

External validation of nomograms including MRI features for the prediction of side-specific extraprostatic extension

J. G. Heetman 1, E. J. R. J. van der Hoeven², P. Rajwa 3, F. Zattoni 4, C. Kesch 5, S. Shariat^{3,6,7,8,9,10}, F. Dal Moro 4, G. Novara 4, G. La Bombarda⁴, F. Sattin⁴, N. von Ostau⁵, N. Pötschl¹, P. A. T. Baltzer¹¹, L. Wever¹, J. P. A. Van Basten¹², H. H. E. Van Melick¹, R. C. N. Van den Bergh¹, G. Gandaglia¹³, T. F. W. Soeterik^{1,14} ✉ and on behalf of the European Association of Urology Young Academic Urologists Prostate Cancer Working Party*

BACKGROUND: Prediction of side-specific extraprostatic extension (EPE) is crucial in selecting patients for nerve-sparing radical prostatectomy (RP). Multiple nomograms, which include magnetic resonance imaging (MRI) information, are available to predict side-specific EPE. It is crucial that the accuracy of these nomograms is assessed with external validation to ensure they can be used in clinical practice to support medical decision-making.

METHODS: Data of prostate cancer (PCa) patients that underwent robot-assisted RP (RARP) from 2017 to 2021 at four European tertiary referral centers were collected retrospectively. Four previously developed nomograms for the prediction of side-specific EPE were identified and externally validated. Discrimination (area under the curve [AUC]), calibration and net benefit of four nomograms were assessed. To assess the strongest predictor among the MRI features included in all nomograms, we evaluated their association with side-specific EPE using multivariate regression analysis and Akaike Information Criterion (AIC).

RESULTS: This study involved 773 patients with a total of 1546 prostate lobes. EPE was found in 338 (22%) lobes. The AUCs of the models predicting EPE ranged from 72.2% (95% CI 69.1–72.3%) (Wibmer) to 75.5% (95% CI 72.5–78.5%) (Nyarangi-Dix). The nomogram with the highest AUC varied across the cohorts. The Soeterik, Nyarangi-Dix, and Martini nomograms demonstrated fair to good calibration for clinically most relevant thresholds between 5 and 30%. In contrast, the Wibmer nomogram showed substantial overestimation of EPE risk for thresholds above 25%. The Nyarangi-Dix nomogram demonstrated a higher net benefit for risk thresholds between 20 and 30% when compared to the other three nomograms. Of all MRI features, the European Society of Urogenital Radiology score and tumor capsule contact length showed the highest AUCs and lowest AIC.

CONCLUSION: The Nyarangi-Dix, Martini and Soeterik nomograms resulted in accurate EPE prediction and are therefore suitable to support medical decision-making.

INTRODUCTION

Accurate prediction of extraprostatic extension (EPE) of prostate cancer (PCa) is crucial for preoperative risk assessment, especially in case nerve-sparing surgery is desired. Several previous studies showed that combining multi-parametric resonance imaging (MRI) features with other clinical parameters such as prostate-specific antigen (PSA) serum values and biopsy information, improves the accuracy of EPE risk prediction [1–3]. To further individualize surgical planning, the risk of EPE can be established for both prostate lobes separately; in a side-specific manner. A number of nomograms for side-specific EPE prediction, including MRI

parameters, have been recently developed previously [4–7]. The variables used as inputs for these four nomograms are detailed in Supplementary Table 1. To determine whether these nomograms can be safely applied in daily clinical practice, their diagnostic performance should be established using contemporary patient cohorts other than the ones used for model development [8]. As the prediction formula is tailored to the development data, a nomogram may show excellent performance in the development population but can perform poorly in an external cohort. Preferably, a nomogram is externally validated in different cohorts [9]. The nomograms described by Soeterik et al. and Martini et al.,

1Department of Urology, St. Antonius Hospital, Utrecht, The Netherlands. 2Department of Radiology, St. Antonius Hospital, Utrecht, The Netherlands. 3Department of Urology, Medical University of Vienna, Vienna, Austria. 4Department of Surgery, Oncology and Gastroenterology, University of Padua, Padua, Italy. 5Department of Urology, University Hospital Essen, Essen, Germany. 6Institute for Urology and Reproductive Health, Sechenov University, Moscow, Russia. 7Department of Special Surgery, The University of Jordan, Amman, Jordan. 8Department of Urology, University of Texas Southwestern Medical Center, Dallas, USA. 9Department of Urology, Second Faculty of Medicine, Charles University, Prague, Czechia. 10Department of Urology, Weill Cornell Medical College, New York, USA. 11Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Vienna, Austria. 12Department of Urology, Canisius Wilhelmina Hospital, Nijmegen, The Netherlands. 13Unit of Urology/Division of Oncology, San Raffaele Hospital, Milan, Italy. 14Department of Radiation Oncology, University Medical Center Utrecht, Utrecht, The Netherlands. *A list of authors and their affiliations appears at the end of the paper.

✉email: t.soeterik@antoniusziekenhuis.nl

include EPE risk prediction respectively in trichotomous (no tumor present on MRI, suspicious lesion on present on MRI, and EPE present on MRI) and dichotomous fashion (EPE present yes or no) [4, 6]. Both nomograms have been externally validated, showing moderate to good discrimination and moderate to strong calibration respectively [10–14]. The nomograms developed by Nyarangi-Dix et al. and Wibmer et al. include other MRI features such as tumor capsule contact length (TCCL) on MRI and the European Society of Urogenital Radiology (ESUR) score for EPE

[5, 7]. In prior studies, these quantification methods for establishing EPE risk have been shown to improve diagnostic accuracy of MRI [15–18]. However, it is unclear if the incorporation of these promising MRI features into nomograms leads to improved EPE risk prediction, as both Nyarangi-Dix and Wibmer nomograms have not yet been externally validated. It is crucial that the accuracy of these nomograms is assessed in patients that underwent diagnostic evaluation according to the contemporary guidelines, minimizing the risk of discordance between biopsy and surgical pathology [19]. In addition, to minimize inter-reader variability of MRI interpretation, reporting should be done according to the most recent prostate imaging reporting and data system (PI-RADS) version 2.1 [20]. Therefore, the aim of this study is to externally validate four available side-specific EPE nomograms including MRI parameters, by using an international multi-center contemporary cohort of patients with prostate cancer undergoing radical prostatectomy.

MATERIALS/SUBJECTS AND METHODS

Patient population and study data

Data of consecutive patients undergoing radical prostatectomy at four high-volume European tertiary referral centers, from 2017 to 2021, were used for the analyses. All clinically relevant variables in addition to those included in the four nomograms were retrospectively collected. Prostate biopsy evaluation and histopathological evaluation of the surgical specimens was done according to the International Society of Urogenital Pathology (ISUP) guidelines [21]. MRI prostate (either biparametric [bp] or multiparametric [mp]) reading and reporting was done according to PI-RADS version 2.1 [22]. If missing, the ESUR score for EPE and tumor capsule contact length were retrospectively determined by experienced uro-radiologists in a side-specific manner, according to the ESUR guidelines [23].

Model discrimination, calibration and clinical usefulness External validation of the four nomograms was done according to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines [24]. Discrimination, which refers to the ability of the nomogram 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 (AUC) [24]. Furthermore, the AUC of a logistic regression model, comprising PSA and the presence of EPE on MRI (no lesion, no EPE, equivocal, and EPE), was provided to demonstrate the additional value of the nomograms compared to generally used clinical parameters. The Model calibration, which refers to the agreement between observed endpoints and predictions, was assessed using calibration slopes [24]. The net benefit per risk threshold was determined using decision-curve analysis (DCA). 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) [25].

Association between the MRI features and EPE

The predictive value of individual MRI variables included in the nomograms were assessed by multivariate regression analysis. In

Table 1. Descriptive characteristics on patient level in the overall population.

Overall	
No. of patients	773
Age (median, IQR)	67 (62, 71)
PSA (ng/ml)	7.5 (5.5, 11.0)
Mean (IQR)	
PSA density (ng/ml/ml)	0.19 (0.12, 0.28)
Mean (IQR)	
Clinical T stage N (%) ^a	
T1	424 (55)
T2	284 (38)
T3	45 (6)
Missing	20 (3)
MRI T stage N (%) ^a	
T0	44 (6)
T2	493 (64)
T2/T3 (uncertain EPE)	126 (16)
T3	107 (14)
T4	2 (0)
Missing	1 (0)
Biopsy type N (%) ^a	
TRUS-guided systematic	168 (22)
MRI-guided	47 (6)
TRUS + MRI-guided	552 (71)
Missing	4 (1)
Biopsy ISUP Grade Group N (%) ^a	
Benign	5 (1)
1	126 (16)
2	252 (33)
3	171 (22)
4	150 (19)
5	67 (9)
missing	2 (0)
Pathological stage N (%) ^a	
T0	1 (0)
T2	460 (60)
T3	301 (40)
T4	1 (0)
missing	10 (1)
Radical prostatectomy ISUP grade Group N (%) ^a	
1	60 (8)
2	313 (40)
3	224 (29)
4	79 (10)
5	80 (10)
missing	17 (2)
Postive surgical margin N (%) ^a	246 (32)
missing	14 (2)

EPE extraprostatic extension, PSA prostate specific antigen, PI-RADS Prostate Imaging-Reporting and Data System, ISUP International Society of Urologic Pathology, TRUS transrectal ultrasound.

^aPercentages may not sum to 100% due to rounding.

Table 2. Descriptive characteristics on a per lobe level in the total cohort, divided by the presence of extra prostatic extension.

Characteristic	No EPE, (n = 1204)	EPE (n = 338)	p value
Age(years) (IQR)	67 (62, 71)	68 (63, 72)	0.067
PSA (ng/ml) (IQR)	7.2 (5.3, 10.5)	8.6 (5.9, 13.0)	<0.001
Prostate volume(ml) (IQR)	40 (30, 55)	39 (30, 53)	0.3
PSA density (ng/ml/ml) (IQR)	0.18 (0.12, 0.27)	0.22 (0.15, 0.34)	<0.001
Clinical T stadium N (%a)			<0.001
cT1/cT2a	1138 (95)	280 (83)	
cT2b/c	38 (3)	41 (12)	
cT3/4	19 (2)	14 (4)	
Unknown	9(1)	3 (1)	
PI-RADS 4 or 5, N (%a) missing	602 (50) 8 (1)	264 (78) 1 (0)	<0.001
EPE on MRI, N (%a)			<0.001
No visible lesion	571 (47)	74 (22)	
No EPE	499 (41)	143 (42)	
Equivocal	90 (7)	58 (17)	
EPE	38 (3)	61 (18)	
Unknown	6 (0)	2 (1)	
ESUR score (IQR)	0 (0, 1)	1 (1, 5)	<0.001
TCCL(mm) (IQR)	0 (0,10)	12 (5,21)	<0.001
ISUP Grade Group, N (%a)			
Benign	332 (28)	38 (11)	
1	279 (23)	33 (10)	
2	263 (22)	64 (19)	
3	145 (12)	62 (18)	
4	124 (10)	81 (24)	
5	35 (3)	57 (17)	
Unknown	26 (2)	3 (1)	
Percentage positive SB (%a) (IQR) missing	20 (0, 50) 52	50 (17, 83) 15	<0.001
Tumor extent in SB (mm) (IQR) Missing	2 (0, 10) 179	9 (0, 13) 65	<0.001
Tumor involvement in biopsy (%a) (IQR) missing	15 (0, 50) 80	50 (15,87) 27	<0.001
Cohort, N (%)			
1	460 (38)	112 (33)	
2	337 (28)	113 (33)	
3	329 (27)	81 (24)	
4	78 (6)	32 (9)	

EPE extraprostatic extension, PSA prostate specific antigen, PI-RADS Prostate Imaging-Reporting and Data System, MRI magnetic resonance imaging, ESUR European Society of Urogenital Radiology, TCCL tumor capsule contact length, ISUP International Society of Urologic Pathology, SB systematic biopsy.

aPercentages may not sum to 100% due to rounding.

multivariable analysis, including PSA density and biopsy ISUP Grade Group, value of the five different MRI variables included in the four nomograms (supplement Table 1), was assessed using an AIC (Akaike Information Criterion) to determine the features with the best fit. In addition, the AUC of the ROC was established.

Missing data

Missing data patterns were explored using response matrix and correlation plots. Missing data were handled by using multivariate imputation by chained equations including pooling using Rubin's rules [26].

Table 3. Discrimination of all four nomograms in the overall population.

	Overall	
	Lobes	AUC (95% CI)
Soeterik	1546 (100%)	74.6% (71.6–77.7%)
Martini	1150 (74%)	74.3% (71.1–77.6%)
Wibmer	1469 (95%)	72.2% (69.1–75.3%)
Nyarangi-Dix	1546 (100%)	75.5% (72.5–78.5%)

AUC Area Under the Curve.

RESULTS

Baseline characteristics

A total of 773 patients were included, representing a total of 1546 prostate lobes. Descriptive characteristics of the total cohort are presented in Table 1 and per cohort in supplementary Table 2. A bpMRI was used in 288 (37%) patients, and a mpMRI in 485 (63%). The characteristics per lobe regarding the covariates used in the different nomograms are presented in Table 2. Of all the lobes 338 (22%) had EPE in prostatectomy specimens. In the per lobe analysis, presence of EPE was associated with relatively higher absolute serum PSA levels and PSA density measured in the patient. The lobes with EPE had more PI-RADS 4 or 5 lesions, a higher ESUR score, more tumor involvement in the biopsy cores, a higher percentage of positive systematic biopsy cores, more tumor extent in the systematic biopsy cores, and a higher ISUP Grade Group. In one of the four cohorts, all data regarding tumor core involvement (%) was not available. Due to the extensive amount of missing data, we decided not to impute this variable. Therefore, the prostate lobes containing cancer of this cohort were excluded per analysis. Missing data patterns of other variables showed data to be either missing completely at random or missing at random and were therefore imputed.

Model discrimination

The AUCs for the four nomograms are shown in Table 3. The AUC values are comparable between the four nomograms, ranging from 72.2% (95% CI 69.1–75.3%) for the Wibmer nomogram (lowest) to 75.5% (95% CI 72.5–78.5%) for the Nyarangi-Dix nomogram (highest). All nomograms exhibited a higher AUC than the 70.4% (95% CI 67.2–73.6%) of the model with PSA and EPE on MRI. AUCs of all four nomograms per individual hospital are presented in the Supplemental section (Supplementary Table 3); showing in- between-hospital differences of AUC values of all four nomograms.

Model calibration

The agreement between predicted and observed probabilities of all four nomograms are shown in Fig. 1. For the clinically most relevant thresholds for the risk of EPE of 0 to 40%, calibration was fair to good for the Soeterik, Martini and Nyarangi-Dix nomograms. The Soeterik and Martini nomogram showed slight overestimation, whereas the Nyarangi-Dix nomogram showed slight underestimation of EPE probability. For the Wibmer nomogram, substantial overestimation of the predicted risk was shown for the thresholds of 25% and above. Overall, the Soeterik nomogram showed the highest agreement of the predicted and observed probabilities for thresholds 0–90%.

Clinical utility

The DCA of the four nomograms are shown in Fig. 2. All four nomograms can be regarded as clinically useful for risk thresholds 9–30%. The Nyarangi-Dix nomogram showed slightly lower net benefit compared with the “treat all” approach for risk thresholds 3–11%, respectively. The Wibmer nomogram showed a no benefit from risk threshold 40% and above, leading to a negative net benefit for risk thresholds 40% and above. The Soeterik and

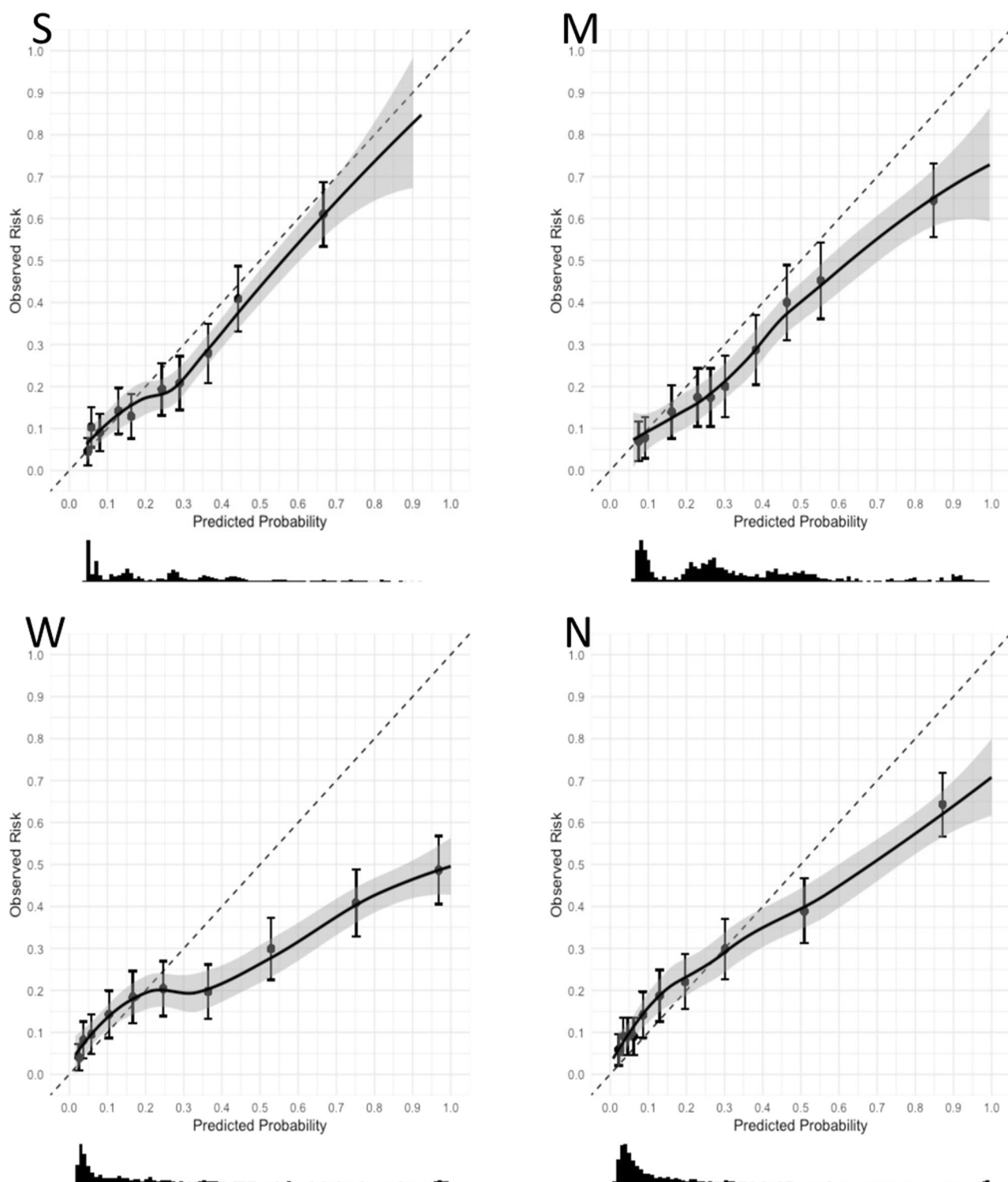


Fig. 1 Calibration slope for all four models. The Soeterik (S) nomogram, located on the top left, demonstrated fair to good calibration for the clinically most relevant thresholds concerning the risk of extraprostatic extension (EPE) from 0 to 40%. In addition, it showcased the highest concordance between its predictions and observed probabilities across the range of 0–90%. The Martini (M) nomogram, positioned on the top right, also exhibited a fair to good calibration for the 0–40% EPE risk thresholds. The Wibmer (W) nomogram, situated on the bottom left, displayed a more pronounced discrepancy. It substantially overestimated the predicted risk from a threshold of 25% and higher. Lastly, the Nyarangi-Dix (N) nomogram, located on the bottom right, was calibrated as fair to good for the clinically significant 0–40% EPE risk thresholds. Notably, in contrast to the Soeterik and Martini nomograms, it tended to slightly underestimate the EPE probability within this range.

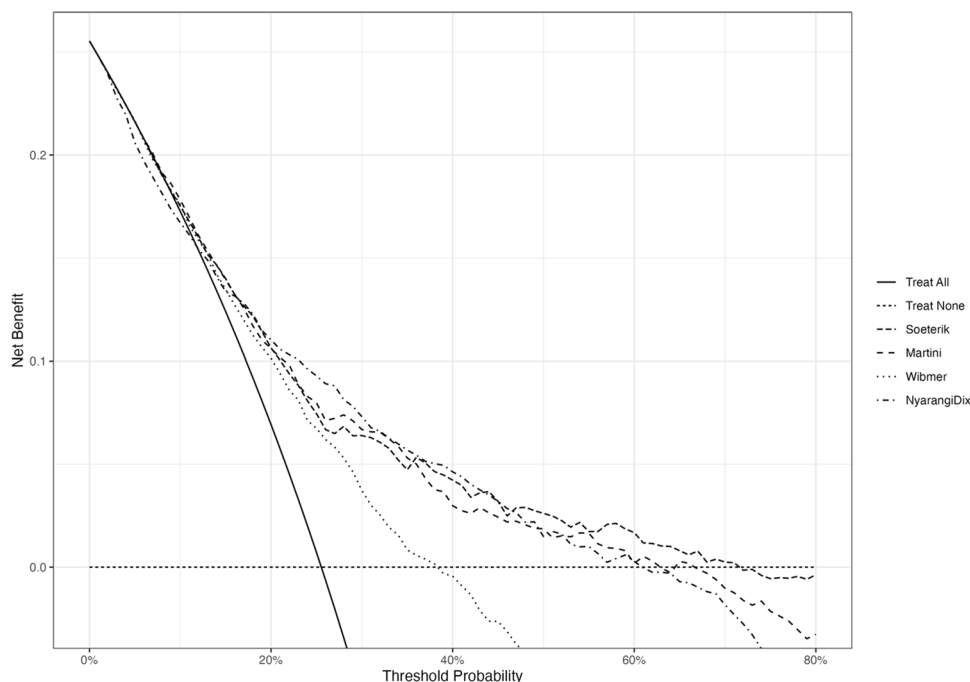


Fig. 2 Decision-curve analysis for the four models. The Soeterik nomogram showed a net benefit for risk thresholds ranging from 0 to 70%. The Martini nomogram performed comparable to the Soeterik nomogram, in the range from 0 to 35%. The Wibmer nomogram offered no benefit from a risk threshold of 40% and above. The Nyarangi-Dix nomogram while it exhibited slightly lower net benefit than the “treat all” approach for risk thresholds between 3 and 9%, it surpassed the other three models by offering a slightly higher net benefit for thresholds between 20 and 30%.

Martini nomogram showed comparable net benefit for risk thresholds 0–35%. The Nyarangi-Dix nomogram was associated with a slightly higher net benefit for the risk thresholds 20–30%, compared with the other three nomograms.

Predictive value of side-specific MRI features

In Table 4, an analysis of the different MRI features incorporated in the four models is shown. The dichotomous classification for EPE had a lower AUC 74.2 (95% CI 71.1–77.3) and a higher AIC of 1426 than the trichotomous classification for EPE used by Soeterik (AUC of 75.1 [95% CI 72.0–78.1] and an AIC of 1415) and the classification used by Wibmer 75.2 (95% CI 72.2–78.3) AIC 1414. The continuous variables used by Nyarangi-Dix had the highest AUCs and lowest AIC with an AUC of 76.5 (95% CI 73.6–79.5) and an AIC of 1396 for TCCL and an AUC of 76.3 (95% CI 73.3–79.3) and an AIC 1379 for the ESUR score, respectively.

DISCUSSION

In this study, we present the results of the external validation of four MRI-based nomograms for the prediction of side specific EPE in a European dataset consisting of 773 patients with a total of 1546 prostate lobes. We observed a fair discriminative ability of all four nomograms, with AUC’s ranging from 72.2 to 75.5%. The calibration of the Soeterik, Martini, and Nyarangi-Dix nomograms was fair to good for the clinically most relevant risk thresholds of 0–40%. The Wibmer nomogram showed substantial overestimation of the predicted EPE risk for risk thresholds from 25% and above. DCA showed that the Soeterik, Martini and the Nyarangi-Dix nomograms are all clinically useful for risk thresholds 8 to 40%. We conclude that the Soeterik, Martini and the Nyarangi-Dix nomograms are well suitable for use in clinical practice. Based on this study, the Wibmer nomogram should be used with caution due to substantial miscalibration and limited clinical usefulness for risk thresholds above 25%.

Our findings regarding model performance of the Soeterik and Martini nomograms are consistent to those reported in previous external validation studies. The study of Blas et al. reported an AUC of 81% for the Soeterik nomogram and 75% for the Martini nomogram, respectively [13]. Another external validation study of the Martini nomogram reported an AUC of 78% [14]. The study of Veerman et al. reported an AUC of 80% for the Soeterik nomogram [12]. A different study by Diamand et al. presented an AUC of 71% for the Soeterik nomogram and 73% for the Martini nomogram [11]. With regard to calibration, these prior studies all reported moderate to good agreement of predicted and observed probabilities for both the Martini and the Soeterik nomogram. To our knowledge, this is the first study in which the Wibmer and Nyarangi-Dix nomograms are externally validated. On external validation, they both showed substantially lower AUCs compared to the AUCs reported for the development cohorts; respectively 76% versus 87% for Nyarangi-Dix and 72.2% versus 82.8% for the Wibmer nomogram [5, 7]. In this study, the Wibmer nomogram showed the relatively lowest AUC of all validated nomograms of respectively 72.2%, substantial underestimation of the predicted EPE risk from thresholds above 25% and a negative net benefit on DCA for thresholds above 40%. The Nyarangi-Dix nomogram showed more favorable results, with an AUC of 75.5%, fair agreement between predicted and observed probabilities and the highest net benefit on DCA compared with the other nomograms (for the clinically most relevant risk thresholds between 10 to 40%). The reason the Nyarangi-Dix nomogram showed slightly better model performance compared to the other nomograms could be due to the inclusion of the potentially more robust MRI predictors: TCCL and the ESUR score. Due to the scaling of these variables, they may have the potential to explain more variance compared to MRI predictors including solely two or three subclasses. This hypothesis is supported by our multivariable analysis for the

GG4	3.5 (2.3, 5.6) < 0.001
GG5	7.4 (4.2, 13.2) < 0.001

Table 4. Model discrimination of the multivariable logistic regression models including different MRI variables in the overall population.

OR 95% CI p value	AUC 95% CI	AIC
Model 1	74.2 (71.1–77.3)	1426
PSAD	2.0(1.2, 3.3) 0.007	
ISUP		
Benign	Reference	
GG1	1.0 (0.6, 1.6) > 0.9	
GG2	2.1 (1.4, 3.2) < 0.001	
GG3	3.2 (2.0, 5.0) < 0.001	
GG4	5.0 (3.3, 7.8) < 0.001	
GG5	11.0 (6.4, 19.1) < 0.001	
MRI		
No EPE	Reference	
EPE	4.0 (2.5, 6.3) < 0.001	
Model 2	75.1 (72.0–78.1)	1415
PSAD	1.8 (1.1, 3.0) 0.02	
ISUP		
Benign	Reference	
GG1	0.9 (0.5, 1.4) 0.6	
GG2	1.6 (1.0, 2.5) 0.048	
GG3	2.4 (1.5, 3.9) < 0.001	
GG4	3.9 (2.5, 6.2) < 0.001	
GG5	8.7 (5.0, 15.3) < 0.001	
MRI		
No lesion	Reference	
No EPE	1.8 (1.3, 2.5) < 0.001	
EPE	6.0 (3.6, 10.1) < 0.001	
Model 3	75.2 (72.2–78.3)	1414
PSAD	1.8 (1.1, 3.0) 0.02	
ISUP		
Benign	Reference	
GG1	1 (0.6, 1.6) > 0.9	
GG2	1.9 (1.2, 2.9) 0.004	
GG3	2.9 (1.8, 4.6) < 0.001	
GG4	4.6 (2.9, 7.1) < 0.001	
GG5	9.5 (5.5, 16.7) < 0.001	
MRI		
No EPE	Reference	
Equivocal	2.1 (1.4, 3.1) < 0.001	
EPE	4.5 (2.8, 7.3) < 0.001	
Model 4	76.5 (73.6–79.5)	1396
PSAD	1.7 (1.0, 2.9) 0.03	
ISUP		
Benign	Reference	
GG1	0.8 (0.5, 1.4) 0.5	
GG2	1.4 (0.9, 2.2) 0.14	
GG3	2.4 (1.5, 3.8) < 0.001	

Table 4. continued

	OR 95% CI p value	AUC 95% CI	AIC
MRI	1.1 (1.0, 1.1) < 0.001		
TCCL			
Model 5		76.3 (73.3–79.3)	1379
PSAD	1.6 (1.0, 2.7) 0.08		
ISUP	Reference		
Benign			
GG1	0.9 (0.6, 1.5) 0.8		
GG2	1.6 (1.0, 2.5) 0.045		
GG3	2.6 (1.7, 4.2) < 0.001		
GG4	3.7 (2.3, 5.8) < 0.001		
GG5	7.1 (4.0, 12.6) < 0.001		
MRI			
ESUR	1.3 (1.2, 1.4) < 0.001		

PSAD prostate specific antigen density, ISUP International Society of Urologic Pathology, MRI magnetic resonance imaging, TCCL tumor capsule contact length, ESUR European Society of Urogenital Radiology, EPE extraprostatic extension, Area Under the Curve, AIC Akaike Information Criterion.

prediction of EPE; showing that inclusion of TCCL and the ESUR score on multivariable logistic regression leads to overall best model fit in terms of AIC as well as most favorable discrimination in terms of AUC. The suggested higher predictive potential is countered by the additional effort required to document these features (in a side specific manner) during routine clinical care - making the nomogram potentially less easy-to-use in daily practice. Besides TCCL and the ESUR score, other methods have been proposed to improve EPE risk prediction. For instance, the use of artificial intelligence and radiomics features could potentially further improve EPE risk prediction. Hou et al. showed an excellent AUC of 86% for their developed artificial intelligent model, showing the outperform the radiologist (AUC of 72%) for the prediction of EPE [27]. In another study, combined use of MRI index lesion radiomics in a machine learning model was demonstrated to have a high accuracy for EPE detection, reaching an overall accuracy of 83% in the training set [28]. In addition, a prior study by Solari and colleagues showed that PSMA PET/MR radiomics could further improve prostate cancer staging in addition to MRI radiomics. The authors evaluated 9 support vector machine models with PET and/or MRI radiomics features including the apparent diffusion coefficient (ADC). The authors concluded that the best performing model included both PET and ADC radiomics; suggesting their complementary value [29].

Moving forward, it is also crucial to evaluate if the use of side-specific EPE nomograms leads to improved patient selection for nerve sparing RP. Such an approach could potentially enhance functional outcomes owing to the benefits of more nerve preservation without risking a PSM [30, 31]. However, studies on

this topic are scarce. To our best knowledge, one prior single-center prospective study was performed on this subject and showed that the use of a side-specific EPE MRI-based nomogram for preoperative planning results in comparable rates of full nerve-sparing (45% vs. 30%; $p = 0.083$), but relatively lower rates of PSM on lobes with histological EPE (45% vs. 85%; $p < 0.05$) [32]. Future prospective multicenter trials are needed to further evaluate if the use of

nomograms for preoperative planning improve clinical outcomes for the patient.

Although our study has a number of strengths such as being a multicenter international study including a contemporary population of patients treated at tertiary referral centers, it is not exempt

from limitations. First, due to the retrospective collection of data there is a risk of information bias. In addition, although MRI review was performed by dedicated high-volume uro-radiologists, the lack of central review is a limitation. However, interobserver variability is unavoidable in daily clinical practice and thus, on the other hand, our study reflects a real-world clinical situation. It should also be noted that we used both bpMRI and mpMRI in this study. However, we do not consider this as a major limitation as both modalities have been shown to be comparably effective in detecting EPE [33].

CONCLUSION

The external validation of four side-specific nomograms including MRI features showed that three of four nomograms (Nyarangi-Dix, Soeterik and Martini) showed fair to good model discrimination, calibration, and net benefit. Based on this study data, these nomograms can be used in clinical practice to support medical decision-making.

DATA AVAILABILITY

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

REFERENCES


- Feng TS, Sharif-Afshar AR, Wu J, Li Q, Luthringer D, Saouaf R, et al. Multi-parametric MRI improves accuracy of clinical nomograms for predicting extra- capsular extension of prostate cancer. *Urology*. 2015;86:332–7.
- Rayn KN, Bloom JB, Gold SA, Hale GR, Baiocco JA, Mehralivand S, et al. Added value of multiparametric magnetic resonance imaging to clinical nomograms for predicting adverse pathology in prostate cancer. *J Urol*. 2019;200:1041–7.
- Gandaglia G, Ploussard G, Valerio M, Mattei A, Fiori C, Roumiguié M, et al. The key combined value of multiparametric magnetic resonance imaging, and magnetic resonance imaging–targeted and concomitant systematic biopsies for the pre- diction of adverse pathological features in prostate cancer patients undergoing radical prostatect. *Eur Urol*. 2020;77:733–41.
- Martini A, Gupta A, Lewis SC, Cumarasamy S, Haines KG, Briganti A, 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.
- Nyarangi-Dix J, Wiesenfarth M, Bonekamp D, Hithaler B, Schütz V, Dieffenbacher S, 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*. 2020;6:1205–12.
- Soeterik TFW, Van Melick HHE, Dijkman LM, Küsters-Vandeveld H, Stomps S, Schoots IG, et al. Development and external validation of a novel nomogram to predict side-specific extraprostatic extension in patients with prostate cancer undergoing radical prostatectomy. *Eur Urol Oncol*. 2020;S2588-9311.
- Wibmer AG, Kattan MW, Alessandrino F, Baur ADJ, Boesen L, Franco FB, et al. International multi-site initiative to develop an mri-inclusive nomogram for side-specific prediction of extraprostatic extension of prostate cancer. *Cancers*. 2021;13:2627.
- Bleeker SE, Moll HA, Steyerberg EW, Donders ART, Derksen-Lubsen G, Grobbee DE, et al. External validation is necessary in prediction research: A clinical example. *J Clin Epidemiol*. 2003;56:826–32.
- Siontis GCM, Tzoulaki I, Castaldi PJ, Ioannidis JPA. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. *J Clin Epidemiol* 2015;68:25–34.
- Soeterik TFW, van Melick HHE, Dijkman LM, Küsters-Vandeveld H, Stomps S, Schoots IG, et al. Development and external validation of a novel nomogram to predict side-specific extraprostatic extension in patients with prostate cancer undergoing radical prostatectomy. *Eur Urol Oncol*. 2022;5:328–37.
- Diamand R, Roche JB, Lievore E, Lacetera V, Chiacchio G, Beatrice V, et al. External validation of models for prediction of side-specific extracapsular extension in prostate cancer patients undergoing radical prostatectomy. *Eur Urol Focus*. 2023;9:309–16.
- Veerman H, Heymans MW, van der Poel HG. External validation of a prediction model for side-specific extraprostatic extension of prostate cancer at robot- assisted radical prostatectomy. *Eur Urol Open Sci*. 2022;37:50–52.
- Blas L, Shiota M, Nagakawa S, Tsukahara S, Matsumoto T, Lee K, et al. Validation of user-friendly models predicting extracapsular extension in prostate cancer patients. *Asian J Urol*. 2023;10:81–88.

1. Soeterik TFW, van Melick HHE, Dijkman LM, Küsters-Vandeveldel HVN, Biesma DH, Witjes JA, 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. *Urologic Oncol: Semin Original Investig.* 2020;38:372–8.
2. Baco E, Rud E, Vlatkovic L, Svindland A, Eggesbø HB, Hung AJ, et al. Predictive value of magnetic resonance imaging determined tumor contact length for extracapsular extension of prostate cancer. *J Urol.* 2015;193:466–72.
3. Eurboonyanun K, Pisuchpen N, O'Shea A, Lahoud RM, Atre ID, Harisinghani M. The absolute tumor-capsule contact length in the diagnosis of extraprostatic extension of prostate cancer. *Abdom Radiol.* 2021;46:4014–24.
4. Li W, Dong A, Hong G, Shang W, Shen X. Diagnostic performance of ESUR scoring system for extraprostatic prostate cancer extension: a meta-analysis. *Eur J Radiol.* 2021;143:109896.
5. Asfuroğlu U, Asfuroğlu BB, Özer H, Gönül İI, Tokgöz N, İnan MA, et al. Which one is better for predicting extraprostatic extension on multiparametric MRI: ESUR score, Likert scale, tumor contact length, or EPE grade? *Eur J Radiol.* 2022;149:110228.
6. Ahdoot M, Wilbur AR, Reese SE, Lebastchi AH, Mehralivand S, Gomella PT, et al. MRI-targeted, systematic, and combined biopsy for prostate cancer diagnosis. *NEngl J Med* 2020;382:917–28.
7. Padhani AR, Weinreb J, Rosenkrantz AB, Villeirs G, Turkbey B, Barentsz J. Prostate imaging-reporting and data system steering committee: PI-RADS v2 status update and future directions. *Eur Urol.* 2019;75:385–96.
8. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. 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.
9. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, et al. Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2. *Eur Urol.* 2019;76:340–51.
10. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR prostate MR guidelines 2012. *Eur Radiol.* 2012;22:746–57.
11. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *Eur Urol.* 2015;67:1142–51.
12. 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.
13. Van Buuren S, Groothuis-Oudshoorn K. Multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1–67.
14. Hou Y, Zhang YH, Bao J, Bao ML, Yang G, Shi HB, et al. Artificial intelligence is a promising prospect for the detection of prostate cancer extracapsular extension with mpMRI: a two-center comparative study. *Eur J Nucl Med Mol Imaging.* 2021;48:3805–16.
15. Cuocolo R, Stanzione A, Faletti R, Gatti M, Callaris G, Fornari A, et al. MRI index lesion radiomics and machine learning for detection of extraprostatic extension of disease: a multicenter study. *Eur Radiol.* 2021;31:7575–83.
16. Solari EL, Gafita A, Schachoff S, Bogdanović B, Villagrán Asiares A, Amiel T, et al. The added value of PSMA PET/MR radiomics for prostate cancer staging. *Eur J Nucl Med Mol Imaging.* 2022;49:527–38.
17. Nguyen LN, Head L, Wituk K, Punjani N, Mallick R, Cnossen S, 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.
18. Soeterik TFW, van Melick HHE, Dijkman LM, Stomps S, Witjes JA, van Basten JPA. Nerve sparing during robot-assisted radical prostatectomy increases the risk of ipsilateral positive surgical margins. *J Urol.* 2020;204:91–95.
19. Heetman JG, Soeterik TFW, Wever L, Meyer AR, Nuininga JE, van Soest RJ, et al. Side-specific nomogram for extraprostatic extension may reduce the positive surgical margin rate in radical prostatectomy. *World J Urol.* 2022;40:2919–24.
20. Christophe C, Montagne S, Bourrelie S, Roupert M, Barret E, Rozet F, et al. Prostate cancer local staging using biparametric MRI: assessment and comparison with multiparametric MRI. *Eur J Radiol.* 2020;132:109350.

AUTHOR CONTRIBUTIONS

Conception and design: TFWS. Acquisition of data: JGH, EJRVan derH, PR, FZ, CK, SS, FDM, GN, GLB, FS, NvonO, NP, PATB, LW, HHEvM, RCNVandenB, GG, TFWS. Statistical analysis: JGH. Analysis and interpretation of data: JGH. Drafting of the manuscript: TFWS, JGH. Critical revision of the manuscript for important intellectual content: JGH, EJRVanderH, PR, FZ, CK, SS, FDM, GN, GLB, FS, NvonO, NP, PATB, LW, JPAVB, HHEvanM, RCNVandenB, GG, TFWS. Supervision: TFWSoeterik.

ON BEHALF OF THE EUROPEAN ASSOCIATION OF UROLOGY YOUNG ACADEMIC UROLOGISTS PROSTATE CANCER WORKING PARTY

BOARD MEMBERS G. Gandaglia¹³, L. Bianchi¹⁵, F. Ceci¹⁶, P. K-F. Chiu¹⁷, F. Giganti^{18,19}, I. Heidegger²⁰, V. Kasivisvanathan²¹, C. V. Kesch⁵, G. Marra²², A. Martini²³, J. Olivier²⁴, F. Preisser²⁵, P. Rajwa and F. Zattoni ⁴

ASSOCIATES K. Aas^{26,27}, U. G. Falagario²⁸, V. Fasulo²⁹, M. Maggi³⁰, I. Puche Sanz³¹, M. C. Roesch³², A. Sigle³³, T. Soeterik¹ and L. F. Stolzenbach³⁴

¹⁵Division of Urology, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy. ¹⁶Division of Nuclear Medicine, IEO European Institute of Oncology, IRCCS, Milan, Italy. ¹⁷Department of Surgery, SH Ho Urology Centre, The Chinese University of Hong Kong, Hong Kong, China. ¹⁸Department of Radiology, University College London Hospital NHS Foundation Trust, London, UK. ¹⁹Division of Surgery & Interventional Science, University College London, London, UK. ²⁰Department of Urology, Medical University of Innsbruck, Innsbruck, Austria. ²¹Division of Surgery and Interventional Science, University College London, London, UK. ²²Department of Urology, Città della Salute e della Scienza, University of Turin, Turin, Italy. ²³Department of Urology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. ²⁴Department of Urology, Lille University Hospital, Lille, France. ²⁵Martini-Klinik Prostate Cancer Center, University Hospital Hamburg Eppendorf, Hamburg, Germany. ²⁶Akershus University Hospital, Oslo, Norway. ²⁷Faculty of Medicine, University of Oslo, Oslo, Norway. ²⁸Department of Urology and Organ Transplantation, University of Foggia, Foggia, Italy. ²⁹Department of Urology, IRCCS Humanitas Research Hospital, Milan, Italy. ³⁰Department of Maternal-Infant and Urological Sciences, Sapienza University, Policlinico Umberto I Hospital, Rome, Italy. ³¹Department of Urology, Instituto Investigación Biosanitaria ibs.Granada, Hospital Universitario Virgen de las Nieves (HUVN), Granada, Spain. ³²Department of Urology, University Hospital Schleswig-Holstein Campus Lübeck, Lübeck, Germany. ³³Department of Urology, University Hospital of Freiburg, Freiburg am Breisgau, Germany. ³⁴Martini-Klinik Prostate Cancer Center, University Hospital Hamburg-Eppendorf, Hamburg, Germany.