Multicenter Optimization and Validation of a 2-Gene mRNA Urine Test for Detection of Clinically Significant Prostate Cancer before Initial Prostate Biopsy

Alexander Haese, Geert Trooksees,* Sandra Steyaert, Daphne Hessels,* Michael Brawer,* Virginie Vlaeminck-Guillen, Alain Ruffion, Derya Tilki, Jack Schalken,* Jack Groskopf* and Wim Van Criekinge*†

From the Martini-Klinik Prostate Cancer Center (AHI) and Department of Urology, University Hospital Hamburg-Eppendorf (UDE), Hamburg, Germany; MDxHealth, Inc., 750 Long Ridge Road, Norwalk, Connecticut 06850, USA; Division of Urology, Vanderbilt University, Nashville, Tennessee, USA; iGT, SS, DH, MB, JG, Irvine, California, Medical Unit of Molecular Oncology and Transfer, Department of Biochemistry and Molecular Biology (IV-G) and Urology Department, University Lyon Sud Hospital; Pierre Bénite (AR) and Cancer Research Centre of Lyon, U1052 INSERM, CNRS 5286, Claude Bernard University Lyon 1, Lyon Bichat Centre (V-G, AHI), Lyon, France; Department of Urology, Radboud University Medical Centre (US), Nijmegen, The Netherlands, and Department of Bioinformatics, Ghent University (VHC), Ghent, Belgium.

Abbreviations and Acronyms

DCA = decision curve analysis
DRE = digital rectal examination
GI = International Society of Urological Pathology Grade Group
mpMRI = multiparametric magnetic resonance imaging
NVP = negative predictive value
PCA = prostate cancer
PCPTRC = Prostate Cancer Prevention Trial Risk Calculator; version 2.0
PSA = prostate specific antigen
PV = prostate volume

Purpose: A 2-gene, urine based molecular test that combines mRNA biomarkers with clinical factors can risk stratify patients for clinically significant prostate cancer. To ensure the generalizability of assay results we optimized and validated the clinical model for men with serum prostate specific antigen less than 10 ng/ml who were undergoing initial prostate biopsy.

Materials and Methods: Urine samples were collected from 1,955 men from The Netherlands, France and Germany prior to an initial prostate biopsy and study subjects were divided into training and validation cohorts. Urinary HOXC6 and DLX1 mRNA levels were quantified and RNA results were then combined with other risk factors in a clinical model optimized to detect ISUP (International Society of Urological Pathology) Grade Group 2 or greater prostate cancer in men with prostate specific antigen less than 10 ng/ml. Results in the validation cohort were compared with the PCPTRC (Prostate Cancer Prevention Trial Risk Calculator), version 2.0.

Results: The optimal clinical model included urinary HOXC6 and DLX1 mRNA levels, patient age, digital rectal examination and prostate specific antigen density (serum prostate specific antigen/prostate volume). In the 715 validation cohort subjects with prostate specific antigen less than 10 ng/ml the AUC was 0.82 with 89% sensitivity, 53% specificity and 95% negative predictive value. The PCPTRC AUC was 0.70. The full validation cohort of 916 men including all prostate specific antigen levels yielded an AUC of 0.85 with 93% sensitivity, 47% specificity and 95% negative predictive value. The PCPTRC AUC was 0.76.

Conclusions: The 2-gene based urine assay, which is optimized for biopsy naive patients with serum prostate specific antigen less than 10 ng/ml, demonstrated high sensitivity and negative predictive value to detect clinically significant prostate cancer. These data support using the test to help guide initial prostate biopsy decisions.

Key Words: prostatic neoplasms; RNA, messenger; prostate specific antigen; biopsy; risk
There is a persistent unmet diagnostic need for better methods to identify clinically significant PCa. Despite substantial evidence of the survival benefit associated with the treatment of intermediate and high risk, early stage disease, the low specificity of PSA screening has led to excess prostate biopsies and over detection of lower risk cancers, treatment of which lacks proven benefit. Prostate biopsy procedures and interventionl cancer treatment may carry significant risks, especially in patients with comorbidities. PCa specific biomarkers used in conjunction with PSA and other clinical risk factors can improve the detection of PCa with specific aggressive features such as Gleason score 7 (GG2) or higher grade, thereby helping to maximize the cancer survival benefit while also avoiding prostate biopsies which would detect indolent or no PCa.

A 2-gene urine based molecular test targeting mRNA known to be overexpressed in aggressive PCa, the cell proliferation gene HOXC6 and the progression gene DLX1, has been validated for the detection of GG2 and higher PCa. Urinary HOXC6 and DLX1 mRNA levels are measured following DRE and the RNA results are combined with clinical risk factors to determine the individualized patient risk of GG2 or higher PCa histopathology at a subsequent biopsy. The 2-gene test was clinically validated for the detection of GG2 or greater PCa in a prospective multicenter cohort of 386 subjects, including 342 (89%) undergoing initial biopsy and 44 (11%) with prior PCa negative biopsies, and the test had 98% NPV. Analytical validation studies have demonstrated the robustness and reproducibility of the 2-gene urine test results. Risk scores have also been associated with the mpMRI outcome, suggesting a potential role for identifying the patients most likely to benefit from mpMRI or for risk stratification following indeterminate mpMRI results. The cost-effectiveness of the 2-gene test in men with elevated PSA was assessed in European and American modeling studies. In each study using the 2-gene test decreased prostate biopsy with its associated over diagnosis and/or overtreatment compared to the current standard of care, which resulted in cost savings and increased quality adjusted life years per patient. In addition, the impact on initial biopsy decision making was demonstrated in a multicenter study of community urology practices in the United States. The initial biopsy rate was fivefold higher in test positive vs negative patients.

In this multicenter study we optimized the 2-gene urinary biomarker based clinical model in men with PSA levels less than 10 ng/ml who were undergoing initial prostate biopsy. We then validated its clinical performance in all subjects regardless of serum PSA level.

Materials and Methods

Study Population
The study population consisted of 1,955 sequentially enrolled men, including 750 from Radboud University Medical Centre, Nijmegen, The Netherlands; 762 from Lyon Sud Hospital, Lyon, France; and 443 from the Martini-Klinik Prostate Cancer Center, Hamburg, Germany. These men underwent an initial prostate biopsy for suspected PCa between December 2007 and December 2014. All subjects underwent extended 10 to 12-core transrectal ultrasound guided biopsy. PV was determined by ultrasound at biopsy. In 295 of the 1,955 subjects mpMRI was done prior to biopsy to help target supplemental cores during the transrectal ultrasound guided procedure. None of these men underwent mpMRI fusion biopsy. Subjects were divided into 1,039 in the training cohort and 916 in the validation cohort. The Dutch training and validation cohorts were previously described. The French and German cohorts were evenly split by randomization. Table 1 shows a summary of study population demographic and clinical characteristics in the 805 and 715 men in the training and validation cohorts, respectively, with PSA less than 10 ng/ml as well as the entire cohort.

Study exclusion criteria were a prior diagnosis of PCa, medical therapy known to affect serum PSA levels within 6 months prior to urine sample collection, such as finasteride (Proscar® or Propecia®), dutasteride (Avodart®), antiandrogen therapies (Lupron® or Eligard®), and invasive treatment of BPH within 6 months of urine sample collection.

Sample Collection and Processing
Prior to biopsy first catch urine was collected after a standardized DRE consisting of 3 sweeps per prostatic lobe. The urine was mixed with a RNA preservative (Hologic®) and stored at −70°C. Urine samples were tested elsewhere with 2-gene reverse transcriptase-polymerase chain reaction. Laboratory technicians were blinded to biopsy outcomes.

Briefly, HOXC6 and DLX1 mRNA were quantified using reverse transcriptase-polymerase chain reaction and then normalized for prostate RNA recovery using urinary KLK3 mRNA levels. KLK3 is the gene that encodes for PSA. The normalized HOXC6 and DLX1 levels were combined into a single RNA value, which was then used with other risk factors in a clinical model, resulting in a continuous risk score of −6 to 6. To report results the risk score was then converted to the percent likelihood that subsequent biopsy would identify GG2 or greater PCa.

Statistical Analysis
The clinical model was designed to discriminate GG2 or higher PCa from GG1 or no PCa at biopsy. Model variables included the urinary RNA value, patient age, PSA density calculated as serum PSA/prostate volume and the DRE result (normal or suspicious). Although the AUC often serves as a metric for clinical performance, only a small part of the ROC curve is relevant in the initial prostate biopsy setting, in which the patient risk
Table 1. Subject demographic and clinical characteristics in training and validation cohorts

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Validation</th>
<th>All Subjects</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. subjects</td>
<td>805</td>
<td>715</td>
<td>1,039</td>
<td>916</td>
<td></td>
</tr>
<tr>
<td>Median age (IQR)</td>
<td>64 (59–68)</td>
<td>64 (59–69)</td>
<td>64 (59–69)</td>
<td>65 (60–70)</td>
<td></td>
</tr>
<tr>
<td>Median ng/ml serum PSA (IQR)</td>
<td>5.5 (4.1–9.9)</td>
<td>5.4 (4.1–7.2)</td>
<td>6.2 (4.6–9.4)</td>
<td>5.4 (4.5–9.2)</td>
<td></td>
</tr>
<tr>
<td>No. DRE result (%)</td>
<td>Normal</td>
<td>635 (79)</td>
<td>585 (82)</td>
<td>759 (73)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Suspicious</td>
<td>168 (21)</td>
<td>124 (17)</td>
<td>280 (27)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not available</td>
<td>1 (0.1)</td>
<td>6 (1)</td>
<td>1 (0.1)</td>
<td></td>
</tr>
<tr>
<td>No. PCa family history (%)</td>
<td>Yes</td>
<td>128 (16)</td>
<td>110 (15)</td>
<td>149 (14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No available</td>
<td>648 (80)</td>
<td>499 (70)</td>
<td>845 (81)</td>
<td></td>
</tr>
<tr>
<td>Biopsy outcome 1%</td>
<td>No PCa</td>
<td>455 (56.5)</td>
<td>388 (54.3)</td>
<td>518 (49.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG1</td>
<td>195 (24.3)</td>
<td>175 (24.5)</td>
<td>239 (23.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG2</td>
<td>99 (12.3)</td>
<td>97 (13.6)</td>
<td>145 (14.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG3</td>
<td>26 (3.2)</td>
<td>27 (3.8)</td>
<td>54 (5.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG4</td>
<td>15 (1.9)</td>
<td>17 (2.4)</td>
<td>38 (3.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG5</td>
<td>14 (1.7)</td>
<td>11 (1.5)</td>
<td>45 (4.3)</td>
<td></td>
</tr>
</tbody>
</table>

associated with false-negative results is higher than that of false-positive findings. Therefore, we defined the cost function to optimize the model as the average false-positive rate (1 – specificity) of the ROC curve interval between 93% and 96% sensitivity. The coefficients of the 4 variables were exhaustively determined by minimizing the cost function. To allow for cases in which PV was not available we generated an alternate model to maintain similar sensitivity and NPV by allowing variable specificity. In the absence of PV PSA density was replaced by serum PSA as a stand-alone variable.

Statistical analyses were performed with R, version 3.4.0 (R Foundation for Statistical Computing, Vienna, Austria). The clinical performance of RNA biomarkers, clinical variables, and models was assessed based on the ROC AUC. The 95% CI and AUC comparisons were determined by the DeLong method as implemented in the R package pROC. The Mann-Whitney test was applied to assess differences in quantitative measurements. The PCPT/RC was used for comparison. DCA was performed in R as previously described. Calibration curves were generated in R using the val.prob function.

RESULTS

Combining Risk Factors to Detect High Grade Prostate Cancer

Clinical models were developed using 805 training cohort subjects with serum PSA less than 10 ng/ml (table 1). This yielded an AUC of 0.80 (95% CI 0.76–0.84). At the optimal risk score cutoff of −2.8 sensitivity was 93%, specificity was 52% and NPV was 96%. Figure 1 shows the relationship between the risk score and the likelihood of GG2 or higher PCa, GG1 PCa and no PCa at biopsy using data on the full sample set of 1,955 primary biopsies. Higher risk scores were associated with an increased likelihood of GG2-5 PCa at biopsy, providing an individualized risk assessment to assist with biopsy decision making.

Clinical Validation

In the training cohort of 805 men with PSA less than 10 ng/ml the AUC was 0.80 (95% CI 0.76–0.84). At the optimal risk score cutoff of −2.8 sensitivity was 93%, specificity was 52% and NPV was 96%. The optimized clinical model was applied to the 715 validation cohort subjects with serum PSA less than 10 ng/ml. This yielded an AUC of 0.82 (95% CI 0.79–0.86), 89% sensitivity, 53% specificity and 95% NPV. When PV was omitted from the model, sensitivity was 87%, specificity was 38% and NPV was 92%.

Because many clinicians use suspicious DRE results as a stand-alone indication for biopsy, we also evaluated performance in subjects with normal DRE results. Table 2 summarizes performance

Figure 1. Percent probability of GG2 or greater (green curve), no PCa (blue curve) and GG1 PCa (yellow curve) at biopsy in relation to risk score. Vertical line at risk score −2.8 represents cutoff point of overall positive vs negative result.
Table 2. Optimized clinical model performance characteristics in validation cohort stratified by serum PSA and digital rectal examination results

<table>
<thead>
<tr>
<th></th>
<th>PSA Less Than 10 ng/ml</th>
<th>All PSA Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All DRE Results</td>
<td>Normal DRE Results</td>
</tr>
<tr>
<td>No. subjects</td>
<td>715</td>
<td>591</td>
</tr>
<tr>
<td>ROC AUC (95% CI)</td>
<td>0.82 (0.79–0.86)</td>
<td>0.83 (0.78–0.85)</td>
</tr>
<tr>
<td>% Sensitivity</td>
<td>89</td>
<td>84</td>
</tr>
<tr>
<td>% Specificity</td>
<td>53</td>
<td>57</td>
</tr>
<tr>
<td>% Pos predictive value</td>
<td>34</td>
<td>29</td>
</tr>
</tbody>
</table>

Negative predictive value remained constant at 95% across all patient subgroups.

Characteristics in the validation cohort, stratified by PSA (PSA less than 10 ng/ml or all values) and the DRE result. For all conditions the NPV was 95%. The optimized model was also applied to the validation cohort from the previous clinical validation study. The AUC was 0.88 (0.84–0.92), sensitivity was 98%, specificity was 42% and NPV was 98%.

Figure 2 shows ROC curves of the optimized clinical model and the PCPTRC in the subject groups. In validation cohort subjects with PSA less than 10 ng/ml the AUC was 0.82 (95% CI 0.79–0.86) compared to the PCPTRC AUC of 0.70 (95% CI 0.65–0.74, p < 0.001). In the entire validation cohort regardless of PSA level the optimized clinical model AUC was 0.85 (95% CI 0.83–0.88), with 93% sensitivity, 47% specificity and 95% NPV while the PCPTRC AUC was 0.76 (95% CI 0.72–0.80, p < 0.001).

To further characterize the optimized model we analyzed the incremental benefit of the urinary biomarkers relative to clinical risk factors and assessed the model calibration. Figure 3 shows urinary HOXC6 and DLX1 mRNA levels normalized to KLK3 as the prostate specific control in validation cohort subjects with no PCa, GG1 PCa, or GG2 or greater PCa at biopsy. For each mRNA the relative level was significantly higher in men diagnosed with GG2 or greater PCa vs GG1 PCa (p < 0.0001).

DCA was performed using the optimized model with and without HOXC6 and DLX1 mRNA measurements (fig. 4). The DCA showed a net benefit of the full model compared to the no RNA model to detect GG2 or higher PCa while avoiding excess biopsies, particularly at lower risk thresholds. Results were similar in men with PSA less than 10 ng/ml and in the full validation cohort (fig. 4).

Figure 5 shows calibration plots of the risk score and PCPTRC in the full validation cohort of 916 subjects. The optimized model showed good

---

![Figure 2](image)

**Figure 2.** ROC curves of optimized clinical model (red curves) and PCPTRC (blue curves) in 715 validation cohort subjects with PSA less than 10 ng/ml (A) and full validation cohort of 916 with all PSA levels (B).
calibration with a slight trend toward overestimating risk at probabilities below approximately 25% and underestimating risk at higher probabilities (fig. 5, A). As a point of reference the risk score threshold used to determine overall positive vs negative assay results and determine performance characteristics (table 2) corresponded to a probability of 12%. In contrast, the PCPTRC consistently underestimated risk in the validation cohort (fig. 5, B).

To assess the potential impact of the optimized model on patient treatment the validation cohort of 715 men with PSA less than 10 ng/ml was evaluated when assuming that assay results were the sole determinant for the biopsy decision on study. Based on the patients with negative test results 314 of all 715 biopsies (44%) and 297 of 563 biopsies (53%) which identified GG1 or no PCa would have been avoided or delayed. In patients with negative test results who were found to have PCa at biopsy 61 of 175 (35%) with GG1 PCa, 12 of 97 (12%) with GG2, 4 of 27 (15%) with GG3, 1 of 17 (5.9%) with GG4 and 0 of 11 (0%) with GG5 may have had a delayed diagnosis. In the overall validation cohort 17 of 715 men (2.4%) had negative test results and were diagnosed with GG2 or greater PCa at biopsy.

Figure 3. Box plots of urinary HOXC6 (A) and DLX1 (B) mRNA in subjects with no PCa, low grade PCa or clinically significant PCa at biopsy. Horizontal line indicates median. x indicates average.

Figure 4. Decision curve analysis of full clinical model (red curves) vs clinical model without RNA markers (blue curves) in validation cohort subjects with PSA less than 10 ng/mL (A) and in full validation cohort (B). Gray curve: all subjects have GG2 or higher PCa, horizontal line: no subjects have GG2 or higher PCa.
**DISCUSSION**

Accurately identifying clinically significant PCa remains one of the key unmet needs in the diagnosis and management of this pervasive disease. An approach to this dilemma is to use PCa specific biomarkers which can be combined with clinical risk factors to improve the detection of GG2 and higher PCa. We previously described the development and validation of the 2-gene urinary biomarker based clinical model to discriminate GG2 and higher PCa from GG1 and benign disease in men undergoing prostate biopsy. In this study we confirmed and expanded these findings in a larger, multinational population of men who were undergoing an initial prostate biopsy with an emphasis on patients with PSA less than 10 ng/ml.

The clinical model was maintained to maintain high sensitivity and NPV to detect GG2 or greater PCa in the initial biopsy setting. In the validation cohort of 715 subjects with PSA less than 10 ng/ml the AUC was 0.82, sensitivity was 89%, specificity was 53% and NPV was 95%. If the 2-gene test had been used for the initial biopsy decision, 53% of excess biopsies (ie biopsies which would have identified GG1 or no cancer) could have been avoided and a diagnosis of clinically significant PCa may have been delayed in 5% of patients.

Because PV is not always available in biopsy naive men, the model was designed to compensate and yielded a NPV of 92% when this parameter was not included. In all 916 validation cohort subjects with any PSA level sensitivity was 93%, specificity was 47% and NPV was 95%. The NPV of 95% was maintained when analysis was limited to men with normal DRE results, who therefore might be more likely to benefit from additional testing. Importantly, when the optimized model was applied to the validation cohort in the initial validation study, results were equivalent to previously reported findings with 98% NPV. Taken together these results demonstrate the robustness of the clinical model across different patient cohorts.

Characterization of the optimized model confirmed that HOXC6 and DLX1 mRNA measurements provide a significant improvement in test performance compared to clinical risk factors alone. Urinary levels of the individual mRNAs were significantly associated with clinically significant PCa, consistent with previous reports. DCA demonstrated net benefit of the RNA biomarkers, particularly at lower risk thresholds where the test may be most useful in clinical practice. Finally, the optimized clinical model was found to be well-calibrated for predicting the probability of detecting GG2 or higher PCa at biopsy.

Several other options are available to clinicians to help guide initial biopsy decisions. Increasingly mpMRI is performed for PCa detection. However, a recent meta-analysis highlighted mpMRI variability with 64% to 88% NPVs, lower than in our report. In addition, the AUA (American Urological Association) currently recommends mpMRI only in the repeat biopsy setting.

Blood based PSA derivative assays and other urinary markers have shown potential to improve the prediction of the initial biopsy outcome. The
results of this clinical validation compare favorably with those of published studies using these other methodologies.11

The strengths of this multicenter validation study include the relatively large training and validation cohorts, and the assessment of clinical performance in all patients with an initial biopsy using a model optimized for men with PSA less than 10 ng/ml. Limitations include the fact that the study was retrospective and there was a lack of racial/ethnic background information on the study population, which was presumed to be primarily Caucasian. This likely does not reflect the diversity of potential patients who undergo an initial biopsy in the United States.

CONCLUSIONS
The 2-gene urine biomarker based assay was optimized for biopsy naive patients with serum PSA less than 10 ng/ml and clinically validated in a multicenter study. The 2-gene test demonstrated high sensitivity and NPV to detect CG2 or greater PCa with the potential to improve the detection of clinically significant disease by avoiding approximately half of excess prostate biopsies. These data support the use of this test to help guide initial prostate biopsy decisions.

ACKNOWLEDGMENT
Reverse transcriptase-polymerase chain reaction was performed at MDxHealth, Nijmegen, The Netherlands.

REFERENCES


EDITORIAL COMMENT

In the United States a shotgun approach to screening and treatment led to a 52% reduction in age adjusted prostate cancer specific mortality but this was tempered by the pandemic of men who never needed to be diagnosed, let alone treated. This mRNA urine test (SelectMDx®) is scientifically sound, now has 2 validation studies and is capable of decreasing the biopsy rate by 40% to 50% while delaying the diagnosis in 2% to 5% of men known to have GG2 or greater (Gleason 7 or greater cancer). Interestingly, similar to other biomarkers, no consistent relationship was observed between the biomarker risk score and GG1 (Gleason 6) (fig. 1 in article), further evidence corroborating its inert nature.

About 10 biomarkers are available or in development with similar intent, of which most are costly. It is a busy clinical space but collectively a much welcomed advance. At a bare minimum when urologists screen for prostate cancer, they should capitalize on the powerful information embedded in inexpensive tests such as total PSA, percent free PSA, PSA density and digital rectal examination. Use of secondary markers (magnetic resonance imaging, or urine, blood or tissue based markers) is data based and sensible but pricey. Reflex use in all men is mindless and should be discouraged. When used appropriately, added value is provided, particularly when scores are low or high.

Scott Eggener
Departments of Surgery and Radiology
University of Chicago Medicine
Chicago, Illinois

Disclaimer: Opinions expressed by the authors and advertisers are not necessarily those of the American Urological Association, its affiliates, or of the Publisher. The American Urological Association, its affiliates, and the Publisher disclaim any liability to any party for the accuracy, completeness, efficacy, or availability of the material contained in this publication (including drug dosages) or for any damages arising out of the use or non-use of any of the material contained in this publication.