Title
Integrative Radiomics Models To Predict Biopsy Results For Negative Prostate MRI

Authors
Zheng, Haoxin
Miao, Qi
Raman, Steven S
et al.

Publication Date
2021-04-01

DOI
10.1109/isbi48211.2021.9433879

Peer reviewed

INTEGRATIVE RADIOMICS MODELS TO PREDICT BIOPSY RESULTS FOR NEGATIVE PROSTATE MRI

Haoxin Zheng[a,b], Qi Miao[a], Steven S. Raman[a], Fabien Scalzo[b,c], Kyunghyun Sung[a]

[a] Radiological Sciences, University of California – Los Angeles, Los Angeles, CA 90095, USA
[b] Computer Science, University of California – Los Angeles, Los Angeles, CA 90095, USA
[c] Neurology, University of California – Los Angeles, Los Angeles, CA 90095, USA

ABSTRACT

Multi-parametric MRI (mpMRI) is a powerful non-invasive tool for diagnosing prostate cancer (PCa) and is widely recommended to be performed before prostate biopsies. Prostate Imaging Reporting and Data System version (PI-RADS) is used to interpret mpMRI. However, when the pre-biopsy mpMRI is negative, PI-RADS 1 or 2, there exists no consensus on which patients should undergo prostate biopsies. Recently, radiomics has shown great abilities in quantitative imaging analysis with outstanding performance on computer-aid diagnosis tasks. We proposed an integrative radiomics-based approach to predict the prostate biopsy results when pre-biopsy mpMRI is negative. Specifically, the proposed approach combined radiomics features and clinical features with machine learning to stratify positive and negative biopsy groups among negative mpMRI patients. We retrospectively reviewed all clinical prostate MRIs and identified 330 negative mpMRI scans, followed by biopsy results. Our proposed model was trained and validated with 10-fold cross-validation and reached the negative predicted value (NPV) of 0.99, the sensitivity of 0.88, and the specificity of 0.63 in receiver operating characteristic (ROC) analysis. Compared with results from existing methods, ours achieved 11.2% higher NPV and 87.2% higher sensitivity with a cost of 23.2% less specificity.

Index Terms— Computer-aided diagnosis, prostate cancer, MRI, radiomics

1. INTRODUCTION

Prostate cancer (PCa) is the most common solid organ malignancy among men in the United States [1]. Multi-parametric MRI (mpMRI) gains clinical acceptance as the preferred imaging technique for diagnosing PCa, and mpMRI is widely recommended, prior to prostate biopsies, to improve the detection rate of clinically significant PCa (csPCa) [2]. Prostate Imaging Reporting and Data System: Version 2.1 (PI-RADS v2.1) [3] is commonly used to interpret mpMRI. When mpMRI is negative, where PI-RADS is either 1 or 2, there is no consensus about which patients can forgo biopsies with current medical practice.

Several studies have investigated systematic strategies to predict risks associated with csPCa among patients with negative mpMRI [4-7]. The studies suggested performing prostate biopsy among negative mpMRI patients if the prostate-specific antigen density (PSAD) level is higher than 0.15/ng/ml/ml [4-6]. Moreover, another study stated that patients' age showed a significant difference between negative and positive biopsy groups among negative mpMRI patients [7]. However, the current prediction models based on clinical information are limited by marginal improvement in prediction abilities [6].

Radiomics is an emerging field in quantitative imaging that aims to associate large-scale radiomic features with specific clinical endpoints [8]. The radiomic features, extracted from medical images, are assumed to provide large-scale imaging information, such as intensity, shape, size, and texture features within regions of interest (ROIs). Many studies have shown its outstanding performances on computer-aid diagnosis tasks, especially for the detection and classification of the aggressiveness PCa [9-12], due to its unique ability to extract and assess the spatial arrangement within the tumor regions on medical images.

In this study, we proposed an integrative radiomics-based machine learning approach that predicts biopsy results when pre-biopsy mpMRI is negative. Specifically, we combined radiomics features within the whole prostate, extracted from functional and anatomical information of mpMRI, with routinely used clinical information to separate negative and positive biopsy groups among negative mpMRI patients. The efficacy of the integration between radiomics features and clinical information was evaluated by comparing machine learning models with individual feature groups alone. In addition, the proposed model was also compared with existing conventional strategies that were proposed to predict risks associated with csPCa [4, 5], evaluated by negative predictive value (NPV), sensitivity, specificity, and area under the curve (AUC).

2. MATERIALS AND METHODS

2.1. MRI Data and Clinical Information

We retrospectively identified negative prostate MRI cases by reviewing all clinical prostate MRI scans, acquired from January 2016 to December 2018, at a single academic institution and classified them as negative MRI if the PI-RADS v2.1 score was either one or two. All scans were performed with a standard protocol using one of the three 3 Tesla scanners: Siemens Magnetom Trio, Skyra, and Verio scanner (Siemens Medical Systems, Malvern, Pennsylvania, USA). The MRI scans with the following conditions were excluded: 1) patients with prior treatment for PCa and benign prostatic hyperplasia (BPH), 2) patients who did not undergo prostate biopsies within six months after MRI, 3) patients who were undergoing active surveillance, and 4) patients who did not include prostate-specific antigen density (PSAD) on their medical records.

We included the following clinical information: patients' age, family history of PCa, prostate biopsy history, prostate volume, PSA, and PSAD. Other clinical information was not included in the study to avoid potential selection bias [13] as it was partially available to certain patient cohorts. Patients with negative mpMRI and high suspicion of PCa (family history of PCa, elevated PSA, abnormal DRE, etc.) underwent standardized 12-14 core systematic transrectal ultrasound-guided (TRUS) biopsy. All biopsy cores were immediately fixed in formalin, stained with hematoxylin and eosin (H & E), and microscopic evaluation was performed by dedicated pathologists as part of the routine histopathological evaluation.

In all, 330 men were included in the study. A total of 24 patients were categorized as having positive biopsies, defined as patients who had at least one positive biopsy core, Gleason Score (GS) ≥ 7 per biopsy session, and the rest (n=306) was categorized as the negative biopsy (GS=6 or benign condition).

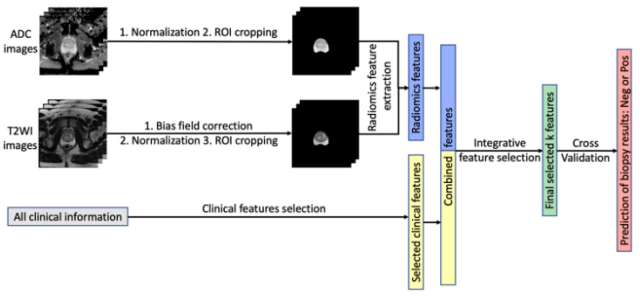

Fig. 1: Workflow of building the integrative radiomics model for predicting biopsy results. The three inputs of the model were the patient's clinical information, T2WI, and ADC images. First, clinical features were selected from all clinical information, and radiomics features were extracted from the T2WI and ADC images that have been pre-processed and cropped based on ROI. Then, integrative feature selection was made based on the combination of the two categories of features. Finally, with the selected features, cross-validation was performed to evaluate the model's predictability.

2.2. Imaging Pre-processing

The workflow for our proposed radiomics model is shown in Figure 1. For a patient-basis prediction of positive or negative biopsy results, we used both apparent diffusion coefficient maps (ADC) and T2-weighted images (T2WI) from mpMRI [3]. ADC images were first registered to T2WI through rigid spatial transformation using voxel size and real-world coordinates information for each patient [14]. We did not apply for additional non-rigid registration since only the minimal patient motion was observed between two imaging sequences as the time between DWI and T2WI was minimal [15]. The prostate area was manually segmented on T2W images slice-by-slice by an experienced radiologist (4 years of post-fellowship experience) using OsiriX MD (ver. 11.0.3). We then applied N4 bias field correction to T2WI to compensate for the low-frequency intensity non-uniformities [16] and applied z-score normalization [17] to T2WI and ADC images to minimize the potential intensity differences among different scans and other physiological variations.

2.3. Radiomics Features

Radiomics features were extracted from ADC and T2WI after cropping the whole prostate, as shown in Fig. 1. All the slices containing ROIs of the whole prostate were used for feature extraction, and the mid-prostate slice was separately used to extract additional radiomics features. Among texture features, Gray-Level Cooccurrence Matrix (GLCM) and Gray-Level Run Length Matrix (GLRLM) were included using Pyradiomics package based on Python [18]. We extracted a total of 300 radiomics features for each patient, including 32 shape-based, 38 first-order, and 80 texture (or second-order) features from each of the T2WI and ADC images.

2.4. Feature Selection

We calculated the significance level of all routinely-used clinical features between prostate biopsy positive group and negative group in order to pre-select important clinical information. Given the six initial clinical characteristics, we used the Mann-Whitney U test for continuous-valued features (e.g., age, PSA, PSAD, prostate volume) and the Chi-Square test for categorized features (e.g., family PCa history, prostate biopsy history). We selected the features with a significant difference ($p<0.05$) between the biopsy positive and negative groups. Finally, age, prostate volume, and PSAD were pre-selected.

We first combined the pre-selected clinical features and all radiomics features and then applied the Sequential Floating Forwarding Selection (SFFS) algorithm [19] for integrative feature selection (Fig. 1). The number of integrative features, k, is a hyperparameter, which was decided by investigating the relationship between k and the prediction performances

Table 1: Selected nine integrative features.

Selected Features	Type	Imaging Sequence
Gray Level Nonuniformity	GLRLM	ADC
Run Length Nonuniformity	GLRLM	ADC
Sum Squares	GLCM	T2WI
Least Axis Length	Shape	ADC/T2WI
Major Axis Length	Shape	ADC/T2WI
Minor Axis Length	Shape	ADC/T2WI
Age	Clinical Info.	--
PSAD	Clinical Info.	--
Prostate Volume	Clinical Info.	--

Table 2: Comparison of different machine learning approaches.

Models	AUC	Sens.	Spec.	NPV
Clinical-only	0.70	0.82	0.53	0.97
Radiomics-only	0.71	0.84	0.53	0.98
Integrative Approach	**0.80**	**0.88**	**0.63**	**0.99**

Table 3: Comparisons of the biopsy prediction results. Conducted by other existing studies, using the PSAD level within our dataset, and the results generated by our integrative approach.

	Sens.	Spec.	NPV
Oishi et al. [5]	0.47	**0.82**	0.90
Distler et al. [4]	N/A	N/A	0.89
Our data + PSAD>0.15	0.50	0.73	0.95
Integrative Approach	**0.88**	0.63	**0.99**

during training. We ran the SFFS algorithm with different values of k and selected k with the highest AUC.

2.5. Model Evaluation and Comparison

We used a quadratic-kernelized SVM classifier with a class-balanced weight for building our proposed model. In order to minimize potential overfitting, the model was validated by 100-times iterative 10-fold random-split cross-validation. The data were randomly split into ten folds at each time, and the model was trained on nine of the ten folds (n=297) and validated on the remaining fold (n=33).

We first investigated the value of the integrative radiomic-feature-based model by comparing the performance of the integrative radiomics model with the radiomics-features-only and the clinical-features-only models. All models were based on the same classifier, the quadratic-kernelized SVM with the class-balanced weight, for a fair comparison. We then compared the proposed model's prediction performance with conventional strategies [4, 5] that predict prostate cancer risks using routinely available clinical features. The method was based on a threshold of the PSAD level, suggesting patients with PSAD>0.15ng/ml/ml are high risks of having prostate cancer among mpMRI negative patients [4, 5].

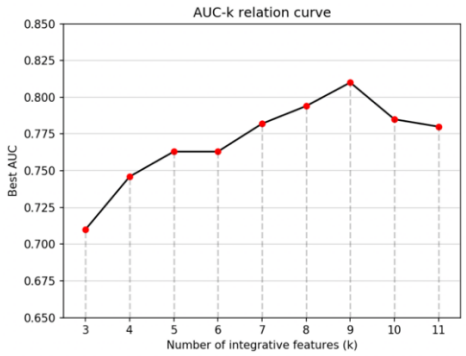

Fig. 2: Relationship between the number of integrative features and their corresponding AUCs generated by the prediction model

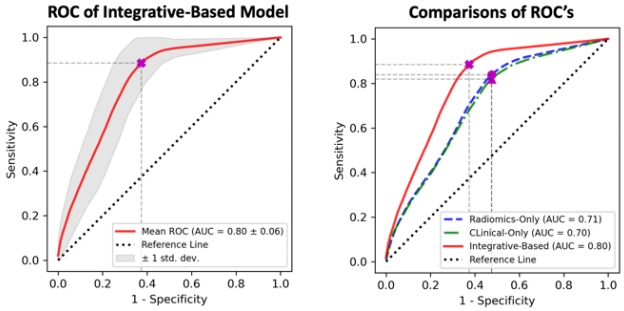

Fig. 3: Comparisons between the integrative-based model and machine learning models with individual feature groups. Red solid, blue dash and green dot-dash curves are the ROC curves of the integrative-features-based, radiomics-features-only, and clinical-features-only model. Points represented with the cross, dot, and triangle shapes and labeled with magenta color are the optimal points for the three ROC curves. The shadow area visualizes the variability of ROC curves generated from each iteration of the 100 times 10-fold cross-validation using the proposed integrative-features-based model, and is ranged using ± 1 standard deviation. Horizontal and vertical gray dash lines of each optimal point aim to visualize sensitivity and "1-specificity" value on the figures.

For each experiment, we identified the optimal cut point for the prediction of biopsy results by maximizing Youden's index value (sensitivity+specificity-1) [20]. The sensitivity, specificity and NPV were calculated based on the optimal cut point. All the model comparisons were evaluated based on AUC, sensitivity, specificity and NPV.

3. RESULTS

3.1. Integrative Radiomics Model

The total number of the integrative features was nine (three clinical and six radiomics features), and their corresponding AUC values with different numbers of selected features are shown in Figure 2. The selected integrative features are listed in Table 1. Radiomics features extracted from both imaging sequences (ADC and T2WI) and certain clinical features helped improve the predictability of the prostate biopsy results, consistent with previous studies [4-7].

3.2 Model Comparisons

Figure 3 shows the comparisons between different machine learning models. Based on the ROC curves of the three models, the integrative approach not only reached the highest AUC, compared with the models using an individual group of radiomics or clinical features but also showed the highest sensitivity and specificity at the optimal cut point. Table 2 shows the comparisons of the AUC, sensitivity, specificity and NPV at the optimal point of the clinical-only, radiomics-only and integrative-based models. The integrative model outperformed the other two models in all measurements. The integrative-based model improved the AUC of 14.3% and 12.6%, the sensitivity of 7.3% and 4.8%, the specificity of 18.9% and 18.9% and NPV of 2.1% and 1.0%, compared with the clinical-only and radiomics-only models.

Figure 4 shows representative examples of negative MRI (T2WI and ADC) scans with positive and negative biopsy results. The proposed model cam achieve a high averaged accuracy of 1.0 and 0.79, calculated by the total number of the correct prediction / the total number of the prediction with the 100 iterative random-split 10-fold cross-validations.

Table 3 shows the comparison results between the proposed model and existing conventional strategies [4, 5], designed to predict prostate biopsy results among patients with negative mpMRI. Only NPV was included for Distler. et al. since the specificity and sensitivity were not available. We observed that in terms of sensitivity and NPV, our proposed integrative approach got better performances compared with not only the reported results from the two studies but also the results conducted by applying their method onto our dataset. Our model outperformed both Oishi [5] and Distler [4], largely in the sensitivity of 87.2%. Comparing with both studies, we also outperformed the NPV of 10.0% and 11.2%, respectively. With the identical patient cohort, our proposed model improved 76% of the sensitivity and 4.2% of NPV.

4. DISCUSSIONS

We proposed an integrative radiomics-based approach to predict the biopsy results when mpMRI was negative. The proposed model was built based on SVM and the integration of radiomics features extracted from mpMRI and clinical features. We trained and validated our model with 330 patients with negative mpMRI, where the ground truth for the biopsy results was confirmed by a systematic TRUS biopsy within six months.

In current medical practice, to stratify which patient with negative prostate MRI should undergo prostate biopsy, both high NPV and sensitivity are desired since we want to avoid as much negative biopsies as possible, and meanwhile limit the number of false-negative prediction cases, which are measured by both NPV and sensitivity. Compared with other models, our proposed model achieved much higher NPV and sensitivity, and thus can help improve the stratification results for the patients with negative prostate mpMRI. In addition,

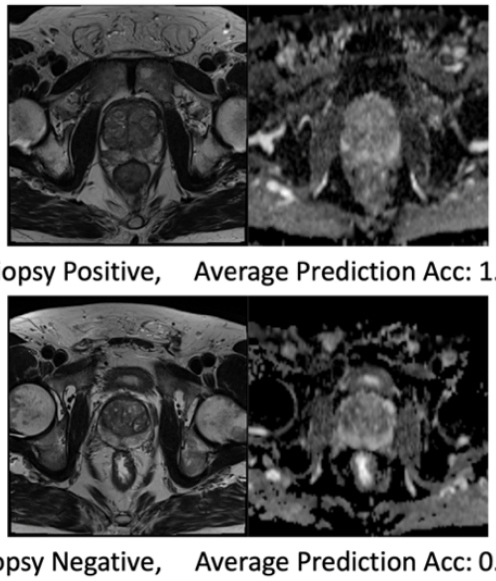

Fig. 4: Examples of T2WI (left) and ADC (right) images with negative mpMRI and the averaged prediction accuracy. A): a patient with a positive biopsy. B): a patient with a negative biopsy.

the results conducted by the proposed model showed that both radiomics features and clinical information helped build the prediction model. The investigation of important radiomics features may also provide more insights for radiologists to interpret mpMRI in the future. In all, our proposed integrative radiomics-based model outperformed the other methods conducted largely on the predictability of prostate biopsy results among patients with negative prostate mpMRI, especially on sensitivity (improved by 87.2%) and NPV (improved by 11.2%).

Our next step would include adapting our methods onto other datasets from multiple institutions to further investigate the generalization of our prediction model.

5. CONCLUSIONS

The study showed an integrative radiomics-based machine learning model to predict biopsy results among patients with negative prostate MRI. We presented that integration of radiomics and clinical features helps predict biopsy results, compared with other approaches using an individual group of features. Moreover, our proposed model outperformed other existing strategies on the prediction of prostate biopsy results among patients with negative prostate mpMRI.

6. ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health (NIH) R01-CA248506 and funds from the Integrated Diagnostics Program, Department of Radiological Sciences & Pathology, David Geffen School of Medicine at UCLA. The authors have no conflict of interest to declare.

7. COMPLIANCE WITH ETHICAL STANDARDS

The study was carried out in compliance with the United States Health Insurance Portability and Accountability Act (HIPAA) of 1996. The protocol was approved by the UCLA institutional review board (IRB) with a waiver of the requirement for informed consent.

8. REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: A Cancer Journal for Clinicians,* vol. 68, no. 1, pp. 7-30, 2018, doi: 10.3322/caac.21442.

[2] M. Brizmohun Appayya *et al.*, "National implementation of multi-parametric magnetic resonance imaging for prostate cancer detection – recommendations from a UK consensus meeting," *BJU International,* vol. 122, no. 1, pp. 13-25, 2018, doi: 10.1111/bju.14361.

[3] B. Turkbey *et al.*, "Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2," *Eur Urol,* vol. 76, no. 3, pp. 340-351, Sep 2019, doi: 10.1016/j.eururo.2019.02.033.

[4] F. A. Distler *et al.*, "The Value of PSA Density in Combination with PI-RADS for the Accuracy of Prostate Cancer Prediction," *J Urol,* vol. 198, no. 3, pp. 575-582, Sep 2017, doi: 10.1016/j.juro.2017.03.130.

[5] M. Oishi *et al.*, "Which Patients with Negative Magnetic Resonance Imaging Can Safely Avoid Biopsy for Prostate Cancer?," *J Urol,* vol. 201, no. 2, pp. 268-276, Feb 2019, doi: 10.1016/j.juro.2018.08.046.

[6] M. A. Pagniez, V. Kasivisvanathan, P. Puech, E. Drumez, A. Villers, and J. Olivier, "Predictive Factors of Missed Clinically Significant Prostate Cancers in Men with Negative Magnetic Resonance Imaging: A Systematic Review and Meta-Analysis," *The Journal of urology,* vol. 204, no. 1, pp. 24-32, 2020, doi: 10.1097/JU.0000000000000757.

[7] V. Panebianco *et al.*, "Negative Multiparametric Magnetic Resonance Imaging for Prostate Cancer: What's Next?," *Eur Urol,* vol. 74, no. 1, pp. 48-54, Jul 2018, doi: 10.1016/j.eururo.2018.03.007.

[8] A. Zwanenburg, S. Leger, M. Vallières, and S. Löck, "Image biomarker standardisation initiative," *CancerData,* vol. 41, no. 2, pp. 366-373, 2016, doi: 10.17195/candat.2016.08.1.

[9] C. Yan, Y. Peng, and X. Li, "Radiomics analysis for prostate cancer classification in multiparametric magnetic resonance images," 2019, pp. 247-250.

[10] A. Algohary *et al.*, "Radiomic features on MRI enable risk categorization of prostate cancer patients on active surveillance: Preliminary findings," *Journal of Magnetic Resonance Imaging,* vol. 48, no. 3, pp. 818-828, 2018, doi: 10.1002/jmri.25983.

[11] T. J. M. Castillo, M. P. A. Starmans, W. J. Niessen, I. Schoots, S. Klein, and J. F. Veenland, "Classification of prostate cancer: High grade versus low grade using a radiomics approach," *Proceedings - International Symposium on Biomedical Imaging,* vol. 2019-April, no. Isbi, pp. 1319-1322, 2019, doi: 10.1109/ISBI.2019.8759217.

[12] B. Varghese *et al.*, "Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images," *Scientific Reports,* vol. 9, no. 1, pp. 1-10, 2019, doi: 10.1038/s41598-018-38381-x.

[13] G. Tripepi, K. J. Jager, F. W. Dekker, and C. Zoccali, "Selection bias and information bias in clinical research," *Nephron Clin Pract,* vol. 115, no. 2, pp. c94-9, 2010, doi: 10.1159/000312871.

[14] C. L. Novak *et al.*, "A prostate cancer computer-aided diagnosis system using multimodal magnetic resonance imaging and targeted biopsy labels," presented at the Medical Imaging 2013: Computer-Aided Diagnosis, 2013.

[15] S. Wang, K. Burtt, B. Turkbey, P. Choyke, and R. M. Summers, "Computer aided-diagnosis of prostate cancer on multiparametric MRI: a technical review of current research," *Biomed Res Int,* vol. 2014, p. 789561, 2014, doi: 10.1155/2014/789561.

[16] N. J. Tustison *et al.*, "N4ITK: improved N3 bias correction," *IEEE Trans Med Imaging,* vol. 29, no. 6, pp. 1310-20, Jun 2010, doi: 10.1109/TMI.2010.2046908.

[17] S. J. Hectors *et al.*, "Radiomics Features Measured with Multiparametric Magnetic Resonance Imaging Predict Prostate Cancer Aggressiveness," *J Urol,* vol. 202, no. 3, pp. 498-505, Sep 2019, doi: 10.1097/JU.0000000000000272.

[18] J. J. M. van Griethuysen *et al.*, "Computational Radiomics System to Decode the Radiographic Phenotype," *Cancer Res,* vol. 77, no. 21, pp. e104-e107, Nov 1 2017, doi: 10.1158/0008-5472.CAN-17-0339.

[19] D. Zongker and A. Jain, "Algorithms for Features Selection: An Evaluation," presented at the International Conference on Pattern Recognition, Vienna, Austria, Austria, 25-29 Aug. 1996, 1996.

[20] N. J. P. Marcus D. Ruopp, Brain W. Whitcomb, and Enrique F. Schisterman, "Youden Index and Optimal Cut-Point Estimated from Observations Affected by a Lower Limit of Detection," *Biometrical Journal,* vol. 50(3), pp. 419-430, 2008 Aug 13 2008, doi: 10.1002/bimj.200710415.