Development and Validation of an 18-Gene Urine Test for High-Grade Prostate Cancer

Jeffrey J. Tosoian, MD, MPH; Yuping Zhang, PhD; Lanbo Xiao, PhD; Cassie Xie, MS; Nathan L. Samora, MD; Yashar S. Niknafs, PhD; Zoey Chopra, MA; Javed Siddiqui, MS; Heng Zheng, MD; Grace Herron, BA; Neil Vaishampayan, BS; Hunter S. Robinson, MD; Kumaran Arivoli, BS; Bruce J. Trock, PhD; Ashley E. Ross, MD; Todd M. Morgan, MD; Ganesh S. Palapattu, MD; Simpa S. Salami, MD, MPH; Lakshmi P. Kunju, MD; Scott A. Tomlins, MD, PhD; Lori J. Sokoll, PhD; Daniel W. Chan, PhD; Sudhir Srivastava, PhD; Ziding Feng, PhD; Martin G. Sanda, MD; Yingye Zheng, PhD; John T. Wei, MD; Arul M. Chinnaiyan, MD, PhD; for the EDRN-PCA3 Study Group

IMPORTANCE Benefits of prostate cancer (PCa) screening with prostate-specific antigen (PSA) alone are largely offset by excess negative biopsies and over-detection of indolent cancers resulting from the poor specificity of PSA for high-grade PCa (ie, grade group [GG] 2 or greater).

OBJECTIVE To develop a multiplex urinary panel for high-grade PCa and validate its external performance relative to current guideline-endorsed biomarkers.

DESIGN, SETTING, AND PARTICIPANTS RNA sequencing analysis of 58,724 genes identified 54 markers of PCa, including 17 markers uniquely overexpressed by high-grade cancers. Gene expression and clinical factors were modeled in a new urinary test for high-grade PCa (MyProstateScore 2.0 [MPS2]). Optimal models were developed in parallel without prostate volume (MPS2) and with prostate volume (MPS2+). The locked models underwent blinded external validation in a prospective National Cancer Institute trial cohort. Data were collected from January 2008 to December 2020, and data were analyzed from November 2022 to November 2023.

EXPOSURE Protocolized blood and urine collection and transrectal ultrasound-guided systematic prostate biopsy.

MAIN OUTCOMES AND MEASURES Multiple biomarker tests were assessed in the validation cohort, including serum PSA alone, the Prostate Cancer Prevention Trial risk calculator, and the Prostate Health Index (PHI) as well as derived multiplex 2-gene and 3-gene models, the original 2-gene MPS test, and the 18-gene MPS2 models. Under a testing approach with 95% sensitivity for PCa of GG 2 or greater, measures of diagnostic accuracy and clinical consequences of testing were calculated. Cancers of GG 3 or greater were assessed secondarily.

RESULTS Of 761 men included in the development cohort, the median (IQR) age was 63 (58-68) years, and the median (IQR) PSA level was 5.6 (4.6-7.2) ng/mL; of 743 men included in the validation cohort, the median (IQR) age was 62 (57-68) years, and the median (IQR) PSA level was 5.6 (4.1-8.0) ng/mL. In the validation cohort, 151 (20.3%) had high-grade PCa on biopsy. Area under the receiver operating characteristic curve values were 0.60 using PSA alone, 0.66 using the risk calculator, 0.77 using PHI, 0.76 using the derived multiplex 2-gene model, 0.72 using the derived multiplex 3-gene model, and 0.74 using the original MPS model compared with 0.81 using the MPS2 model and 0.82 using the MPS2+ model. At 95% sensitivity, the MPS2 model would have reduced unnecessary biopsies performed in the initial biopsy population (range for other tests, 15% to 30%; range for MPS2, 35% to 42%) and repeat biopsy population (range for other tests, 9% to 21%; range for MPS2, 46% to 51%). Across pertinent subgroups, the MPS2 models had negative predictive values of 95% to 99% for cancers of GG 2 or greater and of 99% for cancers of GG 3 or greater.

CONCLUSIONS AND RELEVANCE In this study, a new 18-gene PCa test had higher diagnostic accuracy for high-grade PCa relative to existing biomarker tests. Clinically, use of this test would have meaningfully reduced unnecessary biopsies performed while maintaining highly sensitive detection of high-grade cancers. These data support use of this new PCa biomarker test in patients with elevated PSA levels to reduce the potential harms of PCa screening while preserving its long-term benefits.

JAMA Oncol. doi:10.1001/jamaoncol.2024.0455
Published online April 18, 2024.
Prostate cancer (PCa) remains the most commonly diagnosed malignancy and a leading cause of cancer death worldwide.1 The European Randomized Study of Screening for PCa and Göteborg Randomized Prostate Cancer Screening trial showed significant reductions in cancer mortality for men participating in prostate-specific antigen (PSA)-based screening.2,3 At the same time, these studies confirmed that PSA screening leads to unnecessary invasive biopsies in men without cancer and frequent overdiagnosis of low-grade, indolent cancers (grade group [GG] 1).4 In response to this, current clinical guidelines offer that men with an elevated PSA level undergo multiparametric magnetic resonance imaging (mpMRI), if available, or biomarker testing for risk stratification prior to biopsy.5,6

Indeed, use of prostate mpMRI with targeted biopsy has improved detection of clinically significant, high-grade cancer (ie, cancer of GG 2 or greater) in men with tumors visible on mpMRI.7 While these data support prebiopsy mpMRI in patients requiring biopsy, the use of negative findings on mpMRI to rule out high-grade cancers in men with elevated PSA levels is not well supported. Population-level data spanning academic and community settings reveal a negative predictive value (NPV) of only 77% for high-grade cancers,8 and subjective interpretation of mpMRI is highly problematic, with NPVs as low as 63% by site and 40% among radiologists.9,10 Thus, even following negative findings on mpMRI, its limited sensitivity merits biopsy in a substantial proportion of men. Moreover, there are practical reasons mpMRI may not be feasible for population-wide use after PSA, including its resource burden and limited availability in the community setting.11,12

Objective, noninvasive biomarker tests could be a more practical option. Current National Comprehensive Cancer Network (NCCN) guidelines offer 6 blood-based and urine-based biomarker tests, each including 3 or fewer markers of PCa (ie, cancer of any grade).5 While consistently outperforming PSA alone,13 these assays have not evolved to reflect current understanding of PCa biology. For one, given the minimal metastatic potential of low-grade cancers, contemporary practice is focused on detecting high-grade cancers, while reducing overdiagnosis of low-grade disease.4 Thus, assays based solely on markers associated with cancer of any grade have limited biologic specificity for high-grade cancers. Moreover, assays including only 2 to 3 biomarkers simply cannot capture the multitude of diverse molecular pathways driving lethal disease.14,15

We hypothesized that augmenting the prior generation of cancer-associated biomarkers with novel molecules selectively expressed by high-grade, aggressive cancers would improve testing accuracy. Leveraging multi-institutional transcriptomic data,14,16,17 we identified novel genes specifically overexpressed by high-grade cancers. We then adopted multiplex polymerase chain reaction (PCR)-based technology to evaluate 54 candidate markers in a development cohort, deriving an optimal 18-gene assay for standard clinical use. Finally, we performed blinded external validation of the new assay, including direct comparison with currently endorsed biomarker tests.

Key Points

**Question** Can a new 18-gene urinary test for high-grade prostate cancer (ie, grade group [GG] 2 or greater) improve prostate-specific antigen (PSA) screening outcomes relative to existing biomarker tests?

**Findings** In this diagnostic study including 761 men in the development cohort and 743 men in the validation cohort, novel cancer-specific and high-grade cancer-specific genes were identified from RNA sequencing data and optimally modeled in a development cohort, yielding an 18-gene test for high-grade prostate cancer. Applying a testing approach with 95% sensitivity for high-grade prostate cancer to an external validation population, use of the 18-gene test would have reduced the number of unnecessary biopsies performed relative to current guideline-endorsed tests.

**Meaning** The new 18-gene prostate cancer test may reduce more burdensome additional testing (eg, imaging and biopsy) while maintaining highly sensitive detection of high-grade cancer in patients undergoing PSA screening.

Methods

Institutional review board approval was obtained from the University of Michigan Institutional Review Board and at each site, and all participants provided written informed consent. This study followed the Standards for Reporting of Diagnostic Accuracy (STARD) reporting guideline.18

Biomarker Discovery

The original MyProstateScore (MPS) test incorporates prostate cancer antigen 3 (PCA3) and TMPRSS2:ERG gene fusion expression with serum PSA level to estimate risk of high-grade cancers and is endorsed by NCCN guidelines for prebiopsy risk stratification.5,19 To derive a gene panel for high-grade cancers, we performed differential expression analysis of 58 724 genetic targets in multi-institutional RNA sequencing data (Figure 1; eFigures 1 and 2 in Supplement 1 and the eTable in Supplement 2). A total of 72 genes met predefined nomination criteria for cancer (n = 50) or high-grade cancer (n = 22) (eTable 1 in Supplement 1). Removal of collinear genes and those without PCR primers resulted in 44 candidate markers (eFigures 1 to 3 in Supplement 1). These were supplemented with 10 previously described PCA-associated or reference genes, yielding a 54-gene candidate panel.

Model Development

**Development Cohort**

Prebiopsy urine has been prospectively collected at the University of Michigan Prostate Specialized Program of Research Excellence under a National Cancer Institute (NCI) Early Detection Research Network (EDRN) protocol approved by the University of Michigan Institutional Review Board since 2008. First-catch urine was obtained following digital rectal examination and was mixed with RNA stabilization buffer and frozen at −70°C.20 The development cohort included patients presenting for 12-core or greater prostate biopsy due to elevated
**Figure 1. Biomarker Discovery and Development of the MyProstateScore 2.0 Urinary Test for High-Grade Prostate Cancer**

**A** Discovery and nomination of candidate biomarkers

| Tissue-RNA sequencing discovery cohort | 220 Normal (benign) prostate tissue | 71 Low-grade cancer | 484 High-grade cancer |

28 High-grade cancer-associated genes

50 Cancer-associated genes not meeting high-grade criteria

10 Curated prostate cancer targets and candidate reference genes

44 Candidate genes

| 775 | Tissue-RNA sequencing discovery cohort |

| 220 | Normal (benign) prostate tissue |

| 71 | Low-grade cancer |

| 484 | High-grade cancer |

**B** Development of the optimal 18-gene model for high-grade cancer

761 Participants included in the development cohort

**Partition 1**

**Partition 2**

**Partition 3**

**Partition 4**

Repeated ×10

| Model 1 | (+4n) |

| Model 2 | (+4n) |

| Model 3 | (+4n) |

| Model 4 | (+4n) |

18 Markers

TMRPSS2-ERG

SLC3A1

OR51E2

APOC1

PCAT14

CAMKK2

PC3

KN1

B3GNT5

PFN1

TFF3

SPON2

ERG

VSTM2L

DLX1

ERG

TIM15A

KLK4

HOXC6

54 Final target list

ACSM1

ETV1

LRRN1

SCHLAP1

AMACR

F5

MIPEP

SPDEF

APOC1

GAPDH

MS4A8

SNRK1

AR-V7

GLI1

NKAIN1

T2-ERG

B3GNT6

GOLM1

NUDT7B

TDG2

CAMKK2

GRN3A

OR51E2

TFF3

COL5A2

HIC6

PC3

TX1

CASP3

HPN

PCAT14

TMEFF2

CST2

KLK2

PCGEM1

TMS15A

CYP51A1

KLK3

PDLIM5

TRCV9

DLX1

KLK4

PEX10

VSTM2L

EEF1A2

LBH

PLA1A

ERG

LINC00993

PLAG2

**A. Discovery and nomination of candidate biomarkers for the multiplex urinary panel.** Biomarker discovery was performed using RNA sequencing data from 220 benign prostates, 71 with cancers of grade group 1, and 484 with cancers of grade 2 or greater available through the Cancer Genome Atlas, the Genotype-Tissue Expression portal, and the University of Michigan. A total of 72 markers met predefined criteria. Of these, quantitative polymerase chain reaction probe could not be successfully designed for 19, and 9 genes were highly cross-correlated, resulting in exclusion from the candidate panel. The remaining 44 transcripts meeting nomination criteria were supplemented with 10 curated genes to yield a 54-gene candidate panel.

**B. Development of the optimal 18-gene model for high-grade cancer.** To avoid multicollinearity in regression models, highly correlated variables were identified and removed with a stepwise procedure. We assessed 3 model-building approaches: (1) logistic regression with stepwise feature selection, (2) logistic regression with recursive feature elimination, and (3) regularized logistic regression with elastic net. Performance of each model-building approach was quantified as the area under the receiver operating characteristic curve on repeated cross-validation (10-fold cross-validation repeated 3 times) with up sampling of the minor class to yield balanced classes. Elastic net modeling yielded the highest median area under the curve and was used for development. Using an ensemble approach, the development set was randomly divided into 4 partitions, and the model yielding the highest area under the curve was identified for each partition. This approach was repeated 10 times with different random seeds, yielding 40 elastic net models in total. For each candidate gene, the frequency of model inclusion and importance to high-grade prostate cancer detection was tabulated across models. Based on analysis of optimal feature size and technical features of the OpenArray platform (Thermo Fisher Scientific), the 17 biomarkers providing optimal discriminative accuracy for prostate cancer of grade group 2 or greater were included with standard clinical variables and the normalization gene KLK3 in the MyProstateScore 2.0 model (without prostate volume) and MyProstateScore 2.0 Plus model (with prostate volume). Models were calibrated and internally cross-validated prior to external validation.
PSA levels (3-10 ng/mL; to convert to micrograms per liter, multiply by 1) from 2008 to 2020. In accordance with guidelines, we excluded patients with PCAs. Based on proposed use of this test as a pre-mpMRI, prebiopsy test to rule out the need for mpMRI or biopsy, we excluded men with a history of prostate mpMRI and targeted biopsy.

**Multiplex Quantitative PCR OpenArray Profiling**
OpenArray technology (Thermo Fisher Scientific) is a high-throughput real-time quantitative PCR (qPCR) method for rapid screening of multiple TaqMan assays. RNA isolation, extraction, and complementary DNA synthesis were performed (eFigure 4 in Supplement 1).

**Model Building and Calibration**
We assessed the 54-gene candidate panel using multiple model-building strategies (Figure 1). Clinical factors consistently associated with PCAs (age, race, digital rectal examination findings, PSA level, family history of prostate cancer, and prior negative biopsy) were locked into models a priori. Because prostate volume improves predictive value but is not available for all patients, we developed a second model including volume for use when volume is clinically available (ie, previous biopsy or mpMRI). Test outputs were standardized to represent the percentage likelihood of detecting high-grade cancers (0% to 100%). The optimal 18-gene model without prostate volume (MPS2) and with prostate volume (MPS2+) were calibrated (eFigure 5 in Supplement 1) to account for differences in outcome prevalence between cohorts, but is not available for all patients, we developed a second model including volume for use when volume is clinically available (ie, previous biopsy or mpMRI). Test outputs were standardized to represent the percentage likelihood of detecting high-grade cancers (0% to 100%). The optimal 18-gene model without prostate volume (MPS2) and with prostate volume (MPS2+) were calibrated (eFigure 5 in Supplement 1) to account for differences in outcome prevalence between cohorts, because prostate volume improves predictive value but is not available for all patients, we developed a second model including volume for use when volume is clinically available (ie, previous biopsy or mpMRI). Test outputs were standardized to represent the percentage likelihood of detecting high-grade cancers (0% to 100%). The optimal 18-gene model without prostate volume (MPS2) and with prostate volume (MPS2+) were calibrated (eFigure 5 in Supplement 1) to account for differences in outcome prevalence between cohorts, because prostate volume improves predictive value but is not available for all patients, we developed a second model including volume for use when volume is clinically available (ie, previous biopsy or mpMRI). Test outputs were standardized to represent the percentage likelihood of detecting high-grade cancers (0% to 100%).

**Model Validation**
**External Validation Cohort**
The validation cohort consisted of patients in the prospective NCI EDRN PCA3 Evaluation Trial. This trial enrolled consecutive patients presenting for biopsy across 11 academic centers, primarily due to elevated PSA levels or abnormal digital rectal examination findings, PSA level, family history of prostate cancer, and prior negative biopsy were locked into models a priori. Because prostate volume improves predictive value but is not available for all patients, we developed a second model including volume for use when volume is clinically available (ie, previous biopsy or mpMRI). Test outputs were standardized to represent the percentage likelihood of detecting high-grade cancers (0% to 100%). The optimal 18-gene model without prostate volume (MPS2) and with prostate volume (MPS2+) were calibrated (eFigure 5 in Supplement 1) to account for differences in outcome prevalence between cohorts, because prostate volume improves predictive value but is not available for all patients, we developed a second model including volume for use when volume is clinically available (ie, previous biopsy or mpMRI). Test outputs were standardized to represent the percentage likelihood of detecting high-grade cancers (0% to 100%). The optimal 18-gene model without prostate volume (MPS2) and with prostate volume (MPS2+) were calibrated (eFigure 5 in Supplement 1) to account for differences in outcome prevalence between cohorts, because prostate volume improves predictive value but is not available for all patients, we developed a second model including volume for use when volume is clinically available (ie, previous biopsy or mpMRI). Test outputs were standardized to represent the percentage likelihood of detecting high-grade cancers (0% to 100%). The optimal 18-gene model without prostate volume (MPS2) and with prostate volume (MPS2+) were calibrated (eFigure 5 in Supplement 1) to account for differences in outcome prevalence between cohorts, because prostate volume improves predictive value but is not available for all patients, we developed a second model including volume for use when volume is clinically available (ie, previous biopsy or mpMRI). Test outputs were standardized to represent the percentage likelihood of detecting high-grade cancers (0% to 100%).

**Specimens and Laboratory Analysis**
Deidentified urine specimens were shipped to the University of Michigan for OpenArray profiling. Laboratory procedures were conducted per the identical protocol used in development. We derived a multiplex 2-gene model (HOXC6 and DLXI) and a multiplex 3-gene model (PCA3, ERG, and SPDEF). These genes are measured in the commercially available SelectMDx and ExoDx Prostate Intelliscore (EPI) tests, respectively. The multiplex models considered herein were independently derived based on gene expression measured using the OpenArray qPCR platform and are not proposed to represent the commercial products. Serum PSA, free PSA, and [−2]proPSA were measured using the Access 2 Immunoassay System (Beckman Coulter) at the Johns Hopkins EDRN Laboratory.

**Blinded Validation and Comparative Analysis**
Expression data and model coefficients were available to 2 investigators (C.X. and Y. Zheng) at the NCI EDRN for predefined validation. Locked model coefficients from development were used to generate outputs of the derived multiplex 2-gene model, derived multiplex 3-gene model, MPS2, and MPS2+. The original MPS was calculated using qPCR-based PCA3 and TMPRSS2:ERG scores; a subset of data were previously described. Prostate Health Index (PHI) was calculated using the formula ([−2]proPSA/free PSA) × √(PSA).

Comparative analysis included PSA, the Prostate Cancer Prevention Trial risk calculator, PHI, the derived multiplex 2-gene and 3-gene models, MPS, MPS2, and MPS2+. The primary outcome was cancer of GG 2 or greater on biopsy; cancer of GG 3 or greater was secondarily assessed. Diagnostic potential was visualized by receiver operating characteristic (ROC) curves and quantified by area under the ROC curve (AUC) using R package pROC. For the development cohort, mean AUC from repeated 10-fold cross-validation was reported. For the validation cohort, 95% CIs of AUCs were computed with 2000 stratified bootstrap.

We sought to illustrate test performance using a straightforward, clinically applicable approach. As described, considering prevalence of high-grade cancers in testing populations (7% to 31%), relative harms of false-negative and false-positive testing results, and the proposed role of biomarkers for rule-out testing, we assessed thresholds providing 95% sensitivity for high-grade cancer. Performance measures were calculated using the confusionMatrix function of R package caret. Given disparate risk profiles of initial and repeat biopsy populations, analyses were carried out in each subpopulation.

Decision curve analysis (DCA) was used to quantify net benefit of each biomarker on the decision to undergo biopsy compared with (1) biopsying all patients and (2) biopsying no patients. Considering a more than 20% risk of high-grade cancer justifies performing biopsy and a less than 5% risk justifies foregoing biopsy in most patients, we assessed threshold probabilities spanning this clinically relevant range. DCA was performed using dca in the R package dcrules.

**Statistical Analysis**
Statistical analyses were performed using R version 4.1.1 (The R Foundation). Two-tailed tests were used for all comparisons, and P values less than .05 were considered statistically significant.

**Results**
**Model Development**
Among 815 participants in the development cohort, qPCR yielded valid results in 761 (93.4%) (eFigure 4 in Supplement 1).
ment 1). The median (IQR) age was 63 (58-68) years, and the median (IQR) PSA level was 5.6 (4.6-7.2) ng/mL (Table 1). On study biopsy, 293 men (38.5%) had cancer of GG 2 or greater. The contribution of candidate genes to model predictions was quantified across elastic net models (Figure 1; eTable 3 in Supplement 1). The final MPS2 model included clinical variables and the 17 most informative markers, including 13 from the discovery analysis (4 high-grade cancer-specific genes [APOCI, B3GNT6, NKAIN1, and SCHLAP1] and 9 cancer-specific genes [PCGEM1, SPON2, TRGV9, PCA3, ORS1E2, CAMKK2, TFF3, PCAT14, and TMSB15A]), 4 curated markers (HOXC6, ERG, TMPRSS2:ERG, and KLK4), plus the reference gene KLK3 (eTable in Supplement 3). Model coefficients were determined in the overall cohort (eTable 4 in Supplement 1). Calibration and internal cross-validation were performed (eFigures 5 and 6 in Supplement 1), and the MPS2 models were locked for external validation.

**External Validation and Comparative Analysis**

**Overall Study Population**

Of 813 patients in the validation cohort (eFigure 7 in Supplement 1), qPCR was successful in 743 (91.4%). The median (IQR) age was 62 (57-68) years, and the median (IQR) PSA level was 5.6 (4.1-8.0) ng/mL. A total of 95 men (12.8%) were Black and 648 (87.2%) were another race, and 247 men (33.2%) had a previous negative biopsy (Table 1). On study biopsy, 151 men (20.3%) had high-grade PCa. Median (IQR) MPS2 values were significantly higher in men with cancer of GG 2 or greater (0.44 [0.23-0.69]) than in men with negative biopsies (0.08 [0.03-0.19]; P < .001) and in men with cancer of GG 1 (0.20 [0.08-0.43]; P < .001) (Table 1; eFigure 8 in Supplement 1). Similarly, median (IQR) MPS2+ values were significantly higher in men with PCa of GG 2 or greater (0.54 [0.27-0.79]) relative to those with negative biopsies (0.08 [0.03-0.21]; P < .001) or those with cancer of GG 1 (0.25 [0.09-0.48]; P < .001). The AUC values for high-grade cancer were 0.60 (95% CI, 54.7-64.6) for PSA alone, 0.66 (95% CI, 61.1-70.7) for the Prostate Cancer Prevention Trial risk calculator, 0.77 (95% CI, 73.0-81.3) for PHI, 0.76 (95% CI, 71.9-80.3) for the derived multiplex 2-gene model, 0.72 (95% CI, 67.0-76.1) for the derived multiplex 3-gene model, and 0.74 (95% CI, 69.4-78.0) for MPS compared with 0.81 (95% CI, 76.9-84.6) for MPS2 and 0.82 (95% CI, 78.1-85.5) for MPS2+ (eFigure 9 in Supplement 1). The observed prevalence of high-grade cancer closely approximated MPS2 and MPS2+ predicted probabilities (Figure 2), reflecting good calibration. Critically, the models were particularly well-calibrated for predicted probabilities less than 30%, which are most clinically pertinent.

We assessed clinical consequences of prebiopsy biomarker testing. At a 95% sensitivity testing threshold, the proportions of unnecessary biopsies that would have been avoided using each test were 11% for PSA alone, 20% for the Prostate Cancer Prevention Trial risk calculator, 26% for PHI, 27% for the derived multiplex 2-gene model, 17% for the derived multiplex 3-gene model, and 23% for MPS compared with 37% for MPS2 and 41% for MPS2+. Full performance measures and the estimated numbers of unnecessary biopsies avoided per 1000 patients are listed in Table 2. Critically, MPS2 and MPS2+ each provided 99% sensitivity and 99% NPV for cancer of GG 3 or greater.

**Initial Biopsy Subpopulation**

The initial biopsy population included 496 patients with a median (IQR) PSA level of 5.0 (3.8-6.6) ng/mL (eTable 5 in Supplement 1). On study biopsy, 133 (26.8%) had high-grade cancer. Using a 95% sensitivity threshold, the proportions of unnecessary biopsies avoided were 15% for PSA alone, 27% for the Prostate Cancer Prevention Trial risk calculator, 30% for PHI, 30% for the derived multiplex 2-gene model, 17% for the derived multiplex 3-gene model, and 27% for MPS compared with 35% for MPS2 (Table 2; eTable 6 in Supplement 1). Although patients undergoing initial biopsy often may not have prostate volume available, use of MPS2+ would have avoided 42% of unnecessary biopsies. Performance of MPS2 models with and without clinical factors are provided by subpopulation in eTables 7 and 8 in Supplement 1. An alternative initial biopsy model was developed in the initial biopsy population of the development cohort and similarly validated (eTables 9 and 10 and eFigure 10 in Supplement 1).

**Repeat Biopsy Subpopulation**

The repeat biopsy population included 247 men with median (IQR) PSA level of 7.2 (5.5-9.8) ng/mL, of which 18 (7.3%) were found to have high-grade cancer (eTable 5 in Supplement 1). At 95% sensitivity, the proportions of unnecessary biopsies that would have been avoided were 15% for PSA alone, 8.7% for PHI, 14% for the derived multiplex 2-gene model, 16% for the derived multiplex 3-gene model, 15% for MPS, 46% for MPS2, and 51% for MPS2+ (Table 2). Accordingly, MPS2 testing would have avoided approximately one-half of unnecessary biopsies while maintaining detection of 94.4% of high-grade cancers.

**DCA**

DCA was used to evaluate the net benefit of biomarker testing relative to performing biopsy in all patients and performing no biopsies. Across the clinically pertinent threshold probabilities spanning 5% to 20%, use of the MPS2 models would have provided the highest net clinical benefit across all tests (Figure 3A). Expressing benefit as net reduction in unnecessary biopsies, use of the MPS2 models would have provided the greatest net reduction in unnecessary biopsies without failing to biopsy a single patient with high-grade cancer (Figure 3B).

**Discussion**

Acknowledging the indolent nature of low-grade PCa, contemporary guidelines emphasize a narrowed diagnostic focus on high-grade cancers.5,6,44 Existing biomarkers expressed by all PCa—including low-grade, indolent tumors—therefore offer limited potential to selectively detect high-grade disease. Translating sequencing-based discovery to an expandable qPCR platform, we developed and validated a new urinary test incorporating 17 markers of cancer, and—for the first time, to our knowledge—novel markers uniquely...
Table 1. Characteristics of the Development and Validation Populations Overall and Stratified by Pathologic Findings on Prostate Biopsy

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Median (IQR)</th>
<th>Development cohort</th>
<th>External validation cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total (n = 761)</td>
<td>Negative (n = 362)</td>
</tr>
<tr>
<td>Age, y</td>
<td>63 (58-68)</td>
<td>62 (57-67)</td>
<td>64 (57-68)</td>
</tr>
<tr>
<td>Race, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>33 (4)</td>
<td>12 (3)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>Other race</td>
<td>728 (96)</td>
<td>350 (97)</td>
<td>102 (96)</td>
</tr>
<tr>
<td>Positive family history, No. (%)</td>
<td>163 (21)</td>
<td>105 (29)</td>
<td>22 (21)</td>
</tr>
<tr>
<td>Abnormal DRE, No. (%)</td>
<td>104 (14)</td>
<td>34 (9)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>Prostate volume, mL (b)</td>
<td>5.6 (4.6-7.2)</td>
<td>5.6 (4.6-6.8)</td>
<td>5.5 (4.6-7.0)</td>
</tr>
<tr>
<td>PSA, ng/mL</td>
<td>0.12 (0.08-0.16)</td>
<td>0.10 (0.07-0.14)</td>
<td>0.12 (0.09-0.16)</td>
</tr>
<tr>
<td>PHI</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Derived multiplex 2-gene model</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Derived multiplex 3-gene model</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>MPS</td>
<td>37 (20-58)</td>
<td>26 (14-42)</td>
<td>42 (24-63)</td>
</tr>
<tr>
<td>MPS2+ (0.05-0.39)</td>
<td>0.16</td>
<td>0.07</td>
<td>0.15</td>
</tr>
<tr>
<td>MPS2+ (0.05-0.42)</td>
<td>0.14</td>
<td>0.06</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Abbreviations: DRE, digital rectal examination; GG, grade group; MPS, MyProstateScore; MPS2, MyProstateScore 2.0; MPS2+, MyProstateScore 2.0 plus prostate volume; NA, not applicable; PHI, Prostate Health Index; PSA, prostate-specific antigen.

For conversion factors, to convert PSA to \(\mu\)g/L, multiply by 1.

a Race was self-reported by participants via a questionnaire. Black race was pertinent to the current study due to the well-established association of race with prostate cancer incidence, outcomes, and tumor molecular subtypes. 26 The other race category includes American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, White, other, and unknown race.

b Measured by transrectal ultrasound.

c PSA density equals serum PSA divided by prostate volume.

d MPS2 and MPS2+ values are reported on a continuous scale as the likelihood of cancer of GG 2 or greater detection on biopsy.
overexpressed by high-grade cancers relative to low-grade cancers. On validation, MPS2 testing with 95% sensitivity for high-grade cancer had 95% to 99% NPV and 35% to 51% specificity across subgroups. For individual patients, NPVs approaching 100% provide clear guidance for confident decision-making. For clinicians, uniform use of MPS2 could avoid unnecessary biopsies while preserving immediate detection of 95% of cancers of GG 2 or greater diagnosed using the biopsy all approach. Critically, MPS2 had 99% sensitivity and 99% NPV for cancers of GG 3 or greater, meaning the rare false-negative MPS2 results were almost uniformly more favorable cancers of GG 2 least likely to metastasize.

The 2023 NCCN guidelines for PCa early detection propose consideration of prebiopsy risk stratification with validated biomarker tests, including PHI, SelectMDx, 4Kscore, EPI, MPS, and IsoPSA. These tests have consistently outperformed PSA alone, with aggregate data approximating 90% to 95% sensitivity and 30% to 40% specificity for high-grade cancer. However, heterogeneity of published data and a lack of head-to-head comparisons have precluded recommendations of any particular testing approach. Using an NCI.

Table 2. Performance of Prostate-Specific Antigen (PSA) Alone, Prostate Cancer Prevention Trial Risk Calculator, Prostate Health Index (PHI), Derived Multiplex 2-Gene and 3-Gene Models, MyProstateScore (MPS), MPS2, and MPS2 Plus Prostate Volume (MPS2+) in the Validation Cohort

<table>
<thead>
<tr>
<th>Model</th>
<th>% Sensitivity</th>
<th>% Specificity</th>
<th>NPV</th>
<th>PPV</th>
<th>Estimated unnecessary biopsies avoided per 1000 patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall (n = 743)</td>
<td>95</td>
<td>11</td>
<td>90</td>
<td>21</td>
<td>108</td>
</tr>
<tr>
<td>PSA</td>
<td>95</td>
<td>20</td>
<td>94</td>
<td>23</td>
<td>198</td>
</tr>
<tr>
<td>Prostate Cancer Prevention Trial risk calculator</td>
<td>95</td>
<td>26</td>
<td>96</td>
<td>25</td>
<td>258</td>
</tr>
<tr>
<td>PHI</td>
<td>95</td>
<td>27</td>
<td>96</td>
<td>25</td>
<td>270</td>
</tr>
<tr>
<td>Derived multiplex 2-gene model</td>
<td>95</td>
<td>27</td>
<td>96</td>
<td>25</td>
<td>270</td>
</tr>
<tr>
<td>Derived multiplex 3-gene model</td>
<td>95</td>
<td>17</td>
<td>94</td>
<td>23</td>
<td>171</td>
</tr>
<tr>
<td>MPS</td>
<td>95</td>
<td>23</td>
<td>94</td>
<td>24</td>
<td>230</td>
</tr>
<tr>
<td>MPS2</td>
<td>95</td>
<td>37</td>
<td>97</td>
<td>28</td>
<td>370</td>
</tr>
<tr>
<td>MPS2+</td>
<td>95</td>
<td>41</td>
<td>97</td>
<td>29</td>
<td>405</td>
</tr>
<tr>
<td>Initial biopsy (n = 496)</td>
<td>95</td>
<td>15</td>
<td>89</td>
<td>29</td>
<td>152</td>
</tr>
<tr>
<td>PSA</td>
<td>95</td>
<td>27</td>
<td>94</td>
<td>32</td>
<td>267</td>
</tr>
<tr>
<td>Prostate Cancer Prevention Trial risk calculator</td>
<td>95</td>
<td>30</td>
<td>95</td>
<td>33</td>
<td>295</td>
</tr>
<tr>
<td>PHI</td>
<td>95</td>
<td>30</td>
<td>95</td>
<td>33</td>
<td>303</td>
</tr>
<tr>
<td>Derived multiplex 2-gene model</td>
<td>95</td>
<td>17</td>
<td>91</td>
<td>30</td>
<td>168</td>
</tr>
<tr>
<td>Derived multiplex 3-gene model</td>
<td>95</td>
<td>27</td>
<td>93</td>
<td>32</td>
<td>270</td>
</tr>
<tr>
<td>MPS</td>
<td>95</td>
<td>35</td>
<td>95</td>
<td>35</td>
<td>347</td>
</tr>
<tr>
<td>MPS2</td>
<td>95</td>
<td>42</td>
<td>96</td>
<td>37</td>
<td>419</td>
</tr>
<tr>
<td>MPS2+</td>
<td>95</td>
<td>46</td>
<td>99</td>
<td>12</td>
<td>462</td>
</tr>
<tr>
<td>Repeat biopsy (n = 247)</td>
<td>94.4</td>
<td>15</td>
<td>97</td>
<td>8.0</td>
<td>148</td>
</tr>
<tr>
<td>PSA</td>
<td>94.4</td>
<td>21</td>
<td>98</td>
<td>8.6</td>
<td>210</td>
</tr>
<tr>
<td>Prostate Cancer Prevention Trial risk calculator</td>
<td>94.4</td>
<td>8.7</td>
<td>95</td>
<td>7.5</td>
<td>87</td>
</tr>
<tr>
<td>PHI</td>
<td>94.4</td>
<td>14</td>
<td>97</td>
<td>8.0</td>
<td>144</td>
</tr>
<tr>
<td>Derived multiplex 2-gene model</td>
<td>94.4</td>
<td>16</td>
<td>97</td>
<td>8.1</td>
<td>162</td>
</tr>
<tr>
<td>Derived multiplex 3-gene model</td>
<td>94.4</td>
<td>15</td>
<td>97</td>
<td>8.0</td>
<td>148</td>
</tr>
<tr>
<td>MPS</td>
<td>94.4</td>
<td>46</td>
<td>99</td>
<td>12</td>
<td>462</td>
</tr>
<tr>
<td>MPS2</td>
<td>94.4</td>
<td>51</td>
<td>99</td>
<td>13</td>
<td>511</td>
</tr>
</tbody>
</table>

Abbreviations: NPV, negative predictive value; PPV, positive predictive value.
cohort clinically indicated for biomarker testing, we directly compared the new 18-gene test for high-grade PCa with existing guideline-endorsed tests. Broadly, AUC values for MPS2 models were associated with meaningful improvement compared with currently available options. Using a testing approach with 95% sensitivity for high-grade cancer, MPS2 would have meaningfully reduced unnecessary biopsies performed relative to other tests. These data support use of MPS2 to mitigate the potential harms of screening while preserving its long-term benefits.

Patients with a prior negative biopsy pose a unique challenge. Because most patients undergo initial biopsy due to PSA elevation, the value of repeated PSA testing is particularly limited in this population. In one prior study, among 229 patients undergoing repeat biopsies, the EPI test provided 82% sensitivity and 27% specificity for high-grade cancer. Among 268 patients undergoing repeat biopsies, MPS provided 100% sensitivity and 23% specificity. In the current analysis including 247 patients at 94.4% sensitivity, MPS2+ provided 51% specificity, compared with 8.7% with PHI, 14% with the derived multiplex 2-gene model, 16% with the derived multiplex 3-gene model, and 15% with MPS. While striking, these findings are plausible, as most current assays include PSA and PSA isoforms, underscoring a continued dependence on PSA. Second, existing assays measure 3 or fewer non-PSA markers. Given the multiple pathways driving lethal PCa, it is difficult to conceive that most aggressive cancers would overexpress one of so few markers early in the disease course. By capturing 17 cancer-associated, PSA-independent markers, MPS2 provides roughly 5-fold the breadth of previous tests and offers promise of a new generation of biomarkers not reliant on PSA.

The ideal diagnostic test has been described as safe, accurate, available, and actionable and providing a favorable benefit-to-harm ratio. While PSA alone offers favorable practical attributes, its lack of cancer specificity has driven the need for a complementary test to improve screening outcomes. While prebiopsy mpMRI improves detection of high-grade cancer in men with positive findings on mpMRI, data describing the use of negative mpMRI findings to rule out significant cancer merit concern. Findings from a statewide collaborative revealed an NPV of only 77% across diverse settings. Even at experienced centers, subjective MRI interpretation yields significant variability, with NPVs as low as 63% at one center and 40% among individual radiologists. Moreover, MRI bears an extensive time and resource burden, is not widely available in community settings, and is not an option for some patients, posing critical barriers to widespread use. While a valuable component of the diagnostic armamentarium, practical limitations and suboptimal rule-out performance sug-
gest MRI may be best used later in the diagnostic pathway, eg, to improve the yield of biopsy in men most likely to benefit from invasive testing.

The accuracy of MPS2 offers potential for straightforward application at the primary care level (ie, negative test rules out high-grade disease; positive test prompts specialist referral). For specialists, providing patients with early noninvasive molecular tumor data could enable more informed, individualized cancer care. For example, in patients indicated for biopsy, the relationship of tumor subtypes with MRI visibility could help identify patients likely to benefit from pre-biopsy mpMRI and those better served by immediate biopsy.56 In men with PCAs of GG 1, high-grade markers could signal the presence of occult aggressive tumors, while their absence could obviate the need for scheduled surveillance biopsies.57 Finally, while biopsy and tissue-based assays rely on the specific tumor foci sampled,58,59 urine provides a comprehensive assessment of prostatic gene expression—an ideal complement to mitigate the sampling limitations of biopsy.

Limitations

The current study has limitations. For one, there was limited racial diversity in the study population. Thus, it is unclear how our findings could differ in Black men, and we are currently pursuing analyses to ensure optimal testing for all patients. Second, the reference standard was systemic biopsy, which is subject to undersampling that could increase NPV and decrease positive predictive value relative to surgical pathology.60-62 Nonetheless, sampling misclassification would be expected to impact all tests equally, and we uniquely performed head-to-head comparison of MPS2 with existing biomarker tests. Furthermore, we repeated model development in patients with more definitive pathologic data (eg, radical prostatectomy), and prostatectomy-derived MPS2 models did not differ substantially (eTable 1 in Supplement 1). Notably, the current analysis used the Prostate Cancer Prevention Trial risk calculator due to its extensive validation and recognition by clinicians63; other risk calculators could have performed differently.64

We acknowledge the limitations of deriving molecular models developed on other platforms. Although the derived multiplex models capture the components of other commercially available tests, these models should not be interpreted as equivalent to the commercial assays, just as no conclusions can be drawn regarding biomarkers not assessed. Still, external comparison of a newly validated test with guideline-endorsed tests has not previously been performed, to our knowledge, and the 18-gene test would have yielded clinically meaningful improvement in accuracy for high-grade PCa relative to current testing options. While encouraging, these findings do not rule out disparate findings in additional cohorts. Moreover, the 95% sensitivity threshold is a single data point that, while illustrative and clinically applicable, may not be ideal for all populations; decision curves presented herein provide a greater breadth of information regarding utility. Finally, this study population was not suitable for comparing biomarkers with mpMRI, which remains a critical knowledge gap. We are currently conducting a prospective multicenter trial for this assessment.65 Regardless, the externally validated performance of MPS2 supports its effectiveness in accurately ruling out the need for mpMRI and biopsy altogether. Additional studies are needed to corroborate these data and confirm the observed positive impact of MPS2 testing on longer-term outcomes.

Conclusions

In this study, within an external validation population referred for prostate biopsy, an 18-gene urinary test had higher diagnostic accuracy for high-grade PCa beyond currently available testing options. Clinically, use of this test would have safely avoided unnecessary additional testing with imaging or biopsy in 35% to 51% of patients while maintaining high sensitivity for high-grade cancers that stand to benefit from early detection. These findings suggest that use of the test in patients with elevated PSA levels can reduce the potential harms of prostate cancer screening while preserving its long-term benefits.
Supervision: Tosoian, Morgan, Palapattu, Srivastava, Feng, Y. Zheng, Chinnaiyan.

Conflict of Interest Disclosures: Dr Tosoian reported personal fees from Lynx Dx and equity interest from Lynx Dx outside the submitted work; and has a patent for a novel multiplex urine test for high-grade prostate cancer pending. Dr Zhang reported personal fees from Lynx Dx outside the submitted work and has a patent for a novel multiplex urine test for high-grade prostate cancer pending. Dr Xiao reported grants from Prostate Cancer Foundation as well as personal fees from Lynx Dx during the conduct of the study; and has a patent for a novel multiplex urine test for high-grade prostate cancer pending. Dr Nilakans reported personal fees from Lynx Dx during the conduct of the study; and personal fees from Lynx Dx outside the submitted work; and has a patent for use of some biomarkers as diagnostic tools issued. Dr Trock reported personal fees from Artera during the conduct of the study as well as personal fees from Myriad Genetics outside the submitted work. Dr Salami reported personal fees from Bayer and Novartis during the conduct of the study. Dr Tomlins reported grants and personal fees from Astellas as well as equity interest from Strata and NRich during the conduct of the study. Dr Feng reported grants and personal fees from AstraZeneca as well as pending. Dr Zhang reported grants from Prostate Cancer Foundation as well as personal fees from Limpixivix during the conduct of the study. Dr Minner reported grants and personal fees from Limpixivix as well as pending. Dr Feng reported grants and personal fees from Limpixivix.

Additional Contributions: We thank Stephanie Miner, PhD (Michigan Center for Translational Pathology, University of Michigan, Ann Arbor), for her help in the editing and file preparation of this study. Dr Minner was not compensated for her work.

REFERENCES


Development and Validation of an 18-Gene Urine Test for High-Grade Prostate Cancer


