biopsy. Among 1225 informative validation cohort

Please cite this article in press as: Tomlins SA, et al. Urine *TMPRSS2:ERG* Plus PCA3 for Individualized Prostate Cancer Risk Assessment. Eur Urol (2015), http://dx.doi.org/10.1016/j.eururo.2015.04.039

Table 2 – Performance of serum prostate-specific antigen, Prostate Cancer Prevention Trial, and Mi-Prostate Score–based models for predicting cancer and high-grade cancer on biopsy

Model	Prediction	*n*	AUC	*p* value vs PSA (or PCPT)	*p* value vs T2:ERG or PCA3
PSA	Cancer	1225	0.585	NA	NA
PSA plus T2:ERG			0.693	<0.001	NA
PSA plus PCA3			0.726	<0.001	<0.05
PSA plus T2:ERG plus PCA3 (MiPS)			0.751	<0.001	<0.001, <0.001
PCPTrc	Cancer	1166	0.639	NA	NA
PCPTrc plus T2:ERG			0.718	<0.001	NA
PCPTrc plus PCA3			0.739	<0.001	NS
PCPTrc plus T2:ERG plus PCA3 (MiPS_PCPT)			0.762	<0.001	<0.001, <0.001
PSA	HG cancer	1225	0.651	NA	NA
PSA plus T2:ERG			0.729	<0.001	NA
PSA plus PCA3			0.747	<0.001	NS
PSA plus T2:ERG plus PCA3 (MiPShg)			0.772	<0.001	<0.01, <0.001
PCPThg	HG cancer	1218	0.707	NA	NA
PCPThg plus T2:ERG			0.754	<0.01	NA
PCPThg plus PCA3			0.752	<0.01	NS
PCPThg plus T2:ERG plus PCA3 (MiPS_PCPThg)			0.779	<0.001	<0.05, <0.001

AUC = area under the curve; HG = high grade; MiPS = Mi-Prostate Score; MiPShg = Mi-Prostate Score, high grade; NA = not available; NS = not significant; PCPThg = Prostate Cancer Prevention Trial high-grade risk calculator; PCPTrc = Prostate Cancer Prevention Trial risk calculator; PSA = prostate-specific antigen; T2:ERG = *TMPRSS2:ERG*.

Trained logistic regression models incorporating urine T2:ERG and/or PCA3 scores along with serum PSA, the PCPTrc or PCPThg for predicting cancer or HG cancer (Gleason score >6) were evaluated in a separate validation cohort of 1225 evaluable patients. The number of patients with clinicopathologic information needed for the given model and informative urine samples are given, along with AUCs for each model. Significant differences in AUCs (for combined models compared with PSA (or PCPTrc-based models) are indicated. Significant differences for PSA plus T2:ERG vs PSA plus PCA3 models or MiPS vs PSA plus T2:ERG and PSA plus PCA3 models (or equivalent PCPTrc-based models) are also given.

patients, AUCs for PSA, PSA plus T2:ERG score, PSA plus PCA3 score, and PSA plus T2:ERG plus PCA3 (MiPShg) were 0.651, 0.729, 0.747, and 0.772, respectively (Table 2). Among 1218 informative validation cohort patients with calculable PCPThg scores, AUCs for PCPThg, PCPThg plus T2:ERG score, PCPThg plus PCA3 score, and PCPThg plus T2:ERG plus PCA3 (MiPS_PCPThg) were 0.707, 0.754, 0.752, and 0.779, respectively (Table 2). As shown in Table 2, all MiPS models showed significantly greater AUC than single-biomarker (PSA or PCPT) or two-biomarker (PSA or PCPT plus T2:ERG or PCA3) models. Calibration plots demonstrate that MiPS_PCPThg modestly overestimated high-grade cancer risk (most notably at observed vs predicted risks likely to trigger biopsy regardless), and the MiPS_PCPThg model showed enhanced risk stratification compared with the PCPThg model (Fig. 2B; Supplementary Fig. 1b).

3.3. *Impact of MiPS-based models on biopsies averted and high-grade cancer diagnoses delayed*

Clinical consequences of using various cut-offs of the MiPS and PCPT–based model risk predictions (compared with the strategy of biopsying all patients), including the number of biopsies that could have been avoided and the number of high-grade cancer diagnoses that would have been delayed, is shown in Table 3. To further assess potential clinical benefit of incorporating T2:ERG and PCA3 scores into the PCPTrc and PCPThg, we performed DCA using predicted risk probabilities in the validation cohort. DCA sums net benefits (true positives) and subtracts harms (false positives) to determine net clinical benefit. False positives are weighted by a factor related to the relative harm of a missed cancer compared with a negative biopsy; this factor is derived from the probability of PCa (or high-grade cancer) at which a patient would choose to be biopsied (threshold probability). As shown in Figure 2C and 2D, compared with the PCPT-based models or the strategy of biopsying all patients, the MiPS-based models resulted in net clinical benefit across a wide range of relevant threshold probabilities. Likewise, across relevant threshold probabilities, DCA demonstrated that biopsying based on MiPS-based models would reduce the number of biopsies compared with PCPT-based models or biopsying all patients (Fig. 2E and 2F). Together, these results demonstrate that at relevant threshold probabilities, MiPS-based strategies are clinically superior to serum PSA (or PCPT)-based strategies for making biopsy decisions.

3.4. *MiPS performance across clinical subsets*

Although urine PCA3 is FDA approved in the repeat biopsy setting, recent reports demonstrate utility at initial biopsy [10–12,21,23]. Hence, we assessed the performance of MiPS-based models in comparison to serum PSA and PCPT-based models in subsets of our validation cohort stratified by presentation for initial versus repeat biopsy, normal versus abnormal DRE status, and serum PSA (<3, ≥3 and ≤10, and >10 ng/ml). As shown in Supplementary Table 4, in 55 of 56 comparisons (98%) from 14 subsets with >50 patients, MiPS-based models had higher AUC than serum PSA or PCPT-based models for predicting cancer or high-grade cancer on biopsy. For example, among 557 patients who presented for initial biopsy with serum PSA ≥3 and ≤10 ng/ml and normal DRE, AUCs for PCPThg and

Please cite this article in press as: Tomlins SA, et al. Urine *TMPRSS2:ERG* Plus PCA3 for Individualized Prostate Cancer Risk Assessment. Eur Urol (2015), http://dx.doi.org/10.1016/j.eururo.2015.04.039

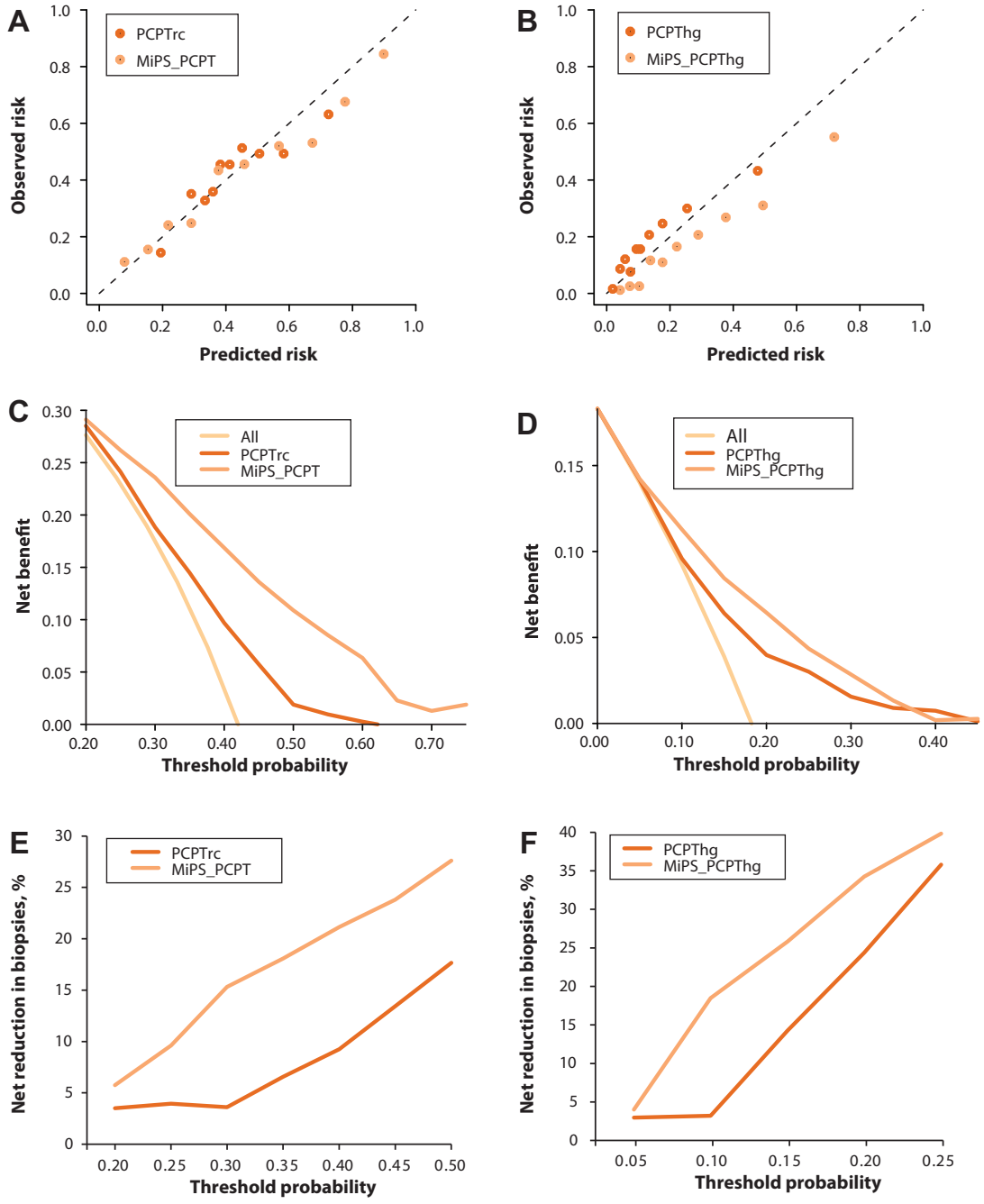

Fig. 2 – Calibration plots and decision curve analysis (DCA) demonstrate enhanced risk stratification and net clinical benefit of Mi-Prostate Score (MiPS) models compared with Prostate Cancer Prevention Trial risk calculators for predicting prostate cancer (PCPTrc) or high-grade cancer (PCPThg) on biopsy. Logistic regression models incorporating serum prostate-specific antigen (or PCPT risk calculators), urine *TMPRSS2:ERG* (T2:ERG) score, and urine PCA3 scores were trained on a cohort of 711 evaluable patients for predicting prostate cancer on biopsy. Trained models were then evaluated in a separate validation cohort of 1225 evaluable patients. (A) Calibration plot (observed vs predicted risks of prostate cancer on biopsy) using predicted probabilities from the PCPTrc version 1.0 (PCPTrc, dark orange points) and the trained MiPS_PCPT model (PCPTrc plus T2:ERG plus PCA3, medium orange points) in the validation cohort. Plots from groups of $n = 10$ are shown (Supplementary Fig. 1 shows plots from group of $n = 5$ and $n = 15$). Perfect calibration is indicated by the dashed 45° line. (B) As in panel (A) but predicting high-grade cancer (Gleason score >6) on biopsy and using the PCPThg (dark orange points) and the trained MiPS_PCPThg model (PCPThg plus T2:ERG plus PCA3, medium orange points). (C–F) Decision curve analysis (DCA) demonstrated net clinical benefit of biopsying patients in the validation cohort based on MiPS- versus PCPT-based models across a range of clinically relevant threshold probabilities (the risk of cancer [or high-grade cancer] on biopsy that a patient would choose to undergo biopsy based on their weight of relative harms of false-positive and false-negative predictions). (C) Net clinical benefit of the PCPTrc (dark orange line) and the MiPS_PCPT model (medium orange line) are shown (using 5% increments) compared with strategies of biopsying everyone (light orange line) and biopsying no one (*x*-axis). The MiPS_PCPT-based strategy shows net clinical benefit compared with PCPTrc across a range of relevant threshold probabilities. (D) As in panel (C) but comparing the PCPThg (dark orange line) and the MiPS_PCPThg model (medium orange line). (E, F) DCA can also be used to visualize the percentage of biopsies avoided (compared with biopsying all patients) without missing any events across threshold probabilities. (E) Percentage of biopsies avoided without missing any cancers using the PCPTrc (dark orange line) and the MiPS_PCPT model (medium orange line) in the validation cohort. (F) Percentage of biopsies avoided without missing any high-grade cancers using the PCPThg (dark orange line) and the MiPS_PCPThg model (medium orange line). As an example, at a threshold probability of 10% chance of high-grade cancer on biopsy, biopsying patients on the basis of MIPS_PCPThg would result in an 18.5% reduction in biopsies without missing any high-grade cancers compared to only 3.2% for PCPThg. MiPS = Mi-Prostate Score; PCPTrc = Prostate Cancer Prevention Trial risk calculator; PCPThg = PCPT high grade cancer risk calculator.

Please cite this article in press as: Tomlins SA, et al. Urine *TMPRSS2:ERG* Plus PCA3 for Individualized Prostate Cancer Risk Assessment. Eur Urol (2015), http://dx.doi.org/10.1016/j.eururo.2015.04.039

Table 3 – Impact of using Prostate Cancer Prevention Trial or Mi-Prostate Score–based risk predictors on biopsies avoided and all and high-grade cancers detected or delayed

Model	Predicted risk cut-off, %	Biopsies performed, *n*	Biopsies avoided, *n* (%)	Cancer		HG cancer	
				Detected, *n* (%)	Missed, *n* (%)	Detected, *n* (%)	Delayed, *n* (%)
PCPT or MiPS_PCPT	0	1166	0 (0)	491 (42)	0 (0)	208 (17)	0 (0)
PCPT	≥20	1115	51 (4.4)	489 (42)	2 (0.2)	208 (17)	0 (0)
MiPS_PCPT	≥20	909	257 (22)	453 (39)	38 (3.3)	202 (17)	6 (0.5)
PCPT	≥30	974	192 (16)	446 (38)	45 (3.9)	198 (17)	10 (0.9)
MiPS_PCPT	≥30	754	412 (35)	421 (36)	70 (6.0)	196 (17)	12 (1.0)
PCPT	≥40	583	583 (50)	301 (26)	190 (16)	152 (13)	56 (4.8)
MiPS_PCPT	≥40	617	549 (47)	370 (32)	121 (10)	181 (16)	27 (2.3)
PCPThg or MiPS_PCPThg	0	1218	0 (0)	515 (42)	0 (0)	223 (18)	0 (0)
PCPThg	≥5	1002	216 (17)	472 (39)	43 (3.5)	214 (18)	9 (0.7)
MiPS_PCPThg	≥5	1149	69 (5.9)	506 (42)	9 (0.7)	222 (18)	1 (0.1)
PCPThg	≥10	639	579 (48)	321 (26)	194 (16)	169 (14)	54 (4.4)
MiPS_PCPThg	≥10	932	286 (23)	469 (39)	46 (3.8)	217 (18)	6 (0.5)
PCPThg	≥15	391	827 (68)	212 (17)	303 (25)	125 (10)	98 (8)
MiPS_PCPThg	≥15	777	441 (36)	430 (37)	85 (7.0)	204 (17)	19 (1.6)

HG = high-grade; MiPS = Mi-Prostate Score; PCPT = Prostate Cancer Prevention Trial; PCPThg = Prostate Cancer Prevention Trial high-grade risk calculator. The number and percentage of biopsies that would be avoided and the number and percentage of biopsies with all and HG (Gleason score >6) cancers detected and delayed from applying the indicated predicted risk cut-off (from PCPT or MiPS-based models) as the threshold for undergoing biopsy are shown.

MiPS_PCPThg prediction of high-grade cancer were 0.637 and 0.752, respectively ($p = 0.0004$).

4. Discussion

We developed and validated risk models, termed MiPS, combining urine T2:ERG and PCA3 scores with serum PSA or the PCPTrc for predicting the presence of PCa (or high-grade cancer) on biopsy. T2:ERG and PCA3 represent the most advanced urine biomarkers for PCa [11,12]. The Progensa PCA3 test (used in this study) is FDA approved in the setting of a prior negative biopsy, and recent reports support utility in the initial biopsy setting [10–13,21,23]. Of note, in our large validation cohort, MiPS showed significantly increased AUC compared with serum PSA (or PCPT-based models) for predicting cancer or high-grade cancer in the initial biopsy setting and in subsets of patients presenting for repeat biopsy (Supplementary Table 4).

T2:ERG is one of the most well-characterized tissue-based PCa biomarkers, with independent reports of >99.99% specificity using ERG immunohistochemistry as a surrogate for T2:ERG fusions [14]. Importantly, in addition to studies using the quantitative T2:ERG TMA assay (or earlier versions) evaluated in this study to assess urine specimens, studies have also used assays reporting binary T2:ERG status [11,12]. In a prospective study, for example, Leyten et al used whole transcriptome amplification of urine sediment followed by quantitative polymerase chain reaction to assess T2:ERG status (positive or negative) for men presenting for biopsy [24]. They found that incorporating T2:ERG status (along with PCA3, as assessed in this study) increased the European Randomised Study of Screening for Prostate Cancer risk calculator (ERSPCrc) from 0.799 to 0.842; T2:ERG status (unlike PCA3 score) also added significant predictive value to the ERSPCrc for predicting Gleason score and clinical stage. Because T2:ERG fusions are present in only about 50% of PCa, this binary approach for early detection is appealing. However, in a study of 41 men from our training cohort who later underwent prostatectomy, approximately 75% of prostates harbored at least one T2:ERG-positive cancer focus, and urine T2:ERG scores and total T2:ERG-expressing cancer tissue volume were highly correlated ($r_s = 0.68$) [19]. Likewise, in our current study, biopsy-detectable cancer risk is continuously associated with T2:ERG score (Supplementary Fig. 2). Together, these findings support quantitative urine T2:ERG assays as providing more individualized risk assessment.

DCA demonstrates that across the range of clinically relevant threshold probabilities (the probability of cancer or high-grade cancer at which a patient would elect to undergo biopsy), biopsy decision making based on MiPS models shows increased net clinical benefit compared with serum PSA or PCPT-based approaches. Likewise, as shown in Table 3, using various MiPS_PCPT (≥30% or ≥40%) or MiPS_PCPThg (≥15%) cut-offs for biopsying patients in our validation cohort would avoid 35–47% of biopsies while delaying the diagnosis of only 1.0–2.3% of high-grade cancers. Although cross-study comparisons are challenging, MiPS results compare favorably to those from other early detection tests, including 4Kscore and PHI [25,26].

A complete comparison of MiPS and other early detection biomarkers (including multiparametric magnetic resonance imaging [MRI]) and cost–benefit analysis has not been performed. However, our current study validating the MiPS models will enable head-to-head (and combined) assessment with 4Kscore, PHI, and multiparametric MRI, as required for a true cost–benefit analysis, to determine the optimal early detection approach for various clinical scenarios. MiPS requires a DRE prior to urine collection, unlike 4Kscore and PHI, which require a blood draw. Unlike the binary urine T2:ERG assay described by Leyten et al that

Please cite this article in press as: Tomlins SA, et al. Urine *TMPRSS2:ERG* Plus PCA3 for Individualized Prostate Cancer Risk Assessment. Eur Urol (2015), http://dx.doi.org/10.1016/j.eururo.2015.04.039

requires sedimented urine [24], both the T2:ERG and PCA3 assay in MiPS are performed on the same whole urine specimen.

Multiple studies have demonstrated that *TMPRSS2:ERG* gene fusions are more common at the tissue level in early onset PCa and in men presenting with low serum PSA [27,28]. In an exploratory analysis shown in Supplementary Table 5, we found that PSA or (PCPT) plus T2:ERG models showed nearly equivalent or greater AUC than full MiPS models for predicting cancer and high-grade cancer in men with serum PSA <3 ng/ml. Hence, MiPS (or models including only T2:ERG, given its cancer specificity) may have particular utility in this setting, supporting formal assessment in future studies.

A limitation of our study was the use of PCPTrc_v1 and PCPThg_v1, rather than updated version 2 (v2) risk calculators, because our MiPS models were locked for subsequent validation studies prior to PCPT_v2 risk calculator development. Of note, PCPTrc_v2 and PCPThg_v2 were poorly calibrated in our validation cohort (Supplementary Table 6), with no significant difference in AUCs compared with version 1 (PCPThg_v1 showed significantly increased AUC compared to PCPThg_v2). In addition, we observed greater improvement for predicting all cancers, compared with high-grade cancer only, when incorporating T2:ERG plus PCA3 scores. Although overdiagnosis of low-grade cancer drives overtreatment, whether our models show utility in identifying the subset of patients with low-grade cancer who harbor undiagnosed higher grade cancer (approximately 20–40%) or can be combined with novel imaging or tissue-based prognostic tests should be investigated. Of note, tissue and urine assessment of PCA3 and/or T2:ERG have been variably associated with significant disease and progression [10–14,29,30], supporting the need for additional investigation in these settings. Last, our validation cohort consisted of men without cancer undergoing biopsy based on current standard of care (ie, elevated serum PSA), so no conclusions can be drawn from this study regarding performance in active surveillance or screening cohorts.

5. Conclusions

In summary, we reported validated individualized risk models (MiPS) incorporating serum PSA (or the PCPTrc) and urine T2:ERG and PCA3 scores for predicting PCa and high-grade PCa risk on needle biopsy. By AUC, assessment of unnecessary biopsies avoided, and DCA, MiPS models significantly outperformed serum PSA (or PCPTrc)-based strategies, supporting the use of the MiPS test as a decision-making aide for men (and their physicians) concerned about serum PSA test results, particularly in the initial biopsy setting. The MiPS test, which uses these validated models to report quantitative risk assessments for PCa and high-grade PCa on biopsy, is clinically available through a College of American Pathology/Clinical Laboratory Improvement Amendments–certified laboratory. Additional studies will be needed to compare MiPS performance with other early detection–based strategies and to determine costs and benefits of various early detection approaches.

Author contributions: Scott A. Tomlins and Arul M. Chinnaiyan 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: Tomlins, Groskopf, Wei, Chinnaiyan.
Acquisition of data: Day, Meyer, Hodge, Groskopf.
Analysis and interpretation of data: Tomlins, Day, Lonigro, Hovelson, Kunju, Dunn, Groskopf, Wei, Chinnaiyan.
Drafting of the manuscript: Tomlins, Wei, Chinnaiyan.
Critical revision of the manuscript for important intellectual content: Tomlins, Day, Lonigro, Hovelson, Dunn, Groskopf, Wei, Chinnaiyan.
Statistical analysis: Tomlins, Lonigro, Hovelson, Dunn.
Obtaining funding: Wei, Chinnaiyan.
Administrative, technical, or material support: Siddiqui.
Supervision: None.
Other (specify): None.

Financial disclosures: Scott A. Tomlins and Arul M. Chinnaiyan certify 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: The University of Michigan has been issued a patent on the detection of ETS gene fusions in prostate cancer, on which A.M.C. and S.A.T. are listed as coinventors. The University of Michigan licensed the diagnostic field of use to Hologic/Gen-Probe, Inc, which sublicensed some rights to Ventana Medical Systems, Inc. A.M.C. has served as consultant to Gen-Probe, Inc. and Ventana Medical Systems. S.A.T. has received honoraria from, and served as a consultant to Ventana Medical Systems. Gen-Probe provided material support for a clinical trial evaluating PCA3, on which J.T.W. was the primary investigator. J.R.D., S. M., P.H., and J.G. are employees of Hologic/Gen-Probe. The remaining authors declare no conflicts of interest.

Funding/Support and role of the sponsor: Supported in part by Hologic/Gen-Probe Inc., the Early Detection Research Network (U01 CA111275 and U01 CA113913), NIH S.P.O.R.E. (P50 CA186786), and R01 CA132874. A.M.C. and S.A.T. are supported by the Prostate Cancer Foundation and the A. Alfred Taubman Medical Research Institute. A.M.C. is supported by the Doris Duke Foundation and is an American Cancer Society Clinical Research Professor. Hologic/Gen-Probe was involved in the design and conduct of the study, funded specimen collection, performed T2:ERG and PCA3 assays, assisted in the preparation of and reviewed and approved the manuscript. All data analysis was performed at the University of Michigan. S.A.T. 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.

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.2015.04.039.

References

[1] Andriole GL, Crawford ED, Grubb III RL, et al. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. J Natl Cancer Inst 2012;104:125–32.

[2] Etzioni RD, Thompson IM. What do the screening trials really tell us and where do we go from here? Urol Clin North Am 2014;41:223–8.
[3] Schroder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. Lancet 2014;384:2027–35.
[4] Thompson IM, Ankerst DP, Chi C, et al. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. J Natl Cancer Inst 2006;98:529–34.
[5] Shariat SF, Kattan MW, Vickers AJ, Karakiewicz PI, Scardino PT. Critical review of prostate cancer predictive tools. Future Oncol 2009;5:1555–84.
[6] Zhu X, Albertsen PC, Andriole GL, Roobol MJ, Schröder FH, Vickers AJ. Risk-based prostate cancer screening. Eur Urol 2012;61:652–61.
[7] Bryant RJ, Lilja H. Emerging PSA-based tests to improve screening. Urol Clin North Am 2014;41:267–76.
[8] Catalona WJ, Partin AW, Sanda MG, et al. A multicenter study of [-2]pro-prostate specific antigen combined with prostate specific antigen and free prostate specific antigen for prostate cancer detection in the 2.0 to 10.0 ng/ml prostate specific antigen range. J Urol 2011;185:1650–5.
[9] Vickers A, Cronin A, Roobol M, et al. Reducing unnecessary biopsy during prostate cancer screening using a four-kallikrein panel: an independent replication. J Clin Oncol 2010;28:2493–8.
[10] Cary KC, Cooperberg MR. Biomarkers in prostate cancer surveillance and screening: past, present, and future. Ther Adv Urol 2013;5:318–29.
[11] Truong M, Yang B, Jarrard DF. Toward the detection of prostate cancer in urine: a critical analysis. J Urol 2013;189:422–9.
[12] Salagierski M, Schalken JA. Molecular diagnosis of prostate cancer: PCA3 and TMPRSS2:ERG gene fusion. J Urol 2012;187:795–801.
[13] Auprich M, Bjartell A, Chun FK, et al. Contemporary role of prostate cancer antigen 3 in the management of prostate cancer. Eur Urol 2011;60:1045–54.
[14] Pettersson A, Graff RE, Bauer SR, et al. The TMPRSS2:ERG rearrangement, ERG expression, and prostate cancer outcomes: a cohort study and meta-analysis. Cancer Epidemiol Biomarkers Prev 2012;21:1497–509.
[15] Warrick JI, Tomlins SA, Carskadon SL, et al. Evaluation of tissue PCA3 expression in prostate cancer by RNA in situ hybridization–a correlative study with urine PCA3 and TMPRSS2-ERG. Mod Pathol 2014;27:609–20.
[16] Tomlins SA, Aubin SM, Siddiqui J, et al. Urine TMPRSS2:ERG fusion transcript stratifies prostate cancer risk in men with elevated serum PSA. Sci Transl Med 2011;3:94ra72.
[17] Cornu JN, Cancel-Tassin G, Egrot C, Gaffory C, Haab F, Cussenot O. Urine TMPRSS2:ERG fusion transcript integrated with PCA3 score, genotyping, and biological features are correlated to the results of prostatic biopsies in men at risk of prostate cancer. Prostate 2012;73:242–9.
[18] Stephan C, Jung K, Semjonow A, et al. Comparative assessment of urinary prostate cancer antigen 3 and TMPRSS2:ERG gene fusion with the serum [-2]proprostate-specific antigen-based Prostate Health Index for detection of prostate cancer. Clin Chem 2013;59:280–8.
[19] Young A, Palanisamy N, Siddiqui J, et al. Correlation of urine TMPRSS2:ERG and PCA3 to ERG+ and total prostate cancer burden. Am J Clin Pathol 2012;138:685–96.
[20] Lin DW, Newcomb LF, Brown EC, 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;19:2442–50.
[21] Wei JT, Feng Z, Partin AW, et al. Can urinary PCA3 supplement PSA in the early detection of prostate cancer? J Clin Oncol 2014;32:4066–72.
[22] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.
[23] Hansen J, Auprich M, Ahyai SA, et al. Initial prostate biopsy: development and internal validation of a biopsy-specific nomogram based on the prostate cancer antigen 3 assay. Eur Urol 2013;63:201–9.
[24] Leyten GH, Hessels D, Jannink SA, et al. Prospective multicentre evaluation of PCA3 and TMPRSS2-ERG gene fusions as diagnostic and prognostic urinary biomarkers for prostate cancer. Eur Urol 2014;65:534–42.
[25] 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. In press. http://dx.doi.org/10.1016/j.eururo.2014.10.021
[26] de la Calle C, Patil D, Wei JT, et al. Multicenter evaluation of the Prostate Health Index for detection of aggressive prostate cancer in biopsy naive men. J Urol. In press. http://dx.doi.org/10.1016/j.juro.2015.01.091
[27] Steurer S, Mayer PS, Adam M, et al. TMPRSS2-ERG fusions are strongly linked to young patient age in low-grade prostate cancer. Eur Urol 2014;66:978–81.
[28] Schaefer G, Mosquera JM, Ramoner R, et al. Distinct ERG rearrangement prevalence in prostate cancer: higher frequency in young age and in low PSA prostate cancer. Prostate Cancer Prostatic Dis 2013;16:132–8.
[29] Tosoian JJ, Loeb S, Kettermann A, et al. Accuracy of PCA3 measurement in predicting short-term biopsy progression in an active surveillance program. J Urol 2010;183:534–8.
[30] Berg KD, Vainer B, Thomsen FB, et al. ERG protein expression in diagnostic specimens is associated with increased risk of progression during active surveillance for prostate cancer. Eur Urol 2014;66:851–60.

Please cite this article in press as: Tomlins SA, et al. Urine *TMPRSS2:ERG* Plus PCA3 for Individualized Prostate Cancer Risk Assessment. Eur Urol (2015), http://dx.doi.org/10.1016/j.eururo.2015.04.039