Clinical-Prostate cancer

Reducing unnecessary biopsies while detecting clinically significant prostate cancer including cribriform growth with the ERSPC Rotterdam risk calculator and 4Kscore

Jan F.M. Verbeek, M.D., M.Sc.[a,*], Chris H. Bangma, M.D., Ph.D.[a],
Charlotte F. Kweldam, M.D., Ph.D.[b], Theodorus H. van der Kwast, M.D., Ph.D.[c],
Intan P. Kümmerlin, M.D.[b], Geert J.L.H. van Leenders, M.D.[b], Monique J. Roobol, Ph.D.[a]

[a] Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands
[b] Department of Pathology, Erasmus University Medical Center, Rotterdam, The Netherlands
[c] Department of Pathology, Princess Margaret Cancer Center, University Health Network, Toronto, Ontario, Canada

Abstract

Introduction: The use of risk calculators predicting clinically significant prostate cancer (csCaP) on biopsy reduces unnecessary biopsies and overdiagnosis of indolent disease compared to a Prostate Specific Antigen (PSA) strategy. Updating these tools using more specific outcome measures and contemporary predictors could potentially lead to further reductions. Our objective was to assess clinical impact of the 4 kallikrein (4K) score, the Rotterdam Prostate Cancer Risk Calculator (RPCRC), and the combination of both for predicting csCaP based on the latest International Society of Urological Pathology grading system and cribriform growth pattern.

Materials and methods: Our prospective cohort consisted of 2,872 men from the first screening round in the European Randomized Study of Screening for Prostate Cancer Rotterdam; biopsy indication PSA ≥ 3.0. The predictive performance of the 4Kscore, RPCRC, and the combination of RPCRC with 4Kscore were assessed with area under the receiver operator characteristic curve (AUC) and calibration plots. Decision curve analysis was used to evaluate the reduction of unnecessary biopsy and indolent CaP.

Results: The csCaP was present in 242 (8%) men, and indolent CaP in 578 (20%). The 4Kscore and RPCRC had similar high AUCs (0.88 vs. 0.87; $P = 0.41$). The 4Kscore-RPCRC combination improved AUC to 0.89 compared to 4Kscore ($P < 0.01$) and RPCRC ($P < 0.01$). The RPCRC and 4Kscore reduced the number of Bx with 42 and 44, respectively, per 100 men at risk compared to a ≥PSA 3.0 strategy without increasing missed csCaP. The RPCRC-4Kscore combination resulted in a slight additional net reduction of 3.3 biopsies per 100 men.

Conclusions: The RPCRC and 4Kscore had similar reductions of unnecessary biopsies and overdiagnosis of indolent disease. Combination of both models slightly reduced unnecessary biopsies further. Gain in net benefit must, however, be weighed against additional costs and availability of tests. © 2018 Elsevier Inc. All rights reserved.

1. Introduction

Prostate Specific Antigen (PSA)-based prostate cancer (CaP) screening is beneficial in terms of mortality reduction, however its main drawbacks are overdiagnosis and overtreatment of indolent CaP [1]. A more and more used strategy to limit overtreatment of indolent CaP cancer is the

use of active surveillance as initial treatment [2]. To improve screening efforts further, a balance must be found between minimizing overdiagnosis, and optimizing the timely detection of potentially deadly disease [3]. Preferably only men with a clinically significant CaP (csCaP) should be identified and diagnosed. In addition, the therapeutic intervention of choice should have minimal impact on quality of life, to maximize the gain in quality adjusted life years.

For several decades, multivariable prediction tools have been constructed with the aim of selectively predicting the presence of csCaP on biopsy [4−6]. The Rotterdam Prostate Cancer Risk Calculator (RPCRC) is an example of these and is available as an app. It was developed on the basis of the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer (ERSPC). The RPCRC has demonstrated its value in predicting CaP risk and in avoiding unnecessary testing [4,7,8]. However, risk calculators are not perfect and with the development of new biomarkers, potential updates, and adaptations to the models remain a topic of ongoing research [9]. Last year, we improved the RPCRC with a better understanding of disease burden by incorporating the latest International Society of Urological Pathology (ISUP) grading system, and invasive cribriform tumor (CR) and intraductal carcinoma (IDC) growth patterns. We found a considerable decrease in unnecessary prostate biopsies and overdiagnosis of potentially indolent disease [10]. These growth patterns subdivide indolent and csCaP within the Gleason score 7 (3 + 4) CaP [11]. Also the 2014 ISUP grading of CaP provided more accurate stratification of tumors than the previous system [12]. The ISUP grading Gleason score 6 and below CaP were condensed in a single ISUP grade 1 category and Gleason score 7 CaP were split into 2 categories, ISUP grade 2 for Gleason score 7 (3 + 4) and ISUP grade 3 for Gleason score 7 (4 + 3). Furthermore, now all CR and glomeruloid pattern carcinoma are included in the Gleason grade 4 pattern, whereas IDC has been excluded from grading.

In the field of novel biomarkers for CaP detection, the 4-kallikrein (4K) panel comes forth in outperforming total PSA and consists of total PSA, fPSA, intact PSA, and kallikrein-related peptidase 2 (hK2) [13]. The 4K panel, together with clinical parameters (age and outcome of digital rectal examination [DRE]) have been combined into a multivariable model (4Kscore). In several large prospective studies, the 4Kscore considerably reduced unnecessary biopsies without missing many csCaP [13−16]. These findings and continuous improvements may further improve the promoted risk-adapted strategy in the prediction of csCaP for an individual patient [17]. The aim of this study was to evaluate the clinical impact of the 4Kscore, RPCRC, and the combination of both in predicting the improved definition of csCaP based on the latest ISUP grading system and inclusion of cribriform growth pattern in Gleason 4 CaP.

2. Materials and methods

2.1. Patients

We studied men who were biopsied due to a PSA ≥ 3.0 ng/ml in the first round of the screening program of the European Randomized Study of Screening for Prostate Cancer (ERSPC) section Rotterdam [11]. The inclusion period for the randomized screening trial was November 1993 to March 2000; the detailed protocol and ethical approval has been described previously [4]. We included 3,028 men who received sextant biopsies solely on the basis of an elevated PSA (PSA ≥3.0 ng/ml). We did not include men with PSA <3.0 ng/ml who had abnormal DRE or hypoechogenic lesion on transrectal ultrasound (TRUS). This cohort was also used in a previously published 4Kscore validation study [15]. Clinical measurements (age, PSA, DRE, and prostate volume) were obtained prospectively.

In 2015, the prostate biopsy slides were re-evaluated in order to adapt to the 2014 ISUP grading system and to record CR and/or IDC components [10,11]. The urogenital pathologist and 2 pathologists in training were blinded to patient information and biopsy outcome. The primary outcome was detection of csCaP on biopsy defined as ISUP grade 2 CaP with CR/IDC plus all CaP with ISUP grade ≥3. Secondary outcome was the detection of indolent CaP defined as ISUP grade 1 or 2 without CR/IDC [10]. We excluded 114 men whose 4K panel could not be measured due to insufficient frozen blood samples, and 42 men whose histopathology slides could not be re-evaluated. Total of 2,872 men could be used in the analysis.

2.2. The 4K panel

The 4K panel (total PSA, free PSA, intact PSA, and human kallikrein-related 2) was measured in frozen serum samples at the Wallenberg Research Laboratories, Department of Laboratory Medicine, Lund University, University Hospital UMAS in Malmö, Sweden, using the dual-label DELFIA Prostatus total/free PSA-Assay (Perkin−Elmer, Turku, Finland) [13].

2.3. Risk prediction models

The RPCRC was applied to provide the probabilities of csCaP and indolent CaP on a biopsy. These probabilities were calculated by the developed multinomial logistic regression analysis with PSA, prostate volume estimated through DRE, DRE abnormalities, and age as predictors [10]. Since no data on prostate volume assessed with DRE was available in this validation cohort, the available TRUS-assessed prostate volume was recoded into 3 volume classes as can be estimated by DRE [7]. TRUS-assessed volumes <30 cc were recoded as 25 cc, volumes between 30 and 50 cc as 40 cc, and

volumes >50 cc as 60 cc. The effect of interobserver variability of DRE volume estimation on the performance of the RPCRC was externally validated and was negligible due to substantial agreement in DRE volume estimation. In addition, there was a good correlation between the TRUS-assessed volume and DRE-estimated volume [18].

The 4Kscore is an algorithm constructed from the 4K panel and clinical parameters. Originally, a laboratory model and clinical model were developed by Vickers et al. [13]. The clinical model consists of the 4K panel, age, and DRE outcome (4Kscore). A blinded dataset for outcome has been sent to the 4Kscore developers in order to receive the 4Kscore probability of csCaP on biopsy. To update the RPCRC with the 4Kscore, we first recalibrated the 4Kscore by re-estimating the intercept and slope of the linear predictor, and, subsequently, we added the 4Kscore as a predictor [19].

2.4. Statistical analysis

The predictive performance of each risk model was evaluated according to the area under the receiver operator characteristic curve (AUC). Correction for overestimation was done by bootstrapping techniques using 1,000 samples. Differences in AUCs were tested after calculation of the standard error of the AUC with the DeLong method [20]. Calibration was assessed by grouping men by deciles of absolute risk. The observed and expected counts of incident CaP cases in each decile were compared for deviance and significance calculated according to the Hosmer−Lemeshow statistic [21]. Finally, reduction of unnecessary biopsy by 4Kscore, RPCRC, and combination of both was assessed with decision curve analysis and the net benefit (NB) formula [22]. The additional reduction of unnecessary biopsy

without missing any csCaP for the 4Kscore, RPCRC, and the 4Kscore-RPCRC combination compared to the defaults strategy (biopsy PSA ≥ 3.0 ng/ml) was calculated with the net avoided false positives formula per 100 men at risk: $\frac{\Delta NB}{odds(T)} * 100$, where ΔNB is the NB difference between the models and the default strategy; T = risk threshold. By inversing $\frac{\Delta NB}{odds(T)}$, it was possible to assess whether the reduction of unnecessary biopsies would outweigh the costs associated using the models [23]. Baseline characteristics were tested with the nonparametric Kruskal−Wallis test for continuous variables, and with chi-square for categorized values. P values < 0.05 were taken to indicate statistical significance. Statistical computations were performed with R, version 3.4.0 (R Foundation for Statistical Computing, Vienna, Austria).

3. Results

Baseline characteristics of the 2,872 men stratified by biopsy outcome are presented in Table 1. A total of 578 (20%) indolent CaP cases and 242 (8%) csCaP cases were detected. Men with csCaP had higher 4K values ($P < 0.01$), smaller prostate volume ($P < 0.01$), and more abnormal digital rectal findings ($P < 0.01$) compared to men without CaP. The 4Kscore and RPCRC predicted csCaP with a similar high AUC (0.88 vs. 0.87; $P = 0.41$). The 4Kscore-RPCRC combination improved AUC to 0.89 compared to 4Kscore ($P < 0.01$) and RPCRC ($P < 0.01$), as shown in Fig. 1. At a 5% risk threshold, the 4Kscore-RPCRC combination had a sensitivity of 88% and specificity of 71%. Calibration was assessed with the Hosmer−Lemeshow test and indicated a good calibration of the combined 4Kscore with RPCRC ($P = 0.09$), see Appendix Fig. A.1.

Table 1
Patient and tumor characteristics stratified to groups without and with indolent or clinically significant prostate cancer

Characteristic	No CaP $n = 2,052$ (72%)	Indolent CaP $n = 578$ (20%)	Clinically significant CaP $n = 242$ (8%)
Age, y, median (IQR)	66 (62−70)	66 (62−70)	69 (64−72)
Kallikrein panel, ng/ml, median (IQR)			
PSA	4.8 (3.9−6.4)	5.5 (4.0−7.9)	9.8 (4.9−17.0)
Free PSA	1.1 (0.8−1.5)	1.0 (0.7−1.4)	1.1 (0.7−1.9)
Intact PSA	0.5 (0.4−0.7)	0.5 (0.4−0.7)	0.7 (0.4−1.1)
hK2	0.07 (0.05−0.09)	0.08 (0.05−0.11)	0.11 (0.07−0.18)
Prostate volume, ml, median (IQR)	48 (38−63)	39 (29−48)	36 (29−48)
Abnormal DRE, n (%)	424 (21)	201 (35)	167 (69)
ISUP, n (%)			
Grade 1	−	364 (63)	−
Grade 2	−	214 (37)	48 (20)
Grade 3	−		82 (34)
Grade 4 and 5	−		112 (46)
CR/IDC presence, n (%)	−	−	158 (65)

IQR = Interquartile range; Indolent CaP = International Society of Urological Pathology (ISUP) grade 1 or ISUP grade 2 without presence of CR/IDC; clinically significant CaP = ISUP grade ≥3 or ISUP grade 2 with presence of CR/IDC; DRE = digital rectal examination
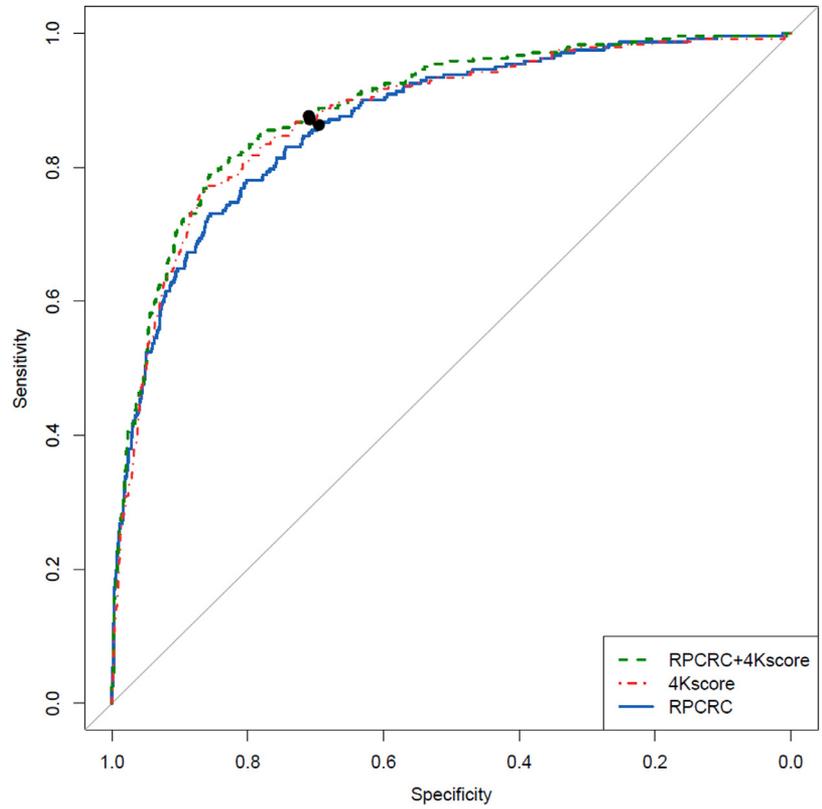
Models	Threshold 5%			AUC model comparison	
	Sensitivity	Specificity	AUC	RPCRC p-value	4Kscore p-value
RPCRC	0.86	0.70	0.868	-	0.43
4Kscore	0.87	0.71	0.876	0.43	-
RPCRC + 4Kscore	0.88	0.71	0.888	<0.01	<0.01

Fig.1. Receiver operating characteristic for RPCRC, 4Kscore, and the combined risk calculators predicting csCaP; the black dots display the 5% risk thresholds; sensitivity, and specificity are displayed. AUC = area under the curve of the receiver operating characteristic; 4K = 4-kallikrein; RPCRC = Rotterdam Prostate Cancer Risk Calculator.

3.1. Decision curve analysis

The 4Kscore and the 4Kscore-RPCRC combination have a higher NB compared to the RPCRC at threshold probabilities ≥5%, Fig. A.2. Applying a risk threshold of 5% with the use of the RPCRC, 4Kscore, and 4Kscore-RPCRC models would result in 65%, 66%, and 66% less biopsies respectively, and would miss 16%, 14%, and 14% indolent CaP, respectively in those men not biopsied. All risk calculators would equally miss the diagnosis of high grade cancers with 2% in those men not receiving prostate biopsy compared to a PSA ≥ 3.0 ng/ml strategy. Table 2 shows the numbers of reduced biopsies, missed CaP diagnoses, and net true reduction of biopsies for the different risk calculators using a risk threshold range of 4% to 10%. In terms of NB, at 5% threshold, the RPCRC reduced the number of Bx with 42 per 100 men at risk compared to a PSA ≥3.0 ng/ml strategy, without missing more csCaP cases.

For 4Kscore, the net true biopsy reduction was 44 per 100 men at risk. By combining the 4Kscore with RPCRC, an additional 3.3 biopsies per 100 men at risk could have been avoided, i.e., to avoid 1 unnecessary biopsy, the 4Kscore should be applied to 30 men. The price of the 4Kscore is estimated to range from $400 to $1,100 (personal communication with OPKO Health, Inc). In our study, setting the costs of the prediagnostic work-up would increase with $12,000 to $33,000 to avoid 1 additional unnecessary biopsy.

4. Discussion

As compared to a "biopsy all men with a PSA ≥ 3.0 ng/ml" strategy, our results show that basing the decision to biopsy on the RPCRC or the 4Kscore gives an equal reduction (approximately two-third) in the number of biopsies and the rate of overdiagnosis. The combination of 4Kscore

Table 2

Performed and reduced numbers of biopsy, delayed prostate cancer diagnosis, and net true reduction in biopsies for the RPCRC, 4Kscore, and combination of both models compared to the standard strategy (PSA ≥ 3.0 ng/ml), with different thresholds for csCaP

Threshold	Number of biopsies		Number of indolent cancers		Number of high-grade cancers		Additional net true reduction of biopsies compared to	
Clinical significant CaP	Performed	Reduced	Detected	Not detected (% reduced biopsies)	Detected	Missed diagnosis (% reduced biopsies)	PSA ≥ 3.0 (×100 men)	RPCRC (×100 men)
Biopsy all	2,872	0	578	0	242	0		
4%								
RPCRC-Cribriform	1,142	1,730 (68%)	317	261 (15%)	213	29 (2%)	35.0	Ref.
4Kscore	1,196	1,676 (58%)	370	208 (12%)	218	24 (1%)	37.5	2.5
4K+RPCRC-Cribriform	1,151	1,721 (60%)	361	217 (13%)	219	23 (1%)	39.9	4.9
5%								
RPCRC-Cribriform	1,011	1,861 (65%)	285	293 (16%)	209	33 (2%)	41.9	Ref.
4Kscore	978	1,894 (66%)	322	256 (14%)	211	31 (2%)	44.4	2.5
4K+RPCRC-Cribriform	975	1,897 (66%)	318	260 (14%)	212	30 (2%)	45.2	3.3
7.5%								
RPCRC-Cribriform	772	2,100 (73%)	226	352 (17%)	189	53 (3%)	48.5	Ref.
4Kscore	724	2,148 (75%)	246	332 (15%)	198	44 (2%)	53.4	4.9
4K+RPCRC-Cribriform	711	2,161 (75%)	245	333 (15%)	200	42 (2%)	55.7	7.2
10%								
RPCRC-Cribriform	613	2,259 (79%)	182	396 (18%)	178	64 (3%)	56.4	ref
4Kscore	582	2,290 (80%)	203	375 (16%)	187	55 (2%)	60.6	4.2
4K+RPCRC-Cribriform	576	2,296 (80%)	190	388 (17%)	191	51 (2%)	62.2	5.8

Bold values indicates the risk thresholds for detection of csCaP with the different models.

4K = 4-kallikrein; PSA = Prostate Specific Antigen; RPCRC = Rotterdam Prostate Cancer Risk Calculation.

and RPCRC slightly further reduced the number of unnecessary biopsies without missing additional csCaP diagnoses.

As both the 4Kscore and the RPCRC showed similar reduction of unnecessary biopsies and overdiagnosis of indolent CaP in our study cohort, other aspects should be considered when evaluating clinical usefulness. Factors that are important to consider are e.g., burden to the patient, availability of the test, and costs. Both risk calculators use PSA, age, and DRE findings. The risk calculators differ in estimation of prostate volume: the 4Kscore uses the 4K panel (total PSA, free PSA, intact PSA, and human kallikrein-related 2) as a proxy for prostate volume [15], where the RPCRC estimates the prostate volume based on DRE (prostate volume is categorized as <30, 30−49 and ≥50 cm^3) [8]. Here, it should be noted that a DRE volume estimation is a subjective procedure while PSA (sub forms) measurements are objective and in principle independent of the experience of the executor. Since both risk calculators require blood and DRE, the burden to the patient is therefore similar for both approaches and outpatient clinic costs are comparable to both risk calculator approaches. In terms of test-availability the 4K panel is not available in certain areas (e.g., Europe) and requires sending the blood to an external laboratory for analysis, which involves logistics and additional costs. The price of the 4K panel is estimated to range from $400 to $1100. The RPCRC is freely available on the internet or purchasable as an app for 1.99 dollar.

Besides comparison of the models we fused both models to optimize csCaP prediction, and found that the 4Kscore needs to be applied to 30 men to avoid 1 man getting an unnecessary biopsy compared to the RPCRC alone. A detailed cost effectiveness study was outside the scope of this research, however, on estimation prediagnostic work-up would increase up to $12,000 to $33,000 to avoid 1 additional unnecessary biopsy. A recent cost-effectiveness study in the United States showed a reduction in health care costs by applying 4Kscore before biopsy compared to biopsy all patients referred to a urologist with suspicion of having CaP in a theoretical cohort [24]. Most likely, a larger cost reduction could be achieved when applying a freely available risk stratification tool. In addition, it would be of value to perform cost-effectiveness studies within a primary care setting where initial risk stratification with the RPCRC could be (and actually is in the Netherlands [25]) applied.

Strengths of our study are the large population-based sample size, prospectively collected measurements of risk factors allowing for updating, and a contemporary pathology review including the latest ISUP scoring system. The ERSPC section Rotterdam cohort on which both models were originally developed was used to enable a fair comparison between the two risk prediction models [13]. To our knowledge, there is no other large population-based sample size available including the detailed pathological grading and data on the 4K pattern. Limitations of our study are that we did not have the actual DRE estimated prostate

volume available in this cohort, the DRE estimated prostate volume was derived using the TRUS based volume. The DRE based volume estimate approach is externally validated in a clinical setting and showed good concordance [18]. In addition, using volume classes has no effect in discriminatory ability as was confirmed in the development of models with a urinary molecular biomarker-based score [26]. The ERSPC section Rotterdam cohort represents two decades ago first-time screened men, who are predominantly white and received sextant biopsies. Sextant biopsies are known to detect less CaP than the present standard of 12-core based transrectal ultrasound biopsy procedures [27], thus it might have under-represented the CaP detection rate. However, in multiple contemporary clinical settings with 12-core TRUS biopsies, the RPCRC showed good predictive performance [28,29]. A comparable situation holds for the 4Kscore which is also partly developed on similar cohorts reflecting old practices, including, next to the currently used Dutch ERSPC cohort, the Swedish ERSPC data [13]. Also the 4Kscore performs well after external validation in a contemporary clinical setting [30]. Although our comparison is based on prediction tools that are (partly) based on cohorts from two decades ago, this comparison is still relevant for today's clinical practice. This is even confirmed by external validation in cohorts where multi-parametric Magnetic resonance imaging (mpMRI) targeted biopsies were applied [31]. Individualized risk assessment on having a biopsy detectable CaP including mpMRI Prostate Imaging Reporting and Data System (PI-RADS) score and clinical data can result in a considerable reduction of unnecessary biopsies [32]. Hence, biomarkers and mpMRI results could be combined to optimize upfront risk prediction. It must however be noted that long-term outcomes like e.g., metastatic disease and/or CaP death are not available for a detection pathway driven on individual risk and mpMRI targeted biopsy. The European Association of Urology guidelines recommend that a clinical risk prediction tool should be incorporated in the decision-making process, as the "one-size-fits-all" approach with a PSA cutoff does not provide a good balance between reduction of CaP morbidity, mortality, and overdiagnosis of indolent CaP [17]. A recent head-to-head comparison of the most well-known prediction tools showed that the ERSPC RPCRC is superior in identifying those men at risk for csCaP compared to Prostate Cancer Prevention Trial and Sunnybrook risk calculators [33]. Continuous updating of these existing prediction models will refine the balance between harms and benefit, as previously demonstrated with further refining the definition of csCaP [10], by incorporating the latest ISUP grading system [12]. Still future research is necessary by combining novel biomarkers, MRI findings, and the latest Gleason grading modifications (cribriform architecture) to fully assess the potential of the currently available prediction tools in contemporary clinical trials to reduce the detection of those CaP that will never become life-threatening as well as those CaP deemed suitable for active surveillance [11].

5. Conclusion

The RPCRC and 4Kscore had similar reductions of unnecessary biopsies and overdiagnosis of indolent CaP. Combination of both models slightly further reduced unnecessary biopsies. Given that the improvement in clinical impact was marginal, adding additional risk factors and biomarkers associated with CaP risk remains a tradeoff where aspects such as costs and patient burden are important considerations. More research is needed to validate the updated model in independent, contemporary, and various populations.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.1016/j.urolonc.2018.11.021.

References

[1] Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Zappa M, Nelen V, 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.

[2] Bruinsma SM, Bangma CH, Carroll PR, Leapman MS, Rannikko A, Petrides N, et al. Active surveillance for prostate cancer: a narrative review of clinical guidelines. Nat Rev Urol 2016;13:151–67.

[3] Esserman LJ, Thompson IM, Reid B, Nelson P, Ransohoff DF, Welch HG, et al. Addressing overdiagnosis and overtreatment in cancer: a prescription for change. Lancet Oncol 2014;15:e234–42.

[4] Roobol MJ, Steyerberg EW, Kranse R, Wolters T, van den Bergh RC, Bangma CH, et al. A risk-based strategy improves prostate-specific antigen-driven detection of prostate cancer. Eur Urol 2010;57:79–85.

[5] Ankerst DP, Hoefler J, Bock S, Goodman PJ, Vickers A, Hernandez J, 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.

[6] 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.

[7] Roobol MJ, van Vugt HA, Loeb S, Zhu X, Bul M, Bangma CH, 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.

[8] Roobol MJ, Schroder FH, Hugosson J, Jones JS, Kattan MW, Klein EA, et al. Importance of prostate volume in the European Randomised Study of Screening for Prostate Cancer (ERSPC) risk calculators: results from the prostate biopsy collaborative group. World J Urol 2012;30:149–55.

[9] Pruthi DK, Ankerst DP, Liss MA. Novel definitions of low-risk and high-risk prostate cancer: implications for the European Randomized Study of Screening for Prostate Cancer risk assessment tool. Eur Urol 2017;72:52–3.

[10] Roobol MJ, Verbeek JFM, van der Kwast T, Kummerlin IP, Kweldam CF, van Leenders G. Improving the Rotterdam European Randomized Study of Screening for Prostate Cancer risk calculator for initial prostate biopsy by incorporating the 2014 International Society of Urological Pathology Gleason Grading and Cribriform growth. Eur Urol 2017;72:45–51.

[11] Kweldam CF, Kummerlin IP, Nieboer D, Verhoef EI, Steyerberg EW, van der Kwast TH, et al. Disease-specific survival of patients with invasive cribriform and intraductal prostate cancer at diagnostic biopsy. Mod Pathol 2016;29:630–6.

[12] Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA, et al. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. Am J Surg Pathol 2016;40:244–52.

[13] Vickers AJ, Cronin AM, Aus G, Pihl CG, Becker C, Pettersson K, et al. A panel of kallikrein markers can reduce unnecessary biopsy for prostate cancer: data from the European Randomized Study of Prostate Cancer Screening in Goteborg. BMC Med 2008;6:19.

[14] Carlsson S, Maschino A, Schroder F, Bangma C, Steyerberg EW, van der Kwast T, et al. Predictive value of four kallikrein markers for pathologically insignificant compared with aggressive prostate cancer in radical prostatectomy specimens: results from the European Randomized Study of Screening for Prostate Cancer section Rotterdam. Eur Urol 2013;64:693–9.

[15] Vickers AJ, Cronin AM, Roobol MJ, Savage CJ, Peltola M, Pettersson K, et al. A four-kallikrein panel predicts prostate cancer in men with recent screening: data from the European Randomized Study of Screening for Prostate Cancer, Rotterdam. Clin Cancer Res 2010;16:3232–9.

[16] Benchikh A, Savage C, Cronin A, Salama G, Villers A, Lilja H, et al. A panel of kallikrein markers can predict outcome of prostate biopsy following clinical work-up: an independent validation study from the European Randomized Study of Prostate Cancer screening, France. BMC Cancer 2010;10:635.

[17] Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. Eur Urol 2017;71:618–29.

[18] Pereira-Azevedo N, Braga I, Verbeek JF, Osorio L, Cavadas V, Fraga A, et al. Prospective evaluation on the effect of interobserver variability of digital rectal examination on the performance of the Rotterdam prostate cancer risk calculator. Int J Urol 2017;24:826–32.

[19] Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Stat Med 2004;23:2567–86.

[20] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38.

[21] Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. Stat Med 1997;16:965–80.

[22] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.

[23] Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, et al. Reporting and interpreting decision curve analysis: a guide for investigators. Eur Urol 2018;74:796–804.

[24] Voigt JD, Dong Y, Linder V, Zappala S. Use of the 4Kscore test to predict the risk of aggressive prostate cancer prior to prostate biopsy: overall cost savings and improved quality of care to the us healthcare system. Rev Urol 2017;19:1–10.

[25] Osses DF, Alberts AR, Bausch GCF, Roobol MJ. Multivariable risk-based patient selection for prostate biopsy in a primary health care setting: referral rate and biopsy results from a urology outpatient clinic. Transl Androl Urol 2018;7:27–33.

[26] Van Neste L, Hendriks RJ, Dijkstra S, Trooskens G, Cornel EB, Jannink SA, et al. Detection of high-grade prostate cancer using a urinary molecular biomarker-based risk score. Eur Urol 2016;70:740–8.

[27] Scattoni V, Maccagnano C, Capitanio U, Gallina A, Briganti A, Montorsi F. Random biopsy: when, how many and where to take the cores? World J Urol 2014;32:859–69.

[28] van Vugt HA, Kranse R, Steyerberg EW, van der Poel HG, Busstra M, Kil P, et al. Prospective validation of a risk calculator which calculates the probability of a positive prostate biopsy in a contemporary clinical cohort. Eur J Cancer 2012;48:1809–15.

[29] Gayet M, Mannaerts CK, Nieboer D, Beerlage HP, Wijkstra H, Mulders PFA, et al. Prediction of prostate cancer: external validation of the ERSPC risk calculator in a contemporary dutch clinical cohort. Eur Urol Focus 2018;4:228–34.

[30] Parekh DJ, Punnen S, Sjoberg DD, Asroff SW, Bailen JL, Cochran JS, 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.

[31] Alberts AR, Roobol MJ, Verbeek JFM, Schoots IG, Chiu PK, Osses DF, et al. Prediction of high-grade prostate cancer following multi-parametric magnetic resonance imaging: improving the Rotterdam European Randomized Study of Screening for Prostate Cancer risk calculators. Eur Urol 2018;pii: S0302-2838(18)30553-0. [Epub ahead of print].

[32] van Leeuwen PJ, Hayen A, Thompson JE, Moses D, Shnier R, Bohm M, et al. A multiparametric magnetic resonance imaging-based risk model to determine the risk of significant prostate cancer prior to biopsy. BJU Int 2017;120:774–81.

[33] Pereira-Azevedo N, Verbeek JFM, Nieboer D, Bangma CH, Roobol MJ. Head-to-head comparison of prostate cancer risk calculators predicting biopsy outcome. Transl Androl Urol 2018;7:18–26.