



Development and external validation of a multivariable [⁶⁸Ga] Ga-PSMA-11 PET-based prediction model for lymph node involvement in men with intermediate or high-risk prostate cancer

Urs J. Muehlematter^{1,2} · Lilith Schweiger³ · Daniela A. Ferraro^{1,4} · Thomas Hermanns⁵ · Tobias Maurer^{6,7} · Matthias M. Heck⁶ · Niels J. Rupp⁸ · Matthias Eiber³ · Isabel Rauscher³ · Irene A. Burger^{1,9}

Received: 9 January 2023 / Accepted: 19 May 2023 / Published online: 1 June 2023

© The Author(s) 2023

Abstract

Purpose To develop and evaluate a lymph node invasion (LNI) prediction model for men staged with [⁶⁸Ga]Ga-PSMA-11 PET. **Methods** A consecutive sample of intermediate to high-risk prostate cancer (PCa) patients undergoing [⁶⁸Ga]Ga-PSMA-11 PET, extended pelvic lymph node dissection (ePLND), and radical prostatectomy (RP) at two tertiary referral centers were retrospectively identified. The training cohort comprised 173 patients (treated between 2013 and 2017), the validation cohort 90 patients (treated between 2016 and 2019). Three models for LNI prediction were developed and evaluated using cross-validation. Optimal risk-threshold was determined during model development. The best performing model was evaluated and compared to available conventional and multiparametric magnetic resonance imaging (mpMRI)-based prediction models using area under the receiver operating characteristic curves (AUC), calibration plots, and decision curve analysis (DCA). **Results** A combined model including prostate-specific antigen, biopsy Gleason grade group, [⁶⁸Ga]Ga-PSMA-11 positive volume of the primary tumor, and the assessment of the [⁶⁸Ga]Ga-PSMA-11 report N-status yielded an AUC of 0.923 (95% CI 0.863–0.984) in the external validation. Using a cutoff of $\geq 17\%$, 44 (50%) ePLNDs would be spared and LNI missed in one patient (4.8%). Compared to conventional and MRI-based models, the proposed model showed similar calibration, higher AUC (0.923 (95% CI 0.863–0.984) vs. 0.700 (95% CI 0.548–0.852)—0.824 (95% CI 0.710–0.938)) and higher net benefit at DCA. **Conclusions** Our results indicate that information from [⁶⁸Ga]Ga-PSMA-11 may improve LNI prediction in intermediate to high-risk PCa patients undergoing primary staging especially when combined with clinical parameters. For better LNI prediction, future research should investigate the combination of information from both PSMA PET and mpMRI for LNI prediction in PCa patients before RP.

Keywords Prostate cancer · Lymph node invasion · Prostate-specific membrane antigen positron emission tomography · Prediction

Introduction

Accurate primary staging of prostate cancer (PCa) is important for individualized treatment planning. Current guidelines recommend a bone scan and an abdominopelvic computed tomography (CT) or magnetic resonance imaging (MRI) for non-invasive initial staging [1, 2]. However,

novel and potentially more reliable diagnostic procedures are evolving rapidly [3]. [⁶⁸Ga]Gallium-Prostate-specific membrane antigen 11 positron-emission tomography (PET) CT or MRI (further referred to as “PSMA PET” showed a promising diagnostic accuracy for primary staging [4–9].

Despite recent advances in imaging, pelvic lymph node dissection (PLND) during radical prostatectomy (RP) represents the gold standard for nodal staging in PCa. However, the therapeutic and prognostic benefits of extended PLND (ePLND) and PLND still remain controversial [10–13]. PSMA-PET might impact the indication for PLND and its extent but the oncologic benefit is not yet known [14]. The European Association of Urology (EAU) recommended in 2022 ePLND in patients with a risk of lymph node invasion

Isabel Rauscher and Irene A. Burger are joint last authors with equal contribution.

Extended author information available on the last page of the article

(LNI) $\geq 7\%$ using the Briganti 2019 nomogram [15]. However, since up to 20% of the patients suffer a complication after PLND, there is a strong need to improve patient selection for PLND [16].

More recently, incorporation of quantitative imaging data from mpMRI or PSMA PET has been proposed to further improve LNI prediction [17–20]. However, there is still only limited data especially on the added value of PSMA PET for LNI prediction.

We aimed to develop and externally evaluate a prediction model using a combination of clinical and qualitative/quantitative information from PSMA PET/CT for prediction of LNI at RP in patients with intermediate to high-risk PCa.

Material and methods

Study design

This is a retrospective, dual-center study reported according to the current guidelines [21]. We developed three LNI prediction models in a training cohort, selected the best performing model (including a clinically meaningful risk threshold), and applied it to a validation cohort. The performance of the model was compared to the performance of available LNI prediction models [17, 18, 22–25]

Source of data and study population

Patient data from two tertiary referral centers served as data source for the training cohort (Klinikum Rechts der Isar, Technical University of Munich, Munich, Germany) and the validation cohort (University Hospital Zurich, University of Zurich, Zurich, Switzerland).

For the training cohort, the retrospective analysis was approved by the Ethics Committee of the Technical University Munich (permit 5665/13). For the validation cohort, all patients gave a general written informed consent for retrospective use of their data (Ethics Commission of the Canton of Zurich, Switzerland, BASEC Nr. 2018-01284).

We used pre-existing cohorts at both centers that were collected for works on T- and N staging in PCa regarding the training cohort and patient selection for ePLND in the validation cohort, respectively. We extended these cohorts with consecutive new patients. In both cohorts, consecutive PCa patients with histologically proven (D'Amico criteria) intermediate or high-risk PCa who underwent PSMA PET for primary staging followed by RP and ePLND were retrospectively identified (training cohort $n = 192$ between January 2013 and June 2017, validation cohort $n = 96$ between April 2016 and July 2019). Patients with missing biopsy data (training cohort $n = 19$) and without written consent for retrospective use of their data (validation cohort $n = 6$)

were excluded, leading to a final training cohort of 173 patients and validation cohort of 90 patients. Ninety-four of 173 patients of the training cohort were part of published works on T- and N staging in PCa patients [6, 8]. Sixty of 90 patients of the validation cohort were part of a published work on patient selection for ePLND in PCa [19].

In the training cohort, all included patients underwent mpMRI-targeted /standard 12-core biopsy followed by PSMA PET and RP with ePLND. ePLND was performed according to a predefined template including bilateral, separate dissection of the obturator fossa, external iliac, internal iliac, and common iliac vessels with the femoral canal and the aortic bifurcation as proximal and distal limits, respectively.

In the validation cohort, all patients with intermediate and high-risk PCa underwent a combined mpMRI-targeted/saturation biopsy (min. 40 cores) followed by PSMA PET and RP with ePLND. ePLND was performed as previously reported [26].

Because of the different biopsy approaches, difference in pathological upgrading at RP between the two cohorts was assessed.

In both cohorts, patients underwent PSMA PET according to standard procedure guidelines and no therapy has been initiated between PET and RP [8, 19].

Outcome

The predicted outcome was LNI at RP with ePLND. In the training cohort, RP was performed mainly open or robotic ($< 10\%$) with ePLND to a template of 8 predefined anatomical fields. For the validation cohort, all surgical procedures were performed via robot-assisted transperitoneal laparoscopic RP with ePLND as described earlier [26]. In both cohorts, the removed LN were assessed for LNI by specialized urologists.

Predictors

The following data was collected for each patient: Clinical parameters: age [years], PSA value at the time of PSMA PET [$\mu\text{g/l}$], highest WHO/International Society of Urological Pathology (WHO/ISUP) grade group (grade groups 1–5) [27] at systematic/targeted biopsy; Quantitative ^{68}Ga -PSMA-11-PET parameters of the primary tumor of the prostate: maximum standard uptake value (SUV_{max}), volume-based PSMA PET parameters were assessed using an absolute cut-off at $\text{SUV} \geq 4$, yielding PSMA positive volume (PSMA_{vol} , [cm^3]), and total PSMA accumulation ($\text{PSMA}_{\text{total}} = \text{PSMA}_{\text{vol}} \times \text{SUV}_{\text{mean}}$) as described earlier [19]. Qualitative PET information: the conclusion of the interpreting physician regarding LNI, i.e., PSMA PET report N-status (0, LNI negative; 0.5, equivocal for LNI; 1, LNI positive, unitless) according

to Fanti et al. [28]. Since PET reporting was not standardized during the inclusion time, all PSMA PET at both centers were reassessed by two nuclear medicine physicians in consensus and blinded regarding the outcome.

Data for comparison with published models

Additional data was extracted to compare the model's performance with published models in the validation cohort as listed in Supplemental Table S1.

Missing data

Cases with missing data were omitted (i.e., complete-case analysis).

Model development and selection

We developed three models for LNI prediction. In the first model, we combined all clinical and quantitative PET information (Model_Clinical_PET). In the second model, we added the PSMA PET report N-status as an additional variable to the first model (Model_Clinical_PET_Report). In the third model, we added the PSMA PET report N-status to the first model as a combined (ensemble) model (Model_Clinical_PET/Report) [29]. WHO/ISUP grade groups (i.e., grade groups 1–5) and the PSMA PET report N-status were treated as continuous predictor. In an internal validation, we assessed the model's discrimination ability. We chose the model with the highest internal discrimination ability for external validation. A probability threshold for clinical application was estimated from the training cohort and applied to the validation cohort.

External validation

In the validation cohort, we assessed the model's performance in terms of model calibration, model discrimination ability, and clinical application.

Model comparison with published prediction models

In the validation cohort, the selected PSMA PET model was compared with six prediction models in clinical use (mpMRI-based models, 2019 Briganti nomogram [17], Draulans et al. nomogram [18]; conventional models, MSKCC Pre-Radical Prostatectomy nomogram [22], the updated Partin tables (v.2016) [23], the Roach formula [24], and the Winter nomogram [25]). For the probability of LNI of the MSKCC Pre-Radical Prostatectomy nomogram, we used the model properties published online (https://www.mskcc.org/nomograms/prostate/pre_op/coefficients, Model

N 6599/11816, updated 01/2020). For all other prediction models, the probability for LNI was calculated using the published model formulas. The final selected model was compared to the published prediction models regarding calibration, discrimination, and clinical application.

Statistical analysis

The patient's characteristics were summarized using the mean, median, standard deviation, and interquartile range (IQR), as appropriate. Comparison of patients' characteristics was conducted by a two-sample *t*-test or Mann-Whitney *U* Test for continuous variables and χ^2 test for categorical variables.

Predictors were investigated for linearity/multicollinearity using scatter plots/generalized-variance-inflation calculations, respectively. We used a multivariable logistic regression model for the Model_Clinical_PET and Model_Clinical_PET_Report model, and two separate logistic regression models for the ensemble model (Model_Clinical_PET/Report) that were averaged using weights that were optimized using nonlinear optimization [29, 30]. Model calibration was assessed at mean, weak, and moderate level including the Brier score and Spiegelhalter's *z* [31]. Discrimination ability was assessed using AUC and clinical application using decision curve analysis, (DCA) respectively. Combined model calibration and discrimination was assessed using the index of prediction accuracy (IPA) [32]. AUC were compared using the Delong method [33]. For the internal validation, we applied a 10-times repeated tenfold cross-validation.

The probability threshold for the final model was selected using a 10-times repeated tenfold cross-validated DCA in the training cohort and by trying to match the reported spared ePLND (65.5%) and missed LNI (12.2%) for the 5% threshold for the 2012 Briganti model development [34].

For the external validation, the final model was trained on the training cohort and was applied to the validation cohort.

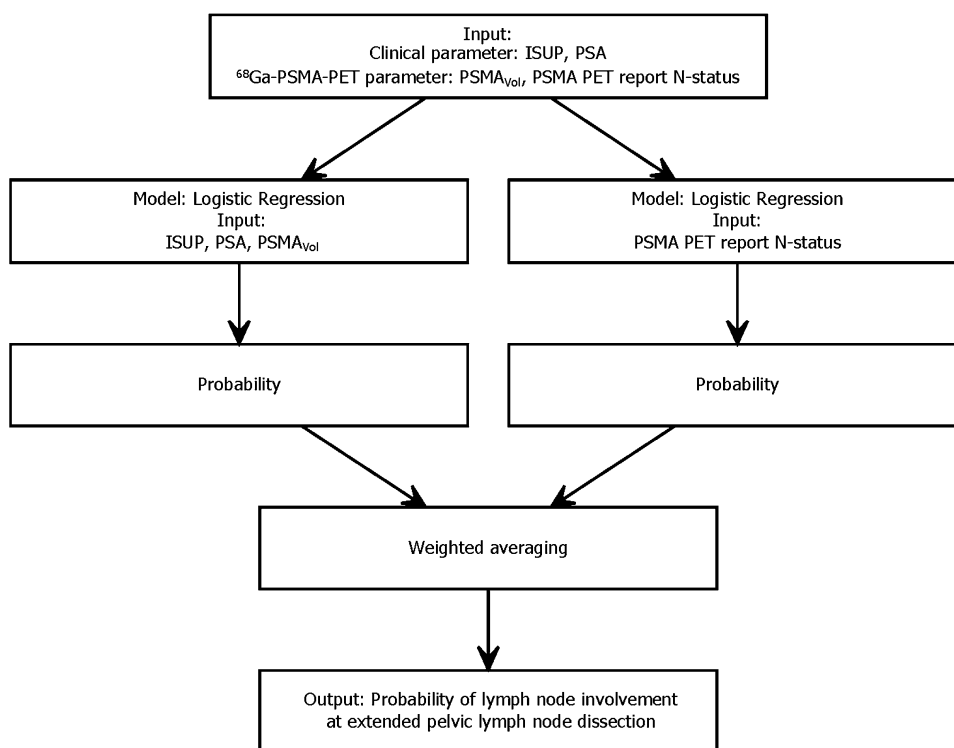
A 2-tailed *P* value of <0.05 was used to determine the statistical significance. We performed all statistical analysis in R version 4.0.5 (R Core Team (2021) R: A language and environment for statistical computing, Vienna, Austria).

Results

Patients' characteristics and qualitative PSMA-11 PET performance

The data assembly process is demonstrated in Fig. 1. A total of 173 patients were available for the training cohort and 90 patients for the validation cohort. Table 1 lists all

Fig. 1 Flow-chart of the ensemble (combined) model (Model_Clinical_PET/Report). The input (predictors) is processed within two separate logistic regression models and combined using averaged weights that were optimized using nonlinear optimization



patient's characteristics. The patients' age was significantly higher in the training cohort (mean age 71.2 vs. 64.7 years, $t = -7.34$, $df = 212.84$, $P < 0.001$). Furthermore, the biopsy WHO/ISUP grade group distribution differed between the two cohorts with more ISUP grade 1 biopsies in the training group (ISUP grade 1, 8 vs. 0%, $\chi^2 = 11.196$, $P = 0.02$). There was no significant difference regarding pathological upgrading after RP (17 vs. 16%, $\chi^2 = 0.0376$, $P = 0.85$). The number of removed lymph nodes during ePLND did not differ between the training and validation cohort (mean 24.1 vs. 23.7, $P = 0.76$).

Eighteen patients of the validation cohort had missing data concerning the 2019 Briganti model and 4 patients for the Draulans et al. model. Patient characteristics of the complete case cohorts of these models are demonstrated in the Supplementary Table S2.

Supplementary Table S3 lists the qualitative PSMA PET results.

Model development and selection

Age and SUV_{max} showed a non-linear relationship with the logit of the outcome and $PSMA_{total}$ showed the highest collinearity. Therefore, we excluded these predictors. The AUC for predicting LNI was consistently high with all three models during internal validation (Model_Clinical_PET 0.721 (CI 0.694–0.747) (ISUP, PSA, $PSMA_{vol}$ as predictors), Model_Clinical_PET_Report 0.816 (CI 0.791–0.841) (ISUP, PSA, $PSMA_{vol}$, PSMA PET report N-status as predictors), and

Model_Clinical_PET/Report 0.842 (CI 0.82–0.865) (ISUP, PSA, $PSMA_{vol}$, PSMA PET report N-status as predictors, combined in two models). Supplemental Table S4 lists the full model specifications. Model_Clinical_PET/Report (Fig. 1, Supplemental Table S4) showed the highest internally validated AUC (0.842 CI 0.82–0.865) and was selected for further analysis. The internally cross-validated DCA of this model showed a better net benefit (NB) than either the treatment or no treatment schemes when the threshold probability was ≥ 0.15 (Supplementary Fig. 1). A threshold probability of $\geq 17\%$ with an estimated spared ePLND of 54.3% and missed LNI of 19.1% fitted best the reported corresponding values of the 5% threshold of the 2012 Briganti model (spared ePLND of 65.5% and missed LNI of 12.2%) and was chosen as threshold for external validation.

External validation

In the external validation, the ensemble model Model_Clinical_PET/Report showed good calibration-in-the-large (event rate = 0.30, average predicted risk = 0.28) and there was no evidence of systematic over- or underfitting (Intercept = -0.297, Slope = 1.095, $p = 0.41$; Brier score 0.12, Spiegelhalter's $z = -0.89$). However, calibration curve showed an overestimation of the risk of LNI among patients with observed LNI probability below 0.22 and underestimated the risk of LNI among patients with observed LNI probability above 0.22 (Fig. 2). The model showed a high discrimination ability for LNI (AUC 0.923, 95% CI 0.863–0.984) (Table 2).

Table 1 Patient characteristics

Characteristics		Training cohort (n = 173)			Validation cohort (n = 90)			P
		LNI-	LNI+	P	LNI-	LNI+	P	
Age	Mean	72	74	0.20	65	63.7	0.46	<0.001
	Range	50–85	53–89		51–79	50–76		
	SD	7.9	7.2		6.02	7.30		
PSA	Median	10	14.7	0.02	8.5	14.0	0.03	0.15
	Range	1.82–93.9	0.57–100		1.22–55.0	2.08–143		
	IQR	7.92	18.87		7.9	13.5		
Highest biopsy ISUP ^a	1	12	2	0.15	-	-	0.01	0.02
	2	24	9		9	5		
	3	23	11		18	2		
	4	35	14		32	5		
	5	23	20		10	9		
cT	T1c	-	-	-	53	9	0.01	-
	T2a	-	-		15	11		
	T3a	-	-		1	1		
SUV _{max}	Median	10.7	12.73	0.08	10.2	18.0	<0.001	0.92
	Range	0–122.8	4.1–42.6		0–45.7	6.2–48.4		
	IQR	11.3	11.6		10.6	14.6		
PSMA _{vol}	Median	3.5	8.4	<0.001	3.0	12.4	<0.001	0.97
	Range	0–89.5	0.04–64.5		0–20.4	0.61–39.5		
	IQR	5.2	24.8		5.2	19.0		
PSMA _{total}	Median	21.2	49.3	<0.001	13	89.9	<0.001	0.51
	Range	0.0–863.8	0.2–643.3		0–108.6	3.0–412.1		
	IQR	37.5	166.1		44.5	148.6		
PSMA PET report N-status	LNI negative	116	17	<0.001	64	10	<0.001	0.49
	Equivocal	0	9		3	2		
	LNI positive	1	30		2	9		
LNI	0	117	0	-	69	0	-	0.17
	1	0	56		0	21		

^aAt targeted/systematic biopsy

Values in bold indicate statistical significance

Table 3 lists the results of the model application to the validation cohort according to thresholds between a predicted probability of LNI of 0 and 0.30. By using the previously estimated cut-off of ≥ 0.17 , 45 ePLNDs (45/90, 50%) would have been avoided, 44 of them in patients without LNI (44/69, 63.8%) and one in a patient with LNI (1/21, 4.8%), respectively.

Model comparison with published prediction models

The proposed model showed similar calibration compared to the conventional and mpMRI-based models. Calibration curves/calibration characteristics are depicted in Fig. 2A-E/ Supplementary Table S3 and Supplementary Fig. 2A-D/ Supplementary Table S5/6 for comparison with conventional and mpMRI-based LNI prediction models, respectively.

The proposed model showed significantly higher discrimination (AUC 0.923, 95% CI 0.863–0.984) compared to all conventional prediction models except the MSKCC model (AUC 0.824, 95% CI 0.710–0.938), and non-significant higher AUC compared to the mpMRI-based LNI prediction models (Table 2).

The proposed model showed higher combined discrimination and calibration (IPA 0.35) compared to the conventional models and a combined discrimination and calibration higher than the Draulans et al. model (IPA 0.31) and lower than the Briganti 2019 model (IPA 0.37). All IPA values are depicted in Table S5/6.

DCA revealed a high NB (0.165) of the proposed model compared to the treat-all strategy at the proposed threshold of $\geq 17\%$ (Fig. 3). Of the conventional models, only the MSKCC Pre-Radical Prostatectomy nomogram showed higher

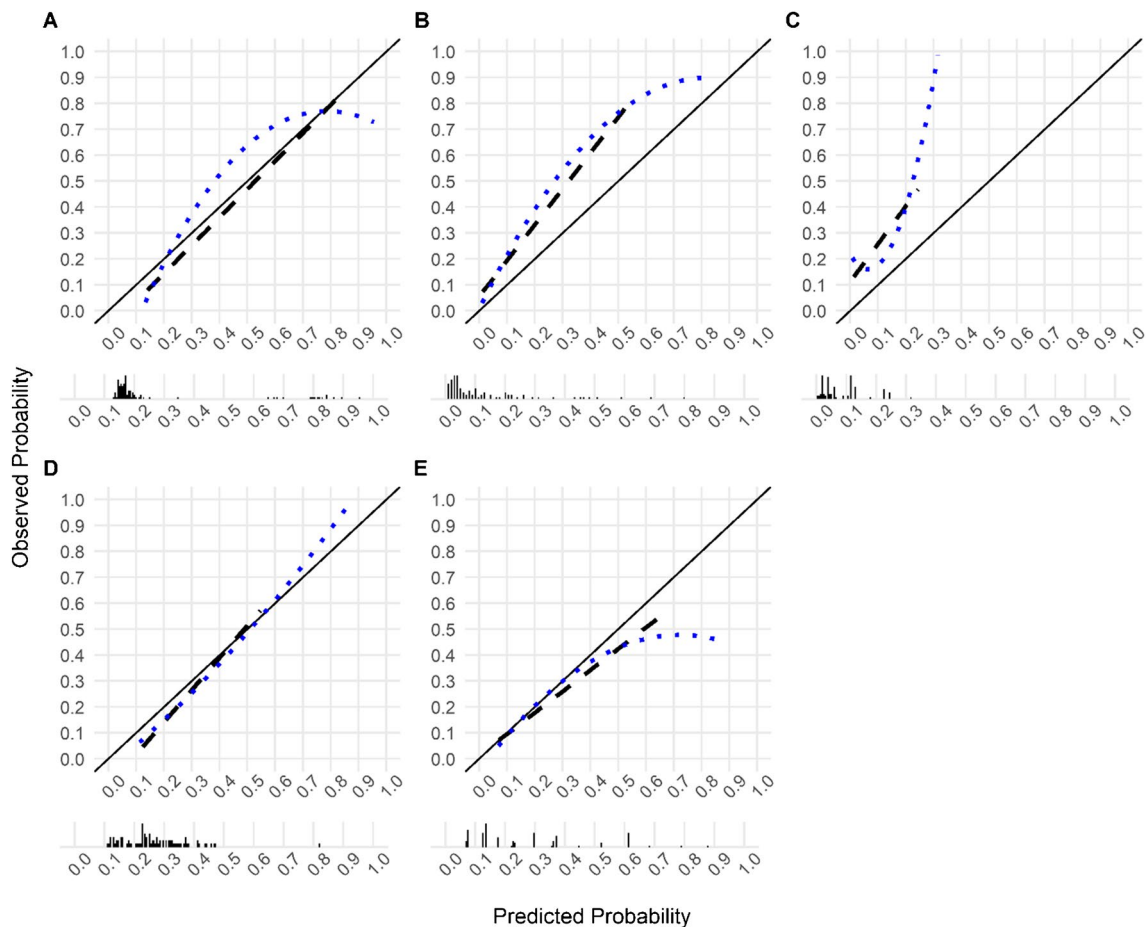


Fig. 2 Model calibration plots of predicted probability versus observed probability of lymph node involvement for (A) the proposed model (Model_Clinical_PET/Report), (B) the MSKCC Pre-Radical Prostatectomy nomogram, (C) the updated Partin tables (v.2016), (D)

the Roach formula, and (E) the Winter nomogram. The dotted lines represent the LOESS fit, the dashed lines represent a straight fit. The black bars denote the distribution of predicted probabilities

NB compared to the treat-all strategy at the recommended threshold of 5% (Fig. 3A). Of the mpMRI-based models, the 2019 Briganti nomogram showed a lower NB compared to the treat-all strategy at the recommended threshold of 7% and the Draulans et al. nomogram (no recommended threshold available) showed a consistently better NB compared to the treat-all strategy at a threshold of 7% (NB 0.189) and above (Fig. 3B, C).

Discussion

In this study, we developed and externally evaluated a multivariable prediction model including quantitative and qualitative information from PSMA PET for predicting LNI at RP with ePLND as reference and assessed its performance against conventional and mpMRI-based prediction models. Our results demonstrated that a model including imaging parameter from PSMA PET might improve models that are

solely based clinical parameters. This is consistent with previous reports that assessed the inclusion of mpMRI parameters into prediction models for LNI [17, 18]. Our study introduces an innovative approach to predict LNI by combining PSMA PET reporting by a nuclear medicine physician with readily available quantitative PSMA PET parameters and clinical parameters. A web-calculator to determine the LNI probability according to the proposed model is available under <https://psma-pet.com/predict> (this calculator should only be used for research purposes).

Our ensemble model incorporating PSA, highest biopsy ISUP, PSMA_{Vol} and the PSMA PET report N-status yielded high sensitivity (0.95) and moderate specificity (0.64) for LNI detection at the proposed threshold of ≥ 0.17 . These results are comparable with the external validation of the 2019 Briganti nomogram (sensitivity of 0.97 and specificity of 0.61) [35]. Our results suggest that especially incorporating the LNI status of the imaging report may improve prediction models, which is in line with previous reports

Table 2 Area under the curve of the proposed model and other lymph node invasion prediction models in the validation cohort

N	Model name	AUC (95% CI)
90	MSKCC	0.824 (0.710–0.938)
	Partin v.2016	0.644 (0.505–0.783)
	Roach	0.700 (0.548–0.852)
	Winter	0.720 (0.590–0.851)
	Model_Clinical_PET/Report	0.923 (0.863–0.984)
67	2019 Briganti	0.786 (0.607–0.966)
	Model_Clinical_PET/Report	0.900 (0.815–0.986)
86	Draulans et al	0.822 (0.700–0.944)
	Model_Clinical_PET/Report	0.921 (0.859–0.983)

regarding esophageal cancer [36]. We suppose that a combination of predictors from both, mpMRI and PSMA PET might be of value in LNI prediction.

The proposed threshold ($\geq 17\%$) appears rather high compared to recommended cut-offs of 5% for conventional nomograms and 7% for the 2019 Briganti nomogram. However, the external validation shows that this threshold led to reliable results (50% of ePLND spared, with a risk of missing only 4.8% LNI) despite very probable differences between the training and validation cohort (e.g., calibration

of the PET scanner). Moreover, a threshold of $\geq 19\%$ in the validation cohort would have yielded an even higher NB (0.185) and would have spared more ePLND (67.8%), with a risk of missing 9.5 LNI. However, the optimal recommended threshold should also be based on clinical reason, and it is questionable if a potentially higher number of missed LNI is clinically acceptable. Therefore, we suppose that the model’s calibration and the threshold should be further investigated in a larger external cohort.

The potential of imaging variables for predicting LNI in PCa has been reported almost 20 years ago using neural networks [37]. Recently, advanced machine learning algorithms have been reported for LNI prediction in PCa [38, 39]. Cysouw et al. reported an internally validated AUC of 0.86 for LNI in intermediate- to high-risk PCa using PSMA PET radiomics [38]. We think that the sophisticated application of more complex models hinders its transition to clinical practice.

Our study has several limitations. Cases for which ePLND was not performed were not included, causing a selection bias. However, all LNI prediction models, which we used for comparison, are also based on PLND. Our proposed model is based on $[^{68}\text{Ga}]\text{Ga-PSMA-11}$ PET, which is costly, not yet a standard procedure at many institutions and must be interpreted with care to avoid false

Table 3 Analyses of the proposed model-derived (Model_Clinical_PET/Report) cut-offs used to discriminate between patients with or without lymph node involvement confirmed at extended pelvic lymph node dissection

Calculated probability of LNI % (cut-off)	Number of patients, n (%)						Sensitivity	Specificity	Net benefit
	Below the cut-off (ePLND not recommended)			Equal to or above the cut-off (ePLND recommended)					
	Total	Without LNI	With LNI	Total	Without LNI	With LNI			
0–13	0 (0)	0 (0)	0 (0)	90 (100)	69 (100)	21 (100)	1.000	0.000	0.119
14	4 (4.4)	4 (5.8)	0 (0)	86 (95.6)	65 (94.2)	21 (100)	1.000	0.058	0.116
15	15 (16.7)	15 (21.7)	0 (0)	75 (83.3)	54 (78.3)	21 (100)	1.000	0.217	0.127
16	28 (31.1)	28 (40.6)	0 (0)	62 (68.9)	41 (59.4)	21 (100)	1.000	0.406	0.147
17	45 (50)	44 (63.8)	1 (4.8)	45 (50)	25 (36.2)	20 (95.2)	0.952	0.638	0.165
18	54 (60)	53 (76.8)	1 (4.8)	36 (40)	16 (23.2)	20 (95.2)	0.952	0.768	0.183
19	61 (67.8)	59 (85.5)	2 (9.5)	29 (32.2)	10 (14.5)	19 (90.5)	0.905	0.855	0.185
20	65 (72.2)	62 (89.9)	3 (14.3)	25 (27.8)	7 (10.1)	18 (85.7)	0.857	0.899	0.181
21	68 (75.6)	63 (91.3)	5 (23.8)	22 (24.4)	6 (8.7)	16 (76.2)	0.762	0.913	0.160
22	69 (76.7)	63 (91.3)	6 (28.6)	21 (23.3)	6 (8.7)	15 (71.4)	0.714	0.913	0.148
23	71 (78.9)	63 (91.3)	8 (38.1)	19 (21.1)	6 (8.7)	13 (61.9)	0.619	0.913	0.125
24	72 (80)	64 (92.8)	8 (38.1)	18 (20)	5 (7.2)	13 (61.9)	0.619	0.928	0.127
25	72 (80)	64 (92.8)	8 (38.1)	18 (20)	5 (7.2)	13 (61.9)	0.619	0.928	0.126
26	73 (81.1)	64 (92.8)	9 (42.9)	17 (18.9)	5 (7.2)	12 (57.1)	0.571	0.928	0.114
27	73 (81.1)	64 (92.8)	9 (42.9)	17 (18.9)	5 (7.2)	12 (57.1)	0.571	0.928	0.113
28	73 (81.1)	64 (92.8)	9 (42.9)	17 (18.9)	5 (7.2)	12 (57.1)	0.571	0.928	0.112
29	73 (81.1)	64 (92.8)	9 (42.9)	17 (18.9)	5 (7.2)	12 (57.1)	0.571	0.928	0.111
30	73 (81.1)	64 (92.8)	9 (42.9)	17 (18.9)	5 (7.2)	12 (57.1)	0.571	0.928	0.110

LNI, lymph node involvement; ePLND, extended pelvic lymph node dissection

Values in bold indicate the proposed threshold

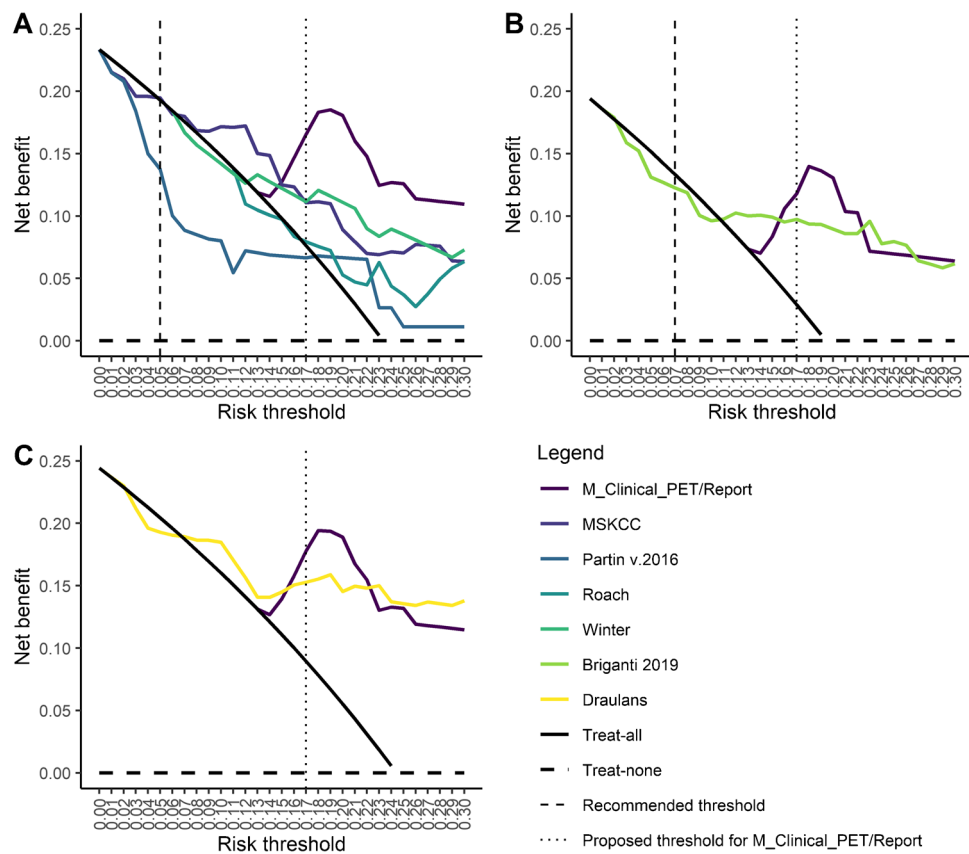


Fig. 3 Decision curve analysis (DCA) of the proposed model (M_clinical_PET/Report) compared to (A) conventional nomograms, (B) the 2019 Briganti nomogram, and (C) the Draulans et al. nomogram. The DCA depicts the net benefit (NB, y-axis) of a model or a strategy (treat-all or treat-none with ePLND) according to a risk threshold (x-axis). Of the conventional models (A), only the MSKCC model showed a higher NB compared to the treat-all strategy at the recommended threshold of 5%. However, at this threshold, the net benefit of the MSKCC nomogram (0.195) was only slightly above the treat-all strategy (0.193), meaning that one can perform 195-192=3 more

beneficial ePLND (out of 1000 patients) when using the MSKCC nomogram). At a threshold of $\geq 17\%$, the net benefit was 0.165 for the proposed model and 0.076 for the treat-all strategy, meaning that one can perform $165-76=89$ more beneficial ePLND (out of 1000 patients) when using the proposed model. Of the mpMRI-based models (B, C), the 2019 Briganti nomogram (B) showed a lower net benefit compared to the treat-all strategy at the recommended threshold of 7% and the Draulans et al. nomogram (C, no recommended threshold available) showed a consistently better net benefit (0.189) compared to the treat-all strategy at a threshold of 7% and above

positive findings. However, it may be included into wide-scale practice soon. Moreover, the proposed model is based on [^{68}Ga]Ga-PSMA-11 and we did not assess its performance with other PSMA tracers. Moreover, we did not investigate potential bias introduced by different PET scanner and both, the training and the validation cohort were of relatively small sample size, which might have led to bias and may limit the generalizability of our results. The two cohorts differentiated regarding preoperative biopsy (mpMRI-targeted with saturation biopsy versus mpMRI-targeted / standard 12-core biopsy) and surgical approach for RP; however, we did not find a significant difference regarding pathological upgrading after RP, number of

removed lymph nodes or LNI rate, respectively. Lastly, because of missing data, we could not directly compare the two mpMRI-based models.

Our results indicate that combining clinical and qualitative/quantitative ^{68}Ga -PSMA-11 information may improve LNI prediction in intermediate to high-risk PCa patients undergoing primary staging. The proposed model with a $\geq 17\%$ threshold yielded a good performance compared to conventional and mpMRI-based models, sparing half of all ePLNDs with a risk of missing only $< 5\%$ LNI. Future research should investigate the combination of information from both PSMA PET and mpMRI for LNI prediction in larger patient cohorts with PCa before RP.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1007/s00259-023-06278-1>.

Acknowledgements The authors acknowledge the technicians Marlena Hofbauer and Josephine Trinckauf and their team for the excellent work on high-quality PET images.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Urs J. Muehlematter, Irene A. Burger, and Isabel Rauscher. The first draft of the manuscript was written by Urs J. Muehlematter and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by University of Zurich

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

Declarations

Ethics approval For the training cohort, the retrospective analysis was approved by the Ethics Committee of the Technical University Munich (permit 5665/13). For the validation cohort, all patients gave a general written informed consent for retrospective use of their data (Ethics Commission of the Canton of Zurich, Switzerland, BASEC Nr. 2018-01284).

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.


References

- Sanda MG, Chen RC, Crispino T, Freedland S, Greene K, Klotz LH, et al. Clinically localized prostate cancer: AUA/ASTRO/SUO Guideline. *Prostate Cancer*. 2017;56.
- EAU Guidelines. Edn. presented at the EAU Annual Congress Amsterdam. Arnhem, the Netherlands: EAU Guidelines Office; 2022.
- EAU Guidelines. Edn. presented at the EAU Annual Congress Amsterdam. Arnhem, the Netherlands: EAU Guidelines Office; 2020.
- Eiber M, Nekolla SG, Maurer T, Weirich G, Wester H-J, Schwaiger M. 68Ga-PSMA PET/MR with multimodality image analysis for primary prostate cancer. *Abdom Imaging*. 2015;40:1769–71.
- Petersen LJ, Zacho HD. PSMA PET for primary lymph node staging of intermediate and high-risk prostate cancer: an expedited systematic review. *Cancer Imaging*. 2020;20:10.
- Thalgott M, Düwel C, Rauscher I, Heck MM, Haller B, Gafita A, et al. One-stop-shop whole-body ⁶⁸Ga-PSMA-11 PET/MRI compared with clinical nomograms for preoperative T and N staging of high-risk prostate cancer. *J Nucl Med*. 2018;59:1850–6.
- Muehlematter UJ, Burger IA, Becker AS, Schawkat K, Hötter AM, Reiner CS, et al. Diagnostic accuracy of multiparametric MRI versus 68Ga-PSMA-11 PET/MRI for extracapsular extension and seminal vesicle invasion in patients with prostate cancer. *Radiology*. 2019;293:350–8.
- Maurer T, Gschwend JE, Rauscher I, Souvatzoglou M, Haller B, Weirich G, et al. Diagnostic efficacy of ⁶⁸Gallium-PSMA positron emission tomography compared to conventional imaging for lymph node staging of 130 consecutive patients with intermediate to high risk prostate cancer. *J Urol*. 2016;195:1436–43.
- Hofman MS, Lawrentschuk N, Francis RJ, Tang C, Vela I, Thomas P, et al. Prostate-specific membrane antigen PET-CT in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (proPSMA): a prospective, randomised, multicentre study. *Lancet*. 2020;395:1208–16.
- Lestingi JFP, Guglielmetti GB, Trinh Q-D, Coelho RF, Pontes J, Bastos DA, et al. Extended versus limited pelvic lymph node dissection during radical prostatectomy for intermediate- and high-risk prostate cancer: early oncological outcomes from a randomized phase 3 trial. *Eur Urol*. 2021;79:595–604.
- Wettstein MS, David LA, Pazhepurackel C, Qureshi AA, Zisman A, Nesbitt M, et al. Benefit of a more extended pelvic lymph node dissection among patients undergoing radical prostatectomy for localized prostate cancer: a causal mediation analysis. *Prostate*. 2021;81:286–94.
- Preisser F, van den Bergh RCN, Gandaglia G, Ost P, Surcel CI, Sooriakumaran P, et al. Effect of extended pelvic lymph node dissection on oncologic outcomes in patients with D'Amico intermediate and high risk prostate cancer treated with radical prostatectomy: a multi-institutional study. *J Urol*. 2020;203:338–43.
- Fossati N, Willemse P-PM, Van den Broeck T, van den Bergh RCN, Yuan CY, Briers E, et al. The benefits and harms of different extents of lymph node dissection during radical prostatectomy for prostate cancer: a systematic review. *Eur Urol*. 2017;72:84–109.
- Perera M, Touijer KA. Pelvic lymph node dissection at the time of radical prostatectomy: extended or not. *The Referee Point of View*. *Eur Urol Open Sci*. 2022;44:24–6.
- EAU Guidelines. Edn. presented at the EAU Annual Congress Milan. Arnhem, the Netherlands: EAU Guidelines Office; 2021.
- Briganti A, Chun FK-H, Salonia A, Suardi N, Gallina A, Da Pozzo LF, et al. Complications and other surgical outcomes associated with extended pelvic lymphadenectomy in men with localized prostate cancer. *Eur Urol*. 2006;50:1006–13.
- Gandaglia G, Ploussard G, Valerio M, Mattei A, Fiori C, Fossati N, et al. A novel nomogram to identify candidates for extended pelvic lymph node dissection among patients with clinically localized prostate cancer diagnosed with magnetic resonance imaging-targeted and systematic biopsies. *Eur Urol*. 2019;75:506–14.
- Draulans C, Everaerts W, Isebaert S, Van Bruwaene S, Gevaert T, Oyen R, et al. Development and external validation of a multiparametric magnetic resonance imaging and International Society of Urological Pathology Based Add-On Prediction Tool to identify prostate cancer candidates for pelvic lymph node dissection. *J Urol*. 2020;203:713–8.
- Ferraro DA, Muehlematter UJ, Garcia Schüler HI, Rupp NJ, Huellner M, Messerli M, et al. 68Ga-PSMA-11 PET has the potential to improve patient selection for extended pelvic lymph node dissection in intermediate to high-risk prostate cancer. *European Journal of Nuclear Medicine and Molecular Imaging [Internet]*. 2019 [cited 2019 Sep 29]; Available from: <https://doi.org/10.1007/s00259-019-04511-4>.
- Meijer D, van Leeuwen PJ, Roberts MJ, Siriwardana AR, Morton A, Yaxley JW, et al. External validation and addition of prostate-specific membrane antigen positron emission tomography to the most frequently used nomograms for the prediction of pelvic

- lymph-node metastases: an international multicenter study. *Eur Urol.* 2021;80:234–42.
21. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* 2015;350:g7594–g7594.
 22. MSKCC Pre-Radical Prostatectomy [Internet]. Available from: https://www.mskcc.org/nomograms/prostate/pre_op. Accessed 5 Feb 2022.
 23. Tosoian JJ, Chappidi M, Feng Z, Humphreys EB, Han M, Pavlovich CP, 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.
 24. Roach M, Marquez C, Yuo H-S, Narayan P, Coleman L, Nseyo UO, et al. Predicting the risk of lymph node involvement using the pre-treatment prostate specific antigen and gleason score in men with clinically localized prostate cancer. *Int J Rad Oncol Biol Phys.* 1994;28:33–7.
 25. Winter A, Kneib T, Rohde M, Henke R-P, Wawroschek F. First nomogram predicting the probability of lymph node involvement in prostate cancer patients undergoing radioisotope guided sentinel lymph node dissection. *Urol Int.* 2015;95:422–8.
 26. Feicke A, Baumgartner M, Talimi S, Schmid DM, Seifert H-H, Müntener M, et al. Robotic-assisted laparoscopic extended pelvic lymph node dissection for prostate cancer: surgical technique and experience with the first 99 cases. *Eur Urol.* 2009;55:876–84.
 27. Epstein JI, Egevad L, Srigley JR, Humphrey PA. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol.* 2016;40:9.
 28. Fanti S, Minozzi S, Morigi JJ, Giesel F, Ceci F, Uprimny C, et al. Development of standardized image interpretation for 68Ga-PSMA PET/CT to detect prostate cancer recurrent lesions. *Eur J Nucl Med Mol Imaging.* 2017;44:1622–35.
 29. LeDell E. Scalable ensemble learning and computationally efficient variance estimation. UC Berkeley; 2015.
 30. Bhadani R. Nonlinear optimization in R using nlopt. arXiv:210102912 [math, stat] [Internet]. 2021 [cited 2021 Apr 5]; Available from: <https://arxiv.org/abs/2101.02912>. Accessed 25 May 2022.
 31. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. *J Clin Epidemiol.* 2016;74:167–76.
 32. Kattan MW, Gerds TA. The index of prediction accuracy: an intuitive measure useful for evaluating risk prediction models. *Diagn Progn Res.* 2018;2:7.
 33. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44:837–45.
 34. Dell'Oglio P, Abdollah F, Suardi N, Gallina A, Cucchiara V, Vizzello D, et al. External validation of the European Association of Urology Recommendations for pelvic lymph node dissection in patients treated with robot-assisted radical prostatectomy. *J Endourol.* 2014;28:416–23.
 35. Gandaglia G, Martini A, Ploussard G, Fossati N, Stabile A, De Visschere P, et al. External validation of the 2019 Briganti Nomogram for the identification of prostate cancer patients who should be considered for an extended pelvic lymph node dissection. *Eur Urol.* 2020;78:138–42.
 36. Shen C, Liu Z, Wang Z, Guo J, Zhang H, Wang Y, et al. Building CT radiomics based nomogram for preoperative esophageal cancer patients lymph node metastasis prediction. *Transl Oncol.* 2018;11:815–24.
 37. Poulakis V, Witzsch U, De Vries R, Emmerlich V, Meves M, Altmannsberger H-M, et al. Preoperative neural network using combined magnetic resonance imaging variables, prostate specific antigen and Gleason score to predict prostate cancer stage. *J Urol.* 2004;172:1306–10.
 38. Cysouw MCF, Jansen BHE, van de Brug T, Oprea-Lager DE, Pfähler E, de Vries BM, et al. Machine learning-based analysis of [18F]DCFPyL PET radiomics for risk stratification in primary prostate cancer. *Eur J Nucl Med Mol Imaging.* 2021;48:340–9.
 39. Hou Y, Bao J, Song Y, Bao M-L, Jiang K-W, Zhang J, et al. Integration of clinicopathologic identification and deep transferrable image feature representation improves predictions of lymph node metastasis in prostate cancer. *EBioMedicine.* 2021;68: 103395.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Urs J. Muehlemaier^{1,2}  · Lilith Schweiger³ · Daniela A. Ferraro^{1,4} · Thomas Hermanns⁵ · Tobias Maurer^{6,7} · Matthias M. Heck⁶ · Niels J. Rupp⁸ · Matthias Eiber³ · Isabel Rauscher³ · Irene A. Burger^{1,9}

✉ Irene A. Burger
irene.burger@usz.ch

¹ Department of Nuclear Medicine, University Hospital Zurich, University of Zurich, Zurich, Switzerland

² Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Zurich, Switzerland

³ Department of Nuclear Medicine, Technische Universität München, Klinikum Rechts Der Isar, Munich, Germany

⁴ Department of Radiology and Oncology, Faculdade de Medicina FMUSP, Universidade de São Paulo, São Paulo, Brazil

⁵ Department of Urology, University Hospital Zurich, University of Zurich, Zurich, Switzerland

⁶ Department of Urology, Technische Universität München, Klinikum Rechts Der Isar, Munich, Germany

⁷ Department of Urology and Martini-Klinik, Universität Hamburg-Eppendorf, Hamburg, Germany

⁸ Department of Pathology and Molecular Pathology, University Hospital Zurich, University of Zurich, Zurich, Switzerland

⁹ Department of Nuclear Medicine, Baden Cantonal Hospital, Baden, Switzerland