

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



Prostate Cancer

External Validation of Urinary PCA3-Based Nomograms to Individually Predict Prostate Biopsy Outcome

Marco Auprich^{b,1,*}, Alexander Haese^{a,1}, Jochen Walz^c, Karl Pummer^b, Alexandre de la Taille^d, Markus Graefen^a, Theo de Reijke^e, Margit Fischⁱ, Paul Kilf^f, Paolo Gontero^g, Jacques Irani^h, Felix K.-H. Chunⁱ

^a Martini Clinic Prostate Cancer Center, University Clinic Eppendorf, Hamburg, Germany

^b Department of Urology, Medical University Graz, Austria

^c Institut Paoli-Calmettes Cancer Center, Marseille, France

^d CHU Henri Mondor, Paris, France

^e Academic Medical Center, Amsterdam, The Netherlands

^f St. Elisabeth Ziekenhuis, Tilburg, The Netherlands

^g Ospedale Molinette, Università di Torino, Turin, Italy

^h CHU "La Milétrie", Poitiers, France

ⁱ Department of Urology, University Clinic Eppendorf, Hamburg, Germany

Article info

Article history:

Accepted June 22, 2010

Published online ahead of
print on July 3, 2010

Keywords:

Prostate biopsy
Prostate cancer gene 3
Biomarker
Prostate cancer
Nomogram
External validation



www.eu-acme.org/
[europeanurology](http://europeanurology.com)

Please visit

www.eu-acme.org/europeanurology to read and
answer questions on-line.
The EU-ACME credits will
then be attributed
automatically.

Abstract

Background: Prior to safely adopting risk stratification tools, their performance must be tested in an external patient cohort.

Objective: To assess accuracy and generalizability of previously reported, internally validated, prebiopsy prostate cancer antigen 3 (PCA3) gene-based nomograms when applied to a large, external, European cohort of men at risk of prostate cancer (PCa).

Design, setting, and participants: Biopsy data, including urinary PCA3 score, were available for 621 men at risk of PCa who were participating in a European multi-institutional study.

Intervention: All patients underwent a ≥ 10 -core prostate biopsy. Biopsy indication was based on suspicious digital rectal examination, persistently elevated prostate-specific antigen level (2.5–10 ng/ml) and/or suspicious histology (atypical small acinar proliferation of the prostate, ≥ 2 cores affected by high-grade prostatic intraepithelial neoplasia in first set of biopsies).

Measurements: PCA3 scores were assessed using the ProgenSA assay (Gen-Probe Inc, San Diego, CA, USA). According to the previously reported nomograms, different PCA3 score codings were used. The probability of a positive biopsy was calculated using previously published logistic regression coefficients. Predicted outcomes were compared to the actual biopsy results. Accuracy was calculated using the area under the curve as a measure of discrimination; calibration was explored graphically.

Results and limitations: Biopsy-confirmed PCa was detected in 255 (41.1%) men. Median PCA3 score of biopsy-negative versus biopsy-positive men was 20 versus 48 in the total cohort, 17 versus 47 at initial biopsy, and 37 versus 53 at repeat biopsy (all $p \leq 0.002$). External validation of all four previously reported PCA3-based nomograms demonstrated equally high accuracy (0.73–0.75) and excellent calibration. The main limitations of the study reside in its early detection setting, referral scenario, and participation of only tertiary-care centers.

Conclusions: In accordance with the original publication, previously developed PCA3-based nomograms achieved high accuracy and sufficient calibration. These novel nomograms represent robust tools and are thus generalizable to European men at risk of harboring PCa. Consequently, in presence of a PCA3 score, these nomograms may be safely used to assist clinicians when prostate biopsy is contemplated.

© 2010 European Association of Urology. Published by Elsevier B.V. All rights reserved.

¹ These authors contributed equally to the manuscript.

* Corresponding author. Department of Urology, Medical University Graz, Auenbruggerplatz 7, A-8036 Graz, Austria. Tel. +43 316 385 2508; Fax: +43 316 385 3550.

E-mail address: marco.auprich@medunigraz.at (M. Auprich).

1. Introduction

Early prostate cancer (PCa) detection has become feasible through the advent of serum prostate-specific antigen (PSA). However, PSA is characterized by a low PCa specificity (especially in the range of 4–10 ng/ml), while only about 25% of men suspicious of harboring PCa have a positive biopsy [1]. Thus, men with benign findings at biopsy possibly have to face further unnecessary biopsies, causing anxiety, discomfort, and complications in patients and increasing health care expenses. New diagnostic markers may circumvent PSA's low specificity. In this context, a promising candidate urinary marker is prostate cancer gene 3 (*PCA3*). In 1999, Bussemakers et al first reported that PCa tissue highly expresses the noncoding differential display code 3 (DD3) messenger ribonucleic acid (mRNA) [2]. A time-resolved fluorescence (TRF) real-time polymerase chain reaction method was subsequently employed by Hessels and coworkers to evaluate the applicability and clinical value as an additional urinary diagnostic test [3]. They measured mRNA of DD3^{PCA3} in urine sediments after digital rectal examination (DRE) prior to prostate biopsy. Using a *PCA3* cut-off (200×10^{-3} copies) to discriminate prostate biopsy outcome, *PCA3* outperformed established clinical parameters such as PSA and yielded a negative predictive value of 90% and an area under the curve (AUC) of 0.72. Subsequently, based on the work of Grosskopf et al, a commercially available *PCA3* assay was developed [4]. Using this assay, multi-institutional European and US studies reported that the probability of a positive repeat biopsy increases with rising *PCA3* score. Furthermore, they demonstrated that the *PCA3* score was superior to PSA and percent free PSA for predicting repeat biopsy outcome and that it may be indicative of clinical stage and PCa significance [5–9].

Moreover, Chun et al recently demonstrated that *PCA3* fulfills most stringent statistical criteria of a novel diagnostic marker [10]. In their analysis, *PCA3* represented an independent, as well as informative, marker capable of increasing multivariable predictive accuracy as postulated by Kattan [11,12]. Consequently, they developed a set of *PCA3* nomograms to predict biopsy outcome. Within these nomograms, combining *PCA3* result with established risk factors is related to a multivariable diagnostic accuracy gain of 3–5%. However, a strong limitation of these new risk stratification tools stems from internal validation. Gold-standard validation within an external patient cohort to truly assess accuracy and calibration of these novel tools has never been performed.

To address this issue, this study evaluates accuracy, performance, and generalizability of previously published *PCA3*-based nomograms using an external multi-institutional patient cohort.

2. Materials and methods

2.1. Patient populations

The study cohort consisted of 805 men from a European multi-institutional trial who were subjected to either initial or repeat biopsy.

After exclusion of 184 patients due to missing clinical variables, 621 patients remained for analyses.

2.2. Clinical evaluation

All men included in our study cohort had been referred for prostatic reevaluation. Biopsy indication was based on suspicious DRE, persistently elevated PSA (2.5–10 ng/ml), and/or suspicious histology (atypical small acinar proliferation, \geq two cores affected by high-grade prostatic intraepithelial neoplasia in first set of biopsies). All men had complete information on age, PSA level, DRE, prostate volume, history of previous biopsy, and urinary *PCA3* score. In all men, \geq 10-core systematic, laterally directed, transrectal ultrasound (TRUS)-guided biopsy was performed.

PSA levels were measured before DRE and TRUS. DRE findings were classified as normal or suspicious. TRUS-derived total prostate volume was calculated using the prolate ellipse formula ($0.52 \times \text{length} \times \text{width} \times \text{height}$), as described earlier [13]. First-catch urine samples were collected following a DRE as described by Grosskopf et al [4]. Urine samples were processed and tested to quantify *PCA3* mRNA and PSA mRNA concentrations using the ProgenSA assay (Gen-Probe Inc, San Diego, CA, USA). The following formula was used to calculate *PCA3* assay score: (mRNA *PCA3*) / (mRNA PSA) \times 1000. Biopsy specimens were evaluated by an experienced uro-pathologist at each participating site. Gleason score was assigned according to the recommendation of Amin et al [14].

2.3. Statistical analyses

To test accuracy and calibration of the previously published *PCA3* nomograms, external validation using the predetermined regression coefficients was performed according to different *PCA3* codings. The AUC of the receiver operator characteristic analysis was used to quantify accuracy of each *PCA3* model in each biopsy setting. The extent of nomogram over- or underestimation was explored graphically within loess calibration plots. All statistical tests were performed using S-PLUS Professional, v.1 (MathSoft Inc, Seattle, WA, USA). All tests were two-sided with a significance level of 0.05.

3. Results

Patient characteristics are shown in Table 1. Overall, PCa was detected in 41.1% ($n = 255$) of patients. As shown in Fig. 1A, median *PCA3* score was 29 (range: 1–301). Median *PCA3* score was significantly higher in patients with PCa versus patients without PCa (48 vs 20, $p < 0.001$). Overall, PSA levels ranged from 1.0 to 82.7 ng/ml (mean and median: 7.3 and 6.1 ng/ml, respectively) and DRE was suspicious in 16.9%. Increasing age and PSA level, suspicious DRE, and decreasing prostate volume were associated with PCa at biopsy with statistically high significance ($p < 0.001$).

Of all men, 75.2% ($n = 467$) compared with 24.8% ($n = 154$) underwent an initial versus repeat biopsy. Initial biopsy revealed PCa in 40.5% ($n = 189$) of patients. Median *PCA3* score on initial biopsy was 25 (range: 1–301). Median *PCA3* score was significantly higher in patients with PCa compared with patients without PCa (47 vs 17, $p < 0.001$) (Fig. 1B). Median PSA level was 5.6 ng/ml (range: 1–57.2 ng/ml). Higher age, increasing PSA level, suspicious DRE, and decreasing prostate volume were statistically significantly associated with PCa at initial biopsy ($p < 0.001$).

At repeat biopsy PCa was diagnosed in 42.9% ($n = 66$) men. Median *PCA3* score was 45 (range: 1–289) and was

Table 1 – Risk factors comparing biopsy-negative and biopsy-positive men

Variables	Validation cohort (n = 621)			p value
	Total	PCa negative	PCa positive	
Patients, No. (%)	621 (100)	366 (58.9)	255 (41.1)	–
Age, yr				<0.001
Mean	63.0	62	64.5	
Median	63	62	65	
Range	35–90	38–83	35–90	
PSA (ng/ml)				<0.001
Mean	7.3	6.5	8.4	
Median	6.1	5.8	6.6	
Range	1–82.7	1.0–22.4	2.4–82.7	
DRE, No. (%)				<0.001
Suspicious	105 (16.9)	40 (10.9)	65 (25.5)	
Unsuspectious	516 (83.1)	326 (89.1)	190 (74.5)	
Total prostate volume (ml)				<0.001
Mean	48.0	52.6	41.3	
Median	44	47.2	38	
Range	10–148	12.4–148	10–132	
Previous biopsy session, No. (%)				0.60
No	467 (75.2)	278 (76.0)	189 (74.1)	
Yes	154 (24.8)	88 (24)	66 (25.9)	
Gleason score	–	–		–
≤3 + 3			146 (57.3)	
3 + 4			67 (26.3)	
4 + 3			21 (8.2)	
≥4 + 4			21 (8.2)	
PCA3 assay score (%)				<0.001
Mean	45.8	32.6	64.7	
Median	29	20	48	
Range	1–301	1–215	4–301	
≤ 17	195 (31.4)	163 (44.5)	32 (12.5)	<0.001
>17	426 (68.6)	203 (55.5)	223 (87.5)	
≤24	264 (42.5)	208 (56.8)	56 (22.0)	<0.001
>24	357 (57.5)	158 (43.2)	199 (78.0)	
≤35	347 (55.9)	256 (69.9)	91 (35.7)	<0.001
>35	274 (44.1)	110 (30.1)	164 (64.3)	<0.001
	–	–	–	

PCa = prostate cancer; PSA = prostate-specific antigen; DRE = digital rectal examination; PCA3 = prostate cancer antigen 3.

significantly higher in patients with PCa versus those without (53 vs 37; $p = 0.002$) (Fig. 1C). Serum PSA levels in men at repeat biopsy ranged from 1.8 to 82.7 ng/ml (mean and median: 7.3 and 6.1 ng/ml, respectively). Except for age and suspicious DRE, higher PSA level, lower prostate volume, and higher PCA3 score were significantly associated with PCa at repeat biopsy ($p < 0.01$).

Table 2 displays accuracy estimates in the external validation cohort ($n = 621$) stratified according to four different PCA3 codings. The AUCs were 0.73 versus 0.75 versus 0.74 versus 0.74 when continuously coded versus cut-off 17 versus cut-off 24 versus cut-off 35 for PCA3 were used.

Fig. 2A–D displays nomogram calibration plots of the previously developed models stratified according to PCA3 coding, as applied to the external validation data set. On each calibration plot, the predicted probability of the previously reported nomogram is represented on the x-axis and the actual proportion of biopsy-proven PCa is repre-

sented on the y-axis. The 45° line indicates perfect agreement between predicted probability and observed proportion of PCa. Within the external validation cohort, equally excellent calibration of all investigated PCA3 codings is demonstrated, virtually overlapping the 45° line of perfect prediction.

Table 2 – External validation accuracy estimates

Variables	Accuracy
Age, PSA, DRE, PV, prevBx + PCA3 (cc)	0.73
Age, PSA, DRE, PV, prevBx + PCA3-17	0.75
Age, PSA, DRE, PV, prevBx + PCA3-24	0.74
Age, PSA, DRE, PV, prevBx + PCA3-35	0.74

PSA = prostate-specific antigen; DRE = digital rectal examination; PV = prostate volume; prevBx = history of previous biopsy (yes vs. no); PCA3 score = prostate cancer antigen 3 score; PCA3-17 = PCA3 assay score threshold 17; PCA3-24 = PCA3 assay score threshold 24; PCA3-35 = PCA3 assay score threshold 35.

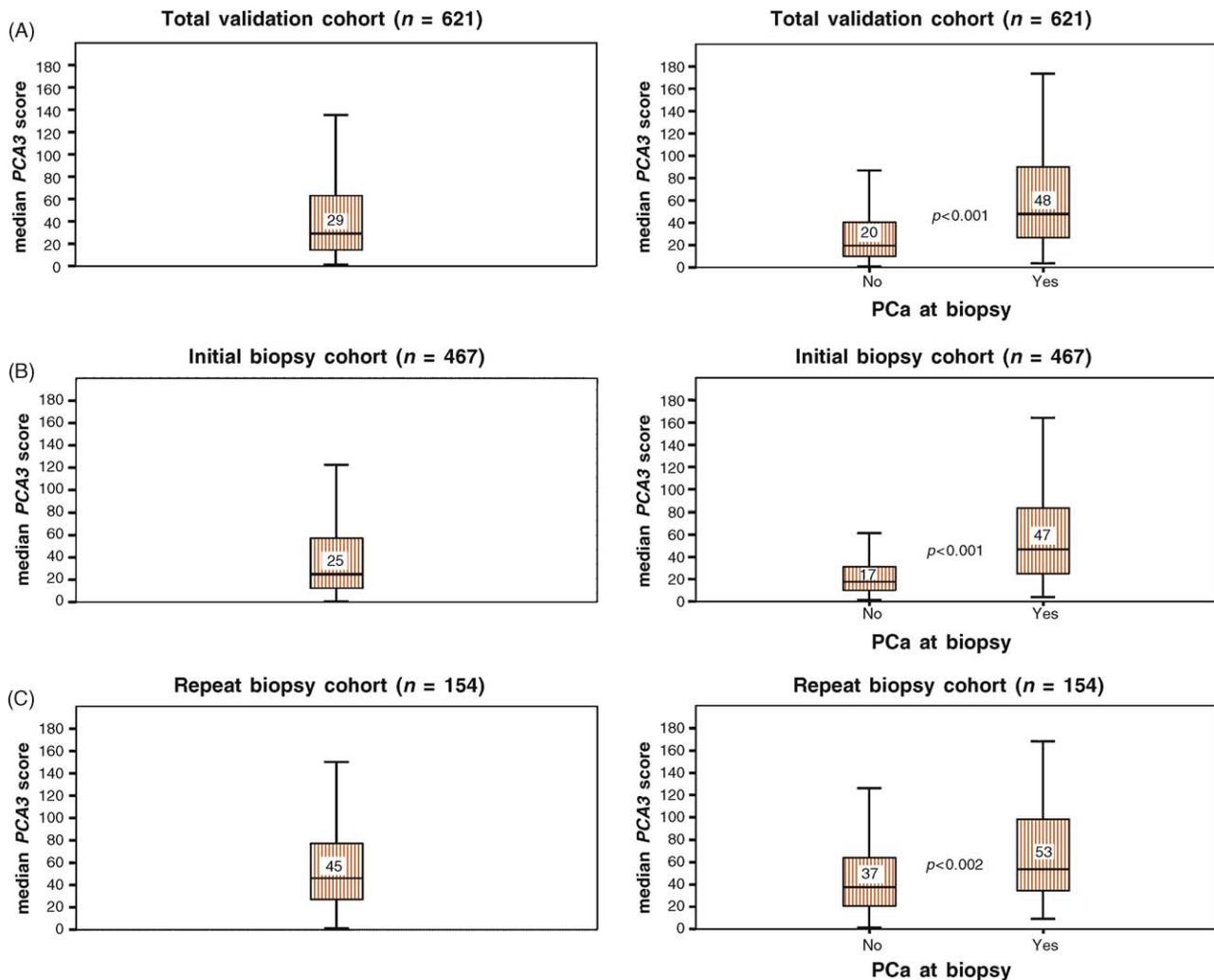


Fig. 1 – Box plots of median *PCA3* assay scores with comparison of biopsy negative and biopsy positive men in total (A), initial biopsy (B) and repeat biopsy (C) cohort.

PCA3 = Prostate cancer gene 3; *PCa* = prostate cancer.

4. Discussion

Novel clinical markers have been intensively studied over the past decade to improve early *PCa* diagnostic specificity. Urinary *PCA3* represents a novel urinary marker that holds great promise. For example, in recent studies *PCA3* demonstrated superior specificity over *PSA* [5–7,15]. Consequently, *PCA3*'s clinical application could potentially reduce the proportion of men exposed to unnecessary biopsy as well as biopsy-related morbidity and health care expenses. Therefore, Chun et al developed novel, internally validated, *PCA3*-based nomograms, where *PCA3* statistically significantly improved accuracy of biopsy outcome prediction [10]. As indicated by work of Steyerberg et al comparing different methods of internal validation with regard to discriminative ability, calibration, and overall accuracy, the bootstrap technique used by Chun et al seems superior to other internal validation techniques such as split sample or cross validation [16]. However, even though the bootstrap validation method has been applied, clinicians may be reluctant to

use these novel *PCA3*-based nomograms because they could still argue that their clinical applicability and generalizability has never been demonstrated in real patients. Therefore, external validation of this set of nomograms is necessary to display their true clinical applicability and generalizability quantified as overall accuracy and calibration.

To address this obvious limitation, we present the first external validation study of these recently developed *PCA3*-based nomograms in a large European multi-institutional patient cohort. Methodologically, each individual patient in the external data set was exposed to the set of previously developed *PCA3* nomograms using the originally derived regression coefficients. External validation calculates risk probabilities of a positive biopsy to assess their accuracy, comparing predicted probability versus observed proportion of *PCa* in each patient. Subsequently, this ratio is graphically explored in a so-called calibration plot, which enables the clinician to investigate a model's weakness expressed as deviations from the 45° line. This line displays perfect agreement between predicted and observed

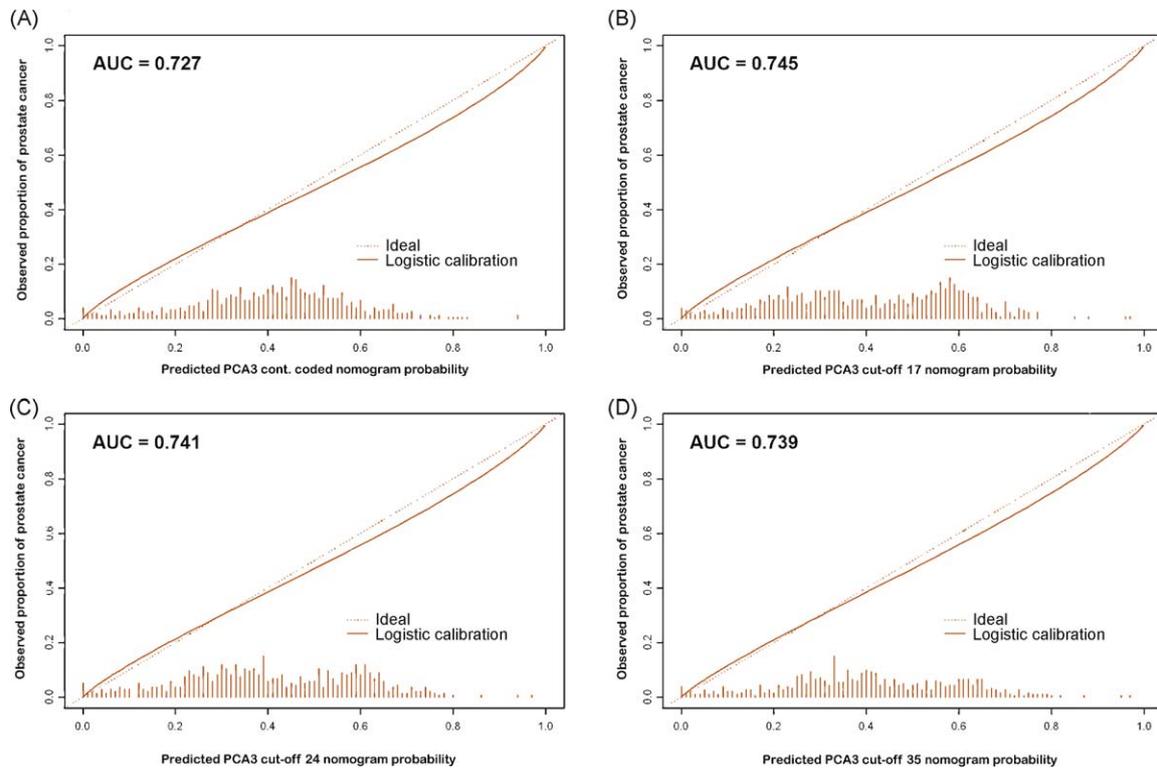


Fig. 2 – Nomogram calibration plots within external validation cohort stratified to various prostate cancer gene 3 (*PCA3*) coding. AUC = area under the curve (= predictive accuracy); *PCA3(cc)* = *PCA3* assay score continuously coded; *PCA3-17* = *PCA3* assay score threshold 17; *PCA3-24* = *PCA3* assay score threshold 24; *PCA3-35* = *PCA3* assay score threshold 35.

probabilities of biopsy outcome. Deviations from this line are interpreted as over- or underestimations of the predictive model, which are crucial to investigate prior to the use of any predictive model in routine clinical practice.

Our external validation findings demonstrate that accuracy depending on *PCA3* coding ranged from 0.73–0.75 (Table 2). These accuracy estimates compare excellently with the original report, in which accuracy ranged from 0.70–0.73. Moreover, external calibration virtually overlaps the 45° line of perfect prediction, indicating an almost perfect agreement between predicted and observed probability of PCa at biopsy (Fig. 2A–D). Taken together, overall external validation in a large European multicenter cohort demonstrates good accuracy and almost perfect calibration.

However, several limitations must be acknowledged. First, despite these excellent external validation findings in European patients, further studies have to be conducted to reconfirm the demonstrated generalizability of this set of diagnostic nomograms. Especially, as Steuber et al demonstrated earlier, PCa clinical characteristics between European and North American men may substantially differ [17]. Therefore, external validation within a North American cohort is still lacking. Moreover, mostly tertiary referral centers have been included in our analyses. Thus, external validation in a community-based cohort may also be relevant. Similar to the original report, it is important to note that all patients included in this study were exposed to early detection and were not part of a screening effort.

Therefore, the findings of this study do not expand to screening cohorts. This truly interesting question must be separately addressed in future trials. Second, from a clinical viewpoint it may be argued that a mixed *PCA3* biopsy model may not be as useful as biopsy-specific models. However, it must be emphasized that the original reported *PCA3* nomograms were based on patients subjected to either an initial or a repeat biopsy. While the diagnostic impact of all implemented risk factors is precisely adjusted for a mixed biopsy population using the original regression coefficients, the impact or weight of each risk factor differs according to biopsy scenario. Therefore, from a clinical and methodologic viewpoint, complementary biopsy scenario-specific *PCA3*-based models may result in even more accurate biopsy outcome predictions as demonstrated earlier [18]. Lastly, despite having been shown to be of diagnostic value as a so-called novel diagnostic marker, it remains under investigation how a clinician should interpret an actual *PCA3* score. For example, whether *PCA3* should be used as a continuous coded variable or whether risk stratification of men into low-, intermediate-, or high-risk categories based on a single *PCA3* cut-off value are important questions.

Our findings clearly demonstrate that *PCA3* represents an important complementary marker to established risk factors. Future studies need to address the multivariable approach in which *PCA3* values are characterized according to clinical parameters (eg, initial vs repeat biopsy or PSA

strata) to identify those men who truly benefit from this additional novel marker.

5. Conclusions

We report the first external validation of *PCA3*-based nomograms demonstrating robustness, accuracy, and applicability in a large European multi-institutional patient cohort. Therefore, the previously reported set of nomograms may help clinicians in counseling and confirming biopsy indication in European men in presence of a *PCA3* value.

Author contributions: Marco Auprich, Alexander Haese and Felix Chun 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: Chun, Haese.

Acquisition of data: Auprich, Haese, Walz, Pummer, de la Taille, Graefen, de Reijke, Fisch, Kil, Gontero, Irani, Chun.

Analysis and interpretation of data: Chun, Haese.

Drafting of the manuscript: Chun, Haese.

Critical revision of the manuscript for important intellectual content: Auprich, Haese, Walz, Pummer, de la Taille, Graefen, de Reijke, Fisch, Kil, Gontero, Irani, Chun.

Statistical analysis: Chun.

Obtaining funding: None.

Administrative, technical, or material support: None.

Supervision: None.

Other (specify): None.

Financial disclosures: I certify that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

Funding/Support and role of the sponsor: Gen-Probe Inc was a funding sponsor of this study and was involved in the design, data collection, data management, data interpretation and analysis, preparation, review, and approval of the manuscript.

References

- [1] Catalona WJ, Partin AW, Slawin KM, et al. Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial. *JAMA* 1998;279:1542–7.
- [2] 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.
- [3] Hessels D, Klein Gunnewiek JMT, van Oort I, et al. DD3 *PCA3*-based molecular urine analysis for the diagnosis of prostate cancer. *Eur Urol* 2003;44:8–16, discussion 15–6.
- [4] 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.
- [5] Haese A, de la Taille A, van Poppel H, et al. Clinical utility of the *PCA3* urine assay in European men scheduled for repeat biopsy. *Eur Urol* 2008;54:1081–8.
- [6] 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.
- [7] Deras IL, Aubin SM, Blase A, et al. *PCA3*: a molecular urine assay for predicting prostate biopsy outcome. *J Urol* 2008;179:1587–92.
- [8] Nakanishi H, Groskopf J, Fritsche HA, et al. *PCA3* molecular urine assay correlates with prostate cancer tumor volume: implication in selecting candidates for active surveillance. *J Urol* 2008;179:1804–9, discussion 1809–10.
- [9] Whitman EJ, Groskopf J, Ali A, et al. *PCA3* score before radical prostatectomy predicts extracapsular extension and tumor volume. *J Urol* 2008;180:1975–8, discussion 1978–9.
- [10] Chun FK, de la Taille A, van Poppel H, et al. Prostate cancer gene 3 (*PCA3*): development and internal validation of a novel biopsy nomogram. *Eur Urol* 2009;56:659–68.
- [11] Kattan MW. Judging new markers by their ability to improve predictive accuracy. *J Natl Cancer Inst* 2003;95:634–5.
- [12] Kattan MW. Evaluating a new marker's predictive contribution. *Clin Cancer Res* 2004;10:822–4.
- [13] Eskew LA, Bare RL, McCullough DL. Systematic 5 region prostate biopsy is superior to sextant method for diagnosing carcinoma of the prostate. *J Urol* 1997;157:199–202, discussion 202–3.
- [14] Amin M, Boccon-Gibod L, Egevad L, et al. Prognostic and predictive factors and reporting of prostate carcinoma in prostate needle biopsy specimens. *Scand J Urol Nephrol Suppl* 2005;216:20–33.
- [15] Van Gils MP, Cornel EB, Hessels D, et al. Molecular *PCA3* diagnostics on prostatic fluid. *Prostate* 2007;67:881–7.
- [16] Steyerberg EW, Harrell Jr FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774–81.
- [17] Steuber T, Graefen M, Haese A, et al. Validation of a nomogram for prediction of side specific extracapsular extension at radical prostatectomy. *J Urol* 2006;175:939–44, discussion 944.
- [18] Chun FK, Briganti A, Graefen M, et al. Development and external validation of an extended repeat biopsy nomogram. *J Urol* 2007;177:510–5.