Accepted Manuscript

Title: Prediction of Prostate Cancer Risk among Men Undergoing Combined MRI-Targeted and Systematic Biopsy Using Novel Pre-Biopsy Nomograms That Incorporate MRI Findings.

Author: Marc A. Bjurlin, Andrew B. Rosenkrantz, Saradwata Sarkar, Herbert Lepor, William C. Huang, Richard Huang, Rajesh Venkataraman, Samir S. Taneja

PII: S0090-4295(17)31185-8
DOI: https://doi.org/10.1016/j.urology.2017.09.035
Reference: URL 20743

To appear in: Urology

Received date: 3-2-2017
Accepted date: 14-9-2017


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Prediction of Prostate Cancer Risk Among Men Undergoing Combined MRI-Targeted and Systematic Biopsy using Novel Pre-Biopsy Nomograms that Incorporate MRI findings.

AUTHORS:

Marc A. Bjurlin DO, MSc¹, Andrew B. Rosenkrantz MD², Saradwata Sarkar PhD³, Herbert Lepor MD¹, William C. Huang MD¹, Richard Huang BS¹, Rajesh Venkataraman PhD³, Samir S. Taneja MD¹

¹Division of Urologic Oncology
Department of Urology
New York University Langone Medical Center
New York, New York

²Department of Radiology
New York University Langone Medical Center
New York, New York

³Eigen
13366 Grass Valley Ave.
Grass Valley, CA 95945

*Corresponding Author:
Samir S. Taneja, MD
The James M. Neissa and Janet Riha Neissa Professor of Urologic Oncology
Professor of Urology and Radiology
Director, Division of Urologic Oncology
Department of Urology
New York University Langone Medical Center
New York, New York
Email: Samir.taneja@nyumc.org
Tel: (646)825-6321
Fax: (646)825-6368

KEY WORDS: nomogram, MRI, prostate cancer, pre-biopsy, fusion, targeted

Acknowledgements: None
ABSTRACT

Objective: To develop nomograms that predict the probability of overall PCa and clinically significant PCa (Gleason ≥7) on MRI targeted, and combined MRI-targeted and systematic, prostate biopsy.

Materials and Methods: From June 2012 to August 2014, MR-US fusion targeted prostate biopsy was performed on 464 men with suspicious regions identified on pre-biopsy 3T MRI along with systematic 12 core biopsy. Logistic regression modeling was used to evaluate predictors of overall and clinically significant PCa, and corresponding nomograms were generated for men who were not previously biopsied or had one or more prior negative biopsies. Models were created with 70% of a randomly selected training sample and bias-corrected using bootstrap resampling. The models were then validated with the remaining 30% testing sample pool.

Results: A total of 459 patients were included for analysis (median age 66 years, PSA 5.2 ng/ml, prostate volume 49 cc). Independent predictors of PCa on targeted and systematic prostate biopsy were PSA density, age, and MRI suspicion score. PCa probability nomograms were generated for each cohort using the predictors. Bias-corrected areas under the receiver-operating characteristic curves for overall and clinically significant PCa detection were 0.82 (0.78) and 0.91 (0.84) for men without prior biopsy and 0.76 (0.65) and 0.86 (0.87) for men with a prior negative biopsy in the training (testing) samples.

Conclusion: PSA density, age, and MRI suspicion score predict prostate cancer on combined MRI-targeted and systematic biopsy. Our generated nomograms demonstrate high diagnostic accuracy and may further aid in the decision to perform biopsy in men with clinical suspicion of PCa.

Introduction

Prostate cancer (PCa) is the most commonly diagnosed non-skin cancer in men in the United States, with a lifetime risk for diagnosis currently estimated at 15%. Contemporary recommendations for PCa screening incorporate the measurement of serum prostate specific antigen (PSA) levels into shared decision making. However, over-diagnosis and overtreatment of PCa, especially of indolent cancer that may never affect a man's longevity, are frequently
attributed to the increased use of PSA screening coupled with its low specificity and positive predictive value for clinically significant PCa. Ultimately, the United States Preventive Services Task Force (USPSTF) recommended against PSA-based screening for PCa (grade D recommendation - moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits). The impact of the USPSTF recommendations has resulted in a decline in PSA screening and prostate cancer apparent incidence, from which one potential consequence may be an increase in prostate cancer mortality over time. We, and others, have proposed that rather than abandoning PSA screening, a more rational paradigm may involve selective use of prostate biopsy among men with elevated PSA levels through further refinement of cancer risk.

Increasing evidence supports the use of Magnetic Resonance Imaging (MRI) in PCa detection when used as a localization tool to guide MRI-targeted biopsy techniques such as MRI to Ultrasound (MRI-US) fusion targeted biopsy. Pre-biopsy MRI not only allows accurate tumor localization, but also provides an assessment of cancer suspicion using an MRI suspicion score, and thus provides accurate prediction of the likelihood of PCa on prostate biopsy. MRI-US fusion targeted prostate biopsy has resulted in increased detection of clinically significant cancer in men with no previous cancer diagnosis, while potentially limiting the detection of indolent disease through avoidance of systematic, random sampling.
Various statistical models have been developed to predict the risk of cancer. Previously proposed nomograms represent a graphical means to further refine risk among men with elevated PSA levels, thereby allowing individualization of the decision to perform a prostate biopsy. The statistical models from which the nomograms are constructed can improve the efficiency of PCa detection, exceeding the performance of experienced clinical experts.\textsuperscript{10,11}

We sought to develop nomograms incorporating the outcomes of pre-biopsy MRI for predicting risk of PCa, and clinically significant PCa, at the time of combined systematic transrectal ultrasound guided biopsy and MRI-US fusion targeted biopsy in men with no known history of PCa presenting for consideration of prostate biopsy. These nomograms allow the use of pre-biopsy MRI data for further individualization of the decision to perform biopsy in men with clinical suspicion of PCa.

**MATERIALS AND METHODS**

*Patient Population*

From June, 2012 to August, 2014, a total of 464 consecutive men with no prior history of PCa underwent standard pre-biopsy 3T MRI at our institution followed by a combined systematic and MRI-US fusion targeted biopsy and were enrolled in a prospectively acquired database registry study. Patients with hip implants were not included. This retrospective analysis of the prospectively recorded data registry was HIPAA compliant and approved by our institutional review board with
a waiver of the requirement for written informed consent. Five patients were excluded due to incomplete records. Before biopsy, MRI results for all patients were reviewed by a single fellowship-trained radiologist with expertise in prostate imaging to identify and score suspicious regions within the prostate using a 5-point Likert scale for cancer suspicion, as previously described. This database has been utilized in multiple studies evaluating the outcomes of MRI-US fusion targeted biopsies.

**MRI**

MRI was performed using a 3-T clinical MRI instrument and an external phased-array coil and included multiplanar T2-weighted images, axial diffusion-weighted imaging using b values of 50 and 1000 s/mm² with generation of the apparent diffusion coefficient map and calculated b-1500 images, and dynamic contrast-enhanced imaging after intravenous administration of a gadolinium chelate. Lesions identified on MRI were scored on the probability of harboring clinically significant cancer as 2 (low probability), 3 (equivocal), 4 (high probability), or 5 (very high probability), as previously described (MRI Suspicion Score: MRIss). Men with MRIss = 1 (no findings suspicious for cancer) were not candidates for targeted biopsy. For men with multiple lesions with differing MRIss values, the highest MRIss for any individual lesion was recorded as representing the MRIss for the patient.
**MRI-US Fusion Targeted Biopsy**

MRI-US fusion targeted biopsies were performed using an Artemis prostate biopsy system and ProFuse (Eigen, Grass Valley, CA, USA) software for MRI segmentation, co-registration of MRI and US images, and three-dimensional biopsy planning, as previously described. Computer-assisted co-registration of segmented MRI and US images of the prostate was performed using manual rigid translation followed by automated elastic deformation. Transrectal biopsies were obtained with the patient in the left lateral decubitus position, up to four biopsy cores targeted to each suspicious lesion identified on MRI and followed by 12 software populated, spatially distributed cores. Sites for 12-core sampling were selected by the Artemis device, not the operating surgeon.

**Statistical Analysis**

Post biopsy, the pathology results were incorporated into our records and the database was split into 2 subcohorts – men who had no prior biopsy and men with a prior negative biopsy. Logistic regression modeling was performed for each subcohort separately to evaluate predictors of overall and clinically significant PCa (Gleason ≥7). A positive or negative coefficient estimate indicates that the individual predictor adds to or lowers the risk of finding overall PCa or clinically significant PCa upon biopsy respectively. Mathematically, the coefficient values represent the change in log odds of finding PCa or clinically significant PCa per unit change in the value of the individual predictors. Each cohort was
randomly split into 70% of the data to form the model training sample and 30% of the data to validate or test the model. Several multivariable logistic regression models were fit to the training data using combinations of clinical characteristics such as PSA level, PSA Density, Age, Prostate Volume, MRIss, Prostate Cancer Antigen 3 (PCA3), family history of PCa, and Body Mass Index (BMI) as predictors. Automated variable selection was not used. The final model selected, which included MRIss, PSA Density and Age as independent predictors, was done through the assistance of Akaike’s Information Criterion and represented a trade-off between the goodness of fit and the complexity of the model. Parameter estimation was done using maximum likelihood. All fitted models were bias-corrected using 300 bootstrap resamples. Corresponding nomograms were generated for men without a prior biopsy and men with a prior negative biopsy, using the optimal models. Additional analysis included LOESS calibration curves and Receiver operating characteristic area under the curve (ROC-AUC) discrimination curves for the full model, MRIss only model, and PSAD only model. The AUCs from the testing sample were obtained by applying the fitted linear predictor from the training sample. All statistical analysis was performed using the regression modeling strategies module in the statistical software package R (https://www.r-project.org/).
RESULTS

A total of 459 men who underwent MRI and MRI-US fusion targeted biopsy along with systematic ultrasound-guided prostate biopsy were evaluated. Table 1 shows patient characteristics of the study cohort, stratified by biopsy indication. Briefly, median age was 66 years for both cohorts (range 39 to 87), median PSA was 4.9 ng/ml and 6.6 ng/ml for no prior biopsy and prior negative biopsy respectively. The median PSA density was 0.1 ng/ml/cc for both cohorts (range 0.0 to 1.89 for men with no prior biopsy and 0.0 to 2.95 for men with prior negative biopsy). Characteristics of men (median PSA, prostate volume, age) in the training and testing datasets for both the no prior biopsy and negative prior biopsy cohorts were not significantly different. Distribution of any PCa and Significant PCa in both cohorts were - No Prior Biopsy: Training (111/201, 70/201), Testing (41/87, 27/87) and Prior Negative: Training (38/119, 26/119), Testing (15/52, 5/52).

Supplemental Figure 1 shows scatterplots of PSA density (y-axis) by age, color-coded for MRIss, in separate panels for those with Gleason scores ≥ 6 and ≥ 7. The plot suggests that patients with PSA Density > 0.25 ng/ml-cc (for a median prostate of size 46cc that implies a PSA > 11.5 ng/ml) are more likely to have PCa detected. However, even for patients with normal PSA Densities, a MRIss of 4 or 5 increases the likelihood of PCa detection.

Table 2 shows predictors of PCa with coefficient estimates stratified by no prior biopsy and prior negative biopsy, and overall and by Gleason score ≥7 cancer, for the training datasets. PSA density, age, and MRIss were statistically
significant independent predictors of PCa in at least one model and were included in all models for consistency. Omnibus p-values of MRIss based on likelihood ratio tests are also included in the table. As per Table 2, an increase of PSA density of 0.1 ng/ml-cc, in a patient with no previous biopsy, approximately doubles [exp(6.3*0.1)=1.9] the odds of finding clinically significant PCa upon targeted and systematic biopsy, when all other factors remain constant. Similarly, all other factors being constant, a man with no prior biopsy with a MRIss of 5 has 16 times [exp(4.68-1.91)=16.1] higher odds of being diagnosed with clinically significant PCa upon biopsy than a patient with MRIss of 3.

**Nomograms to Estimate Risk for PCa and Clinically Significant PCa**

Figure 1 shows nomograms for predicting prostate cancer in men without a prior biopsy and those who had a prior negative biopsy. Nomograms for both overall probability of PCa and clinically significant PCa detection by systematic and targeted biopsy were generated. MRIss demonstrates a substantial influence on cancer detection, especially in men with a prior negative biopsy.

For both patient cohorts, biopsy naïve and prior negative, Supplemental Figure 2 shows the LOESS internal and external calibration plots for the training and testing datasets respectively. The plots represent observed vs model-predicted rate of PCa and Significant PCa upon biopsy and refer to the degree of agreement between observed and predictive probabilities. The relatively better calibration in the biopsy naïve cohort compared to the prior negative cohort may be due to sample size. Receiver operating characteristic discrimination curves
for creating and validating the nomogram in predicting overall cancer and Gleason score ≥7 cancer for men without a prior biopsy and those with a prior negative biopsy are shown in Figure 2. The AUC was high for predicting the probability of PCa. The total AUC for predicting PCa in men with no prior biopsy in the training (testing) sample was 0.82 (0.78) based on the nomogram. The AUC for a MRIss model was 0.80 (0.81) and for a PSAD model was 0.71 (0.61), demonstrating the full model has a 11% (17%) improvement over PSAD alone. The total AUCs for predicting Gleason score ≥7 cancer in the training (testing) sample was 0.91 (0.84), while the AUCs for the MRIss and PSAD models were 0.90 (0.84) and 0.75 (0.69), respectively. In men with a prior negative biopsy, the overall cancer detection AUC in the training (testing) sample were 0.76 (0.65) and the AUC for the MRIss and PSAD models were 0.75 (0.75) and 0.64 (0.65), respectively. The AUC in the training (testing) sample for detecting Gleason score ≥7 PCa was 0.86 (0.87) and 0.83 (0.84) for the MRIss model and 0.76 (0.76) for the PSAD model. AUCs are provided for both the training and testing samples inasmuch as the former are subject to overoptimism bias, despite the bootstrap bias-correction, due to the process of selecting from eight candidate variables, whereas the latter are unbiased but more variable due to smaller sample.

DISCUSSION

In this study we developed novel predictive clinical nomograms,
incorporating pre-biopsy MRI data for assessing risk of PCa prior to biopsy among men with no previous history of PCa presenting for consideration of biopsy. In the face of growing concern about over-utilization of prostate biopsy in urologic practice, these nomograms may provide a clinical instrument that accurately determines an individual’s risk for PCa using readily available clinical information. Furthermore, given the growing desire among urologists to avoid the detection of indolent PCa, unlikely to affect longevity, nomograms specifically evaluating the risk of clinically significant cancer might allow more meaningful risk assessment and patient counseling prior to biopsy.

The accuracy of our nomograms is predicated upon superior MRI quality and may have limited application when high-quality imaging is not available. Although we employed a 5-point Likert scale to prostate MRI interpretation in our study cohort, the recent PI-RADS v2 has been designed to promote global standardization and diminish variation in the acquisition, interpretation, and reporting of prostate MRI examinations. In the context of robust MRI quality, the MRI suspicion score is the single most important determinant of PCa risk. MRI ss appears to be a strong predictor of PCa in our study, corroborating these findings whereby an increase in MRI ss exponentially increases the risk of harboring Gleason score ≥7 PCa, and predicts cancer in all cohorts. In most cases the models using MRI ss alone performs quite well for PCa detection and the incorporation of additional clinical information provides modest incremental improvement. Other factors explored in our study, including Prostate Cancer
Antigen 3 (PCA3), body mass index, and family history of PCa were not found to be predictive.

A number of nomograms have been previously developed to predict the probability of PCa on biopsy and thus help physicians counsel patients regarding the need for biopsy and the risk of cancer. Such nomograms, which are typically based on clinical and ultrasound findings, have shown similar levels of accuracy, even when derived from different patient populations and geographical locations. The application of prior predictive nomograms has focused mainly on overall cancer detection rate, with few studies directed toward detecting clinically significant cancer, and has been limited mainly to the standard systematic template transrectal ultrasound guided biopsy or extended biopsy in the setting of repeat biopsy which may contribute to an under detection of clinically significant disease in the reference test. In one of the few studies evaluating both overall and high-grade PCa, Zaytoun et al. created nomograms using 6 clinical laboratory variables and reported an AUC of 0.73 for all cancer and 0.71 for high-grade cancer. In the setting of a prior negative biopsy, Benecchi et al. developed an accurate model to predict the outcome of repeat prostate biopsy. Adding the free-to-total PSA ratio, digital rectal examination, prostate specific antigen and slope, and history of high grade prostatic intraepithelial neoplasia sharply improves the accuracy of the model to an of AUC 0.856 in the validation sample. Recently biomarkers such as PCA3 and ExoDx Prostate IntelliScore have been incorporated into nomograms and models used to predict the likelihood of overall and high-risk PCa with improved discrimination between
Gleason 6 and Gleason ≥7 cancer. The predictive accuracy of our nomogram compares favorably to these prior studies, highlighting the impact of MRIss in predicting PCa, and performs significantly better than conventional screening with PSA.

In addition to better performance characteristics, our nomograms offer several advantages over other nomograms used for predicting PCa on biopsy. First, our nomograms highlight the probability of detecting any PCa and/or clinically significant PCa on biopsy, whereas the majority of prior nomograms are limited to any PCa including indolent disease. In the era of concern for PCa overdetection and emerging active surveillance, accurately assessing the risk of clinically significant disease while avoiding Gleason score 6 cancer becomes increasingly important. Despite satisfactory predictive accuracy of prior models, our model avoids important limitations of previous nomograms. The predictions of several prior nomograms are only applicable after transrectal ultrasound since transrectal ultrasound variables are necessary for risk estimation. For example, prostate volume is required for the determination of PSA density and is commonly determined on transrectal ultrasound. Use of ultrasound based input is impractical since men who undergo transrectal ultrasound are also likely to undergo ultrasound guided needle biopsy. In our nomogram, prostate volume and PSA density are determined from the prostate MRI allowing incorporation of this valuable information into the nomogram. By avoiding the need for transrectal ultrasound based input, our nomograms can be interpreted to counsel men before planned ultrasound guided biopsy, ultimately allowing counseling patients
on both the need for biopsy as well as the probability of cancer. Internal validation of our nomograms reinforces the usefulness of this clinical decision aid. Consequently, use of our model is substantiated by both the practicality of use and its highly accurate prediction ability.

Our report is not based upon a prospectively enrolled study cohort, but rather a prospective data registry of men consecutively presenting to our institution for prostate biopsy. We have, however, previously reported that 98.9% of men undergoing biopsy at our institution during this period, underwent pre-biopsy MRI and MRI-US fusion biopsy when MRI-suspicious regions were present\(^{28}\), suggesting high levels of adherence to the pre-biopsy MRI paradigm. Potential bias is introduced by the fact that all men presenting for biopsy had pre-existing suspicion for cancer, typically on the basis of elevated PSA or previous biopsy findings, and as such our nomogram tools may be best utilized in this particular clinical setting. This consideration may also be, in part, responsible for the modest contribution of serum PSA to the predictive model. Nonetheless, in our patient cohorts the clinical parameters (PSA Density, age) occupy a wide range of normal and abnormal values. There is also considerable overlap of values in patients with and without cancer. The nature of the cohort is also limited in that men with no suspicious lesion on MRI were not included in the nomogram creation and, thus, we cannot compare the detection of PCa or the grade of PCa between those with and without a suspicious lesion on MRI. As a result, our nomograms are only applicable to those men with MRI abnormal regions. We have, however, previously observed very low rates of clinically significant cancer
detection in men with a negative MRI\textsuperscript{15} suggesting that men without an abnormal lesion on MRI are at an extremely low risk of harboring clinically significant PCa.

MRI with MRI-US fusion targeted biopsy has been shown to detect more Gleason score $\geq 7$ PCa than systematic biopsy while limiting detection of Gleason score 6 cancer in men presenting for prostate biopsy.\textsuperscript{6, 7, 13} The improved accuracy of MRI-targeted biopsy, as compared to systematic biopsy, may have contributed to the relatively high diagnostic accuracy of our predictive nomograms. As such, when employing the proposed nomograms in the setting of systematic biopsy sampling alone, diagnostic accuracy similar to that reported in this study may not be achieved. Our nomograms have the potential to avoid unnecessary biopsies in a considerable number of men at the expense of missing only a few men harboring clinically significant cancer.

Additionally, because our reference standard remains a biopsy rather than a final prostatectomy specimen, we recognize that some men with negative biopsy likely do harbor clinically significant cancer. The relevance of this observation is questionable as men cannot undergo treatment without cancer identified on biopsy. It is noteworthy that predictive models such as nomograms depend strongly on their development data and patient populations. Finally, clinical recommendations derived from our data must be predicated on our considerable experience with MRI of the prostate, its interpretation, and MRI-targeted biopsy techniques. Whether such observations could be duplicated in other centers remains to be determined through additional studies. Although prostate MRI examinations routinely reviewed by a single experienced radiologist
prior to the biopsy may strengthen the internal validity, we acknowledge this may make our results less generalizable given the variation in radiologist’s interpretation of MRIss. Despite its limitations, our study has several strengths, including that, in the absence of contraindication, nearly all men presenting to our center during the study period underwent prebiopsy MRI.28

CONCLUSIONS

PSA density, age, and MRI suspicion score predict PCa on MRI-targeted and systematic biopsy. Our generated nomograms demonstrate high diagnostic accuracy and may further aid in the decision to perform biopsy in men with clinical suspicion of PCa. Although this study reflects our institutional experience, it is important for other centers to confirm and externally validate our findings.
References


**Figure 1.** Nomogram prediction model for predicting overall (A) and Gleason ≥7 PCa (B) in men without a prior biopsy, and with a prior negative biopsy (C, D), respectively. The nomogram is used by first locating a patient’s position for each variable on its horizontal scale and then a point value is assigned according to the points scale (top axis) and summed for all variables. Total points correspond to a probability value for having PCa or Gleason score ≥7. PSA, prostate-specific antigen; MRIss, MRI suspicion score.

**Figure 2.** Receiver operating characteristic discrimination curves for the nomogram in predicting any and Gleason score ≥7 PCa. Men without a prior biopsy predicting any cancer (A) and clinically significant Gleason ≥7 PCa (B), and in men with a prior negative biopsy predicting any cancer (C) and clinically significant Gleason ≥7 PCa (D).

**Supplemental Figure 1. PSA Density, Age, and MRIss, Stratified by Gleason Score ≥ 6 and ≥ 7.** Plot showing distribution of clinical variables in men without prior biopsies with any cancer (A) and Gleason ≥7 PCa (B), and men with a prior negative biopsy with any cancer (C) and Gleason ≥7 PCa (D). The right (left) column labeled TRUE (FALSE) indicates patients who had (had no) any and Gleason ≥7 PCa detected upon biopsy, respectively. The y-axis shows the PSA Density in ng/ml-cc and the x-axis shows patient age in years. The MRI suspicion score is color-coded as yellow (score 2), green (score 3), blue (score 4) and red (score 5) on the plot.

**Supplemental Figure 2.** Loess calibration plots for training (A, C, E, G) and testing (B, D, F, H) datasets for biopsy naive (A, B, C, D) and prior negative cohorts (E, F, G, H). The plots represent observed vs model-predicted rate of PCa (A, B, E, F) and Significant PCa (C, D, G, H) upon biopsy. The sample size (N) and the mean absolute error in predicted and loess-calibrated probabilities (|Eavg|) for the plots are: (A) N=201, |Eavg| = 0.026, (B) N=87, |Eavg| = 0.009, (C) N=201, |Eavg| = 0.017, (D) N=87, |Eavg| = 0.006, (E) N=119, |Eavg| = 0.031, (F) N=52, |Eavg| = 0.068, (G) N=119, |Eavg| = 0.026, (H) N=52, |Eavg| = 0.008.
**Table 1.** Patient characteristics and descriptive statistics by subgroup. Data for MRI Suspicion score is arranged as mode (minimum, median, maximum). All other characteristic variables arranged as mean (minimum, median, maximum)

<table>
<thead>
<tr>
<th>Variable</th>
<th>No (n = 288)</th>
<th>Yes (n = 171)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years</td>
<td>64.58 (39, 66, 87)</td>
<td>66.06 (50, 66, 87)</td>
</tr>
<tr>
<td>PSA in ng/ml</td>
<td>6.21 (0.18, 4.9, 83.1)</td>
<td>13.67 (0.82, 6.6, 827.0)</td>
</tr>
<tr>
<td>Prostate vol in cc</td>
<td>52.93 (5, 46, 207)</td>
<td>72.5 (12, 64, 280)</td>
</tr>
<tr>
<td>PSA density in ng/ml/cc</td>
<td>0.15 (0.0, 0.1, 1.89)</td>
<td>0.17 (0.0, 0.1, 2.95)</td>
</tr>
<tr>
<td>% Positive cores</td>
<td>64.5 (0, 75, 100)</td>
<td>60 (0, 50, 100)</td>
</tr>
<tr>
<td>Total cancer length in mm among patients with any cancer</td>
<td>7.1 (0.1, 6.0, 35) (n=152)</td>
<td>5.97 (0.5, 5.0, 16.0) (n=53)</td>
</tr>
<tr>
<td>MRI suspicion score</td>
<td>3 (2, 3, 5)</td>
<td>3 (2, 3, 5)</td>
</tr>
</tbody>
</table>
Table 2. Predictors of overall and Gleason score ≥7 PCa detection on systematic and MRI-US fusion targeted biopsy. Coefficient estimates provided are parameters of logistic regression models not odds ratios. Distribution of MRIss among Biopsy Naive patients and Prior Negative patients is - MRIss 2: 63 and 38; MRIss 3: 59 and 38; MRIss 4: 37 and 29; MRIss 5: 42 and 14 respectively.

<table>
<thead>
<tr>
<th>All prostate cancer</th>
<th>No Prior Biopsy</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P-value</th>
<th>Prior Negative Biopsy</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRIss 2 Referent</td>
<td>Referent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRIss 3</td>
<td>0.47</td>
<td>0.39</td>
<td></td>
<td>&lt;0.0001</td>
<td>0.13</td>
<td>0.61</td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MRIss 4</td>
<td>2.29</td>
<td>0.54</td>
<td></td>
<td></td>
<td>1.33</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRIss 5</td>
<td>2.48</td>
<td>0.63</td>
<td></td>
<td></td>
<td>4.27</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSA density</td>
<td>6.55</td>
<td>2.35</td>
<td>0.005</td>
<td></td>
<td>0.36</td>
<td>0.96</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.03</td>
<td>0.02</td>
<td>0.19</td>
<td></td>
<td>-0.04</td>
<td>0.0319</td>
<td>0.25</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gleason score ≥7 prostate cancer</th>
<th>No Prior Biopsy</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P-value</th>
<th>Prior Negative Biopsy</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRIss 2 Referent</td>
<td>Referent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRIss 3</td>
<td>1.91</td>
<td>0.84</td>
<td>&lt;0.0001</td>
<td></td>
<td>0.36</td>
<td>0.89</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>MRIss 4</td>
<td>3.63</td>
<td>0.83</td>
<td></td>
<td></td>
<td>2.01</td>
<td>0.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRIss 5</td>
<td>4.68</td>
<td>0.88</td>
<td></td>
<td></td>
<td>5.95</td>
<td>1.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSA density</td>
<td>6.30</td>
<td>2.22</td>
<td>0.005</td>
<td></td>
<td>0.70</td>
<td>1.01</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.04</td>
<td>0.02</td>
<td>0.12</td>
<td></td>
<td>-0.09</td>
<td>0.05</td>
<td>0.04</td>
<td></td>
</tr>
</tbody>
</table>
The authors’ nomograms are based on three risk factors: age, prostate specific antigen density (PSAD), and a five-category ordinal magnetic resonance imaging (MRI) Suspicion Score (MRISS), recommended by a European Society of Urogenital Radiology expert panel for use in conjunction with the Prostate Imaging Reporting and Data System.\textsuperscript{1,2} The current study’s results, particularly for Gleason $\geq 7$ PCa risk, provide substantial encouragement that MRI imaging may contribute to a more efficient approach to PCa biopsy, and an improved benefit to harm ratio from PCa screening.

However, while the authors wisely conclude that “Although this study reflects our institutional experience, it is important for other centers to confirm and externally validate our findings,” the publication of these nomograms could
nevertheless encourage widespread adoption prior to such confirmation. This would be premature.

Although three variables contribute to the nomograms, the subjective MRISS contributes far more to their predictions than the PSAD. For each model, adding the MRISS to the PSAD yielded substantial increases in AUC, while the PSAD added much less or no improvement to predicting from the MRISS. The pivotal MRISS data, however, reflect the clinical judgments and experience of only one experienced radiologist, viewing scans likely of very high quality. The MRISS scale is published without either reference criteria to distinguish between the qualitatively defined probability ranges, or quantitative thresholds through which readers might calibrate their interpretations of the authors’ qualitative descriptors of magnitude (very low, low, equivocal, high, or very high probability).

That health professionals vary on such interpretations of probabilistic language is well documented. It is also likely that radiologists differ in weights assigned to, and abilities to discern, different image features relevant to category assignment. Moreover, a recent review suggests that image quality, specific genitourinary expertise, didactic teaching, and dedicated training sessions all have impact on reader performance using MRI prostate cancer scoring systems. Since the discrimination capacity of the nomograms will depend on readers’ performances in using the MRISS in their own medical settings, comparable nomogram effectiveness in other settings cannot be presumed.

Two recent studies from major urban hospitals have examined both
performance and observer variability of the MRISS in contexts comparable to the current study.\textsuperscript{6,7} Overall training set AUCs using MRISS alone were respectively 0.85 and 0.89, respectively. In the first, comparing three fellowship trained readers with six years of experience, a weighted kappa of 0.56 was obtained, with raw agreement unreported.\textsuperscript{6} In the second, weighted kappa for MRISS results of two readers of unspecified experience was 0.80 in the context of overall 78% agreement.\textsuperscript{7} These and the current study generally may be viewed as mutually confirming of predictive capacity of the MRISS, and of reasonable though imperfect observer agreement within the same urban, research setting.

Two tasks thus remain: i) to develop training and materials to communicate what the MRISS in these studies is conveying, so the diagnostic skill evident in these data can be exported to others, and ii) to demonstrate robustness of such results to diverse settings and varied MRI readers.

References


We are grateful for the editorial comment, and the insightful observations and comments regarding our manuscript predicting prostate cancer risk among men undergoing combined MRI-targeted and systematic biopsy using novel pre-biopsy nomograms that incorporate MRI findings.

We agree with the comment, that despite the existence of PI-RADS for guiding prostate MRI interpretation, such interpretation maintains a subjective component, impacting the generalizability of our nomograms to other centers. In order to reduce this variability, several measures are now underway to further
standardize prostate MRI interpretation. For example, numerous radiological societies provide hands-on courses with expert radiologists in interpretation and reporting. The American College of Radiology is pursuing a prostate MRI certification process to help ensure the quality of centers performing prostate MRI. Lastly, PI-RADS is intended to be a “living” document that will evolve as clinical experience and scientific data accrue. With continued revisions and improvements, PI-RADS will continue to guide MRI interpretation and reporting. Ultimately, reader agreement in interpretation, and concordance of biopsy findings amongst centers, is almost certainly a time-dependent phenomena relying upon a learning curve associated both with the individual components and the collective integration of the MRI-based prostate cancer detection paradigm. It has been demonstrated that experience in MRI interpretation improves reader accuracy and concordance.

Although variability remains in prostate MRI interpretation, other international institutions have assessed the utility of MRI and MRI ultrasound fusion biopsy as a pre-biopsy risk assessment tool, similar to our nomograms. Van Leeuwen et al. developed a predictive model for detecting significant prostate cancer on 393 men and externally validated the model on 198 men from a separate institution. When including age, PSA, DRE, prostate volume, and MRI in predicting the probability of significant prostate cancer in patients undergoing prostate biopsy, MRI continued to have the largest influence. Moreover, the area under the curve for predicting significant prostate cancer is in agreement with our study, suggesting reproducibility. The most critical
observations of these studies are the consistent demonstration that MRI suspicion score (when performed by experienced readers) is the most powerful variable in prediction of risk, and that additional clinical variables, such as PSA density, can be used to further refine that predictive capability. This is critically important in the evolving goals of prostate biopsy – to find clinically significant cancer, with lethal potential, and avoid detection of indolent disease. Ultimately, further evaluation of our nomograms, as well as MRI based risk assessment tools such as apparent diffusion coefficient values and other quantitative imaging metrics, on additional independent patient cohorts is warranted prior to implementation in clinical practice. We would welcome such external validation of the nomograms by other groups or individuals. We suspect that while absolute numbers may vary slightly, according to MRI reader experience, the predictive trends will remain the same.

REFERENCES


Figure 1 Nomograms UROLOGY _bestsetConverted.png
Figure 2 AUC Revision 3 with Cl_bestsetConverted.png