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

European Association of Urology

Development and External Validation of a Novel Nomogram to Predict Side-specific Extraprostatic Extension in Patients with Prostate Cancer Undergoing Radical Prostatectomy

Timo F.W. Soeterik [a,b,*], *Harm H.E. van Melick* [b], *Lea M. Dijksman* [c], *Heidi Küsters-Vandevelde* [d], *Saskia Stomps* [e], *Ivo G. Schoots* [f], *Douwe H. Biesma* [a], *J.A. Witjes* [g], *Jean-Paul A. van Basten* [h]

[a] *Department of Value-Based Healthcare, Santeon Group, Utrecht, The Netherlands;* [b] *Department of Urology, St. Antonius Hospital, Nieuwegein/Utrecht, Netherlands;* [c] *Department of Value-Based Healthcare, St. Antonius Hospital, Nieuwegein/Utrecht, The Netherlands;* [d] *Department of Pathology, Canisius Wilhelmina Hospital, Nijmegen, The Netherlands;* [e] *Department of Urology, Hospital Group Twente, Hengelo/Almelo, The Netherlands;* [f] *Department of Radiology and Nuclear Medicine, Erasmus University Medical Centre, Rotterdam, The Netherlands;* [g] *Department of Urology, Radboud University Medical centre, Nijmegen, The Netherlands;* [h] *Department of Urology, Canisius Wilhelmina Hospital, Nijmegen, The Netherlands*

Article info

Article history:
Received 24 March, 2020
Received in revised form
4 August, 2020
Accepted 18 August, 2020

Associate Editor:
Gianluca Giannarini

Keywords:
Prostate cancer
Nomogram
Extraprostatic extension
Radical prostatectomy
Magnetic resonance imaging
Staging

Abstract

Background: Prediction of side-specific extraprostatic extension (EPE) is crucial in selecting patients for nerve-sparing radical prostatectomy (RP).
Objective: To develop and externally validate nomograms including multiparametric magnetic resonance imaging (mpMRI) information to predict side-specific EPE.
Design, setting, and participants: A retrospective analysis of 1870 consecutive prostate cancer patients who underwent robot-assisted RP from 2014 to 2018 at three institutions.
Outcome measurements and statistical analysis: Four multivariable logistic regression models were established, including combinations of patient-based and side-specific variables: prostate-specific antigen (PSA) density, highest ipsilateral International Society of Urological Pathology (ISUP) biopsy grade, ipsilateral percentage of positive cores on systematic biopsy, and side-specific clinical stage assessed by both digital rectal examination and mpMRI. Discrimination (area under the curve [AUC]), calibration, and net benefit of these models were assessed in the development cohort and two external validation cohorts.
Results and limitations: On external validation, AUCs of the four models ranged from 0.80 (95% confidence interval [CI] 0.68–0.88) to 0.83 (95% CI 0.72–0.90) in cohort 1 and from 0.77 (95% CI 0.62–0.87) to 0.78 (95% CI 0.64–0.88) in cohort 2. The three models including mpMRI staging information resulted in relatively higher AUCs compared with the model without mpMRI information. No major differences between the four models regarding net benefit were established. The model based on PSA density, ISUP grade, and mpMRI T stage was superior in terms of calibration. Using this model with a cut-off of 20%, 1980/2908 (68%) prostatic

* Corresponding author. Koekoekslaan 1, 3435 CM, Nieuwegein, The Netherlands.
Tel. +31 88320 8872.
E-mail address: t.soeterik@antoniusziekenhuis.nl (Timo F.W. Soeterik).

https://doi.org/10.1016/j.euo.2020.08.008

lobes without EPE would be found eligible for nerve sparing, whereas non–nerve sparing would be advised in 642/832 (77%) lobes with EPE.

Conclusions: Our analysis resulted in a simple and robust nomogram for the prediction of side-specific EPE, which should be used to select patients for nerve-sparing RP.

Patient summary: We developed a prediction model that can be used to assess accurately the likelihood of tumour extension outside the prostate. This tool can guide patient selection for safe nerve-sparing surgery.

1. Introduction

A challenging aspect of performing a radical prostatectomy (RP) for prostate cancer includes balancing the risk of positive margins versus optimisation of quality of life by maximising the probability of retaining the patient's erectile function and urinary continence. In 1982, the first purposeful nerve-sparing RP was performed, resulting in normal postoperative sexual function and retained quality of life of the patient [1]. Following the introduction of this technique, its therapeutic effect has been evaluated in several other studies. A recent meta-analysis showed nerve sparing to be associated with a lower risk of postoperative incontinence (relative risk [RR] 0.75, 95% confidence interval [CI] 0.65–0.85) and erectile dysfunction (RR 0.77, 95% CI 0.70–0.85) [2].

Preoperative assessment of extraprostatic extension (EPE) is a long-established strategy to guide patient selection for nerve sparing. If there is a high risk of EPE, nerve sparing should be discouraged due to the increased risk of positive surgical margins [3]. EPE risk prediction is often done by using nomograms such as the Partin tables and the Memorial Sloan Kettering Cancer Center nomogram [4,5]. However, these models do not provide information on the laterality of EPE. Since EPE is mostly one sided (85%), localisation is essential as unilateral nerve-sparing surgery remains possible in the majority of patients [6]. Nomograms predicting side-specific EPE have also been developed. However, these models lack the inclusion of multiparametric magnetic resonance imaging (mpMRI) information [6–8].

Adoption of mpMRI to guide clinical decision making in prostate cancer has increased drastically in recent years [9]. mpMRI alone has limited ability to guide patient selection for nerve sparing, due to a low per-patient sensitivity for the detection of EPE of 57% [10]. However, its predictive potential when combined with other clinical parameters remains poorly understood. Previous studies have shown that the combination of mpMRI information and traditional preoperative clinical parameters, including biopsy information and serum prostate-specific antigen, can improve the prediction of adverse surgical pathology including EPE [11,12]. The number of available nomograms including a combination of both mpMRI and clinical parameters for the prediction of side-specific EPE, however, is scarce. The need for further exploration of the additional value of using mpMRI information for the prediction of side-specific EPE is emphasised by the results of a recent external validation study, showing that mpMRI-naïve nomograms are inaccurate when applied to external populations [13].

Therefore, we aim to develop a nomogram that enables accurate prediction of side-specific EPE, applicable to the current state of clinical practice, including readily available clinical and MRI input parameters. Generalisability of the tool will be assessed by performing external validation using two separate hospital populations.

2. Patients and methods

2.1. Patient population and study data

After receiving institutional review board approval, data from 1871 consecutive patients diagnosed with prostate cancer who underwent robot-assisted radical prostatectomy (RARP) at three teaching hospitals were extracted from prospectively maintained databases. Of these, one patient was excluded due to prior treatment with androgen deprivation therapy. The cohort of patients undergoing RARP from 2014 to 2018 at the Canisius Wilhelmina Hospital (CWH), Nijmegen was used for nomogram development. This cohort was selected for model derivation due to the population size and its multicentre nature. Since 2013, regional prostate cancer surgery has been centralised, and all the patients from two other hospitals (Catharina Hospital Eindhoven and Radboud University Medical Centre) have undergone RARP at CWH. The cohorts of patients undergoing RARP at the Hospital Group Twente in Almelo-Hengelo (validation cohort 1) and St. Antonius Hospital, Nieuwegein-Utrecht (validation cohort 2) from 2015 to 2018 were used for external validation.

2.2. Predictor selection

We used a clinically driven, evidence-based approach for predictor selection. First, a very recent literature review was used to identify significant predictors for side-specific EPE [13]. Second, three consensus meetings were organised with clinical experts including urologists (H.v.M., J.W., S.S., and J.P.v.B.), an expert uroradiologist (I.S.), and a uropathologist (H.K.V.). Predictors were selected based on relevance, availability, and usefulness.

Patient-based (prostate-specific antigen [PSA] density [PSAD]) and side-specific covariates (digital rectal examination [DRE] local staging, mpMRI-based local staging, highest International Society of Urological Pathology biopsy grade, and percentage of positive systematic cores) were included.

2.3. MRI protocol

MRI was performed using 3 Tesla scanners and a body coil. Gadolinium (1 mg/kg) was administrated intravenously. Radiological reporting was

Table 1 – Baseline characteristics of the separate patient cohorts.

	Development *N* (%)	Validation 1 *N* (%)	Validation 2 *N* (%)
No. of patients	887	513	470
Age, median (IQR)	66 (61–69)	66 (61–70)	66 (62–70)
PSA (ng/mL), mean (SD)	7.9 (5.9–11.0)	8.0 (5.9–11)	8.3 (5.9–12.5)
PSA density (ng/mL/mL), mean (SD)	0.18 (0.12–0.27)	0.17 (0.12–0.28)	0.20 (0.13–0.32)
Clinical T stage			
T1c	509 (57)	338 (66)	288 (61)
T2a	240 (27)	93 (18)	148 (32)
T2b	34 (4)	45 (9)	9 (2)
T2c	29 (3)	16 (3)	8 (2)
T3	64 (8)	19 (4)	14 (3)
Unknown	11 (1)	2 (0)	3 (0)
Preoperative MRI			
Yes	879 (99)	496 (97)	387 (82)
No	8 (1)	17 (3)	83 (18)
Radiological T stage			
T0	66 (7)	57 (11)	38 (8)
T2/T2a	285 (32)	130 (25)	216 (46)
T2b	46 (5)	27 (5)	10 (2)
T2c	117 (13)	60 (12)	48 (10)
T2/T3 (uncertain EPE)	94 (11)	62 (12)	12 (3)
T3a	200 (23)	133 (26)	53 (11)
T3b	48 (5)	25 (5)	3 (1)
T4	5 (1)	0 (0)	0 (0)
Unknown	26 (3)	19 (4)	90 (19)
Biopsy type			
TRUS-guided systematic	497 (56)	313 (61)	380 (81)
MRI guided	140 (16)	66 (13)	17 (4)
TRUS + MRI guided	250 (28)	134 (26)	73 (15)
Pathological stage			
T2	36 (4)	1 (0)	35 (8)
T2a	86 (10)	49 (10)	37 (8)
T2b	8 (1)	11 (2)	16 (3)
T2c	303 (34)	256 (50)	237 (50)
T3a	338 (38)	142 (28)	99 (21)
T3b	112 (13)	53 (10)	46 (10)
T4	4 (0)	1 (0)	0 (0)

IQR = interquartile range; EPE = extraprostatic extension; MRI = magnetic resonance imaging; PSA = prostate specific antigen; SD = standard deviation; TRUS = transrectal ultrasonography.

done by dedicated radiologists with at least 2 yr of experience with prostate MRI reading. MRI reporting in 2013 and 2014 was done according to the European Society of Urogenital Radiology guidelines [14]. From 2015 onwards, the principles of Prostate Imaging Reporting and Data System version 2 were followed [15]. Imaging-based T stage was defined according to the American Joint Committee on Cancer TNM classification [16].

2.4. Predictors and outcome definitions

Patient-based PSA and prostate volume, necessary for the calculation of PSAD, were based on the most recent available measurements preoperatively. Prostate volume was measured by transrectal ultrasonography or mpMRI. Side-specific DRE staging information was collected before biopsy by the treating urologist during routine clinical care. Both side-specific DRE and mpMRI staging information were subdivided into three subclasses. These included nonpalpable disease (T1), organ-confined nodal disease (T2), and EPE (T3) for DRE. As for mpMRI, these included nonvisible lesions (T1), organ-confined lesions (T2), and lesions with EPE (T3).

Imaging features used to assess EPE included thickening or suspicion for invasion of the neurovascular bundle, bulging of the prostatic contour, capsule irregularity, obliteration of the rectoprostatic angle, presence of a hypointensive signal in a periprostatic area, and length of tumour contact with the capsule [17–19]. Explicit statements about the presence or absence of EPE in the radiological report were scored accordingly. In less explicit cases, a strong suspicion of EPE was classified as positive. Cases in which EPE could not be ruled out were classified as negative [19].

Side-specific biopsy information, including highest percentage of positive cores on systematic biopsy and highest ISUP grade, was documented during routine clinical care for both the right and the left lobe separately.

Final surgical histopathological information including pathological tumour stage and highest ISUP grade found in the resected prostate specimens were documented on a whole-gland level. If EPE was observed, the laterality (left, right, or both lobes) was reported. EPE was defined as a tumour that bulges beyond the prostate contour, as a tumour that is admixed with periprostatic fat tissue, or, in the posterolateral area, as a tumour within connective tissue or between nerves of the neurovascular bundle. EPE was distinct from microscopic bladder neck invasion (presence of tumour between thick smooth muscle bundles in the absence of benign prostate glands) and seminal vesicle invasion, which were not considered as EPE in our study [20].

The RP specimens were processed with conventional sections in 1810/1870 (97%) cases and using whole-mount sections in 60 (3%) cases.

2.5. Model building

Four models were built according to the "full model" principle [21], including combinations of five predictors corresponding to different staging work-up strategies. Model 1 consisted of PSAD, DRE, ISUP grade, and percentage of positive cores on systematic biopsy. Model 2 included PSAD, mpMRI, and ISUP grade. Model 3 included PSAD, mpMRI, DRE, and ISUP grade. Model 4 included all five predictors (Supplementary Table 1). For analysis purposes, the right and left prostatic lobes of each patient were regarded as separate cases, that is, for calculating the probability of EPE in the right lobe: patient-based PSAD, right-sided biopsy information, right-sided DRE staging information, and right-sided mpMRI staging information were used.

2.6. Model performance, external validation, and clinical usefulness

Performance of all models was assessed in the development cohort and two validation cohorts. Discrimination, which refers to the ability of a model to distinguish a prostate lobe with the endpoint (EPE) from a lobe without EPE, was quantified using the area under the receiver operating characteristic curve [14]. Model calibration, which refers to the agreement between observed endpoints and predictions, was assessed using calibration in the large and calibration slope [14]. The net benefit per risk threshold was determined using decision-curve analysis. The net benefit is calculated as the proportion of "net" true positives (true positives corrected for the false positives weighted by the odds of the risk cut-off, divided by the sample size) [22].

2.7. Missing data

Missing data patterns were explored using response matrix and correlation plots. Missing data were assumed to be missing at random, as their missingness was related to the diagnostic work-up (eg, selection of patients for mpMRI and biopsy protocols) of the hospitals. Missing data were handled by using multivariate imputation by chained equations including pooling using Rubin's rules [23].

3. Results

3.1. Patient populations

Overall, 887 patients were included in the development cohort, 513 in validation cohort 1, and 470 in validation cohort 2. The values of EPE prevalence on prostatic lobe level of these cohorts were, respectively, 458/1774 (26%), 225/1026 (21%), and 148/940 (16%). Baseline characteristics on patient and prostatic lobe levels are presented, respectively, in Tables 1 and 2.

3.2. Performance of the four multivariable logistic regression models in the development cohort

At multivariable analyses, PSAD, DRE staging, mpMRI staging, ISUP grades 3–5, and percentage of positive cores

Table 2 – Baseline characteristics on prostate lobe level.

	Development (*N* = 1774)		Validation 1 (*N* = 1026)		Validation 2 (*N* = 938[a])	
	No EPE at histopathology	EPE at histopathology	No EPE at histopathology	EPE at histopathology	No EPE at histopathology	EPE at histopathology
No. of lobes	1316	458	801	225	790	148
PSA density						
Median (IQR)	0.17 (0.12–0.17)	0.21 (0.14–0.33)	0.16 (0.11–0.25)	0.23 (0.14–0.43)	0.19 (0.13–0.30)	0.21 (0.14–0.33)
Unknown (%)	19 (1)	9 (2)	14 (2)	8 (4)	13 (2)	5 (3)
ISUP grade						
Benign	341 (26)	31 (7)	213 (26)	10 (4)	204 (26)	9 (6)
1	365 (28)	65 (14)	315 (39)	54 (24)	306 (39)	38 (26)
2	296 (22)	147 (32)	119 (15)	79 (35)	151 (19)	41 (28)
3	93 (7)	75 (16)	54 (7)	33 (15)	55 (7)	20 (14)
4	58 (4)	58 (13)	22 (3)	27 (12)	27 (3)	21 (14)
5	36 (3)	71 (16)	12 (2)	13 (6)	17 (2)	17 (11)
Unknown	127 (10)	11 (2)	66 (8)	9 (4)	30 (4)	2 (1)
Percentage of positive cores						
Median (IQR)	0.20 (0–0.50)	0.67 (0.33–1.0)	0.20 (0–0.60)	0.80 (0.40–1.0)	0.33 (0–0.60)	0.75 (0.40–1.0)
Unknown	242 (18)	57 (12)	112 (14)	29 (13)	47 (6)	5 (6)
Clinical stage assessed by DRE						
T1	1107 (84)	236 (51)	697 (87)	134 (60)	644 (81)	85 (57)
T2	134 (10)	133 (29)	71 (9)	61 (27)	101 (13)	39 (26)
T3	20 (2)	54 (12)	7 (1)	14 (6)	6 (1)	7 (5)
Unknown	55 (4)	35 (8)	26 (3)	16 (7)	39 (5)	17 (12)
Clinical stage assessed by MRI						
No lesion visible	622 (47)	66 (14)	395 (50)	28 (13)	315 (40)	23 (16)
Lesion but no EPE	551 (42)	209 (46)	306 (38)	92 (41)	265 (33)	68 (46)
EPE	101 (8)	169 (37)	66 (8)	93 (41)	31 (4)	28 (19)
Unknown	42 (3)	14 (3)	34 (4)	12 (5)	179 (23)	29 (19)

DRE = digital rectal examination; EPE = extraprostatic extension; IQR = interquartile range; ISUP = International Society of Urological Pathology; MRI = magnetic resonance imaging; PSA = prostate-specific antigen.
[a] In this cohort, presence of EPE was unknown in one patient/two prostatic lobes.

Table 3 – Multivariable logistic regression outcomes of four different models.

	Model 1 OR (95% CI)	*p* value	Model 2 OR (95% CI)	*p* value	Model 3 OR (95% CI)	*p* value	Model 4 OR (95% CI)	*p* value
PSA density	1.64 (0.93–2.90)	0.086	2.27 (1.31–3.94)	0.004	1.76 (0.99–3.10)	0.052	1.70 (0.96–3.02)	0.071
Clinical stage at DRE			–	–	–	–		
T1								
T2	Referent						Referent	
T3	2.66 (1.96–3.60)	<0.001					1.97 (1.43–2.72)	<0.001
	5.08 (2.71–9.53)	<0.001					3.32 (1.70–6.48)	<0.001
MRI	–	–						
No lesion			Referent		Referent		Referent	
Lesion, but no EPE			2.36 (1.71–3.25)	<0.001	2.22 (1.60–3.08)	<0.001	1.96 (1.40–2.73)	<0.001
EPE			8.45 (5.79–12.32)	<0.001	6.90 (4.67–10.21)	<0.001	5.22 (3.47–7.87)	<0.001
ISUP grade								
Benign	Referent		Referent		Referent		Referent	
1	0.89 (0.52–1.54)	0.7	1.48 (0.92–2.38)	0.11	0.74 (0.43–1.27)	0.3	0.80 (0.46–1.38)	0.4
2	1.99 (1.17–3.39)	0.012	3.03 (1.94–4.74)	<0.001	1.30 (0.76–2.22)	0.4	1.44 (0.83–2.49)	0.19
3	2.60 (1.44–4.70)	0.001	4.46 (2.66–7.46)	<0.001	1.94 (1.08–3.48)	0.027	1.90 (1.04–3.48)	0.036
4	3.36 (1.86–6.05)	<0.001	5.99 (3.47–10.3)	<0.001	2.79 (1.52–5.14)	0.001	2.62 (1.41–4.86)	0.002
5	4.94 (2.53–9.62)	<0.001	9.92 (5.62–17.5)	<0.001	3.66 (1.88–7.13)	<0.001	3.63 (1.92–7.21)	<0.001
% Positive cores	4.75 (2.83–7.94)	<0.001	–	–	4.77 (2.83–8.00)	<0.001	3.84 (2.25–6.53)	<0.001
AUC	0.80 (0.69–0.87)		0.80 (0.70–0.87)		0.81 (0.71–0.88)		0.82 (0.72–0.89)	
Hosmer and Lemeshow *p* value	0.65		0.59		0.78		0.99	

AUC = area under the curve; CI = confidence interval; DRE = digital rectal examination; EPE = extraprostatic extension; ISUP = International Society of Urological Pathology; MRI = magnetic resonance imaging; OR = odds ratio; PSA = prostate-specific antigen.

were all found to be significant predictors of EPE (Table 3). Model 4, which includes all available predictors, resulted in the highest area under the curve (AUC; 0.82). The AUCs of the other three models ranged from 0.80 to 0.81 (Table 3).

3.3. *Performance of the four models when applied to two external patient cohorts*

Overall, higher AUCs were observed when the models were applied to validation cohort 1 than those observed in cohort 2 (Table 4). As shown in Table 4 and Fig. 1, both discrimination and calibration of all models were excellent when applied to cohort 1. In cohort 2, model 2 had the best performance, with both fair AUC and relatively highest agreement between predicted and observed probabilities in both validation cohorts (Fig. 1 and 2). In cohort 2, substantial miscalibration was observed for the other three models (Fig. 2A–D).

3.4. *Clinical usefulness*

A systematic analysis of the event status of patients who would fall above and below the risk threshold between 5% and 30% is provided in Table 5. At a risk threshold of 20%, a non–nerve-sparing approach would be advised in 642/832 (77%) prostatic lobes with EPE. Nerve sparing would be recommended in 1980/2908 (68%) prostatic lobes without EPE.

Risk thresholds ranging from 0% to 30% were regarded as clinically most relevant, for which net benefits of all four models are presented in Fig. 3. All models can be regarded clinically useful for risk thresholds between 6% and 30%, as net benefits were found to be higher than those of the “treat all” and “treat none” approaches. On external validation, decision-curve analysis revealed relatively lower net benefits for model 1 than for models 2, 3, and 4 (Fig. 3B and 3C).

4. Discussion

Our analysis showed that the three nomograms (models 2, 3, and 4) based on clinical information combined with mpMRI staging information outperformed the nomogram without mpMRI staging information (model 1), in terms of AUC, calibration, and net benefit. Among these three nomograms, discrimination and net benefit were comparable. However, model 2 outperformed both model 3 and model 4 in terms of agreement between predicted and observed probabilities. Therefore, this nomogram should be the preferred tool for side-specific EPE risk prediction. Besides the fact that model 2 outperformed all other models in terms of calibration, it was also the most minimalistic

Table 4 – Performance of all models when applied to two external cohorts.

	Validation cohort 1				Validation cohort 2			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
AUC (95% CI)	0.80 (0.68–0.88)	0.83 (0.71–0.90)	0.83 0.71–0.90)	0.83 (0.72–0.90)	0.77 (0.62–0.87)	0.77 (0.64–0.87)	0.78 (0.64–0.88)	0.78 (0.64–0.88)

AUC = area under the curve; CI = confidence interval.

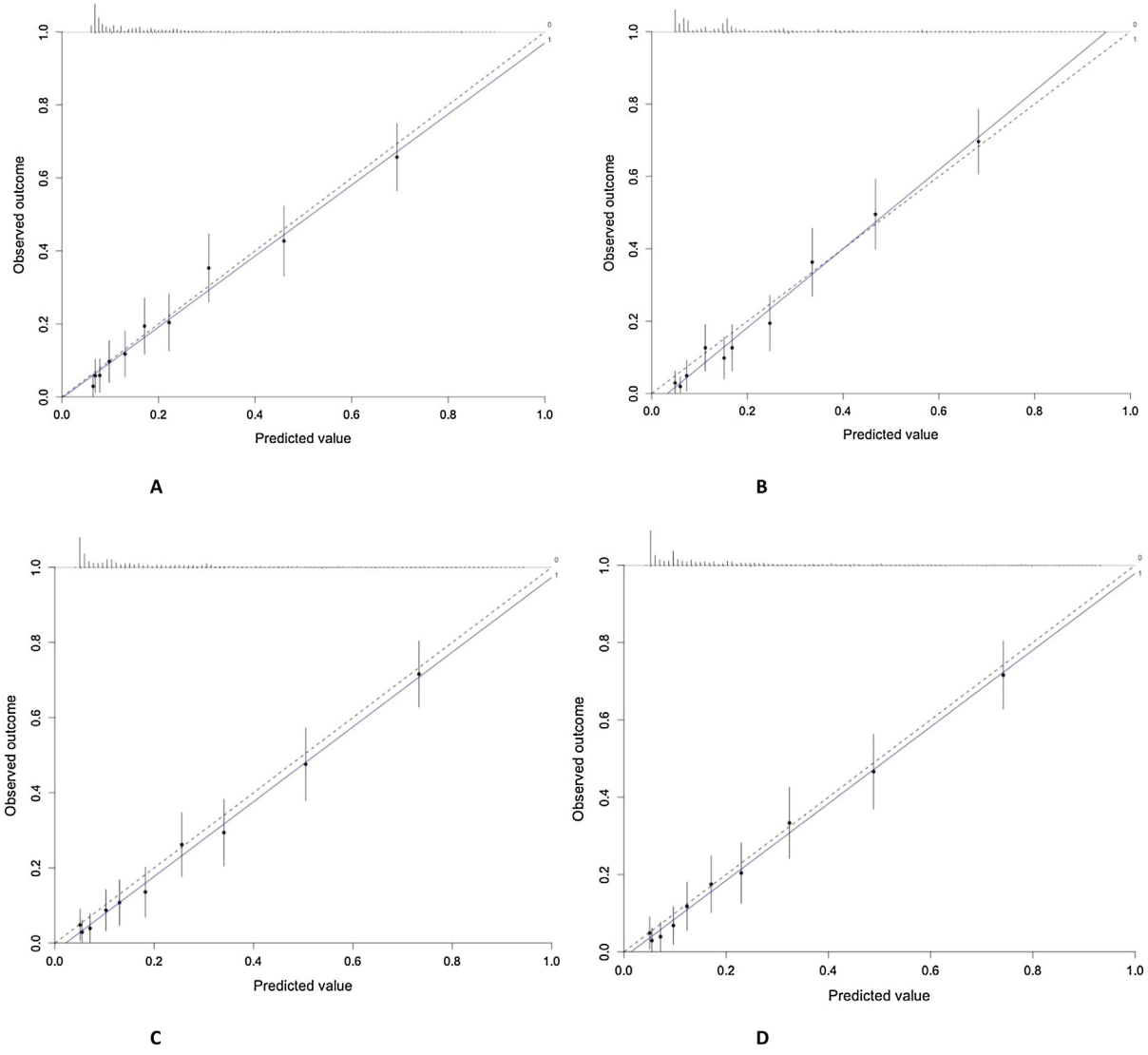

Fig. 1 – Calibration plots of all four models when applied to validation cohort 1: (A) model 1, (B) model 2, (C) model 3, and (D) model 4.

model, since the model consists solely of three predictor variables (PSAD, highest ISUP grade, and mpMRI clinical stage). In addition, this model is applicable to a wide range of clinical situations, independent of the prostate biopsy protocol used. The model can be accessed online at https://www.evidencio.com/models/show/2142.

A common explanation for the miscalibration observed when a nomogram is applied to an external population is the case-mix severity. In this study, this is also the most likely cause of the systematic overestimation of the predicted EPE risk in validation cohort 2. As an example, suspicion of EPE on mpMRI was reported in 15% of the lobes in the development cohort, compared with 6% in validation cohort 2. In addition, the prevalence of EPE on final pathology was substantially higher among cases in the development cohort than among cases in validation cohort 2: 26% versus 16%. As stated previously, the highest agreement between predicted and observed probabilities was achieved using model 2. However, overestimation of the predicted risk was still observed when the nomogram was applied to cohort 2. Overestimation was predominantly observed for predicted risks >30%. This was possibly due to the fact that a substantially lower number of patients with relatively high risk for EPE were selected for RARP in validation cohort 2, compared with the development cohort. For example, only a relative proportion of patients with a low suspicion of (extensive) EPE, on mpMRI or DRE, were selected for RARP. Whereas patients with a high risk of (extensive) EPE were more likely to be treated with radiation therapy. This assumption is supported by positive predictive value (PPV) rates for EPE established by DRE and mpMRI. The PPV for EPE assessed by DRE was 54% in validation cohort 2, whereas this was 73% in the development cohort. PPV of mpMRI for EPE detection was 48% in validation cohort 2, compared with 63% in the development cohort.

Interestingly, model discrimination was found to be higher for models 2, 3, and 4 when applied to validation

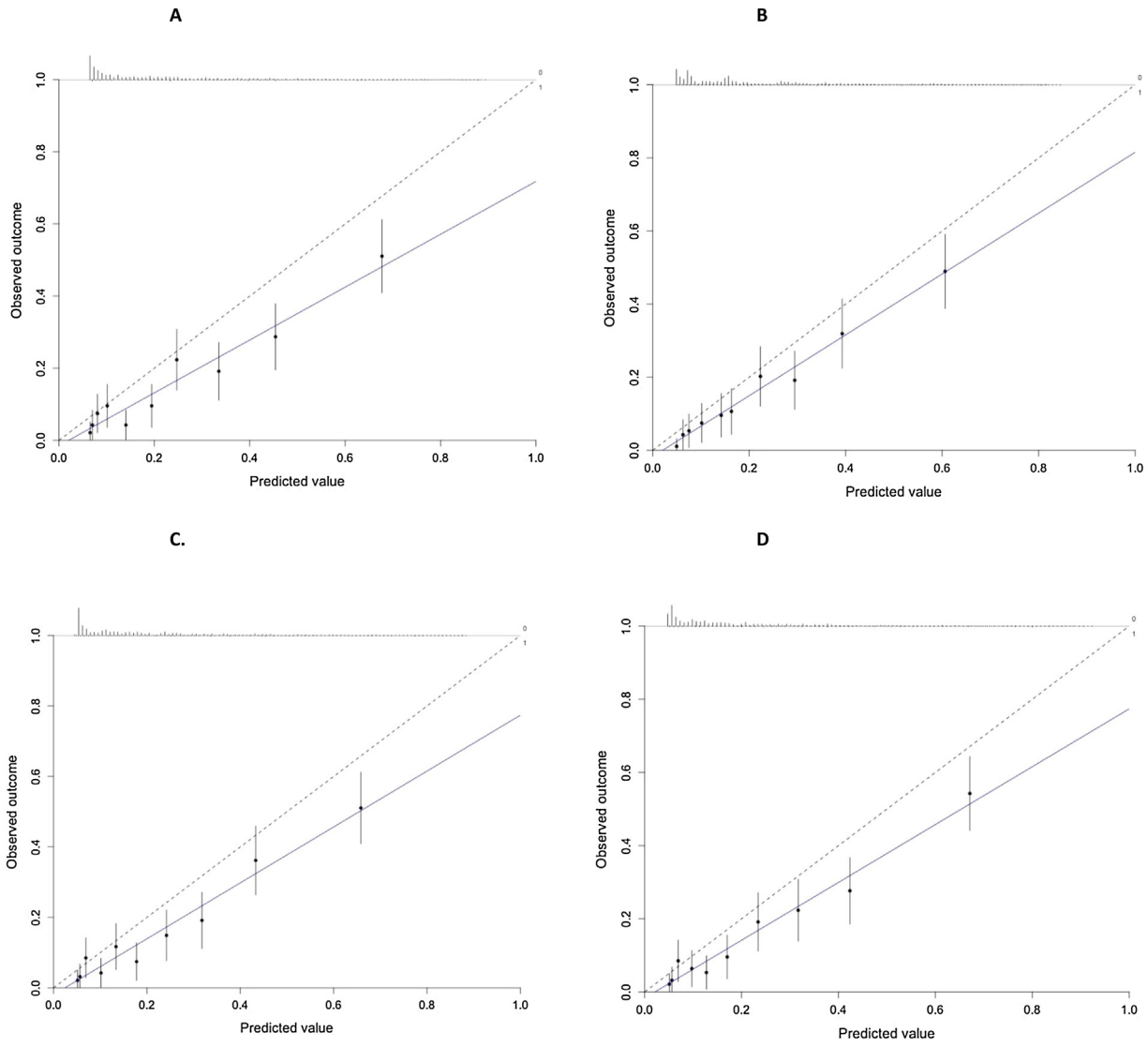

Fig. 2 – Calibration plots of all four models when applied to validation cohort 2: (A) model 1, (B) model 2, (C) model 3, and (D) model 4.

Table 5 – Systematic analysis of outcomes for model 2 per cut-off using all three cohorts.

Threshold (%)	Below the cut-off (Nerve sparing recommended)			Above the cut-off (Nerve sparing not recommended)		
	Total (%)	Without EPE (%)	With EPE (%)	Total (%)	Without EPE (%)	With EPE (%)
5	256 (7)	251 (98)	5 (2)	3484 (93)	2657 (76)	827 (24)
7.5	922 (17)	883 (96)	39 (4)	2818 (83)	2025 (72)	793 (28)
10	1141 (31)	1082 (95)	59 (5)	2599 (69)	1826 (70)	773 (30)
12.5	1346 (36)	1271 (94)	75 (6)	2394 (64)	1637 (68)	757 (32)
15	1612 (43)	1497 (93)	115 (7)	2128 (57)	1411 (66)	717 (34)
17.5	2044 (55)	1876 (92)	168 (8)	1696 (45)	1032 (61)	664 (39)
20	2170 (58)	1980 (91)	190 (9)	1570 (42)	928 (59)	642 (41)
22.5	2235 (60)	2035 (91)	200 (9)	1505 (40)	873 (58)	632 (42)
25	2306 (62)	2085 (90)	221 (10)	1434 (38)	823 (57)	611 (43)
27.5	2508 (67)	2250 (90)	258 (10)	1232 (33)	658 (53)	574 (47)
30	2692 (72)	2379 (88)	313 (12)	1048 (28)	529 (50)	519 (50)

EPE = extraprostatic extension.

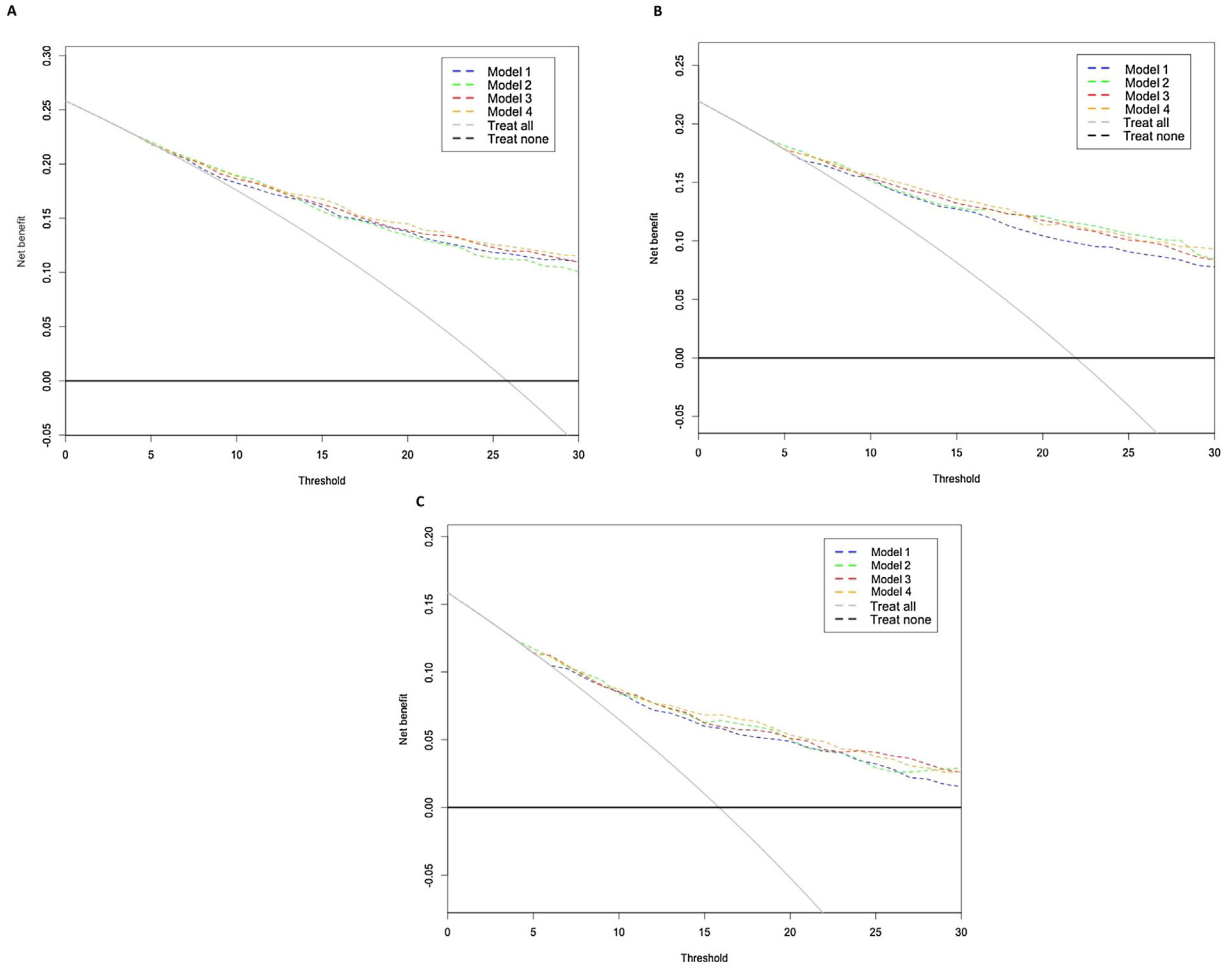

Fig. 3 – Net benefit of model 2 determined in all three cohorts using decision curve analysis: (A) development, (B) validation 1, and (C) validation 2.

cohort 1, compared with the development cohort (0.83 vs 0.80, 0.81, and 0.81). These differences might be explained by the heterogeneity of the patient cohort used for model development. As mentioned previously, a large proportion of patients undergoing RARP at CWH underwent diagnostic staging work-up elsewhere. Owing to the referral pattern, there was a larger variation in used prostate biopsy protocols, mpMRI readings, and histopathological biopsy evaluation as patients came from different hospitals. However, we assume that the multicentre nature of this cohort enabled accurate model estimation leading to robust tools that can be applied to different patient settings. Another explanation for the observed improved discrimination could stem from the fact that a large prospective prostate biopsy trial (4 M study) was on-going in validation cohort 1 during the study period [24]. As part of the protocol, a higher number of patients underwent MRI target biopsy as well as concomitant systematic biopsies in a protocolled trial setting, potentially leading to more accurate tumour sampling.

To our knowledge, two other nomograms for the prediction of side-specific EPE including mpMRI features have been developed previously. One of these was derived using the data of 264 consecutive men undergoing RP between 2012 and 2015. The authors reported excellent model discrimination (AUC: 0.86) and excellent calibration [25]. The model, however, includes a number of complex features, which may not always be readily available in a real-world clinical setting, such as European Society of Urogenital Radiology classification for EPE and capsule contact length on MRI. In addition, this model has not yet been validated externally and thus the performance remains unclear when applied to other populations. The other nomogram, developed by Martini et al [26], was based on data from 589 patients who underwent RARP between February 2014 and October 2015. The authors reported excellent discrimination in terms of AUC (0.82) and high agreement between predicted and observed probabilities. Sighinolfi et al [27] also externally validated the Martini model. In this external validation study, moderate to low discriminative ability (AUC 0.68), and low sensitivity (20%) and specificity (54%) at the 20% cut-off were reported [27]. In another recently published external validation study, good discrimination in terms of AUC (0.78) but poor calibration, even after model updating, was reported [28]. What our study adds to this previous work is that we have shown that our developed

nomogram provides accurate EPE risk prediction not only in the development cohort, but also when applied to external patient populations.

Implementation of tools that facilitate shared decision making may improve the quality of prostate cancer care, as active involvement of patients is associated with less decision conflict and decision regret [29]. Our proposed nomograms can facilitate this, as demonstrated by the net benefit over a range of risk thresholds within a suitable range for clinical decision making. Besides the potential of improving quality of care in terms of patient experience, our nomogram may also improve the quality of care in terms of clinical outcomes. Using the nomogram with a risk threshold set at 20%, accurate patient selection for nerve-sparing RP is possibly leading to the relatively highest clinical benefit. With a 20% risk threshold, 2170/3740 (58%) prostatic lobes of the development cohort would fall below the cut-off and nerve sparing would be advised for these. Of these, however, 190 cases (9%) would have EPE. Although nerve sparing can safely be performed in the majority of patients with a risk of EPE below this threshold, it remains critical to relate these risks to the patient's preferences and willingness to trade-off between the potential quality of life benefit of nerve sparing and the increased risk of positive surgical margins. In addition to optimising preoperative staging, surgeons could consider other tools, such as intraoperative frozen section technology (NeuroSafe), to further optimise surgical outcomes [30]. Moreover, since NeuroSafe is a time-consuming and costly procedure, our nomogram as a triage for NeuroSafe could contribute to the cost-effective deployment of NeuroSafe.

Although this study has a number of strengths, such as a large number of cases and external validation in two separate patient cohorts, some limitations have to be acknowledged. Firstly, the majority of the study data were derived from daily clinical practice, and there was no central histopathological or radiological review. However, the fact that real-world clinical data were used for model development and validation could also be a potential strength, since these features reflect the real-world clinical situation, for which this nomogram is designed. Secondly, although accounted for using multiple imputations, the percentage of prostatic lobes with one or more missing covariates (27%) in this study is a limitation. However, results from our additional analysis performed using complete-case data (data not shown) would not alter the study's main conclusions. Lastly, both model development and external validation were performed by the same study group, which solely concerned Dutch patients. We therefore encourage other (international) study groups to also validate our nomogram externally as well.

5. Conclusions

We developed a simple and robust nomogram, including mpMRI information and readily available clinical parameters, for the prediction of side-specific EPE. This nomogram should be used to optimise patient selection for nerve-sparing RP.

Author contributions: Timo F.W. Soeterik 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: Soeterik, Dijksman, van Melick, van Basten, Schoots, Küsters-Vandevelde.
Acquisition of data: Soeterik, Stomps, Küsters-Vandevelde.
Analysis and interpretation of data: Soeterik, Dijksman, van Melick, van Basten.
Drafting of the manuscript: Soeterik, Dijksman, van Melick, van Basten, Witjes, Biesma, Küsters-Vandevelde, Stomps, Schoots.
Critical revision of the manuscript for important intellectual content: Soeterik, Dijksman, van Melick, van Basten, Witjes, Biesma, Küsters-Vandevelde, Stomps, Schoots.
Statistical analysis: Soeterik.
Obtaining funding: Soeterik, van Basten.
Administrative, technical, or material support: Soeterik, Dijksman, Stomps.
Supervision: Dijksman, van Melick, van Basten, Witjes, Biesma, Küsters-Vandevelde, Stomps, Schoots.
Other: None.

Financial disclosures: Timo F.W. Soeterik 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: This work was supported by Astellas Pharma and Amgen.

Acknowledgements: The authors thank Professor Monique Roobol for providing scientific advice on the study methodology. We thank Judith Wissink (HGT) for the collection of patient data used in this study.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.euo.2020.08.008.

References

[1] Walsh PC. The discovery of the cavernous nerves and development of nerve sparing radical retropubic prostatectomy. J Urol 2007;177:1632–5.
[2] Nguyen LN, Head L, Witiuk K, et al. The risks and benefits of cavernous neurovascular bundle sparing during radical prostatectomy: a systematic review and meta-analysis. J Urol 2017;198:760–9.
[3] Mottet N, Bellmunt J, Bolla 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.
[4] Memorial Sloan Kettering Cancer Center. Pre-radical prostatectomy tool to predict probability of extraprostatic extension in prostate cancer patients. www.mskcc.org/nomograms/prostate/pre_op.
[5] Tosoian JJ, Chappidi M, Feng Z, et al. Prediction of pathological stage based on clinical stage, serum prostate-specific antigen, and biopsy Gleason score: Partin tables in the contemporary era. BJU Int 2017;119:676–83.
[6] Ohori M, Kattan MW, Koh H, et al. Predicting the presence and side of extracapsular extension: a nomogram for staging prostate cancer. J Urol 2004;171, 1844–9; discussion 1849.
[7] Sayyid R, Perlis N, Ahmad A, et al. Development and external validation of a biopsy-derived nomogram to predict risk of ipsilateral extraprostatic extension. BJU Int 2017;120:76–82.

[8] Patel VR, Sandri M, Grasso AAC, et al. A novel tool for predicting extracapsular extension during graded partial nerve sparing in radical prostatectomy. BJU Int 2018;121:373–82.
[9] Caglic I, Kovac V, Barrett T. Multiparametric MRI—local staging of prostate cancer and beyond. Radiol Oncol 2019;53:159–70.
[10] de Rooij M, Hamoen EH, Witjes JA, Barentsz JO, Rovers MM. Accuracy of magnetic resonance imaging for local staging of prostate cancer: a diagnostic meta-analysis. Eur Urol 2016;70:233–45.
[11] Rayn KN, Bloom JB, Gold SA, et al. Added value of multiparametric magnetic resonance imaging to clinical nomograms for predicting adverse pathology in prostate cancer. J Urol 2018;200:1041–7.
[12] Gandaglia G, Ploussard G, Valerio M, et al. The key combined value of multiparametric magnetic resonance imaging, and magnetic resonance imaging-targeted and concomitant systematic biopsies for the prediction of adverse pathological features in prostate cancer patients undergoing radical prostatectomy. Eur Urol 2020;77:733–41.
[13] Rocco B, Sighinolfi MC, Sandri M, et al. Is extraprostatic extension of cancer predictable? A review of predictive tools and an external validation based on a large and a single center cohort of prostate cancer patients. Urology 2019;129:8–20.
[14] Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. Eur Radiol 2012;22:746–57.
[15] Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, version 2. Eur Urol 2016;69:16–40.
[16] Buyyounouski MK, Choyke PL, McKenney JK, et al. Prostate cancer—major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin 2017;67:245–53.
[17] Baco E, Rud E, Vlatkovic L, et al. Predictive value of magnetic resonance imaging determined tumor contact length for extracapsular extension of prostate cancer. J Urol 2015;193:466–72.
[18] Cornud F, Rouanne M, Beuvon F, et al. Endorectal 3D T2-weighted 1mm-slice thickness MRI for prostate cancer staging at 1.5Tesla: should we reconsider the indirect signs of extracapsular extension according to the D'Amico tumor risk criteria? Eur J Radiol 2012;81:e591–7.
[19] Somford DM, Hamoen EH, Futterer JJ, et al. The predictive value of endorectal 3 Tesla multiparametric magnetic resonance imaging for extraprostatic extension in patients with low, intermediate and high risk prostate cancer. J Urol 2013;190:1728–34.
[20] Fine SW, Amin MB, Berney DM, et al. A contemporary update on pathology reporting for prostate cancer: biopsy and radical prostatectomy specimens. Eur Urol 2012;62:20–39.
[21] Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–73.
[22] Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and interpreting decision curve analysis: a guide for investigators. Eur Urol 2018;74:796–804.
[23] Van Buuren S, Groothuis-Oudshoorn K. Multivariate imputation by chained equations in R. J Stat Softw 2011;45:1–67.
[24] van der Leest M, Cornel E, Israel B, et al. Head-to-head comparison of transrectal ultrasound-guided prostate biopsy versus multiparametric prostate resonance imaging with subsequent magnetic resonance-guided biopsy in biopsy-naive men with elevated prostate-specific antigen: a large prospective multicenter clinical study. Eur Urol 2019;75:570–8.
[25] Nyarangi-Dix J, Wiesenfarth M, Bonekamp D, et al. Combined clinical parameters and multiparametric magnetic resonance imaging for the prediction of extraprostatic disease—a risk model for patient-tailored risk stratification when planning radical prostatectomy. Eur Urol Focus. In press. https://doi.org/10.1016/j.euf.2018.11.004.
[26] Martini A, Gupta A, Lewis SC, et al. Development and internal validation of a side-specific, multiparametric magnetic resonance imaging-based nomogram for the prediction of extracapsular extension of prostate cancer. BJU Int 2018;122:1025–33.
[27] Sighinolfi MC, Sandri M, Torricelli P, et al. External validation of a novel side-specific, multiparametric magnetic resonance imaging-based nomogram for the prediction of extracapsular extension of prostate cancer: preliminary outcomes on a series diagnosed with multiparametric magnetic resonance imaging-targeted plus systematic saturation biopsy. BJU Int 2019;124:192–4.
[28] Soeterik TFW, van Melick HHE, Dijksman LM, et al. External validation of the Martini nomogram for prediction of side-specific extraprostatic extension of prostate cancer in patients undergoing robot-assisted radical prostatectomy. Urol Oncol 2020;38:372–8.
[29] van Stam Ma, Pieterse Ah, van der Poel Hg, et al. Shared decision making in prostate cancer care—encouraging every patient to be actively involved in decision making or ensuring the patient preferred level of involvement? J Urol 2018;200:582–9.
[30] Schlomm T, Tennstedt P, Huxhold C, et al. Neurovascular structure-adjacent frozen-section examination (NeuroSAFE) increases nerve-sparing frequency and reduces positive surgical margins in open and robot-assisted laparoscopic radical prostatectomy: experience after 11,069 consecutive patients. Eur Urol 2012;62:333–40.