Original Article

Current Urology

OPEN

Validation of an MRI-based prostate cancer prebiopsy Gleason score predictive nomogram

Adrianna Jiaying Lee[a], Amelia Wnorowski[a], Nancy Ye[a], Linhan Xu[a], Michael Naslund[a,b], Bradford J. Wood[c], Maria J. Merino[d], Baris Turkbey[c], Peter L. Choyke[c], Peter A. Pinto[e], M. Minhaj Siddiqui[a,b,*]

[a]University of Maryland, Baltimore, MD, USA; [b]Baltimore VA Medical Center, Baltimore, MD, USA; [c]Molecular Imaging Program, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; [d]Center for Interventional Oncology, National Cancer Institute & Clinical Center, National Institutes of Health, Bethesda, MD, USA; [e]Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

Abstract

Background: Gleason score grading is a cornerstone of risk stratification and management of patients with prostate cancer (PCa). In this work, we derive and validate a nomogram that uses prostate multiparametric magnetic resonance imaging (MP-MRI) and clinical patient characteristics to predict biopsy Gleason scores (bGS).

Materials and methods: A predictive nomogram was derived from 143 men who underwent MP-MRI prior to any prostate biopsy and then validated on an independent cohort of 235 men from a different institution who underwent MP-MRI for PCa workup. Screen positive lesions were defined as lesions positive on T2W and DWI sequences on MP-MRI. Prostate specific antigen (PSA) density, number of screen positive lesions, and MRI suspicion were associated with PCa Gleason score on biopsy and were used to generate a predictive nomogram. The independent cohort was tested on the nomogram and the most likely bGS was noted.

Results: The mean PSA in the validation cohort was 9.25 ng/mL versus 6.8 ng/mL in the original cohort ($p = 0.001$). The distribution of Gleason scores between the 2 cohorts were not significantly different ($p = 0.7$). In the original cohort of men, the most probable nomogram generated Gleason score agreed with actual pathologic bGS findings in 61% of the men. In the validation cohort, the most likely nomogram predicted bGS agreed with actual pathologic bGS 51% of the time. The nomogram correctly identified any PCa versus non-PCa 63% of the time and clinically significant (Gleason score ≥ 7) PCa 69% of the time. The negative predictive value for clinically significant PCa using this prebiopsy nomogram was 74% in the validation group.

Conclusions: A preintervention nomogram based on PSA and MRI findings can help narrow down the likely pathologic finding on biopsy. Validation of the nomogram demonstrated a significant ability to correctly identify the most likely bGS. This feasibility study demonstrates the potential of a prebiopsy prediction of bGS and based on the high negative predictive value, identification of men who may not need biopsies, which could impact future risk stratification for PCa.

Keywords: Multiparametric magnetic resonance imaging; Nomogram; Prostate cancer

1. Introduction

Prostate cancer (PCa) is the most common type of cancer affecting men (excluding skin cancer)[1] but a diagnosis of PCa is not always an indication for treatment. While PCa is the third most common cause of cancer mortality in men.[2] Some types of PCa grow relatively slowly and can be managed with close follow-up and active surveillance.[3] Other types of PCa are more aggressive and must be diagnosed and treated promptly. Thus, risk stratification of PCa at the time of screening and diagnosis is important in subsequent management of the disease. There are currently several tools clinicians can use to assess the state of the prostate, including prostate specific antigen (PSA), digital rectal exam (DRE), magnetic resonance imaging (MRI), and targeted or random biopsies of the gland itself.[4]

Several prostate nomograms using clinical and pathologic scores have been developed to evaluate PCa throughout progression of the disease.[1] Notably, the Kattan nomograms are valuable predictors of PCa mortality pre- and postradical prostatectomy (RP), postoperative recurrence based on surgeon experience, and biochemical recurrence.[5–8] However, the prostate multiparametric MRI (MP-MRI) based nomogram developed in this study would assist in risk stratification of men with PCa risk factors presenting prior to biopsy. This nomogram outputs the percent probability of a man having clinically significant PCa given a certain set of clinical scores. The ability to stratify patients and quantify their risk will help improve clinician decision making for each patient based on the clinical characteristics of their disease. Quantification of risk will also help patients gain better understanding of their disease based on their imaging and pathological results, especially for high risk patients. The purpose of this study was to create a predictive nomogram for PCa presence and aggressiveness based on demographic and

* Corresponding Author: M. Minhaj Siddiqui, Division of Urology, University of Maryland Medical Center, 29 S Greene St Suite 500, Baltimore, MD 21201, USA. E-mail address: msiddiqui@som.umaryland.edu (M.M. Siddiqui).

Current Urology, (2022) 16, 38–43

Received January 6, 2020; Accepted May 6, 2020.

http://dx.doi.org/10.1097/CU9.0000000000000069

imaging characteristics which was then validated in an independent cohort.

2. Methods

2.1. Patient selection and database

After approval from the Institutional Review Board, a nomogram was derived from 143 biopsy naïve men who underwent MP-MRI at the National Institutes of Health (NIH) between 8/2007 and 12/2012. The nomogram was then validated on an independent cohort of 235 men who were seen at an independent institution, the University of Maryland, between 2/2015 and 12/2018. Men were selected for the study based on accessibility to the patient population at our institution. While demographically distinct from the original cohort, the heterogeneity in patient populations tested on this nomogram would allow applicability of the nomogram to more diverse patient populations in the future. Independent cohort data was prospectively maintained in a research electronic data capture (REDCap) database.[9] Men were excluded from this study if they had PCa recurrence, had previous treatment for PCa, or had previously undergone fusion biopsy.

2.2. Imaging

Prior to biopsy, patients underwent prostate MP-MRI (consisting of small field of view multiplanar T2-weighted, diffusion-weighted, and dynamic-contrast enhanced imaging) with an endorectal coil (BPX-30, Bayer Inc.) on a 3.0T MRI scanner (Achieva, Philips Healthcare) in the original cohort and on either a 1.5T MRI scanner (Magnetom Avanto, Siemens Medical Solutions) with an endorectal coil (MedRad eCoil, Bayer Inc.) or 3.0T MRI scanner (Discovery 750w, GE Healthcare; Magnetom Trio, Siemens Medical Solutions; or Achieva, Philips Healthcare) without an endorectal coil in the independent cohort. MRI images were read by 2 independent readers in the original cohort and 6 independent readers for the independent cohort. Images were segmented and lesion location was recorded (DynaCAD, Invivo). Lesions were scored corresponding to the Prostate Imaging-Reporting and Data System version 2 (PIRADS v2) as low (PIRADS 1–2), moderate (PIRADS 3) or high (PIRADS 4–5) based on the number of positive sequences per lesion in the original cohort. PIRADS v2 was used in the independent cohort. Screen positive lesions (SPLs) were defined as lesions positive on T2-weighted and diffusion weighted imaging sequences on MP-MRI.

2.3. Biopsy session protocol

MRI-transrectal ultrasound (TRUS) targeted biopsy was performed using the UroNav platform (Invivo, Philips Healthcare). Following a B-mode TRUS survey of the prostate, segmentation yielded a 3-dimensional volume that was coregistered and overlaid with T2-weighted images from the preprocedure MRI. Lesions from the MRI were displayed as targets and 1–4 core biopsies were taken for each lesion depending on patient comfort and physician judgement. Immediately following biopsy of MRI lesions, a standard of care 12-core TRUS biopsy was performed. Biopsy needle trajectories were mapped with the electromagnetic tracking included in the platform. Pathologic specimens were reviewed by an experienced genitourinary pathologist. The biopsy Gleason scores (bGS) used for the study was the highest total Gleason score of both the targeted and nontargeted positive biopsy cores.

2.4. Statistical analyses

Logistic regression modeling was used to assess association of the different parameters with bGS (no cancer, Gleason 6, 7, and ≥8). Screening variable selection for multivariate analysis was chosen based on p-values of the variable under consideration on univariate analysis. An optimized model for the prediction of bGS was determined and a nomogram was generated based on PSA density (PSAD), number of SPLs, and MRI suspicion associated with PCa Gleason score on biopsy. These parameters were selected based on sequential analysis of factors that varied based on Gleason scores and those that demonstrated strong predictive association or improvement of the model were included in the nomogram.

Calibration plots were generated using bootstrap analysis to validate nomogram performance. All p-values were acquired from two-sided tests with a significance level of $p < 0.05$. Analyses were performed with JMP v11.0 (SAS Institute, Cary, NC) and RStudio v1.1.453. The RMS package was used to generate the nomogram (https://cran.r-project.org/web/packages/rms/).

3. Results

An original cohort of 143 men and independent cohort of 235 men were examined. The demographics of the cohorts are shown in Table 1. The mean age in the original cohort was 60.7 years and independent cohort was 65.1 years. The average number of lesions found on MRI for the independent cohort was 2.2, with 500 total lesions detected on MRI. Only 1 patient with a negative MRI was biopsied in the independent cohort. The PIRADS distribution of the independent cohort was 62, 149, 237, and 52 of PIRADS 2, 3, 4, and 5 lesions, respectively. Total 199 men in the independent cohort had received at least one prior random prostate biopsy, and of those, 60 patients had a prior positive biopsy. The clinical characteristics of the cohorts are shown in Table 2 with patients divided into 4 groups based on bGS: no cancer (Gleason < 6), Gleason 6, Gleason 7, or Gleason ≥ 8.

The nomogram (Fig. 1A) utilized PSAD, SPLs and MRI suspicion score to derive a predicted bGS. SPLs were defined as lesions with a score of 3 or higher on both T2-weighted and diffusion weighted imaging sequences. PSAD, number of SPLs, and MRI suspicion each received a designated score on the nomogram and the total score was associated with a certain probability of being in each bGS group (benign, Gleason 6, 7, or ≥8) based on the derived probability curve (Fig. 1B). The independent cohort was also tested on the nomogram and the most likely bGS based on the curve was noted.

Table 1

Original cohort and independent cohort patient demographics.

Characteristic	Original cohort	Independent cohort	p
Number of men	143	235	N/A
Age, yr	60.7 ± 7.7	65.1 ± 6.7	<0.001
PSA, ng/mL	6.8 ± 6.5	9.2 ± 6.1	0.001
Prostate volume, cm^3	48.3 ± 25.2	67.9 ± 45.0	<0.001
PSA density, ng/mL/cm^3	0.15 ± 0.14	0.18 ± 0.16	0.06
Race, n (%)			
White	104 (73%)	137 (58%)	
African American	23 (16%)	68 (29%)	
Other	16 (11%)	30 (13%)	

PSA = prostate specific antigen.

Table 2

Cancer related outcomes for original and independent cohorts with corresponding *p*-values.

	No cancer			Gleason 6			Gleason 7			Gleason ≥ 8		
	Original	Validation	*p*	Original	Validation	*p*	Original	Validation	*p*	Original	Validation	*p*
Number of men	59	122		29	44		24	44		31	25	
Age, yr	59.2±7.2	64.5±6.3	<0.001	59.9±7.0	65.0±7.4	0.006	61.1±8.6	65.9±7.1	0.04	64.1±7.7	67.3±5.9	0.091
Prostatic volume, cm^3	53.8±30.8	79.3±48.2	<0.001	44.3±16.6	55.4±32.3	0.064	41.3±15.1	61.4±48.7	0.019	46.6±13.5	45.8±19.8	0.865
Total number of lesions on MRI	2.1±1.3	2.1±0.94	1	2.9±1.4	2.4±1.0	0.107	3.0±1.3	2.3±1.0	0.031	3.5±1.0	2.0±0.91	<0.001
Number of screen positive lesions	1.6±1.1	1.1±1.0	0.005	2.9±1.3	1.2±1.1	<0.001	2.8±1.2	1.6±1.1	<0.001	3.4±1.1	1.6±1.1	<0.001
PSA	4.6±2.7	9.3±5.8	<0.001	4.4±3.0	8.8±6.1	<0.001	6.2±2.7	8.8±7.0	0.039	13.5±10.3	10.5±5.9	0.184
PSAD	0.09±0.05	0.16±0.17	<0.001	0.10±0.06	0.19±0.13	<0.001	0.16±0.07	0.17±0.12	0.669	0.31±0.20	0.27±0.17	0.427
MRI suspicion												
Low (PIRADS 1–2)	16 (27%)	10 (8.2%)		8 (28%)	0 (0%)		1 (4%)	2 (4.54%)		0 (0%)	1 (4%)	
Moderate (PIRADS 3)	43 (73%)	32 (26.23%)		20 (69%)	9 (20.45%)		15 (63%)	3 (6.82%)		15 (48%)	0 (0%)	
High (PIRADS 4–5)	0 (0%)	80 (65.57%)		1 (3%)	35 (79.55%)		8 (33%)	39 (88.63%)		16 (52%)	24 (96%)	

MRI = magnetic resonance imaging; PIRADS = Prostate Imaging-Reporting and Data System; PSA = prostate specific antigen; PSAD = prostate specific antigen density.

In order to obtain a predicted bGS using this tool, each patient's characteristics must be matched on the nomogram to obtain a point value. The point values are summed to give a total number of points on the nomogram, which then outputs certain probabilities of each bGS on the probability curve. For example, a 64-year-old male with PSAD 0.2 ng/mL/cm^3, 3 SPLs, and high MRI suspicion would earn 22, 15, and 13 points for those characteristics, respectively (Fig. 2A). His characteristic combination would sum to a total of 50 points on the nomogram, which could then be matched on the probability curve (Fig. 2B). On the derived probability curve, this patient would have a 42% chance of biopsy Gleason ≥ 8, 35% chance of Gleason 7, 15% chance of Gleason 6, and 8% chance of no cancer.

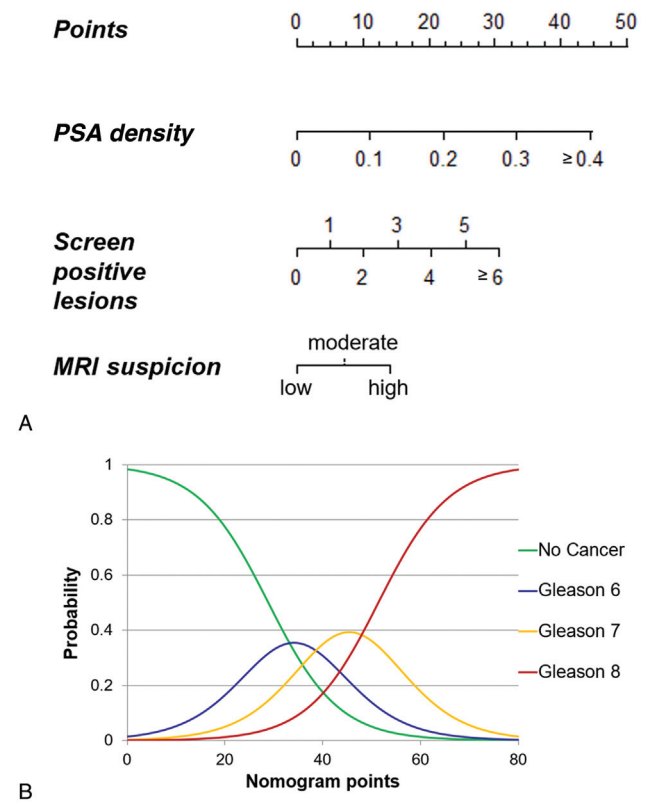

Figure 1. Nomogram and probability curve. (A) Nomogram derived from NIH cohort of 143 patients for validation of biopsy Gleason score; (B) Probability curve showing probability of no cancer, Gleason 6, 7, or ≥8 given total points on nomogram. MRI = magnetic resonance imaging; PSA = prostate specific antigen.

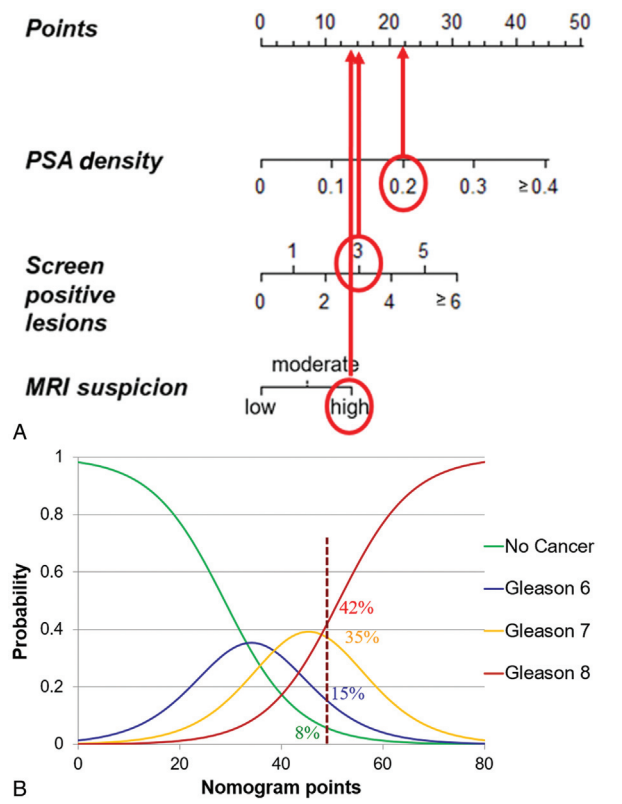

Figure 2. Example patient scoring on nomogram and derived probability. (A) Derivation of total points on nomogram for hypothetical patient; (B) Nomogram point total matched on probability curve for hypothetical patient. MRI = magnetic resonance imaging; PSA = prostate specific antigen.

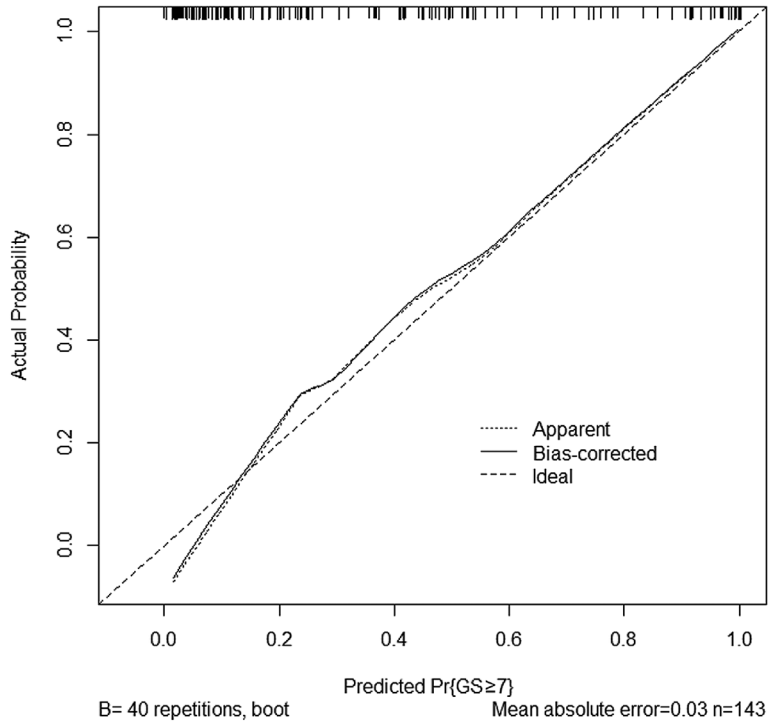

Figure 3. Calibration plot of predicted biopsy Gleason score ≥ 7 where actual probability is the actual outcome. Each tick mark at the top represents a patient.

The nomogram predictions for bGS ≥ 7 were compared with actual outcomes in the original cohort on a calibration plot using bootstrap analysis (Fig. 3), which demonstrated a mean error of 3%. The probability of Gleason ≥ 7 tracked closely with actual outcomes after biopsy. In general, the nomogram slightly overestimated the probability of each Gleason score except at higher probabilities on the probability curve, where it underestimated the probability. This model performs better at higher predicted probabilities of a given bGS as shown by the calibration plot, where it tracks more closely with the actual outcomes at higher probabilities. In the original cohort, the nomogram correctly predicted the bGS 61% of the time. The top two most likely Gleason scores (ie, the two with the highest probabilities on the probability curve) were correctly predicted 83% of the time.

In the independent cohort, the nomogram correctly predicted the bGS 51% of the time out of 4 possible outcomes and correctly predicted the two most likely Gleason scores 69% of the time. Furthermore, the nomogram correctly predicted cancer versus no cancer 63% of the time and the presence of clinically significant PCa 69% of the time. The negative predictive value for clinically significant PCa using this prebiopsy nomogram in the validation cohort was 74%.

4. Discussion

Although a definitive diagnosis of PCa cannot be made without a tissue specimen, a better predictive tool may offer significant benefit to decrease diagnosis of clinically insignificant disease and thus potential overtreatment after a positive screening test.[10] Current screening methods are often unreliable if taken independently and are imprecise indicators of next steps in management of PCa.[2] There are several components of PCa screening, including DRE, PSA, PSAD, prostate volume, and MRI findings. These methods all provide information to assist in diagnosis, but there is a lack of standardization and unification of these clinical characteristics when recommending follow-up and treatment plans. The nomogram and probability curve constructed and validated in this study will help to improve risk stratification and better determine the most effective treatment based on objective measures of clinical findings prior to biopsy. In particular, there is added value in predicting a biopsy score of Gleason 7 or greater, since these tumors tend to have a less favorable prognosis compared to tumors graded Gleason 6 and below.[11] Our nomogram tracks closely with actual biopsy outcomes of Gleason ≥ 7, which can help guide clinician recommendations for or against biopsy particularly in patients who are young but with significant comorbidities, or conversely older but in excellent health. Further studies are required to determine the impact of our nomogram on treatment decisions – whether a nomogram prediction concordant with biopsy score would mean more or less aggressive treatment.

With the multitude of tests patients must endure during the PCa screening process and prior to receiving a diagnosis, results can often be confusing and difficult to interpret.[12] Because PSA, DRE, and other screening modalities have conflicting recommendations in practice when used alone as a screen,[13] it is crucial for the patient to understand both the value of each test and the risks associated with overdiagnosis and treatment of PCa.[14] This nomogram pulls together several common clinical tests to derive probabilities of clinically significant disease, which helps to quantify each patient's likelihood of high-risk disease. The quantification of disease probability will help to simplify risk

stratification and improve both patient understanding as well as joint patient–clinician decision making.[1,12–14]

Other studies have also used nomograms in the management of PCa, including the Kattan nomograms for pre- and post-RP and salvage radiation therapy.[5–7] These nomograms are used mostly for determining the benefit of a treatment, either prostatectomy or radiation therapy, or evaluating the chances of recurrence following RP. The nomogram validated in this study would be useful for management prior to biopsy and treatment, to better evaluate the need for prompt biopsy. A predictive rate of 51% from this nomogram could be immensely valuable in guiding patient management prior to receiving a definitive bGS. Given a negative predictive value of 74% in the validation cohort, this nomogram would be valuable for identification of patients who may not benefit from further diagnostic studies. Stratification of risk based on the predicted bGS and increased patient understanding of clinical tests will also assist in reducing overdiagnosis and overtreatment of low-risk patients.

There were some limitations to this study. There was a slight decrease in efficacy of the nomogram between the original cohort and the independent cohort. This could have been due to differences in the demographics of the 2 groups. There was a higher proportion of African American patients and a lower proportion of White patients in the independent than the original cohort. The independent cohort was also comprised of older men ($p < 0.001$). Differences in age and race may have contributed to the lower accuracy of this nomogram when applied to the independent cohort. In addition, the original cohort had lower PSA values and smaller prostate volumes. Another limitation is the lack of a single interpreter for MP-MRI interpretation since there could have been inter-observer variability dependent on each reviewer's experience. Specifically, the PRECISION study reports about a 78% agreement between radiologist readings in prostate MP-MRI.[15] The most recent version of PIRADS (v2) was used for rating of MP-MRI lesions, but because of this, there may have been variability in standardization of reading protocol and comfort level of reviewers with using this protocol, especially with patients recruited earlier in the study. In addition, it is possible that the prostatic volume and PSA, may have interesting effects on the probability of having PCa, since both tended to be higher in the lower bGS groups in both the original and independent cohort but were generally higher in the latter. There are other parameters such as free PSA, DRE findings, race, and age that may play a role in risk stratification of PCa that were unable to be incorporated and validated in this nomogram study and would be areas of interest in future studies. Future collaboration using this analysis on larger cohorts of patients would also help to improve accuracy and applicability of this study. This nomogram should be tested in guiding decision making to determine the true clinical impact of this nomogram and to improve the accuracy of nomograms at each step of management of PCa.

5. Conclusion

This nomogram and probability curve based on PSAD, number of screen positive lesions on multiparametric prostate MRI and MRI lesion suspicion score has good predictive value of predicting bGS especially for clinically significant PCa. The combination of this nomogram based on clinical characteristics and the derived probability curve is a useful tool for risk stratification of patients, to guide clinician decision making and to improve patient counseling of potential disease.

Acknowledgments

None.

Statement of ethics

The study was approved by institutional review board and informed consent was exempted for this study. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Conflict of interest statement

The authors report no conflicts of interest.

Funding source

None.

Author contributions

Project conception: MMS, PAPA, PLC, BT;
Data curation: AJL, AW, NY, LX, MN, MJM, BT, PLC, PAP, MMS;
Formal analysis: AJL, MMS, PAP;
Funding aquisition: PAP
Writing: AJL, MMS;
Review and edits: AW, NY, LX, MN, BJW, MJM, BT, PLC, PAP.

References

[1] Shariat SF, Kattan MW, Vickers AJ, Karakiewicz PI, Scardino PT. Critical review of prostate cancer predictive tools. *Futur Oncol* 2009;5 (10):1555–1584.
[2] Niu XK, Li J, Das SK, Xiong Y, Yang CB, Peng T. Developing a nomogram based on multiparametric magnetic resonance imaging for forecasting high-grade prostate cancer to reduce unnecessary biopsies within the prostate-specific antigen gray zone. *BMC Med Imaging* 2017;17(1):11.
[3] Loeb S, Zhou Q, Siebert U, et al. Active surveillance versus watchful waiting for localized prostate cancer: A model to inform decisions. *Eur Urol* 2018;72(6):899–907.
[4] Partin AW, Kattan MW, Subong EN, et al. Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. *JAMA* 1997;277(18):1445–1451.
[5] Brockman JA, Alaneec S, Vickers AJ, et al. Nomogram predicting prostate cancer-specific mortality for men with biochemical recurrence after radical prostatectomy. *Eur Urol* 2015;67(6):1160–1167.
[6] Kattan MW, Vickers AJ, Yu C, et al. Preoperative and postoperative nomograms incorporating surgeon experience for clinically localized prostate cancer. *Cancer* 2009;115(5):1005–1010.
[7] Ondracek RP, Kattan MW, Murekeyisoni C, et al. Validation of the Kattan nomogram for prostate cancer recurrence after radical prostatectomy. *J Natl Compr Canc Netw* 2016;14(11):1395–1401.
[8] Dell'Oglio P, Suardi N, Boorjian SA, et al. Predicting survival of men with recurrent prostate cancer after radical prostatectomy. *Eur J Cancer* 2016;54:27–34.
[9] Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42(2):377–381.

[10] Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: Importance of methods and context. *J Natl Cancer Inst* 2009;101(6):374–383.
[11] Pierorazio PM, Walsh PC, Partin AW, Epstein JI. Prognostic Gleason grade grouping: Data based on the modified Gleason scoring system. *BJU Int* 2013;111(5):753–760.
[12] James LJ, Wong G, Craig JC, et al. Men's perspectives of prostate cancer screening: A systematic review of qualitative studies. *PLoS One* 2017;12 (11):e0188258.
[13] Moyer VA. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2012;157(2): 120–134.
[14] Carter HB, Albertsen PC, Barry MJ, et al. Early detection of prostate cancer: AUA guideline. *J Urol* 2013;190(2):419–426.
[15] Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-targeted or standard biopsy for prostate cancer diagnosis. *N Engl J Med* 2018;378 (19):1767–1777.

How to cite this article: Lee AJ, Wnorowski A, Ye N, Xu L, Naslund M, Wood BJ, Merino MJ, Turkbey B, Choyke PL, Pinto PA, Siddiqui MM. Validation of an MRI-based prostate cancer prebiopsy Gleason score predictive nomogram. *Curr Urol* 2022;16(1):38–43. doi: 10.1097/CU9.0000000000000069.