Association Between Combined TMPRSS2:ERG and PCA3 RNA Urinary Testing and Detection of Aggressive Prostate Cancer

Martin G. Sanda, MD; Ziding Feng, PhD; David H. Howard, PhD; Scott A. Tomlins, MD, PhD; Lori J. Sokoll, PhD; Daniel W. Chan, PhD; Meredith M. Regan, DSci; Jack Groskopf, PhD; Jonathan Chipman, MS; Dattatraya H. Patil, MBBS, MPH; Simpa S. Salami, MD; Douglas S. Scherr, MD; Jacob Kagan, PhD; Sudhir Srivastava, PhD; Ian M. Thompson Jr, MD; Javed Siddiqui, MS; Jing Fan, MS; Aron Y. Joon, MS; Leonidas E. Bantis, PhD; Mark A. Rubin, MD; Arul M. Chinnayian, MD, PhD; John T. Wei, MD; and the EDRN-PCA3 Study Group

IMPORTANCE Potential survival benefits from treating aggressive (Gleason score, \( \geq 7 \)) early-stage prostate cancer are undermined by harms from unnecessary prostate biopsy and overdiagnosis of indolent disease.

OBJECTIVE To evaluate the a priori primary hypothesis that combined measurement of PCA3 and TMPRSS2:ERG (T2:ERG) RNA in the urine after digital rectal examination would improve specificity over measurement of prostate-specific antigen alone for detecting cancer with Gleason score of 7 or higher. As a secondary objective, to evaluate the potential effect of such urine RNA testing on health care costs.

DESIGN, SETTING, AND PARTICIPANTS Prospective, multicenter diagnostic evaluation and validation in academic and community-based ambulatory urology clinics. Participants were a referred sample of men presenting for first-time prostate biopsy without preexisting prostate cancer: 516 eligible participants from among 748 prospective cohort participants in the developmental cohort and 561 eligible participants from 928 in the validation cohort.

INTERVENTIONS/EXPOSURES Urinary PCA3 and T2:ERG RNA measurement before prostate biopsy.

MAIN OUTCOMES AND MEASURES Presence of prostate cancer having Gleason score of 7 or higher on prostate biopsy. Pathology testing was blinded to urine assay results. In the developmental cohort, a multiplex decision algorithm was constructed using urine RNA assays to optimize specificity while maintaining 95% sensitivity for predicting aggressive prostate cancer at initial biopsy. Findings were validated in a separate multicenter cohort via prespecified analysis, blinded per prospective-specimen-collection, retrospective-blinded-evaluation (PRoBE) criteria. Cost effects of the urinary testing strategy were evaluated by modeling observed biopsy results and previously reported treatment outcomes.

RESULTS Among the 516 men in the developmental cohort (mean age, 62 years; range, 33-85 years) combining testing of urinary T2:ERG and PCA3 at thresholds that preserved 95% sensitivity for detecting aggressive prostate cancer improved specificity from 18% to 39%. Among the 561 men in the validation cohort (mean age, 62 years; range, 27-86 years), analysis confirmed improvement in specificity (from 17% to 33%, lower bound of 1-sided 95% CI, 0.73%; prespecified 1-sided \( P = .04 \)), while high sensitivity (93%) was preserved for aggressive prostate cancer detection. Forty-two percent of unnecessary prostate biopsies would have been averted by using the urine assay results to select men for biopsy. Cost analysis suggested that this urinary testing algorithm to restrict prostate biopsy has greater potential cost-benefit in younger men.

CONCLUSIONS AND RELEVANCE Combined urinary testing for T2:ERG and PCA3 can avert unnecessary biopsy while retaining robust sensitivity for detecting aggressive prostate cancer with consequent potential health care cost savings.

Published online May 18, 2017.

© 2017 American Medical Association. All rights reserved.
A multiplex evidence has shown survival benefits associated with treatment of intermediate- and high-risk, early-stage prostate cancer.  

Low specificity of tests for serum levels of prostate-specific antigen (PSA), however, has limited its utility for prostate cancer screening. Strategies are needed for detecting cancers with more aggressive features (Gleason score, ≥7) to enable survival benefits from treatment while limiting unnecessary biopsies and overdetection of indolent disease.

Gene expression alterations in prostate cancer representing an opportunity to refine early detection include overexpression of PCA3, a noncoding RNA, and aberrant TMPRSS2:ERG (T2:ERG) expression, which results from a prostate cancer-specific chromosomal rearrangement on chromosome 21 that juxtaposes the 5' untranslated region of TMPRSS2 to the ERG oncogene. In contrast, PSA is expressed in normal and cancerous prostate cells alike, limiting its specificity for cancer.

Following development of clinical assays for detecting PCA3 and T2:ERG in urine, we sought to determine whether measuring these prostatic RNAs could improve specificity for detecting aggressive prostate cancer. We assembled a prospective multicenter cohort to develop a method of combining PCA3 and T2:ERG urine assay results via an "either-or" logic for a decision algorithm to select men for initial biopsy, and then we validated this decision algorithm in a separate, multicenter cohort via prespecified analysis blinded per prospective-specimen-collection, retrospective-blinded-evaluation (PROBE) criteria.

We then evaluated effects of such multiplex urine testing on costs of prostate cancer early detection.

### Methods

#### Participants

The developmental cohort consisted of prospectively enrolled participants at 4 urology groups affiliated with 3 academic medical centers (Dana Farber Harvard Cancer Center; Cornell University; and University of Michigan) comprising a National Cancer Institute Early Detection Research Network (NCI-EDRN) Clinical Validation Center (eFigure 1 in the Supplement). Eligibility was restricted to men planning to undergo first-time prostate biopsy, who signed informed consent approved by the respective institutional review board and provided post-urinary specimens after digital rectal examination (DRE) but before biopsy. Exclusion criteria were previously diagnosed prostate cancer, prior prostate biopsy, previous prostatectomy, other cancer diagnosis, or inability to provide post-DRE urine sample.

The validation cohort consisted of participants in the NCI-EDRN’s Urinary PCA3 Evaluation Trial that enrolled men prior to prostate biopsy at 11 sites with eligibility as previously described. Additional eligibility criteria for the validation cohort analysis in the present study included completion of prostate biopsy, availability of biorepository sample to assay urinary T2:ERG expression, and no prior prostate biopsy (eFigure 2 in the Supplement). The protocol proscribed routine clinical transrectal ultrasonogram-guided prostate biopsy; 87% of participants had 12 biopsy cores sampled.

#### Urine Assay

First-catch urine was collected after DRE, admixed with RNA stabilization buffer, and frozen to ~80°C within 4 hours of collection. Developmental cohort specimens were transported to the CLIA laboratory (Clinical Laboratory Improvement Amendments) at University of Michigan for transcription-mediated amplification quantitative T2:ERG and US Food and Drug Administration (FDA)-approved Progensa PCA3 assay (Hologic Inc.). Urine specimen collection was performed similarly in the multicenter validation cohort, and T2:ERG and PCA3 assays were performed by the EDRN Biomarker Reference Laboratory at Johns Hopkins. A random sample subset was independently analyzed at Hologic Inc as a quality control. Assay laboratories were blinded to prostate biopsy results.

#### Analysis

For both cohorts, the a priori primary end point was presence vs absence of aggressive prostate cancer (Gleason score, >7) on biopsy, where absence included both indolent (Gleason score, ≤6) and no (negative biopsy findings) cancers. Assignment of Gleason score was by local clinical pathology evaluation; a randomly selected subset of the validation cohort had undergone quality review of the cancer diagnosis by a central pathologist.

A clinical decision algorithm to restrict prostate biopsy was derived from the developmental cohort (n = 516), locked-down, and sensitivity and specificity were then evaluated in the validation cohort (n = 561). In the developmental cohort, we used an "OR" rule; ie, the test result was considered positive if 1 of the 3 biomarkers (serum PSA, T2:ERG, or PCA3) was above its normal threshold. This multiplex decision rule and the cutoff points for PSA, T2-ERG, and PCA3 were determined by a grid search to maximize specificity for the combined group of indolent and no cancers, while setting its sensitivity at 95% or higher for aggressive cancers. Optimal thresholds of urinary T2:ERG and PCA3 scores were rounded to integers (T2:ERG score, >8; PCA3 score, >20). The final threshold for serum PSA of greater than 10 ng/mL incorporated evidence supporting treatment of cancer with PSA higher than 10 ng/mL even though in the developmental cohort, all
patients with PSA higher than 10 ng/mL had either a T2:ERG score higher than 8 or a PCA3 score higher than 20.

After the multiplex decision algorithm was locked down, validation cohort data were accessed to posit the primary hypothesis of whether the predefined decision algorithm would improve specificity for detecting prostate cancer with a Gleason score of 7 or higher and to estimate sensitivity of this decision rule. To test whether the multiplex decision rule has superior specificity to PSA alone, the decision rule from the training set was locked down and evaluated on the validation set. A total of 10 000 bootstrap samples were generated to obtain the variance and 95% confidence interval (CI) for the difference of the specificities between the 2 decision rules on the validation set. For each bootstrap sample, sensitivity and specificity were calculated by the prespecified, locked-down decision rule, then again by PSA alone, and the difference in specificities (multiplex decision algorithm minus PSA alone) were calculated. The variance of this difference and its 95% CI were obtained from the 10 000 bootstrap differences. The null hypothesis of no superiority of the multiplex decision ruler over PSA alone would be rejected at a 1-sided a = .05 if the lowest fifth percentile of 10 000 differences was greater than 0.

We note that the power attained with the available sample size using either the empirical receiver operating curve (ROC) for PSA and a kernel-based ROC for PSA is 83.5% and 85.0%, respectively for detecting significant differences in specificities (eAppendix, A in the Supplement).

Cost Analysis
A decision analytic model was developed to compare the multiplex decision algorithm using serum PSA and urinary markers vs standard care among patients with positive PSA screens under 3 scenarios: (1) no biopsies prompted by abnormal PSA (ie, strict adherence to US Protective Services Task Force [USPSTF] recommendations), (2) biopsy based on abnormal PSA or DRE findings, and (3) biopsy restricted by the multiplex urinary marker decision algorithm. We modeled a 1-time test with the multiplex urinary marker decision algorithm. We did not consider repeated screening or testing. We projected the number of biopsies, the number of patients diagnosed with metastatic cancer, and lifetime health care costs under each strategy. Projecting quality-adjusted life-years is left for future work.

The first scenario represents care among patients with PSA higher than 10 ng/mL who had either a T2:ERG score higher than 8 or a PCA3 score higher than 20.

After the multiplex decision algorithm was locked down, validation cohort data were accessed to posit the primary hypothesis of whether the predefined decision algorithm would improve specificity for detecting prostate cancer with a Gleason score of 7 or higher and to estimate sensitivity of this decision rule. To test whether the multiplex decision rule has superior specificity to PSA alone, the decision rule from the training set was locked down and evaluated on the validation set. A total of 10 000 bootstrap samples were generated to obtain the variance and 95% confidence interval (CI) for the difference of the specificities between the 2 decision rules on the validation set. For each bootstrap sample, sensitivity and specificity were calculated by the prespecified, locked-down decision rule, then again by PSA alone, and the difference in specificities (multiplex decision algorithm minus PSA alone) were calculated. The variance of this difference and its 95% CI were obtained from the 10 000 bootstrap differences. The null hypothesis of no superiority of the multiplex decision ruler over PSA alone would be rejected at a 1-sided a = .05 if the lowest fifth percentile of 10 000 differences was greater than 0.

We note that the power attained with the available sample size using either the empirical receiver operating curve (ROC) for PSA and a kernel-based ROC for PSA is 83.5% and 85.0%, respectively for detecting significant differences in specificities (eAppendix, A in the Supplement).

As detailed and supported in the eAppendix, B in the Supplement, post hoc significant association was observed between results of the urinary PCA3:T2:ERG assay and the dichotomized PSA test (<10 ng/mL or >10 ng/mL) in the developmental cohort (Fisher exact test P = .01) but not in the validation cohort (P = .09).

Cost Analysis
A decision analytic model was developed to compare the multiplex decision algorithm using serum PSA and urinary markers vs standard care among patients with positive PSA screens under 3 scenarios: (1) no biopsies prompted by abnormal PSA (ie, strict adherence to US Protective Services Task Force [USPSTF] recommendations), (2) biopsy based on abnormal PSA or DRE findings, and (3) biopsy restricted by the multiplex urinary marker decision algorithm. We modeled a 1-time test with the multiplex urinary marker decision algorithm. We did not consider repeated screening or testing. We projected the number of biopsies, the number of patients diagnosed with metastatic cancer, and lifetime health care costs under each strategy. Projecting quality-adjusted life-years is left for future work.

The first scenario represents care men would receive under USPSTF prostate cancer screening guidelines, which recommend against PSA screening. In this setting, physicians would not know if a patient had an abnormal PSA level nor if a patient had asymptomatic, early-stage prostate cancer in the first place, and prostate cancer care would be limited to treatment of advanced, symptomatic disease. The second scenario corresponds to common practice prior to the release of the current USPSTF recommendation—ie, biopsy based on abnormal PSA or DRE findings, as in our developmental cohort, consistent with National Comprehensive Care Network guidelines. We assumed that men who undergo biopsy receive effective care to prevent disease progression. In the third scenario, patients found to have elevated serum PSA levels undergo biopsy based on testing positive on either T2:ERG or PCA3 or registering a serum PSA level higher than 10 ng/mL. The model considers how incorporating PCA3 and T2:ERG into the diagnostic pathway affects the decision to undergo biopsy and treatment.

We obtained parameter values for sensitivity and specificity from the developmental cohort and tumor progression rates and cost estimates from the literature. Estimates of the costs incurred by patients diagnosed with cancer represent lifetime, stage-specific, prostate cancer–related costs based on an analysis of Medicare beneficiaries diagnosed with prostate cancer between 1991 and 2002, regardless of treatment approach. These estimates represent costs to the Medicare program and reflect treatment patterns and cure rates for beneficiaries diagnosed during this period. We assumed that the cost of a biopsy, inclusive of biopsy-related complications, is $2300. We use enzalutamide as a stand-in for treatment of castration-resistant prostate cancer (CRPC), and sensitivity analysis estimated the impact on costs if varying percentages of men received systemic therapy for CRPC. The cost of enzalutamide was based on the Medicare reimbursement rate and the median duration of treatment in the drug’s phase 3 trial. Because the cost of a commercially available T2:ERG and PCA3 test has not yet been established, we did not include this cost in our analysis. We discounted costs for late-stage disease by 3% per year, assuming that they occur 5 years in the future. We assumed all other costs are incurred in a short time frame (ie, less than a year). All costs are stated in 2013 dollars (additional model details are provided in eFigure 3 in the Supplement).

Results
A total of 516 participants met inclusion criteria (ie, men presenting for first-time prostate biopsy) in the EDRN Clinical Validation Center, and these make up the developmental cohort (Table 1; eFigure 1 in the Supplement). Clinical factors associated with aggressive prostate cancer (Gleason score, ≥7) included older age (P < .001) and abnormal or suspect findings on DRE (P = .03). Urine T2:ERG and PCA3 scores of greater than 60 were each significantly associated with aggressive prostate cancer (P < .001 for both; see the Figure).

Because T2:ERG gene fusion expression occurs in a subset of prostate cancers, our primary hypothesis posited that combining urine T2:ERG with PCA3 assay in “either-or” multiplex combinatorial logic would improve the specificity of predicting aggressive (Gleason score, ≥7) prostate cancer. Urine T2:ERG testing showed limited utility when optimized for higher than 95% sensitivity, as expected, given that T2:ERG gene fusions are only present in about half of prostate cancers (Figure and Table 2). Assays for PSA and PCA3 yielded specificities of 18% and 17%, respectively, at 95% specificity.
for aggressive prostate cancer (Table 2). Combining urinary T2:ERG with PCA3 testing improved specificity of predicting aggressive prostate cancer to 39% (at optimized cut points of a T2:ERG score of 7.6 and a PCA3 score of 19.1; Table 2). In this cohort, every patient with a PSA level higher than 10 ng/mL had either a T2:ERG score greater than 7.6 or a PCA3 score greater than 19.1.

We next sought to validate whether combining urinary T2:ERG and PCA3 measurement improves specificity of predicting aggressive prostate cancer via analysis of a separate 561-patient multicenter validation cohort assayed at an independent laboratory (Table 1 and eFigure 2 in the Supplement).30 Based on developmental cohort results, thresholds of T2:ERG score greater than 8, PCA3 score greater than 20, or serum PSA level greater than 10 ng/mL as a positive test result were prespecified for the validation analysis. The validation analysis confirmed that overexpression of either urinary T2:ERG or PCA3 or PSA level higher than 10 ng/mL improved prediction of aggressive prostate cancer by increasing specificity from 17% to 33% compared with serum PSA level alone while attaining high sensitivity (Table 3) (difference in specificities, 17%; lower 95% CI boundary, 0.73%; P = .04). Post hoc analysis using a method of kernel estimators, we found 2-sided P = .01 and 1-tailed P = .007 (detailed in eAppendix, C in the Supplement). In exploratory analysis, we evaluated the Prostate Cancer Prevention Trial (PCPT) calculator24 for high-grade disease in this validation cohort, then used logistic regression to add PCA3 and T2:ERG and evaluated differences in consequent areas under the curve (AUCs) by the Delong test. The PCPT model AUC was 0.74 (95% CI, 0.70-0.79), whereas for PCPT plus T2:ERG and PCA3, the AUC was 0.81 (95% CI, 0.77-0.85) (P < .001). At 95% sensitivity, the specificity for PCPT in predicting prostate cancer with a Gleason score of 7 or higher was 31%; it was 37% for PCPT combined with PCA3 and T2:ERG measurement.

Restricting biopsy to participants with positive urinary findings of T2:ERG or PCA3 or PSA level higher than 10 ng/mL in the validation cohort would have avoided 42% of unnecessary biopsies (true negative, 124 of 297 biopsies in men with no cancer found on biopsy) and would have avoided overdetecion of 12% of indolent cancers (true negative, 14 of 116 biopsies in men having a Gleason score of ≥6). Overall, 33% of “excessive” prostate biopsies would be avoided by the multiplex decision algorithm (true negative, 138 of 413 biopsies in men having no cancer or prostate cancer with a Gleason score of ≥6). Among participants with Gleason scores of 7 or higher, only 7% would be missed under the combined thresholds (false negative, 11 of 148) compared with 21% (false negative, 31 of 148) when using a PCA3 threshold greater than 20 alone.

To gauge the potential cost impact of using of PCA3 and T2:ERG urine testing vs standard care, we modeled the costs of incorporating PCA3 and T2:ERG into the clinical pathway of decision to perform prostate biopsy and consequent prostate cancer care (Table 4 and eFigure 3 in the Supplement). The costs associated with prostate cancer detection by urinary T2:ERG and PCA3 testing (among men with abnormal PSA or DRE findings), which include costs of biopsy and of treating men with early-stage disease, exceed the costs associated with no PSA-prompted biopsies (per UTSPSTF recommendations), which include only the costs of treating late-stage disease among men.
whose tumors progress. The difference varies with age and with the use of second-line systemic therapy for patients with castration-resistant, metastatic disease. Conversely, in comparison with biopsy of all men with abnormal PSA screening results, restricting biopsy using urinary T2:ERG and PCA3 testing yielded cost savings in the range of $1200 to $2100 per patient.

Discussion

A pivotal flaw in relying on serum PSA to select men for prostate biopsy is low specificity for detecting aggressive prostate cancer at PSA levels having sufficient sensitivity. Herein, we developed and validated a decision algorithm that uses urinary testing for 2 prostate cancer–associated RNA markers, T2:ERG and PCA3, to improve the specificity for detecting aggressive prostate cancer among men with elevated PSA or abnormal DRE findings. The validated “either/or” algorithm, where biopsy is prompted by high expression of either urinary RNA marker, would reduce excess biopsy while preserving detection of aggressive prostate cancers (Gleason score, ≥7).

The T2:ERG gene fusion is present in two-thirds of prostate cancers13,14 and is consequently suitable for combination with other biomarkers in decision algorithms. Even though the presence of T2:ERG is not associated with prostate cancer aggressiveness,25 prior studies have shown association between urine T2:ERG score and total ERG-positive tumor burden, indirectly reflecting tumor aggressiveness.26 Accordingly, we observed association between urinary T2:ERG score...
Optimizing the potential for urinary PCA3 testing to improve selecting men for initial prostate biopsy may hinge on combining PCA3 with other biomarkers, such as T2:ERG.

We constructed a clinical algorithm combining T2:ERG with PCA3 via either-or combinatorial logic. To accommodate the survival benefit of treating prostate cancers in patients having serum PSA levels greater than 10 ng/mL, this PSA threshold was also incorporated in the predictive models a priori. Our findings confirm the hypothesis that combining T2:ERG, PCA3, and PSA measurement would reduce unnecessary prostate biopsy and overdiagnosis while preserving detection of aggressive cancers. These findings extend those of earlier studies of

and presence of cancer having a Gleason of 7 or higher (Figure). However, a low T2:ERG score does not exclude the possibility of aggressive cancer (since one-third of prostate cancers lack the mutation), and so T2:ERG testing would be more effective if combined with other markers.

It has been shown that PCA3 is a noncoding RNA expressed at higher levels in men with prostate cancer than in those with normal prostate.\(^8,10,27,30^\) The Progensa PCA3 assay has been FDA approved to help identify men who do not require a repeat biopsy. Use of PCA3 testing can improve predictive accuracy for cancer on initial biopsy\(^30,31^\); however, at high sensitivity, PCA3 specificity as a stand-alone test is limited (Table 2).\(^22,23^\)

### Table 2. Combining Urinary TMPRSS2:ERG and PCA3 Measurement to Improve Specificity of Predicting Aggressive Prostate Cancer (Gleason Score ≥7)

<table>
<thead>
<tr>
<th>Diagnostic Biomarker*</th>
<th>Developmental Cohort (n = 516)</th>
<th>Gleason ≥7 Sensitivity, % (95% CI) (n = 156)</th>
<th>No Cancer or Gleason ≤6 Sensitivity, % (95% CI) (n = 360)</th>
<th>PPV, % (95% CI) (Observed Prevalence, 26.38%)</th>
<th>NPV, % (95% CI) (Observed Prevalence, 26.38%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSA</td>
<td>3.0</td>
<td>96.2 (93.2-99.2)</td>
<td>18.1 (14.1-22.1)</td>
<td>29.6 (28.4-30.8)</td>
<td>93.0 (85.4-96.8)</td>
</tr>
<tr>
<td>PCA3</td>
<td>6.3</td>
<td>95.5 (92.2-98.7)</td>
<td>16.9 (13.0-20.8)</td>
<td>29.2 (28.0-30.4)</td>
<td>91.3 (83.1-95.7)</td>
</tr>
<tr>
<td>TMPRSS2:ERG⁶</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PSA or TMPRSS2:ERG</td>
<td>3.1; 289</td>
<td>95.5 (92.2-98.7)</td>
<td>20.0 (15.9-24.1)</td>
<td>29.3 (28.7-31.3)</td>
<td>92.5 (85.4-96.3)</td>
</tr>
<tr>
<td>PSA or PCA3</td>
<td>4.2; 21.2</td>
<td>95.5 (92.2-98.7)</td>
<td>23.6 (19.2-27.9)</td>
<td>30.9 (29.5-32.4)</td>
<td>93.6 (87.4-96.9)</td>
</tr>
<tr>
<td>PCA3 or TMPRSS2:ERG</td>
<td>19.1; 7.6</td>
<td>95.5 (92.2-98.7)</td>
<td>39.4 (34.3-44.4)</td>
<td>36.1 (34.0-38.2)</td>
<td>96.1 (92.1-98.1)</td>
</tr>
<tr>
<td>PSA or PCA3 or TMPRSS2:ERG</td>
<td>10; 19.1; 7.6</td>
<td>95.5 (92.2-98.7)</td>
<td>39.4 (34.3-44.4)</td>
<td>36.1 (34.0-38.2)</td>
<td>96.1 (92.1-98.1)</td>
</tr>
</tbody>
</table>

Abbreviations: NPV, negative predictive value; PPV, positive predictive value; PSA, prostate-specific antigen.

* PSA is reported in nanograms per milliliter; both TMPRSS2:ERG and PCA3 are reported as scores.

* Because the TMPRSS2:ERG mutation is absent in many cancers, sensitivity above 95% for urinary TMPRSS2:ERG as a stand-alone predictive biomarker is possible only if no cases are excluded based on urinary TMPRSS2:ERG value alone; more than 5% of cases had a TMPRSS2:ERG score of 0.

* Including PSA level higher than 10 ng/mL as a priori trigger did not improve performance of the "TMPRSS2:ERG or PCA3" dual biomarker model in the developmental cohort because we found that all cases of PSA level above 10 ng/mL already exceeded the positive thresholds of urinary PCA3 (score >93.1) or TMPRSS2:ERG (score >7.6).

### Table 3. Validation of Multiplex Algorithm Including Urinary TMPRSS2:ERG, PCA3, and Serum PSA Level Higher Than 10 ng/mL to Improve Specificity of Predicting Aggressive Prostate Cancer (Gleason Score ≥7)

<table>
<thead>
<tr>
<th>Diagnostic Biomarker*</th>
<th>Threshold Value (n = 561)</th>
<th>Gleason ≥7 Sensitivity, % (95% CI) (n = 148)</th>
<th>No Cancer or Gleason ≤6 Sensitivity, % (95% CI) (n = 413)</th>
<th>PPV, % (95% CI) (Observed Prevalence, 26.38%)</th>
<th>NPV, % (95% CI) (Observed Prevalence, 26.38%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSA</td>
<td>3</td>
<td>91.2 (86.6-95.8)</td>
<td>16.7 (13.1-20.3)</td>
<td>28.2 (28.9-29.5)</td>
<td>84.1 (75.1-90.3)</td>
</tr>
<tr>
<td>PCA3</td>
<td>7</td>
<td>96.6 (93.7-99.5)</td>
<td>18.4 (14.7-22.1)</td>
<td>29.8 (28.6-30.9)</td>
<td>93.8 (86.2-97.3)</td>
</tr>
<tr>
<td>TMPRSS2:ERG⁶</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PSA or TMPRSS2:ERG</td>
<td>4; 289</td>
<td>85.8 (80.2-91.4)</td>
<td>34.1 (29.5-38.7)</td>
<td>31.8 (29.3-33.9)</td>
<td>87.0 (81.5-91.1)</td>
</tr>
<tr>
<td>PSA or PCA3</td>
<td>5; 22</td>
<td>90.5 (85.8-95.2)</td>
<td>32.2 (27.7-36.7)</td>
<td>32.5 (30.5-34.2)</td>
<td>90.4 (84.9-94.1)</td>
</tr>
<tr>
<td>PCA3 or TMPRSS2:ERG</td>
<td>20; 8</td>
<td>90.5 (85.8-95.2)</td>
<td>35.4 (30.8-40.0)</td>
<td>33.4 (31.5-35.4)</td>
<td>91.2 (86.1-94.6)</td>
</tr>
<tr>
<td>PSA or PCA3 or TMPRSS2:ERG</td>
<td>10; 20; 8</td>
<td>92.6 (88.4-96.8)</td>
<td>33.4 (28.8-37.9)</td>
<td>33.2 (31.4-35.1)</td>
<td>92.6 (87.5-95.8)</td>
</tr>
</tbody>
</table>

Abbreviations: NPV, negative predictive value; PPV, positive predictive value; PSA, prostate-specific antigen.

* PSA is reported in nanograms per milliliter; both TMPRSS2:ERG and PCA3 are reported as scores.

* Threshold set as TMPRSS2:ERG score of 0 or higher to include cases in which TMPRSS2:ERG score is 0.

* The rule combining PSA ≥ 10 ng/mL, PCA3 score ≥ 20, and TMPRSS2:ERG score ≥ 8 was prespecified as the primary decision algorithm to be tested by the validation analyses per prospective-specimen-collection, retrospective-blinded-evaluation (PROBE) criteria.⁶⁰ Observed specificity and sensitivity for the individual markers or dual combinations at these thresholds are shown for completeness and to provide context for interpretation of the 3-biomarker rule. When PCA3 and TMPRSS2:ERG were analyzed as continuous variables in the logistic regression model (separately from clinical decision-targeting "OR" combinatorial logic that was the focus of validating the prespecified thresholds following PROBE design), receiver operating curve plots of the individual biomarkers' accuracy in predicting prostate cancer with Gleason score of 7 or higher showed the following area under the curve values: PSA, 0.67; PCA3, 0.77; and TMPRSS2:ERG, 0.66. The corresponding area under the curve value for multivariable logistic regression combining PCA3 and TMPRSS2:ERG was 0.75, while the area under the curve value for combining PCA3, TMPRSS2:ERG, and PSA was 0.77.
preclinical reverse transcriptase–polymerase chain reaction assays evaluating T2:ERG and PCA3 in cellular fraction of post-DRE urine.34-36 Our group’s prior report, which used an earlier version of the T2:ERG assay35 and the corresponding “MIPS” test, did not discern the specific clinical scenario of decision about who should undergo initial prostate biopsy and did not evaluate assay thresholds to optimize specificity at high sensitivity for prostate cancer with a Gleason score of 7 or higher. Cost analysis suggested that, relative to having all men with abnormal PSA levels undergo biopsy, urine PCA3 and T2:ERG testing to select biopsy candidates could reduce cost of prostate cancer detection and consequent care. Relative to USPSTF recommendations (where men would undergo neither PSA testing nor consequent biopsy), urine RNA testing to select men for biopsy would reduce treatment costs for advanced disease, but these were offset by costs of biopsies and care for early-stage prostate cancer.

**Limitations**

We did not evaluate repeat testing, so our study does not inform the relationship of negative results with cancer detection during subsequent screens. We did not directly evaluate the impact of detection on patients’ length and quality of life. Cost estimates did not include the cost of PCA3 and T2:ERG testing (which does not yet have a designated federal payment rate). These considerations indicate a basis for future studies to examine more recent treatment cost data, the impact of detection on costs for unrelated conditions, repeat screening, and quality-adjusted life years.

Generalizability of the decision algorithm described herein has limitations. The multiplex urinary RNA algorithm was developed and validated in cohorts of men presenting for initial prostate biopsy owing to elevated PSA or abnormal DRE findings, and therefore represents a strategy to refine biopsy decisions after PSA screening rather than replacing PSA testing. In addition, biopsy as primary indicator of cancer aggressiveness is limited by ascertainment bias because a subset of patients may harbor cancer with a Gleason score of 7 or higher missed on biopsy. However, biopsy is the most definitive available prostate cancer diagnostic test. Magnetic resonance imaging (MRI)-guided biopsy represents an opportunity to reduce ascertainment error due to false-negative biopsy37 but was not a care standard when our study began, and consensus is still lacking as to MRI role with initial biopsy.19 Although blood tests such as the Prostate Health Index and 4K Score have been developed to complement total PSA testing,38-40 these were not included in our evaluation, and these may be advantageous over the urine test described in this study by not requiring a DRE. Combining urine RNA testing with novel PSA isoform or combination kallikrein tests represents an opportunity for further study.

**Conclusions**

Nevertheless, our results indicate that urinary PCA3 and T2:ERG testing can improve specificity of predicting aggressive prostate cancer beyond either serum PSA level or either urinary marker alone. Use of these tests to select men for initial biopsy after elevated PSA or abnormal DRE findings showed that 42% of men would have been safely excluded from undergoing unnecessary prostate biopsy, while high sensitivity for aggressive cancer was retained. These findings suggest that urinary RNA testing can mitigate harms of prostate screening while retaining the benefits of identifying aggressive cancers suitable for treatment.
Urology, Weill-Cornell Medical Center, New York, New York (Scher); Division of Cancer Prevention, National Cancer Institute, Bethesda, Maryland (Kagan, Srivastava); University of Texas Health Sciences Center – San Antonio, Texas (Thompson); Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Canada (Fan); Department of Pathology, Weill-Cornell Medical Center, New York, New York (Rubin).

Author Contributions: Dr Sanda 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. Drs Feng and Wei contributed equally to this study.

Concept and design: Sanda, Feng, Howard, Tomlins, Chan, Regan, Groskopf, Kagan, Srivastava, Thompson, Rubin, Chinnayyan, Wei, Bidair, Partin.

Acquisition, analysis, or interpretation of data: Sanda, Feng, Howard, Sokol, Regan, Groskopf, Chipman, Patil, Salami, Scherr, Kagan, Thompson, Siddiqui, Fan, Joon, Bantis, Wei, Bidair, Kibel, Lin, Lotan, Partin, Tanega.

Drafting of the manuscript: Sanda, Feng, Howard, Patil, Kagan, Siddiqui, Wei.

Critical revision of the manuscript for important intellectual content: Sanda, Feng, Howard, Tomlins, Sokoll, Chan, Regan, Groskopf, Chipman, Patil, Salami, Scherr, Kagan, Srivastava, Thompson, Fan, Joon, Bantis, Rubin, Chinnayyan, Wei, Bidair, Kibel, Lin, Lotan, Partin, Tanega.


Administrative, technical, or material support: Sanda, Sokoll, Chan, Groskopf, Patil, Salami, Scherr, Kagan, Thompson, Siddiqui, Rubin, Wei, Bidair, Kibel, Lotan, Tanega.

Additional Contributions: We acknowledge the contributions of the group collaborators J. Erik Busby, MD, University of South Carolina, Greenville, and Rosalia Viterbo, MD, Fox Chase Cancer Center for their assembly of the PCA3 Trial biorepository that was accessioned as the validation cohort in this study. We also acknowledge project management by John Noel, MS (for data collection for the developmental cohort), and Jackie Dahlgren, BS (data collection for the validation cohort). These persons were compensated for their contributions only in the normal course of their employment.

Conflict of Interest Disclosures: The University of Michigan and Brigham and Women’s Hospital, Boston, Massachusetts (Kibel); University of Washington Medical Center, Seattle (Lin); University of Texas Southwestern Medical Center, Dallas (Lotan); Johns Hopkins Hospital, Baltimore, Maryland (Partin); New York University School of Medicine, New York (Tanega).

Additional Contributions: We acknowledge the contributions of the group collaborators J. Erik Busby, MD, University of South Carolina, Greenville, and Rosalia Viterbo, MD, Fox Chase Cancer Center for their assembly of the PCA3 Trial biorepository that was accessioned as the validation cohort in this study. We also acknowledge project management by John Noel, MS (for data collection for the developmental cohort), and Jackie Dahlgren, BS (data collection for the validation cohort). These persons were compensated for their contributions only in the normal course of their employment.

Conflict of Interest Disclosures: The University of Michigan and Brigham and Women’s Hospital have been issued a patent on the detection of ETS gene fusions in prostate cancer, on which Drs Tomlins, Rubin, and Chinnayyan are listed as coinventors. The field of use designation has been licensed to Hologic/GenProbe Inc, which sublicensed rights to Ventana Medical Systems Inc. Dr Tomlins has received honoraria from and has served as a consultant to Ventana Medical Systems. Dr Chinnayyan has served as consultant to GenProbe Inc and Ventana Medical Systems. For Dr Rubin, the invention is disclosed and handled under Cornell University’s financial conflict of interest related to research policy (https://www.dfa.cornell.edu/tools-library/policies/financial-conflict-interest-related-research). Drs Chinnayyan and Tomlins have conflict of interest management plans in place that were developed by the University of Michigan Medical School conflict of interest board. The Medical School conflict of interest board is charged with reviewing all disclosures by faculty members and/or professional staff and recommending appropriate conflict management if appropriate. Details of the University of Michigan Medical School conflict of interest board’s policies and procedures can be found at http://research-compliance.umich.edu/operations-manual/conflicts-interest-and-commitment. Dr Groskopf is an employee of Hologic Inc, the manufacturer of the Progensa PCA3 assay and licensee of the diagnostic field of use for TMPRSS2:ERG. Dr Thompson has served as consultant for Exosome Diagnostics. Dr Wei has research collaborations with NCI, Hologic, and Exosome Inc; he has also been on the speaker bureau for Metamark. Among the EDRN-PCA3 Study Group authors, Dr Bidair is investigator for prostate cancer studies with Argenx, Aragon, Astellas, Bayer, Bavarian Nordic and Nymox; he was previously a consultant on an advisory board for prostate cancer for Astellas. Dr Kibel is a consultant for DenDreon, Sanofi Aventis, Tokai, Profound, and MTG. Dr Lin has research collaborations with NCI, DOD, Hologic, Genomic Health Inc, and GenomedX; he is a consultant for Astellas. Dr Tanega is a consultant for Hitachi-Aloka and advisory board member for Opko; he is also a scientific investigator for Trod Medical. No other disclosures are reported.

Funding/Support: This research was supported by the following National Institutes of Health (NIH) grants: NIH U01 CA159713 (Drs Sanda, Wei, and Regan, Mr Chipman, Drs Patil, Salami, and Scherr), NIH U01 CA121725 (Drs Rubin, Mr Siddiqui, and Drs Chinnayyan and Tomlins), NIH U01 CA151052 (Drs Chan and Sokoll), NIH U24 CA068638 (Drs Feng and Joon), and NIH U01 CA068402 (Dr Thompson). In addition, Hologic provided assay reagents for the study, and Drs Tomlins and Chinnayyan are supported by the A. Alfred Taubman Medical Research Institute. Dr Patil is supported by the Dunwoody Country Club Senior Men’s Foundation.

Role of the Funder/Sponsor: The Study sponsor (NIH) participated in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript, to the extent that the NIH is the employer of coauthors Dr Srivastava and Dr Kagan. Hologic Inc provided assay reagents and is the employer of coauthor Dr Groskopf. The Dunwoody Country Club Senior Men’s Foundation provided funding for data management effort but did not participate in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

REFERENCES


