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

European Association of Urology

Platinum Priority – Prostate Cancer

Editorial by Robert A. Gardiner, Paul Mainwaring and Martin F. Lavin on pp. 749–750 of this issue

Detection of High-grade Prostate Cancer Using a Urinary Molecular Biomarker–Based Risk Score

Leander Van Neste [a,1], Rianne J. Hendriks [b,1], Siebren Dijkstra [b,1], Geert Trooskens [c], Erik B. Cornel [d], Sander A. Jannink [c], Hans de Jong [c], Daphne Hessels [c], Frank P. Smit [c], Willem J.G. Melchers [e], Gisèle H.J.M. Leyten [b,f], Theo M. de Reijke [f], Henk Vergunst [g], Paul Kil [h], Ben C. Knipscheer [i], Christina A. Hulsbergen-van de Kaa [j], Peter F.A. Mulders [b], Inge M. van Oort [b], Wim Van Criekinge [k], Jack A. Schalken [b,]*

[a] *Department of Pathology, Maastricht University Medical Center, Maastricht, The Netherlands;* [b] *Department of Urology, Radboud University Medical Center, Nijmegen, The Netherlands;* [c] *MDxHealth BV, Nijmegen, The Netherlands;* [d] *Department of Urology, ZGT Hospital, Hengelo, The Netherlands;* [e] *Department of Medical Microbiology, Radboud University Medical Center, Nijmegen, The Netherlands;* [f] *Department of Urology, AMC University Medical Centre, Amsterdam, The Netherlands;* [g] *Department of Urology, CWZ Hospital, Nijmegen, The Netherlands;* [h] *Department of Urology, St. Elisabeth Hospital, Tilburg, The Netherlands;* [i] *Department of Urology, Scheper Hospital, Emmen, The Netherlands;* [j] *Department of Pathology, Radboud University Medical Center, Nijmegen, The Netherlands;* [k] *Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Ghent, Belgium*

Article info

Article history:
Accepted April 7, 2016

Associate Editor:
Stephen Boorjian

Keywords:
Prostate neoplasms
Biologic markers
Urine
PSA
Messenger RNA
Logistic models

Abstract

Background: To reduce overdiagnosis and overtreatment, a test is urgently needed to detect clinically significant prostate cancer (PCa).
Objective: To develop a multimodal model, incorporating previously identified messenger RNA (mRNA) biomarkers and traditional risk factors that could be used to identify patients with high-grade PCa (Gleason score $\geq$7) on prostate biopsy.
Design, setting, and participants: In two prospective multicenter studies, urine was collected for mRNA profiling after digital rectal examination (DRE) and prior to prostate biopsy. The multimodal risk score was developed on a first cohort (n = 519) and subsequently validated clinically in an independent cohort (n = 386).
Outcome measurements and statistical analysis: The mRNA levels were measured using reverse transcription quantitative polymerase chain reaction. Logistic regression was used to model patient risk and combine risk factors. Models were compared using the area under the curve (AUC) of the receiver operating characteristic, and clinical utility was evaluated with a decision curve analysis (DCA).
Results and limitations: HOXC6 and DLX1 mRNA levels were shown to be good predictors for the detection of high-grade PCa. The multimodal approach reached an overall AUC of 0.90 (95% confidence interval [CI], 0.85–0.95) in the validation cohort (AUC 0.86 in the training cohort), with the mRNA signature, prostate-specific antigen (PSA) density, and previous cancer-negative prostate biopsies as the strongest, most significant components, in addition to nonsignificant model contributions of PSA, age, and family history. For another model, which included DRE as an additional risk factor, an AUC of 0.86 (95% CI, 0.80–0.92) was obtained (AUC 0.90 in the training cohort). Both models were successfully validated, with no significant change in AUC in the validation cohort, and DCA indicated a strong net benefit and the best reduction in unnecessary

[1] These authors contributed equally to this work.
* Corresponding author. Department of Urology, Radboud University Medical Center, Geert-Grooteplein Zuid 10, PO Box 9101, 6500 HB Nijmegen, The Netherlands. Tel. +31 24 36 14 146; Fax: +31 24 35 41 222.
E-mail address: jack.schalken@radboudumc.nl (J.A. Schalken).

http://dx.doi.org/10.1016/j.eururo.2016.04.012

biopsies compared with other clinical decision-making tools, such as the Prostate Cancer Prevention Trial risk calculator and the PCA3 assay.
Conclusions: The risk score based on the mRNA liquid biopsy assay combined with traditional clinical risk factors identified men at risk of harboring high-grade PCa and resulted in a better patient risk stratification compared with current methods in clinical practice. Therefore, the risk score could reduce the number of unnecessary prostate biopsies.
Patient summary: This study evaluated a novel urine-based assay that could be used as a noninvasive diagnostic aid for high-grade prostate cancer (PCa). When results of this assay are combined with traditional clinical risk factors, risk stratification for high-grade PCa and biopsy decision making are improved.

1. Introduction

Prostate cancer (PCa) is the second most frequently diagnosed cancer among men worldwide, with an estimated 1.1 million new cases and 307 500 deaths in 2012 [1]. With the introduction of serum prostate-specific antigen (PSA) testing in the 1990s, the incidence of PCa has increased. PSA testing has also led to an increased number of unnecessary biopsies and the diagnosis of clinically insignificant tumors that would not have been life threatening (ie, potential overtreatment). This is particularly the case in a PSA gray zone <10.0 ng/ml, at which 65–70% of men have a negative biopsy result [2]. Men with indolent disease who undergo treatment may experience complications without reducing their risk of dying from PCa [3]. Albertsen et al showed that men with Gleason score (GS) 8–10 PCa have a relatively high probability of dying from PCa within 10 yr (12.1%), whereas this risk is minimal for men with low-grade disease [4].

The major challenge is to improve the detection of clinically significant or high-grade PCa in an early stage. Both overdiagnosis and overtreatment could be reduced if PCa-specific biomarkers could accurately distinguish indolent from aggressive tumors. Ideally the biomarkers could be measured in a sample that could be obtained noninvasively (eg, in urine). The urinary test based on the prostate cancer antigen 3 (*PCA3*) gene (Progensa PCA3; Hologic Inc, Marlborough, MA, USA) is the only molecular diagnostic test approved by the US Food and Drug Administration for the detection of PCa in urine [5,6]. *PCA3* was identified as a gene encoding a long noncoding RNA that is consistently upregulated in PCa [7,8]. PCA3 was shown to be of value in PCa detection; however, the relation with tumor aggressiveness and thus prognostic value remains controversial [9–11]. New biomarkers for PCa detection are the blood-based Prostate Health Index (PHI) and the four-kallikrein panel [12–17]. Studies describing head-to-head comparison of these markers showed that PHI outperforms PCA3 in the prediction of significant PCa [14].

Previous studies have shown the potential of noninvasive urinary biomarkers to accurately predict the presence of high-grade disease and thus aid in decision making regarding further diagnostic evaluations (eg, prostate biopsies or imaging) and treatment while avoiding unnecessary biopsies. Leyten et al described a stepwise approach for the identification and selection of new biomarkers using messenger RNA (mRNA) expression profiling [18]. A panel measured in urinary sediments predicted a GS ≥7 on prostate biopsy. The test, based on detecting increased mRNA levels of homeobox C6 (HOXC6), distal-less homeobox 1 (DLX1), and tudor domain containing 1 (TDRD1), was shown to have independent additional value to PSA for predicting high-grade PCa on biopsy. HOXC6, DLX1, and TDRD1 may be involved in the onset of PCa and are associated with high-grade PCa [18]. In this study, homeobox C4 (HOXC4) was also included because it was shown to be overexpressed in urine sediments and is transcribed from the same transcription unit as HOXC6 [18].

The aim of this study was to validate the gene panel–based mRNA test performed on whole urine and to develop a model combining molecular profiling with traditional clinical risk factors that could be used to identify patients accurately with high-grade PCa (GS ≥7) on prostate biopsy. To optimize patient management and clinical utility, all relevant information to make the most accurate assessment for each patient should be taken into account. Multimodal risk assessment approaches have been developed that combine multiple information sources into an overall optimal risk prediction for each individual patient. One such score is the Prostate Cancer Prevention Trial risk calculator (PCPTRC), which combines PSA with digital rectal examination (DRE), race, family history, age, and whether a patient had a previous biopsy [19]. Patient management and risk assessment benefit from combining different complementary information sources into one coherent risk score because no single marker can obtain a similar performance on its own [19,20]. The optimal diagnostic model was validated in urine samples from a second, independent cohort to ensure robustness of the proposed risk score.

2. Materials and methods

2.1. Study population

In two prospective multicenter studies, men who were scheduled for (initial or repeat) prostate biopsies, based on elevated PSA levels (≥3 ng/ml), abnormal DRE, or a family history of PCa, were consecutively included. Urine samples were collected after a standardized DRE consisting of three strokes per lobe [5]. Subjects were enrolled from six urology clinics in the Netherlands (Radboud University Medical Center Nijmegen, ZGT Hospital Hengelo, AMC University Medical Centre Amsterdam, CWZ Hospital

Nijmegen, St. Elisabeth Hospital Tilburg, and Scheper Hospital Emmen) between September 2009 and July 2011 (clinical trial A) and between July 2011 and September 2014 (clinical trial B). Exclusion criteria were a history of PCa, medical therapy known to affect PSA levels, prostate biopsy within 3 mo prior to enrollment, and invasive treatment for benign prostate hyperplasia within 6 mo prior to enrollment. Transrectal ultrasound (TRUS)–guided prostate biopsy was performed, with a median of 10 cores (interquartile range: 10–10) per patient, and evaluated per each hospital's standard procedure and by local pathologists.

The institutional review boards of all of the hospitals approved the study protocols, and written informed consent was obtained from each participant. Test results were not provided to the clinical sites for patient care, and the laboratory technicians who performed the biomarker tests were blinded for patient characteristics. The developmental study and the validation study were both performed in accordance with the Standards for Reporting of Diagnostic Accuracy criteria [21].

2.2. Sample collection and processing

Approximately 30 ml of first voided urine was collected in a collection cup after DRE. Urine was immediately transferred into a urine specimen transport tube (Hologic Inc), and samples were shipped at room temperature to a central laboratory and stored within at −80 °C.

2.3. Laboratory-developed test development

In the discovery and initial validation study, urinary sediments were used as described by Leyten et al [18]. Fixed whole urine was used as substrate to further optimize and standardize the assay. Assays were performed using a prototype amplification kit (Labo Biomedical Products BV, Rijswijk, The Netherlands) and are described in detail in Supplement 1. In short, RNA was isolated out of 1 ml urine using the MagNA Pure 96 instrument (Roche Life Science, Indianapolis, IN, USA). Subsequently, RNA levels of HOXC4, HOXC6, TDRD1, DLX1, KLK3, and PCA3 were determined using one-step reverse transcription quantitative polymerase chain reaction. The *KLK3* gene, encoding for PSA, is a kallikrein serine protease and used as a reference for relative biomarker quantitation using the ΔΔCt method [22].

2.4. Statistical analysis

Statistical analyses were performed with SPSS v.20.0 (IBM Corp., Armonk, NY, USA) and R v.3.2.1 (R Foundation for Statistical Computing, Vienna, Austria). For comparison of continuous variables, the Welch *t* test was used or the Mann-Whitney-Wilcoxon test as a nonparametric alternative. A binomial or Fisher exact test was applied to compare proportions. Because mRNA levels of these biomarkers are continuously increasing with patient risk, their performance was assessed and evaluated as area under the curve (AUC) of the receiver operating characteristic. The 95% confidence intervals (CIs) and comparisons of AUCs were determined using DeLong's method as implemented in the R package pROC [23]. The combination and predictive value of multiple risk factors was modeled by logistic regression analysis, resulting in a continuous risk score that can also be evaluated with the AUC method (Supplement 1). The main models aimed to identify high-grade (GS ≥7) PCa, using low-grade (GS ≤6), likely insignificant, cancer and men with a PCa-negative diagnosis as the control group. The logistic regression model was then applied to estimate the probability of detecting no, low-grade, or high-grade PCa on biopsy. Clinical utility was assessed with decision curve analysis (DCA) in R [24].

3. Results

3.1. Patient characteristics

A total of 905 urine samples were collected in two independent prospective clinical trials (cohort A: $n = 519$; cohort B: $n = 386$). Table 1 summarizes the patient characteristics. In cohort A, 212 of 519 men (40.8%) had a positive biopsy outcome, of which 109 men (51.4%) had high-grade (GS ≥7) PCa, compared with 181 of 386 men (46.9%) and 90 men (50.0%), respectively, in cohort B. More men had undergone at least one prior biopsy, and more men had an abnormal DRE outcome in cohort A. The other baseline characteristics showed no statistically significant differences between both cohorts.

Table 1 – Patient characteristics

Characteristics	Cohort A	Cohort B	*p* value
Patients, *n*	519	386	–
Evaluable samples, *n* (%)[a]	492 (94.8)	371 (96.1)	0.4
Age, yr, mean (median; IQR)	64.7 (65; 60–70)	64.9 (65; 60–70)	0.6
PSA, ng/ml, mean (median; IQR)	15.8 (7.4; 5.5–11.1)	11.9 (7.3; 5.2–10.9)	0.3[a]
Family history of PCa, %; no, yes, NA	71.9, 17.5, 10.6	32.4, 19.1, 48.4	<0.001[b]
First biopsy, *n* (%)	410 (79)	342 (89)	<0.001
TRUS prostate volume, ml, mean (median; IQR)	55 (48; 35–67)	51 (45; 35–62)	0.079
PSAD, ng/ml · ml, mean (median; IQR)	0.34 (0.15; 0.10–025)	0.25 (015; 0.10–0.25)	0.9
DRE (% abnormal)	199 (38.3)	119 (31.3)	0.035
PCa diagnosis, *n* (%)[c]	212 (40.8)	181 (46.9)	0.081
GS ≤6, *n* (%)	103 (48.6)	90 (50.0)	0.8
GS 7, *n* (%)	58 (27.4)	51 (28.3)	
GS 8–10, *n* (%)	51 (24.1)	39 (21.6)	

DRE = digital rectal examination; GS = Gleason score; IQR = interquartile range; NA = not available; PCa = prostate cancer; PSA = prostate-specific antigen; PSAD = prostate-specific antigen density; TRUS = transrectal ultrasound.
Because race data were not recorded, based on general hospital records, it was assumed that >95% of this study population was white.
[a] Number of evaluable samples based on a minimal level of KLK3 reference messenger RNA, as described in the text.
[b] The *p* value when only taking into account those patients for which the information was available. When under the assumption that the vast majority of patients for whom this information was not available, have no family history of PCa, the difference is not statistically different ($p = 0.6$).
[c] For one subject the total GS could not be determined, but at least a Gleason 4 component was present.

Table 2 – Biomarker models used for the development with cut-off and clinical performance

Model	Cut-off	AUC	Se, %	Sp, %	NPV, %	PPV, %
PCA3	35.0	0.65	91	20	89	23
TDRD1	1.0	0.69	90	11	80	21
DLX1	0.5	0.65	83	16	79	21
HOXC4	15.5	0.64	91	22	90	23
HOXC4 and DLX1	26.5	0.70	91	31	93	25
HOXC4 and TDRD1	50.5	0.72	91	30	93	25
HOXC4, DLX1, and TDRD1	57.5	0.73	91	31	93	26
HOXC6	19.5	0.73	91	33	93	26
HOXC6 and DLX1	27.5	0.76	91	36	94	27
HOXC6 and TDRD1	50.5	0.74	91	35	94	27
HOXC6, DLX1, and TDRD1	55.5	0.74	91	34	93	26
HOXC6, HOXC4, DLX1, and TDRD1	85.5	0.74	91	33	93	26

AUC = area under the curve; NPV = negative predictive value; PPV = positive predictive value; Se = sensitivity; Sp = specificity.

3.2. *Selection of the most informative messenger RNA biomarkers*

The performance of individual biomarkers to predict high-grade PCa on biopsy was compared. In addition to the AUC, specificity, negative predictive value (NPV), and positive predictive value (PPV) were determined at a fixed sensitivity of approximately 90%, with HOXC6 as the strongest individual marker (Table 2). Next, the best complementary marker for HOXC6, the strongest performing individual marker in terms of AUC, was identified. HOXC6 and HOXC4 were strongly correlated, with a Pearson correlation coefficient of 0.80, indicating limited complementarity. To determine whether DLX1 or TDRD1 could complement the performance of either HOXC6 or HOXC4, models were generated based on the sum of the ratios. The combination of HOXC6 and DLX1 had the best performance with an AUC of 0.76 (95% CI, 0.71–0.81). The addition of other markers to this model did not result in further performance improvement. This combination was successfully validated in cohort B, with an AUC of 0.73 (95% CI, 0.67–0.78; $p = 0.4$ difference of AUCs) (Fig. 1).

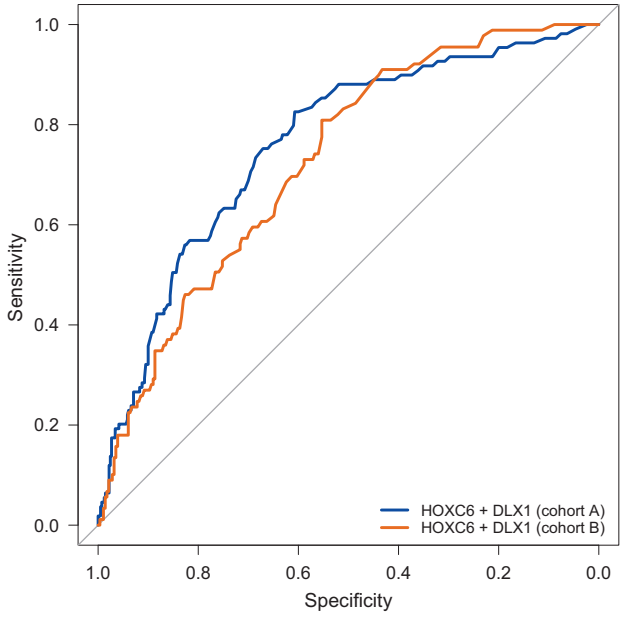

Fig. 1 – Receiver operating characteristic curves for HOXC6 and DLX1 messenger RNA expression levels in urine in the training and validation cohorts.

3.3. *Informative rate*

KLK3 was used as a measure for the presence of prostate-derived transcripts, and the expression level was about 1000-fold higher compared with biomarker mRNAs (Supplementary Table 1). A minimum threshold of 10 000 copies was set for the expression of this gene, and samples without sufficiently high biomarkers signal with fewer copies of the reference gene were considered nonevaluable because of fear of false-negative assay results (Table 1).

3.4. *Combining risk factors for optimized detection of high-grade prostate cancer*

All available molecular and traditional risk factors were combined in a logistic regression model to determine their relative contribution and importance to predict the presence of high-grade PCa on biopsy. A first logistic regression model was built using age, PSA, PSA density (PSAD), family history of PCa, DRE, history of prostate biopsy, and HOXC6 and DLX1 expression levels, and it was evaluated in cohort A. To account for differences in scale, transformations were investigated for some variables. A log-transformation for PSA, HOXC6 and DLX1 expression levels, and PSAD improved the overall model's performance. The output from the model was a risk score based on all available information.

Model 1 included all variables and reached an AUC of 0.90 (95% CI, 0.87–0.93) for high-grade PCa (Fig. 2A and Table 3). The biopsy outcome prediction had a significant contribution from DRE ($p < 0.001$), PSAD ($p = 0.004$), HOXC6 and DLX1 expression levels ($p = 0.003$), history of previous cancer-negative biopsies ($p = 0.02$), however not from PSA ($p = 0.08$), family history ($p = 0.15$), or age ($p = 0.7$). To verify whether the variables that did not have a significant contribution in the model did not result in overfitting, a backward elimination strategy (variables with $p > 0.05$) was applied until the model consisted only of significant variables. This model included DRE, PSAD, previous cancer-negative

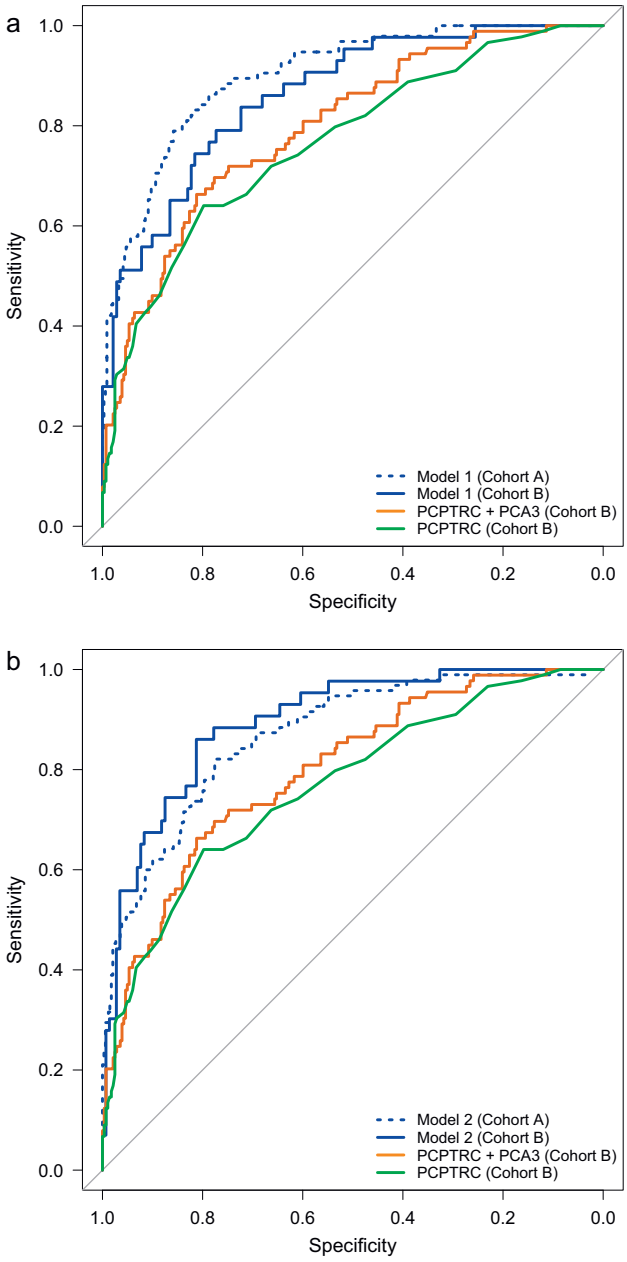

Fig. 2 – Receiver operating characteristic curves comparing (A) model 1 and (B) model 2 in cohorts A and B with the Prostate Cancer Prevention Trial risk calculator (PCPTRC) alone and the combined PCPTRC and PCA3.

biopsies, and *HOXC6* and *DLX1* expression levels and had an AUC of 0.89 (95% CI, 0.86–0.93).

A second model was developed excluding diagnostic DRE results, to avoid variables subject to interobserver variability. Model 2 reached an AUC of 0.87 (95% CI, 0.83–0.91), which was significantly lower than that of model 1 ($p = 0.009$ for comparison of AUCs) (Fig. 2B and Table 3).

3.5. Clinical validation

In cohort B, an AUC of 0.86 (95% CI, 0.80–0.92) (Fig. 2A) was obtained for model 1. The proposed model proved to be a robust predictor for the detection of high-grade PCa, as illustrated by a successful validation in this independent cohort by a direct comparison with cohort A ($p = 0.3$ for the difference between AUCs). In cohort B, model 2 reached an AUC of 0.90 (95% CI, 0.85–0.95), also not significantly different from the AUC obtained in the training cohort ($p = 0.4$) (Fig. 2B). In the validation cohort, model 2 significantly outperformed model 1 ($p = 0.033$ for the difference between AUCs), supporting the interobserver variability hypothesis. To further validate the contribution of HOXC6 and DLX1, model 2 was compared with a model that only incorporated traditional clinical risk factors (AUC: 0.87; 95% CI, 0.81–0.93). The addition of the mRNA markers to the model resulted in a significantly higher AUC ($p = 0.018$). Similarly, PCA3 was added to the clinical risk factors model and trained in cohort A; however, this did not result in a significant improvement of the AUC on validation in cohort B (AUC: 0.88; 95% CI, 0.82–0.94; $p = 0.2$).

Table 3 – Odds ratios, 95% confidence intervals, and *p* values for risk factors as obtained in the training cohort during development of the risk score

Parameter	Model 1, OR; CI (*p*)	Model 2, OR; CI (*p*)
HOXC6 and DLX1	1.68; 1.38–2.05 (0.003)	1.96; 1.40– 2.73 (<0.001)
PSAD	2.91; 1.40–6.06 (0.004)	3.78; 1.89–7.53 (<0.001)
DRE	5.53; 2.89–10.56 (<0.001)	–
Previous biopsy	0.28; 0.09–0.83 (0.02)	0.21; 0.07– 0.63 (0.005)
PSA	5.40; 0.81–35.94 (0.081)	3.42; 0.59–19.98 (0.17)
Family history	1.76; 0.81–3.80 (0.15)	1.56; 0.76–3.18 (0.2)
Age	1.01; 0.97–1.05 (0.7)	1.01; 0.97–1.06 (0.5)

CI = confidence interval; DRE = digital rectal examination; OR = odds ratio; PSA = prostate-specific antigen; PSAD = prostate-specific antigen density.

3.6. Clinical applicability and validity

The performance characteristics of these models were evaluated relative to current, clinically relevant methods in the independent validation cohort B. The PCPTRC v.2, based on a model incorporating PSA with other traditional clinical risk factors, was used as the main benchmark. Merely as a reference for the predictive value and complementarity of models including biomarkers and clinical risk factors, PCPTRC was combined with PCA3. The AUC for the PCPTRC predicting the chance of high-grade PCa was 0.77 (95% CI, 0.71–0.83), indicating that the mRNA-based risk score as incorporated in model 1 or 2 provided a significant improvement ($p = 0.015$ and $p < 0.001$, respectively) (Fig. 2). The AUC for PCPTRC combined with PCA3 was 0.80 (95% CI, 0.74–0.85), which was significantly lower than the AUC of model 2 ($p = 0.18$ and $p = 0.007$ for the comparison of AUCs with models 1 and 2, respectively). Finally, AUCs of 0.90 (95% CI, 0.85–0.96) and 0.93 (95% CI, 0.89–0.97) with model 1 and 2, respectively, to detect high-grade PCa were observed when potentially undergraded GS ≤6 PCa samples were removed from cohort B, only using the PCa-negative men as controls.

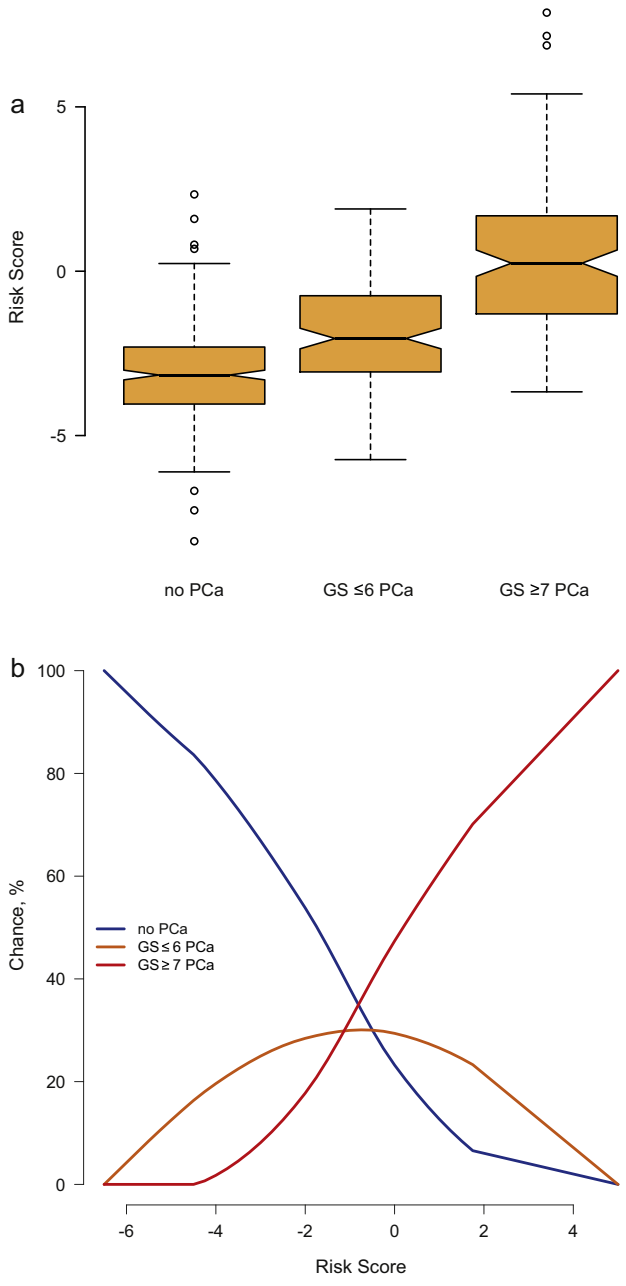

Fig. 3 – Relation between the risk score and biopsy outcome. (A) Box plot shows increasing levels of the risk score when finding no prostate cancer (PCa), low-grade PCa, or high-grade PCa on biopsy that was translated into (B) the likelihood of having a particular biopsy outcome in relation to the risk score.

This risk score showed a significant increase across all groups of men with, respectively, no PCa, GS ≤6, and GS ≥7 PCa (all $p < 0.001$) (Fig. 3A). A Spearman correlation coefficient of 0.61 was observed ($p < 0.001$), indicating a strong positive relationship between the risk score and biopsy outcome. For practical purposes, the outcome of the risk score was translated into the chance of observing no PCa, GS ≤6, or GS ≥7 PCa (Fig. 3B), offering a direct relation to the test's PPVs and NPVs.

To further improve the applicability of the model in clinical practice, PSAD was substituted for DRE-based prostate volume, resulting in three classes, that is, small (<30 ml), medium (≥30 and <60 ml), and large (≥60 ml). Medium and large prostates resulted in a decrease of the risk score relative to small prostates. Interestingly, the AUCs were not significantly lower based on these categorical volume assessments (AUC: 0.85, 95% CI, 0.79–0.92, $p = 0.3$ for model 1; AUC: 0.88, 95% CI, 0.82–0.94, $p = 0.092$ for model 2).

3.7. *Performance in the prostate-specific antigen gray zone*

The performance of the risk score was evaluated in the 264 men with low (<10 ng/ml) serum PSA levels from cohort B, of which 226 men had no or low-grade PCa (85.6%). The risk score remained the strongest predictor in this group of men in the PSA gray zone with an AUC of 0.78 (95% CI, 0.68–0.88) for model 1 and an AUC of 0.85 (95% CI, 0.77–0.93) for model 2, compared with PCPTRC with an AUC of 0.66 (95% CI, 0.57–0.75; $p = 0.071$ and $p = 0.001$). Again merely as a reference, the addition of PCA3 to the PCPTRC to compensate for PSA yielded an AUC of 0.72 (95% CI, 0.64–0.80), which was significantly lower than the AUC of model 2 ($p = 0.5$ and $p = 0.033$ for models 1 and 2, respectively).

3.8. *Clinical utility*

To evaluate the clinical utility of the risk score, a DCA was performed on the independent cohort B and compared with other decision-making tools used in clinical practice, merely for reference purposes (Fig. 4). Test harm was incorporated in the DCA for all evaluated tools, assuming that for first-line diagnostics no more than 50 patients should be evaluated to identify one high-grade PCa. Compared with the PCPTRC, and a model combining the PCPTRC with PCA3, the risk score, especially for model 2, clearly resulted in the largest net benefit in terms of accurately detecting men with high-grade PCa, even for those men who are very risk averse (Fig. 4A), while at the same time maximally reducing the unnecessary biopsy rate (Fig. 4B). From a practical point of view, at a cut-off with an NPV of 98% for GS ≥7 PCa, a total reduction of biopsies by 42% and a decrease of the unnecessary biopsies by 53% were obtained.

4. Discussion

New promising PCa-specific biomarkers have been identified in many studies; however, to date, only a few biomarkers have reached clinical practice. The main challenge is to validate the performance of the biomarkers in a clinical cohort independently and to demonstrate the clinical utility clearly. Leyten et al selected a promising urinary mRNA panel for the prediction of high-grade PCa (GS ≥7) on prostate biopsy [18]. In the current prospective multicenter study, a model was developed combining two of the most promising biomarkers, HOXC6 and DLX1, with traditional risk factors, most notably PSAD and DRE but also PSA, family history of PCa, and age, into one logistic

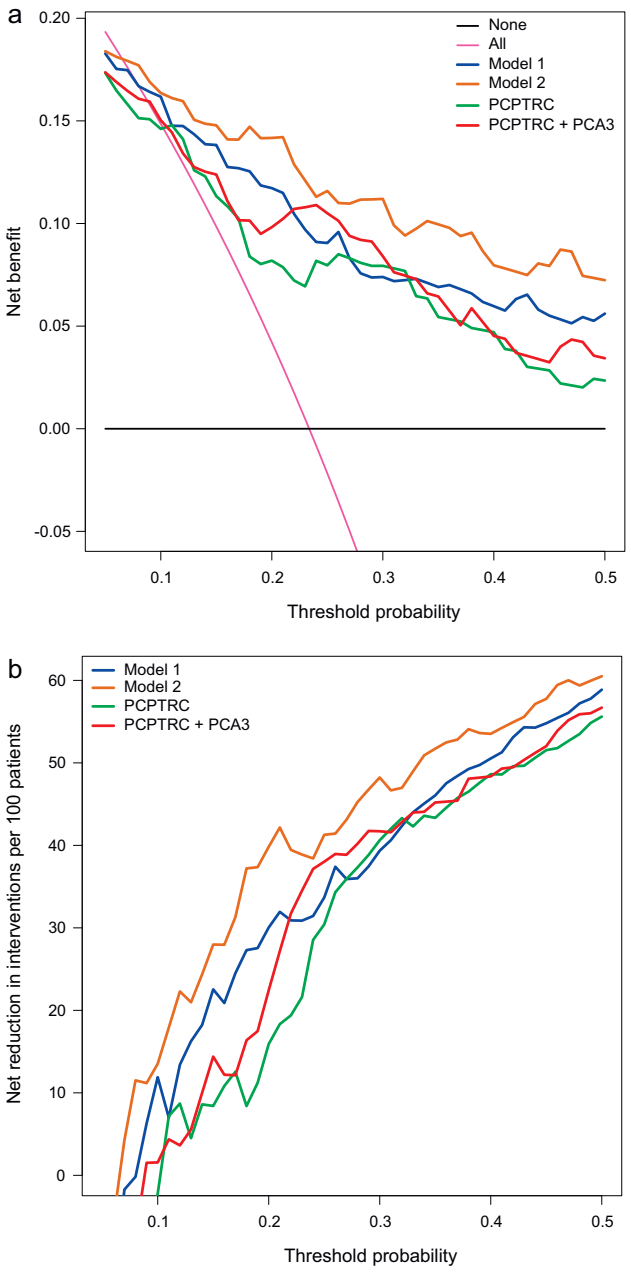

Fig. 4 – Decision curve analysis illustrating the overall clinical utility of models 1 and 2 compared with the Prostate Cancer Prevention Trial risk calculator (PCPTRC) alone and the PCPTRC and PCA3 combined in the validation cohort. Clinical utility of the risk score is demonstrated by (A) the overall net benefit in detecting high-grade prostate cancer without performing unnecessary biopsies and (B) the net reduction in interventions without missing any of these high-grade cancers.

regression model. The risk score derived from this model was the best performing assay to detect high-grade PCa on prostate biopsy and was successfully validated in an independent prospective cohort. A second model, excluding DRE as a risk factor because of potential interobserver variability in its assessment, was also validated successfully. The fact that, compared with the first model, the second model had a higher AUC in the validation cohort, but a lower AUC in the training cohort, is most likely a reflection of this interobserver variability. Hence the inclusion of DRE as a risk factor should be carefully considered. The models significantly outperformed the PCPTRC and PCA3. This was also true for the model that included HOXC6 and DLX1, age, PSA, PSAD, family history of PCa, and a history of prostate biopsy when compared with a combination of PCPTRC and PCA3. The addition of HOXC6 and DLX1 mRNA markers showed an improved patient stratification over the model with only the traditional clinical risk factors, which was not the case for PCA3. Although the traditional clinical risk model resulted in a relatively high AUC by itself, it was mainly driven by PSAD. Interestingly, the model did not depend on PSAD as such, which was illustrated by the similar performance of a model that included categorized DRE volume (small, medium, and large prostate size) rather than PSAD.

In the current study, mRNA assays were performed on whole urine samples, which is preferred for biomarker analysis because it does not require labor-intensive, time-consuming urine-processing procedures, and mRNA yield is not compromised [25].

In the PCPT, Thompson et al reported the diagnosis of PCa in 15.2% of men with a PSA level $\leq$4 ng/ml, of which 14.9% had high-grade disease [26]. This risk was very low for patients with a PSA level $<$1 ng/ml but increased to 9.4% in patients with a PSA between 3 and 4 ng/ml, that is, one could conclude that the currently accepted risk of missing significant cancers using PSA is up to 9.4% when a threshold of 4 ng/ml is used or 5.7% when a threshold of 3 ng/ml is used [27]. The clinical utility of the risk score was vastly beneficial, as illustrated by the potential higher detection rate of high-grade PCa while lowering the number of unnecessary repeat biopsies when adopting this model, in particular when compared with a model including PCPTRC and PCA3; however, it should be noted that the latter was not developed specifically for high-grade cancer.

The risk score enables objective clinical risk assessment and patient management, but it also compensates for risk factors that are, by their very nature, subjective or subject to interobserver variability. In this study, this is particularly true for DRE; however, even when included, the risk score remained the strongest, most significant predictor of patient risk compared with other clinically relevant risk assessment algorithms, such as PCA3 and the PCPTRC.

The main limitations of this study are the lack of centralized pathology and the fact that the gold standard for PCa diagnosis, namely, TRUS-guided biopsy, not only has a false-negative rate of approximately 20% [28,29], but it also has difficulty detecting PCa in the anterior (and apical parts) of the prostate [30]. Because only 16% of the cohort was composed of men with at least one previous cancer-negative biopsy, it would be interesting to study the repeat biopsy setting specifically.

5. Conclusions

The two-gene risk score combining HOXC6 and DLX1 mRNA expression levels with traditional clinical risk factors (ie, PSAD, DRE, PSA, age, history of prostate biopsy, and family

history) is able to detect high-grade, clinically significant PCa accurately and could therefore be used in decision making, reducing the number of unnecessary prostate biopsies and potential overtreatment. This newly developed risk score significantly outperforms the PCPTRC, a multimodal risk assessment approach, and improves the diagnosis and management of PCa patients. Future research might indicate that additional parameters could further optimize the diagnosis of high-grade PCa without contributing to the high unnecessary biopsy rate.

Author contributions: Jack A. Schalken had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Van Neste, Hendriks, Dijkstra, Jannink, de Jong, Hessels, Smit, Melchers, Leyten, Mulders, van Oort, Van Criekinge, Schalken.
Acquisition of data: Van Neste, Hendriks, Dijkstra, Cornel, Jannink, de Jong, Hessels, Smit, Leyten, Reijke, Vergunst, Kil, Knipscheer, Hulsbergen-van de Kaa, Mulders, van Oort, Van Criekinge, Schalken.
Analysis and interpretation of data: Van Neste, Hendriks, Dijkstra, Trooskens, Jannink, de Jong, Hessels, Smit, Melchers, Leyten, Mulders, van Oort, Schalken.
Drafting of the manuscript: Van Neste, Hendriks, Dijkstra, Trooskens, Jannick, de Jong, Schalken.
Critical revision of the manuscript for important intellectual content: Cornel, Hessels, Smit, Melchers, Leyten, Reijke, Vergunst, Kil, Knipscheer, Hulsbergen-van de Kaa, Mulders, van Oort, Schalken.
Statistical analysis: Van Neste, Hendriks, Dijkstra, Jannink, de Jong, Hessels, Smit.
Obtaining funding: None.
Administrative, technical, or material support: None.
Supervision: Van Criekinge, Schalken.
Other (specify): None.

Financial disclosures: Jack A. Schalken certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Jack A. Schalken, Wim Van Criekinge, Geert Trooskens, Sander A. Jannink, Hans de Jong, Daphne Hessels, Frank P. Smit, Leander Van Neste, and Willem J.G. Melchers are employees of or consultants for MDxHealth and may have stock or stock options. The remaining authors have nothing to disclose.

Funding/Support and role of the sponsor: This study was funded partly by grants from EFRO, Ultrasense MR (2010–013177), and PCMM (3-O-203).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eururo.2016.04.012.

References

[1] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin 2015;65:87–108.
[2] Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. J Natl Cancer Inst 2009;101:374–83.
[3] Welch HG, Albertsen PC. Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986–2005. J Natl Cancer Inst 2009;101:1325–9.
[4] Albertsen PC, Hanley JA, Fine J. 20-year outcomes following conservative management of clinically localized prostate cancer. JAMA 2005;293:2095–101.
[5] Groskopf J, Aubin SM, Deras IL, et al. APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis of prostate cancer. Clin Chem 2006;52:1089–95.
[6] Gittelman MC, Hertzman B, Bailen J, et al. PCA3 molecular urine test as a predictor of repeat prostate biopsy outcome in men with previous negative biopsies: a prospective multicenter clinical study. J Urol 2013;190:64–9.
[7] Bussemakers MJ, van Bokhoven A, Verhaegh GW, et al. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Res 1999;59:5975–9.
[8] Hessels D, Klein Gunnewiek JM, van Oort I, et al. DD3(PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. Eur Urol 2003;44:8–15, discussion 16.
[9] Auprich M, Chun FK, Ward JF, et al. Critical assessment of preoperative urinary prostate cancer antigen 3 on the accuracy of prostate cancer staging. Eur Urol 2011;59:96–105.
[10] Marks LS, Fradet Y, Deras IL, et al. PCA3 molecular urine assay for prostate cancer in men undergoing repeat biopsy. Urology 2007;69:532–5.
[11] Ploussard G, Haese A, Van Poppel H, et al. The prostate cancer gene 3 (PCA3) urine test in men with previous negative biopsies: does free-to-total prostate-specific antigen ratio influence the performance of the PCA3 score in predicting positive biopsies? BJU Int 2010;106:1143–7.
[12] Scattoni V, Lazzeri M, Lughezzani G, et al. Head-to-head comparison of prostate health index and urinary PCA3 for predicting cancer at initial or repeat biopsy. J Urol 2013;190:496–501.
[13] Nordstrom T, Vickers A, Assel M, Lilja H, Gronberg H, Eklund M. Comparison between the four-kallikrein panel and Prostate Health Index for predicting prostate cancer. Eur Urol 2015;68:139–46.
[14] Seisen T, Rouprêt M, Brault D, et al. Accuracy of the prostate health index versus the urinary prostate cancer antigen 3 score to predict overall and significant prostate cancer at initial biopsy. Prostate 2015;75:103–11.
[15] Ferro M, Bruzzese D, Perdona S, et al. Predicting prostate biopsy outcome: prostate health index (phi) and prostate cancer antigen 3 (PCA3) are useful biomarkers. Clin Chim Acta 2012;413:1274–8.
[16] Braun K, Sjoberg DD, Vickers AJ, Lilja H, Bjartell AS. A four-kallikrein panel predicts high-grade cancer on biopsy: independent validation in a community cohort. Eur Urol 2016;69:505–11.
[17] Parekh DJ, Punnen S, Sjoberg DD, et al. A multi-institutional prospective trial in the USA confirms that the 4Kscore accurately identifies men with high-grade prostate cancer. Eur Urol 2015;68:464–70.
[18] Leyten GH, Hessels D, Smit FP, et al. Identification of a candidate gene panel for the early diagnosis of prostate cancer. Clin Cancer Res 2015;21:3061–70.
[19] Ankerst DP, Hoefler J, Bock S, et al. Prostate Cancer Prevention Trial risk calculator 2.0 for the prediction of low- vs high-grade prostate cancer. Urology 2014;83:1362–7.
[20] Hansen J, Auprich M, Ahyai SA, et al. Initial prostate biopsy: development and internal validation of a biopsy-specific nomogram based on the prostate cancer antigen 3 assay. Eur Urol 2013;63:201–9.
[21] Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. Clin Chem 2015;61:1446–52.
[22] Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 2008;3:1101–8.

[23] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.
[24] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.
[25] Hendriks RJ, Dijkstra S, Jannink SA, et al. Comparative analysis of prostate cancer specific biomarkers PCA3 and ERG in whole urine, urinary sediments and exosomes. Clin Chem Lab Med 2016;54:483–92.
[26] Thompson IM, Pauler DK, Goodman PJ, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. N Engl J Med 2004;350:2239–46.
[27] Thompson IM, Ankerst DP, Chi C, et al. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. J Natl Cancer Inst 2006;98:529–34.
[28] Wolters T, van der Kwast TH, Vissers CJ, et al. False-negative prostate needle biopsies: frequency, histopathologic features, and follow-up. Am J Surg Pathol 2010;34:35–43.
[29] Stewart GD, Van Neste L, Delvenne P, et al. Clinical utility of an epigenetic assay to detect occult prostate cancer in histopathologically negative biopsies: results of the MATLOC study. J Urol 2013;189:1110–6.
[30] Roehl KA, Antenor JA, Catalona WJ. Serial biopsy results in prostate cancer screening study. J Urol 2002;167:2435–9.