Radboud Repository

Radboud University Nijmegen

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

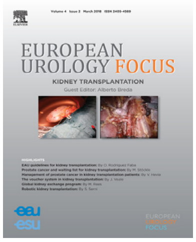
The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/195619

Please be advised that this information was generated on 2020-09-10 and may be subject to change.

EUROPEAN UROLOGY FOCUS 4 (2018) 228–234

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



Prostate Cancer

Prediction of Prostate Cancer: External Validation of the ERSPC Risk Calculator in a Contemporary Dutch Clinical Cohort

Maudy Gayet[a,b,†,*], *Christophe K. Mannaerts*[a,†], *Daan Nieboer*[c], *Harrie P. Beerlage*[a,b], *Hessel Wijkstra*[b,d], *Peter F.A. Mulders*[e], *Monique J. Roobol*[c]

[a] *Department of Urology, Jeroen Bosch Hospital, 's-Hertogenbosch, The Netherlands;* [b] *Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands;* [c] *Department of Urology, Erasmus Medical Centre, Rotterdam, The Netherlands;* [d] *Department of Urology, AMC University Hospital, Amsterdam, The Netherlands;* [e] *Department of Urology, Radboudumc University Hospital, Nijmegen, The Netherlands*

Article info

Article history:
Accepted July 21, 2016

Associate Editor:
James Catto

Keywords:
Nomogram
Decision aids
Risk stratification
Validation
Prostate cancer
Biopsy

Abstract

Background: The validity of prediction models needs external validation to assess their value beyond the original development setting.

Objective: To report the diagnostic accuracy of the European Randomized Study of Screening for Prostate Cancer (ERSPC) risk calculator (RC)3 and RC4 in a contemporary Dutch clinical cohort.

Design, setting, and participants: We retrospectively identified all men who underwent prostate biopsy (PBx) in the Jeroen Bosch Hospital, The Netherlands, between 2007 and 2016. Patients were included if they met ERSPC RC requirements of age (50–80 yr), prostate-specific antigen (PSA) (0.4–50 ng/ml), and prostate volume (10–150 ml). The probability of a positive biopsy for prostate cancer (PCa) and significant PCa (Gleason score $\geq$7 and/or higher than T2b) were calculated and compared with PBx pathology results.

Outcome measurements and statistical analysis: Evaluation was performed by calibration, discrimination, and clinical usefulness using calibration plots, area under the receiver operating characteristic curves (AUCs), and decision curve analyses (DCAs), respectively.

Results and limitations: A total of 2270 PBx sessions were eligible for final analysis. Discriminative ability of RC3 (AUC) was 0.78 and 0.90 for any PCa and significant PCa, respectively. For RC4 the calculated AUCs were 0.62 (any PCa) and 0.76 (significant PCa). The calibration plots of RC3 showed good results for both any PCa risk and significant PCa risk. In the repeat PBx group, RC4 tended to underestimate outcomes for PCa and showed moderate calibration for significant PCa. DCA showed an overall net benefit compared with PSA and digital rectal examination (DRE) alone. Limitations of this study are its retrospective single-institution design, retrospectively assessed DRE outcomes, no time restrictions between the first and repeat biopsy sessions, and no anterior sampling in the repeat PBx protocol.

Conclusions: The ERSPC RCs performed well in a contemporary clinical setting. Most pronounced in the biopsy-naive group, both RCs should be favoured over a PSA plus DRE–based stratification in the decision whether or not to perform PBx.

Patient summary: We looked at the ability of the existing European Randomized Study of Screening for Prostate Cancer risk calculator (RC), using different clinical data to predict the presence of prostate cancer in Dutch men. The RC performed well and should be favoured in the decision of whether or not to perform prostate biopsies over the conventional diagnostic pathway.

© 2016 European Association of Urology. Published by Elsevier B.V. All rights reserved.

[‡] These authors contributed equally to this work.

[*] Corresponding author. Jeroen Bosch Hospital, PO Box 90153 5200 ME, 's-Hertogenbosch, The Netherlands. Tel. +31 73 5532407; Fax: +31 73 5532373.
E-mail address: m.gayet@jbz.nl (M. Gayet).

http://dx.doi.org/10.1016/j.euf.2016.07.007
2405-4569/© 2016 European Association of Urology. Published by Elsevier B.V. All rights reserved.

1. Introduction

An estimated 1.1 million men worldwide were diagnosed with prostate cancer (PCa) in 2012, accounting for 15% of cancers in men, with 70% of them in more developed countries. PCa accounts for 6.6% of the total male cancer mortality. Incidence rates diverge, mainly because of serum prostate-specific antigen (PSA) testing [1]. First described in 1979, PSA made large-scale screening for PCa feasible. However, determination of serum PSA for diagnostic purposes lacks accuracy, with 15–25% false negatives and 60% false positives [2,3]. The likelihood of the presence of PCa is therefore preferably estimated by using additional clinical factors, such as digital rectal examination (DRE) and prostate volume (PV).

Although it has been shown that PCa-specific mortality can be reduced by 20% with PSA-based screening, population-based screening programs are not yet acceptable because of the high number needed to screen and the high number needed to treat to avoid one PCa death. More importantly, PSA-based screening results in a considerable number of unnecessary prostate biopsies (PBx) with potentially serious adverse events and leads to considerable overdiagnosis [4,5]. To achieve higher diagnostic accuracy, several nomograms and artificial neural networks (ANNs) have been developed to predict the outcome of PBx. These models have been shown to improve diagnostic accuracy compared with PSA alone [6,7]. However, it is necessary to assess the validity of these models outside the original development setting. Unfortunately, many of the published nomograms and ANNs lack external validation.

In 2006, different risk calculators (RCs) based on the Dutch section of the European Randomized Study of Screening for Prostate Cancer (ERSPC) were developed using data of men with a purely PSA-driven biopsy indication and a random transrectal ultrasound (TRUS)-guided sextant biopsy scheme [8]. ERSPC RC1 and RC2 are for patient use; RC3 (plus DRE), RC4 (plus DRE), RC5, and RC6 are for use by health care professionals at different stages of the testing process. Several external validation studies have been performed for these RCs. In both European and non-European cohorts, the accuracy of prediction of positive PBx in biopsy-naive or previously biopsied men using the ERSPC RC3 or RC4 was assessed, showing area under the curve (AUC) values in the range of 0.71–0.88 [9–12]. Until now, ERSPC RC3 plus DRE and RC4 plus DRE were externally validated using an extended biopsy scheme instead of a sextant biopsy scheme in both a Swiss and Irish cohort, with AUC for PCa and significant PCa of 0.66–0.77 and 0.85, respectively, and showing sufficient to good calibration [13,14].

The aim of this study was to assess the accuracy of the ERSPC RC3 and RC4 in a contemporary Dutch clinical cohort for which biopsy indications and number of biopsies differed from the development cohort.

2. Material and methods

2.1. Study population

We retrospectively identified all men who underwent PBx due to a clinical suspicion of PCa between January 2007 and December 2015 at the Jeroen Bosch Hospital. In our institution PBx was generally performed in patients with a serum PSA level ≥3.0–4.0 μg/l and/or an abnormal DRE. A standardised 12-core biopsy protocol consisting of two biopsies of each base, mid-gland, and apex in the peripheral zone of the prostate was performed, with additional cores taken when needed (eg, in case of hypoechogenic lesions). We examined patient files and obtained relevant clinical and pathologic data of each patient. Patients were included in our study if PCa risk prediction was considered relevant and possible, thus patients aged 50–80 yr with a PSA level between 0.4 and 50 μg/l, PV between 10 and 150 ml, and no previous positive PBx (ie, under active surveillance). Patients with a history of PCa were excluded. For our analyses, we retrospectively converted the descriptively documented DRE findings in our cohort to clinical T stages.

The patient database was blinded by PCa diagnosis and sent to one of the ERSPC RC designers (M.J.R.) for risk outcome calculations. Probabilities of detection of PCa and significant prostate PCa (Gleason score ≥7 and/or T stage higher than T2b) were calculated for each patient individually using two ERSPC RCs (www.prostatecancer-riskcalculator.com). RC3 was used to calculate probabilities in biopsy-naive patients; RC4 was used for patients with previous negative biopsy sessions undergoing a repeat PBx (Supplementary Table 1). The calculated probabilities were subsequently compared with the actual biopsy results for the entire cohort.

2.2. Statistics

Differences between clinical and pathologic variables in the studied cohort were assessed using the chi-square test for categorical variables and the Mann-Whitney *U* test for continuous variables. The performance of both RCs in the clinical setting was assessed by discrimination, calibration, and clinical usefulness.

Discrimination, that is, predictive accuracy, was quantified using the receiver operating characteristics derived AUC. Calibration refers to the agreement between observed and predicted outcomes with the extent of risk of over- or underestimation of the RCs evaluated graphically using calibration plots [15].

Clinical usefulness of the RCs was evaluated by decision curve analyses (DCAs) as described previously by Vickers and Elkin and by Steyerberg et al [16,17]. DCAs determine the value (net benefit) of a prediction model by examining the theoretical relationship between the threshold probability of an event (eg, PCa at biopsy) and the relative value of false-positive and false-negative results. We compared the RC model with a PSA plus DRE–based model, also developed on original ERSPC data. We also assessed the theoretical number of (significant) cases of PCa missed, numbers of biopsies saved, and number of Gleason score 6 PCa diagnoses saved at different RC thresholds.

Statistical analyses were performed using SPSS v23.0 (IBM Corp, Armonk, NY, USA) and R v3.2.5 (R Foundation for Statistical Computing, Vienna, Austria). A $p < 0.05$ was considered to indicate statistical significance in all analyses.

3. Results

We identified 2862 prostate biopsy sessions in 2124 men. Overall, 426 biopsy sessions were omitted due to the predefined inclusion criteria. In 166 biopsy sessions (<6%), data were incomplete (PSA, $n = 2$; DRE findings, $n = 123$; TRUS PV, $n = 27$; TRUS findings, $n = 40$) and excluded from further analyses. As a result, 2270 prostate biopsy sessions (79.3%) in 1812 different men were eligible for final analysis: 73.0% biopsy-naive men and 27.0% men with a prior negative PBx.

PCa and significant PCa were detected in 44.1% and 20.3% of the biopsy-naive men ($n = 1658$). Men with PCa and

Table 1 – Clinical and pathologic characteristics of the biopsy-naive patient cohort

Variable	Total cohort	Positive biopsy		Negative biopsy
		All PCa, *p* value	Significant PCa, *p* value	
No. of patients, *n* (% total cohort)	1658 (100)	732 (44.1)	337 (20.3)	926 (59.1)
Age, yr, median (IQR)	64 (60–69)	66 (62–71), **<0.001**	68 (63–73), **<0.001**	64 (59–68)
Age, yr, *n* (%)				
50 to <60	369 (22.3)	128 (17.5)	47 (13.9)	241 (26.0)
60 to <70	904 (54.5)	377 (51.5)	148 (44.0)	527 (56.9)
≥70	385 (23.2)	227 (31.0)	142 (42.1)	158 (17.1)
PSA level, μg/l, median (IQR)	7.6 (5.9–11.0)	8.6 (6.2–14.0), **<0.001**	11.0 (7.6–20.0), **<0.001**	7.1 (5.6–9.4)
PSA ranges, μg/l, *n* (%)				
<1	6 (0.4)	1 (0.1)	0 (0.0)	5 (0.5)
1 to <4.0	91 (5.5)	33 (4.5)	9 (2.7)	58 (7.3)
4.0–10.0	1082 (65.2)	416 (56.9)	135 (40.0)	666 (70.9)
≥10.0	479 (28.9)	282 (38.5)	193 (57.3)	197 (21.3)
DRE findings, *n* (%)				
Normal	1198 (72.3)	386 (52.7)	73 (21.7)	812 (87.7)
Abnormal	460 (27.7)	346 (47.3), **<0.001**	264 (78.3), **<0.001**	114 (12.3)
TRUS prostate volume, ml, median (IQR)	40.0 (30.0–55.0)	35.0 (28.0–46.0), **<0.001**	33.0 (27.0–44.0), **<0.001**	46.0 (35.0–60.0)
TRUS findings, *n* (%)				
Normal	1267 (76.4)	461 (63.0)	147 (43.6)	806 (87.0)
Abnormal	391 (23.6)	271 (37.0), **<0.001**	190 (56.4), **<0.001**	120 (13.0)
Total cores taken at biopsy, *n* (%)				
<12	24 (1.4)	13 (1.8)	7 (2.1)	11 (1.2)
12	1379 (83.2)	609 (83.2)	269 (79.8)	770 (83.2)
>12	255 (15.4)	110 (15.0), 0.504	61 (18.1), 0.653	145 (15.7)

DRE = digital rectal examination; IQR = interquartile range; PCa = prostate cancer; PSA = prostate-specific antigen; TRUS = transrectal ultrasound.
The *p* values in bold indicate a statistically significant difference between two groups (positive biopsy; all PCa vs negative biopsy and positive biopsy; significant PCa vs negative biopsy).

significant PCa were significantly older compared with men with no cancer detected, had higher PSA levels, lower PVs, and were more likely to have an abnormal DRE and TRUS (Table 1). In the previously biopsied men ($n = 612$), 25.8% and 7.0% of men were diagnosed with PCa or significant PCa, respectively. Men with PCa and significant PCa detected were significantly older and had lower PVs compared with men in whom no PCa was detected. Their PSA level did not differ from men with no PCa detected. Only men with significant PCa disease were more likely to have an abnormal DRE or TRUS (Table 2).

AUC for the discrimination of (significant) PCa in the biopsy-naive group (RC3) was 0.78 (95% confidence interval [CI], 0.76–9.80) for PCa and 0.91 (95% CI, 0.89–0.92) for significant PCa. The discriminative ability of the RC4 in the repeat PBx group was lower with an AUC of 0.62 (95% CI, 0.56–0.67) for any PCa and 0.74 (95% CI, 0.66–0.81) for significant PCa.

In the biopsy-naive group, mean predicted outcomes were close to the observed outcomes of PCa and significant PCa (Fig. 1). The calibration plots showed good results for both outcomes over the whole prediction range, reflected in the calibration-in-the-large of 0.18 (95% CI, 0.08–0.31) and a calibration slope of 0.92 (95% CI, 0.81–1.02) for PCa and calibration-in-the-large of −0.15 (95% CI, −0.30 to 0.01) and a calibration slope of 1.25 (95% CI, 1.12–1.38) for significant PCa (Fig. 1).

In the repeat PBx group, the RC tended to underestimate outcomes for PCa and showed moderate calibration for significant PCa in the low-risk range between 0 and 0.15. Calibration-in-the-large was 0.49 (95% CI, 0.31–0.67) and 0.31 (95% CI, 0.01–0.62) with a calibration slope of 0.64 (95% CI, 0.37–0.90) and 0.80 (95% CI, 0.55–1.06) for PCa and significant PCa, respectively (Fig. 1).

The net benefit in the biopsy-naive group, assessed with DCA, was highest for the RC over the whole probability range, as compared with a PSA plus DRE–based strategy for PCa and for significant PCa (Fig. 2). A threshold algorithm (≥20.0% for PCa at biopsy or 12.5–20.0% for PCa at biopsy with >4% for significant PCa at biopsy) presented by the ERSPC RC developers (www.prostatecancer-riskcalculator.com) would result in 20% ($n = 337$) fewer biopsies in our cohort. As a consequence we would miss 7% ($n = 52$) of the PCa of which 12% ($n = 6$) is significant PCa. The diagnosis of Gleason score 6 PCa would be spared in 46 men (Supplementary Table 2).

In the repeat PBx group, DCA provided a net benefit for the RC in the threshold probability range from 20% to 35% but also a net harm compared with the "biopsy all" line at the lower risk thresholds for PCa. For significant PCa, the RC provided a small net benefit in the lowest threshold probability range compared with a PSA plus DRE–based strategy (Fig. 2). The threshold algorithm (≥20.0% for PCa at biopsy or 12.5–20.0% for PCa at biopsy with >3% for significant PCa at biopsy) presented by the ERSPC RC developers would result in 47% ($n = 285$) fewer biopsies with 35% ($n = 55$) of the PCa missed of which 18% ($n = 10$) is significant PCa. The diagnosis of Gleason score 6 PCa would be spared in 45 men (Supplementary Table 3).

4. Discussion

The ERSPC RCs quantify the chance of finding PCa on sextant biopsy by translating the presence or absence of abnormal

Table 2 – Clinical and pathologic characteristics of the repeat prostate biopsy patient cohort

Variable	Total cohort	Positive biopsy		Negative biopsy
		All PCa, *p* value	Significant PCa, *p* value	
No. of patients, *n* (% total cohort)	612 (100)	158 (25.8)	43 (7.0)	454 (74.2)
Age, yr, median (IQR)	65 (61–69)	66 (62–71), **0.002**	69 (64–74), **<0.001**	64 (60–69)
Age, yr, *n* (% col)				
50 to <60	122 (19.9)	24 (15.2)	3 (7.0)	98 (21.6)
60 to <70	347 (56.7)	83 (52.5)	19 (44.2)	264 (58.1)
≥70	143 (23.4)	51 (32.3)	21 (48.8)	92 (20.3)
Repeat biopsy session				
First	440 (71.9)	122 (77.2)	32 (74.4)	318 (70.0)
Second	121 (19.8)	26 (16.5)	10 (23.3)	95 (20.9)
Third	34 (5.6)	7 (4.4)	1 (2.3)	27 (5.9)
Fourth or more	17 (2.9)	3 (1.9)	0	14 (3.1)
Time interval, mo*, median (IQR)	16 (7–39)	16 (7–40)	10 (32–51)	16 (7–39)
Time interval, yr*, *n* (%)				
≤1 yr	257 (42.3)	65 (41.1)	12 (27.9)	192 (42.8)
>1 to ≤2 yr	116 (19.1)	31 (19.6)	6 (14.0)	85 (18.9)
>2 to ≤3 yr	72 (11.9)	20 (12.7)	8 (18.6)	52 (11.6)
>3 yr	162 (26.7)	42 (26.6)	17 (39.5)	120 (26.7)
PSA level, µg/l, median (IQR)	9.6 (7.4–14.0)	8.9 (7.2–14.1), **0.369**	10.0 (7.3–18.0), **0.323**	9.6 (7.5–14.0)
PSA ranges, µg/l, *n* (%)				
<1	1 (0.2)	0 (0.0)	0 (0.0)	1 (0.2)
1 to <4.0	7 (1.1)	1 (0.6)	1 (2.3)	6 (1.3)
4.0–10.0	327 (53.4)	91 (57.6)	20 (46.5)	236 (52.0)
≥10.0	277 (45.3)	66 (41.8)	22 (51.2)	211 (46.5)
DRE findings, *n* (%)				
Normal	431 (70.4)	107 (67.7)	16 (37.2)	324 (71.4)
Abnormal	181 (29.6)	51 (32.3), 0.388	27 (62.8), **<0.001**	130 (28.6)
TRUS prostate volume, ml, median (IQR)	48.0 (35.3–66.0)	43.0 (30.0–58.3), **<0.001**	43.0 (30.0–52.0), **0.006**	50.0 (38.0–68.0)
TRUS findings, *n* (%)				
Normal	479 (78.3)	124 (78.5)	28 (65.1)	355 (78.2)
Abnormal	133 (21.7)	34 (21.5), 0.940	15 (34.9), 0.051	99 (21.8)
Total cores taken at biopsy, *n* (%)				
<12	12 (2.0)	3 (1.9)	1 (2.3)	9 (2.0)
12	391 (63.9)	93 (58.9)	26 (60.5)	298 (65.5)
>12	209 (34.1)	62 (39.2), 0.125	16 (37.2), 0.414	147 (32.5)

DRE = digital rectal examination; IQR = interquartile range; PCa = prostate cancer; PSA = prostate-specific antigen; TRUS = transrectal ultrasound.
The *p* values in bold indicate a statistically significant difference between two groups (positive biopsy; all PCa vs negative biopsy and positive biopsy; significant PCa vs negative biopsy).
* Time interval in months/years: Time in months/years between previous biopsy session and the repeat biopsy session used for the analysis.

findings into a probability. PBx is recommended by the ERSPC RC developers at a probability threshold ≥20% for PCa and can be considered in the 12.5–20.0% threshold range for PCa, especially if significant PCa probabilities reach >4% or >3% for the RC3 and RC4, respectively. Both threshold groups were assessed using DCA and showed good to moderate clinical benefit in biopsy-naive and previously biopsied men, respectively. The proposed threshold algorithm for biopsy-naive men (≥20.0% for PCa or ≥12.5% for PCa and >4% for significant PCa) seems to be acceptable. Only 6 significant cases of PCa (2%) are being missed as a result of saving 337 biopsy sessions (20%). However, the threshold algorithm for previously biopsied men (≥20.0% for PCa or ≥12.5% for PCa and >3% for significant PCa) is not optimal for our cohort. With 285 fewer biopsies (47%), 10 significant cases of PCa (23%) are being missed. A threshold algorithm ≥10% for PCa or >2% for significant PCa seems optimal as only 1 significant PCa (2%) is being missed and 131 biopsy sessions (21%) are still saved (Supplementary Table 3).

AUCs of the original sextant biopsy scheme ERSPC data are 0.79 and 0.86 for RC3 and 0.68 and 0.80 for RC4 for all PCa and significant PCa, respectively [18]. In the present study, diagnostic accuracy reaches that of the original ERSPC report, especially for RC3. This is in contrast to previous external validation reports of these RCs that reported various, mostly lower diagnostic accuracy results [9–12]. This may be partially the result of ethnic similarities and (significant) PCa prevalence resemblance between the original development and validation cohort [18]. Previous validation studies assessed a combined AUC for both RC3 and RC4, whereas our study demonstrates that accuracy of both RCs differed substantially. Consequently, a suboptimally performing RC can mask the potentially high accuracy of another RC.

In comparison with the new RC3 plus DRE and RC4 plus DRE, our present study also demonstrates higher diagnostic accuracy [13,14]. One major reason for this may be the use of TRUS-based PV measurements. Predictions of PCa using TRUS-based PV data have been shown to outperform predictions using DRE-based data because PV was demonstrated to be an important predictor in the detection of PCa [18,19]. Redistribution of TRUS-based volume in DRE-stated categories could therefore have negatively influenced

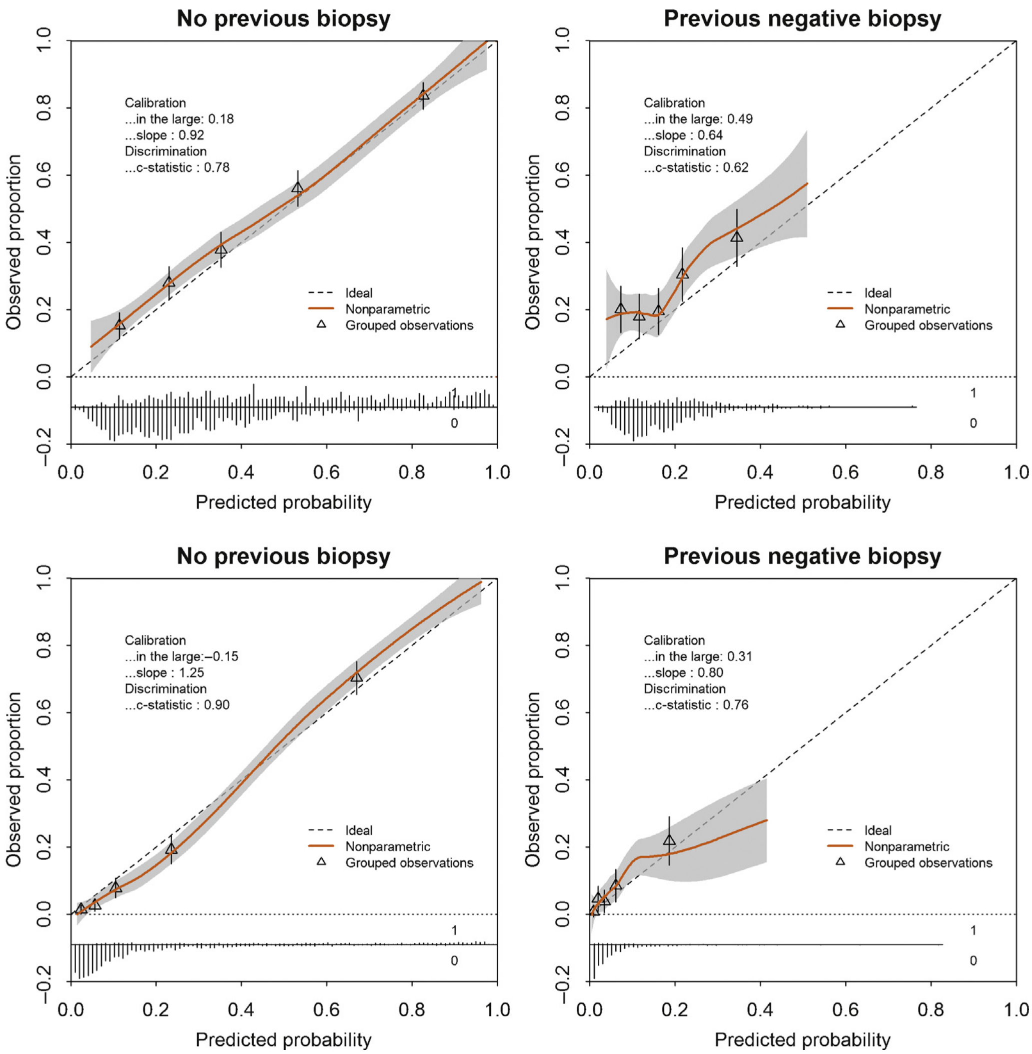


Fig. 1 – Calibration plots for the European Randomised Study of Screening for Prostate Cancer risk calculator (RC) 3 (left) and RC4 (right) demonstrating the agreement between observed and predicted probabilities for prostate cancer (PCa) at biopsy (upper figures) or significant PCa at biopsy (lower figures). The ideal plot is showed with a dashed line through the origin. The solid line reflects the relation between observed and predicted probability with quintiles of grouped patients shown by triangles. The numbers of patients with and without the condition are shown as spikes along the *x*-axis.
PCa = prostate cancer; RC = risk calculator.

predictive accuracy in the external ERSPC validation by Poyet et al [13]. Trottier et al also demonstrated in their external validation that, on multivariate analysis, TRUS-measured PV and TRUS lesion were the most important risk model predictors for a positive (significant) PCa diagnosis at biopsy [12].

In this study, calibration plots showed good outcomes for biopsy-naive men. In the repeat PBx group, mean predicted and mean observed outcome disagreed and tended to underestimate the PBx results. However, there was moderate calibration for significant PCa in the clinically relevant low-risk range between 0 and 0.15. One of the contributing components for this moderate calibration (and the underestimation of PBx results) could be the fact that men with a previous negative PBx were a higher risk cohort compared with the cohort within the ERSPC, in which every man with PSA $\geq$3.0 ng/ml at repeat screening was biopsied again. In addition, RC4 is developed using a cohort with at least 4 yr between initial and repeat PBx. In our analysis, there was no minimal maintained time range between first and repeat PBx with a median time of 16.5 mo (range: 7.0–39.0 mo) between both biopsy sessions.

PSA testing and screening remain a subject of debate. On the one hand, evidence indicates that it reduces PCa mortality, but on the other hand, there is a risk of overdiagnosis and overtreatment. Offering PSA in an organised way and combining it with other relevant risk factors regarding PCa will most likely result in a more beneficial harm–benefit ratio [20–22]. Our study confirms that individual risk assessment using a multivariable prediction model should be used in the consideration whether or not to perform PBx.

Limitations of this study are its retrospective single-institution design. Because of the lack of documentation of

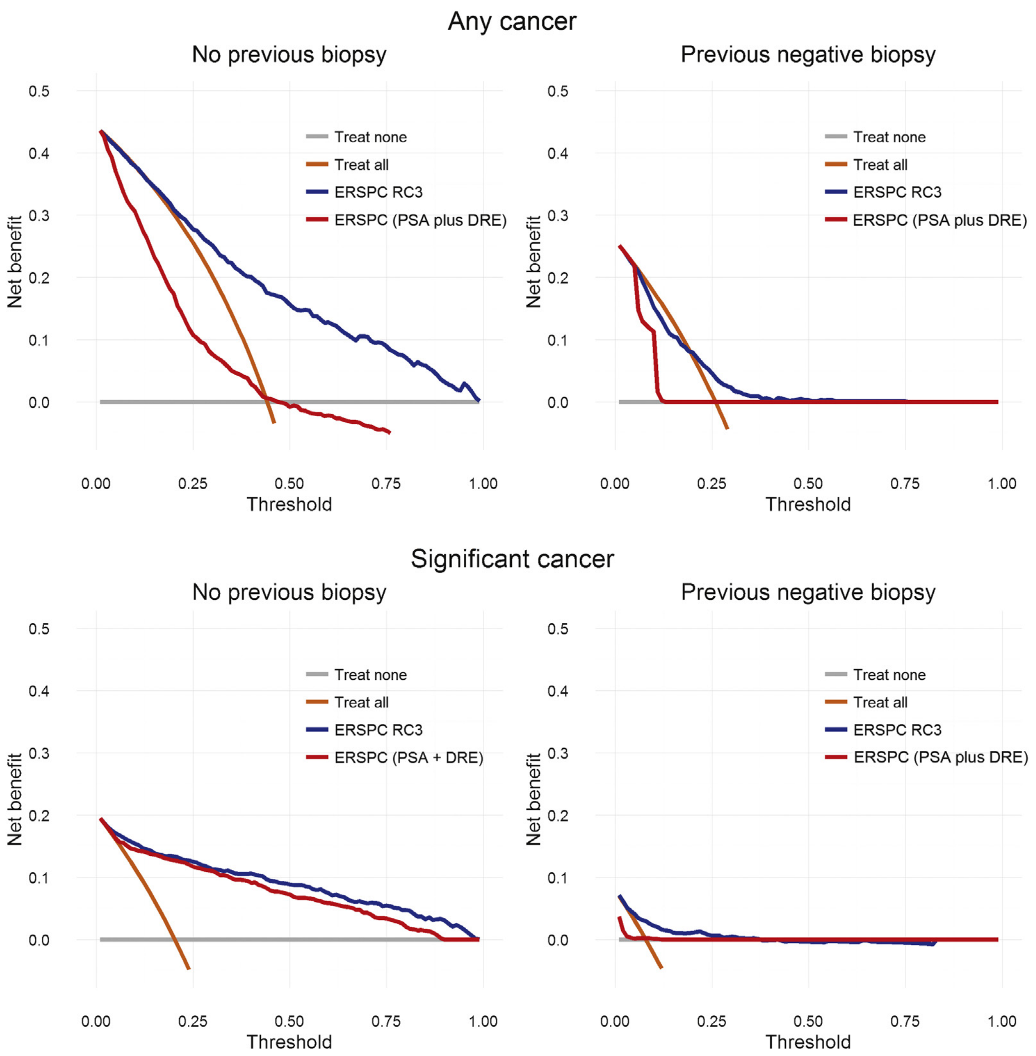


Fig. 2 – Decision curve analysis demonstrating the net benefit of the risk calculator (blue line) and prostate-specific antigen plus digital rectal examination (red line) for prostate cancer (PCa) at biopsy (upper curves) or significant PCa at biopsy (lower curves).
DRE = digital rectal examination; ESRPC = European Randomised Study of Screening for Prostate Cancer; PCa = prostate cancer; PSA = prostate-specific antigen; RC = risk calculator.

T-stage DRE, DRE was retrospectively converted to T stage, which could have created biases in either direction (understaging or overstaging). Also, as mentioned earlier, there was no minimal maintained time range between the first and repeat biopsy sessions. However, most early cases of PCa (T1–T2) have an indolent course for 10–15 yr [23]. The difference in time range between the first and repeat biopsy sessions is unlikely to have influenced our results. Lastly, in repeat PBx sessions, the initial sampling protocol and not an extended sampling protocol including anterior PBx was used.

With new biomarkers and imaging techniques available, the field of diagnosis of PCa is changing. In the last 2 yr, several groups combined new biomarkers and clinical features into prediction models and compared accuracy of these models with ERSPC or Prostate Cancer Prevention Trial RCs, and they concluded that diagnostic accuracy increased compared with the conventional RCs [24–27]. ERSPC RC4 was also used to predict the outcome of multiparametric magnetic resonance imaging (mpMRI) by Alberts et al [28]. Patient selection using a RC can avoid half of the mpMRIs after a prior negative biopsy.

Hence, for future perspectives, it would be of interest to develop and validate RCs that include new diagnostic means, such as biomarkers, or develop RCs that predict outcomes of mpMRI, for example. Such an individualized approach for PCa detection could reduce the adverse effects of our diagnostic approaches and/or treatments.

5. Conclusions

In our external validation of the screening-based ERSPC RCs for both biopsy-naive and previously biopsied men, both RCs showed net benefit in our clinical setting compared with a PSA plus DRE–based strategy that was most pronounced in the biopsy-naive group. Multivariate risk stratification should be favoured in the decision whether or not to perform PBx.

Author contributions: Maudy Gayet 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: Roobol, Beerlage, Wijkstra.
Acquisition of data: Gayet, Mannaerts.
Analysis and interpretation of data: Nieboer, Gayet, Mannaerts.
Drafting of the manuscript: Gayet, Mannaerts.
Critical revision of the manuscript for important intellectual content: Roobol, Beerlage, Wijkstra, Mulders.
Statistical analysis: Nieboer, Roobol.
Obtaining funding: Beerlage.
Administrative, technical, or material support: None.
Supervision: Roobol, Beerlage, Wijkstra, Mulders.
Other (specify): None.

Financial disclosures: Maudy Gayet 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: None.

Funding/Support and role of the sponsor: The present study was undertaken with a research grant from Astellas Pharma Netherlands B.V. Astellas did not influence the content of this manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.euf.2016.07.007.

References

[1] Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–86.
[2] 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.
[3] Gambert SR. Screening for prostate cancer. Int Urol Nephrol 2001;33:249–57.
[4] Hugosson J, Carlsson S, Aus G, et al. Mortality results from the Goteborg randomised population-based prostate-cancer screening trial. Lancet Oncol 2010;11:725–32.
[5] Schroder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. Lancet 2014;384:2027–35.
[6] Schroder F, Kattan MW. The comparability of models for predicting the risk of a positive prostate biopsy with prostate-specific antigen alone: a systematic review. Eur Urol 2008;54:274–90.
[7] Louie KS, Seigneurin A, Cathcart P, Sasieni P. Do prostate cancer risk models improve the predictive accuracy of PSA screening?. A meta-analysis. Ann Oncol 2015;26:848–64.
[8] Steyerberg EW, Roobol MJ, Kattan MW, van der Kwast TH, de Koning HJ, Schroder FH. Prediction of indolent prostate cancer: validation and updating of a prognostic nomogram. J Urol 2007;177:107–12, discussion 112.
[9] van Vugt HA, Roobol MJ, Kranse R, et al. Prediction of prostate cancer in unscreened men: external validation of a risk calculator. Eur J Cancer 2011;47:903–9.
[10] Cavadas V, Osorio L, Sabell F, Teves F, Branco F, Silva-Ramos M. Prostate cancer prevention trial and European randomized study of screening for prostate cancer risk calculators: a performance comparison in a contemporary screened cohort. Eur Urol 2010;58:551–8.
[11] Yoon DK, Park JY, Yoon S, et al. Can the prostate risk calculator based on Western population be applied to Asian population? Prostate 2012;72:721–9.
[12] Trottier G, Roobol MJ, Lawrentschuk N, et al. Comparison of risk calculators from the Prostate Cancer Prevention Trial and the European Randomized Study of Screening for Prostate Cancer in a contemporary Canadian cohort. BJU Int 2011;108:E237–44.
[13] Poyet C, Nieboer D, Bhindi B, et al. Prostate cancer risk prediction using the novel versions of the European Randomised Study for Screening of Prostate Cancer (ERSPC) and Prostate Cancer Prevention Trial (PCPT) risk calculators: independent validation and comparison in a contemporary European cohort. BJU Int 2016;117:401–8.
[14] Foley RW, Lundon DJ, Murphy K, Murphy TB, Galvin DJ, Watson RW. Predicting prostate cancer: analysing the clinical efficacy of prostate cancer risk calculators in a referral population. Irish J Med Sci 2015;184:701–6.
[15] Steyerberg EW. Clinical prediction models. A practical approach to development, validation, and updating. New York, NY: Springer; 2009.
[16] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.
[17] Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38.
[18] Roobol MJ, van Vugt HA, Loeb S, et al. Prediction of prostate cancer risk: the role of prostate volume and digital rectal examination in the ERSPC risk calculators. Eur Urol 2012;61:577–83.
[19] Colleselli D, Bektic J, Schaefer G, et al. The influence of prostate volume on prostate cancer detection using a combined approach of contrast-enhanced ultrasonography-targeted and systematic grey-scale biopsy. BJU Int 2007;100:1264–7.
[20] Roobol MJ. Unorganized prostate-specific antigen-based screening for prostate cancer: more harm than benefit. When will we finally start to implement guidelines and risk assessment tools in clinical practice? Eur Urol 2015;68:363–4.
[21] Loeb S. Prostate biopsy decisions: one-size-fits-all approach with total PSA is out and a multivariable approach with the Prostate Health Index is in. BJU Int 2016;117:383.
[22] Vickers A, Carlsson S, Laudone V, Lilja H. It ain't what you do, it's the way you do it: five golden rules for transforming prostate-specific antigen screening. Eur Urol 2014;66:188–90.
[23] Johansson JE, Andren O, Andersson SO, et al. Natural history of early, localized prostate cancer. JAMA 2004;291:2713–9.
[24] Leyten GH, Hessels D, Jannink SA, et al. Prospective multicentre evaluation of PCA3 and TMPRSS2-ERG gene fusions as diagnostic and prognostic urinary biomarkers for prostate cancer. Eur Urol 2014;65:534–42.
[25] Foley RW, Gorman L, Sharifi N, et al. Improving multivariable prostate cancer risk assessment using the Prostate Health Index. BJU Int 2016;117:409–17.
[26] Van Neste L, Hendriks RJ, Dijkstra S, et al. Detection of high-grade prostate cancer using a urinary molecular biomarker-based risk score. Eur Urol. In press. http://dx.doi.org/10.1016/j.eururo.2016.04.012.
[27] Gronberg H, Adolfsson J, Aly M, et al. Prostate cancer screening in men aged 50-69 years (STHLM3): a prospective population-based diagnostic study. Lancet Oncol 2015;16:1667–76.
[28] Alberts AR, Schoots IG, Bokhorst LP, van Leenders GJ, Bangma CH, Roobol MJ. Risk-based patient selection for magnetic resonance imaging-targeted prostate biopsy after negative transrectal ultrasound-guided random biopsy avoids unnecessary magnetic resonance imaging scans. Eur Urol 2016;69:1129–34.