PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Radiomics analysis potentially reduces over-diagnosis of prostate cancer with PSA levels of 4-10 ng/ml based on DWI data

Zhang, Shuaitong, Qi, Yafei, Wei, Jingwei, Niu, Jianxing, Gu, Dongsheng, et al.

Shuaitong Zhang, Yafei Qi, Jingwei Wei, Jianxing Niu, Dongsheng Gu, Yuqi Han, Xiaohan Hao, Yali Zang, Jie Tian, "Radiomics analysis potentially reduces over-diagnosis of prostate cancer with PSA levels of 4-10 ng/ml based on DWI data," Proc. SPIE 10950, Medical Imaging 2019: Computer-Aided Diagnosis, 109501Z (13 March 2019); doi: 10.1117/12.2511497

SPIE.

Event: SPIE Medical Imaging, 2019, San Diego, California, United States

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 18 Mar 2020 Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Radiomics analysis potentially reduces over-diagnosis of prostate cancer with PSA levels of 4-10ng/ml based on DWI data

Shuaitong Zhang[a,e+], Yafei Qi[b+], Jingwei Wei[a,e+], Jianxing Niu[f, g], Dongsheng Gu[a,e], Yuqi Han[c], Xiaohan Hao[d], Yali Zang[*a], Jie Tian[*a,e]

a CAS Key Laboratory of Molecular Imaging, Institute of Automation, Beijing, 100190, China
b Department of Radiology, Peking Union Medical College Hospital, Beijing, 100006, China
c School of Life Science and Technology, Xidian University, Xi'an, 710126, China
d Centers for Biomedical Engineering, University of Science and Technology of China, Hefei, 230027, China
e University of Chinese Academy of Sciences, Beijing, 100049, China
f Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing,100050, China
g Neurosurgery, The Third Medical Center of Chinese PLA General Hospital, Beijing, 100039, China

ABSTRACT

Prostate specific antigen (PSA) screening is routinely conducted for suspected prostate cancer (PCa) patients. As this technique might result in high probability of over-diagnosis and unnecessary prostate biopsies, controversies on it remains especially for patients with "gray-zone" PSA levels, i.e. 4-10ng/ml. To improve the risk stratification of suspected PCa patients, Prostate Imaging Reporting and Data System version 2 (PI-RADSv2) was released in 2015. Although PI-RADSv2 showed good performance in the detection of PCa, its specificity was relatively low for patients with gray-zone PSA levels. This indicated that over-diagnosis issue could not be dealt well by PI-RADSv2 in the gray zone. Addressing this, we attempted to validate whether radiomics analysis of Diffusion weighted Imaging (DWI) data could reduce over-diagnosis of PCa with gray-zone PSA levels. Here, 140 suspected PCa patients in Peking Union Medical College Hospital were enrolled. 700 radiomic features were extracted from the DWI data. Least absolute shrinkage and selection operator (LASSO) were conducted, and 7 radiomic features were selected on the training set (n=93). Based on these features, random forest classifier was used to build the Radiomics model, which performed better than PI-RADSv2 (area under the curve [AUC]: 0.900 vs 0.773 and 0.844 vs 0.690 on the training and test sets). Furthermore, the specificity values of Radiomics model and PI-RADSv2 was 0.815 and 0.481 on the test set, respectively. In conclusion, radiomics analysis of DWI data might reduce the over-diagnosis of PCa with gray-zone PSA levels.

Keywords: Prostate specific antigen; Prostate cancer; Over-diagnosis; Radiomics; Random forest

+Contribute equally: Shuaitong Zhang, Yafei Qi, Jingwei Wei

*Corresponding authors: Jie Tian, E-mail: tian@ieee.org; Yali Zang, E-mail: yali.zang@ia.ac.cn.

Medical Imaging 2019: Computer-Aided Diagnosis, edited by Kensaku Mori, Horst K. Hahn, Proc. of SPIE Vol. 10950, 109501Z · © 2019 SPIE
CCC code: 1605-7422/19/$18 · doi: 10.1117/12.2511497

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 18 Mar 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

1. INTRODUCTION

Prostate cancer (PCa) in men was the most commonly diagnosed and a main cause of cancer deaths [1, 2]. Prostate specific antigen (PSA) screening is routinely conducted for suspected PCa patients, and when PSA levels of these patients are larger than 4 ng/ml, prostate biopsies are commonly performed [2, 3]. However, this screening technique remains controversial due to the high rate of over-diagnosis and unnecessary prostate biopsies, especially for patients with "gray-zone" PSA levels 4-10 ng/ml [3-5]. Thus, it was crucial to reduce the over-diagnosis of PCa with gray-zone PSA concentrations, which might improve the personalized management.

Currently, Prostate Imaging Reporting and Data System version 2 (PI-RADSv2) released in 2015 was used to evaluate the probability of patients with PCa based on MR images [6]. Several studies had validated the performance of PI-RADSv2 in detecting PCa among patients with all PSA levels [7-9]; however, only a few studies focused on the gray-zone PSA concentrations. In the work of Xu et al. [10], they regarded patients (PSA levels of 4-10 ng/ml) with PI-RADSv2 >= 3 as PCa using the Yonden index and found that the PI-RADSv2 yielded positive predictive and negative predictive values were 243/258 and 122/270. This suggested that 148 of 391 negative PCa patients were over-diagnosed. These results showed that the overdiagnosis among patients with gray-zone PSA levels could not be dealt well using PI-RADSv2, the performance of which were yet to be improved.

Radiomics analysis emerging in recent years provided a reasonable approach [11]. It quantified medical images using quantitative features and then associated these features with the diagnostic and prognostic issues using machine learning algorithms [12]. However, radiomics analysis in the diagnosis of suspected PCa patients with PSA levels of 4-10ng/ml is still lacking.

In this study, we aimed at identifying the negative PCa in gray-zone PSA levels using radiomics analysis based on DWI data. Furthermore, we also compared the performance of the radiomics model with that of PI-RADSv2 which was applied in routine clinic work.

2. MATERIALS AND METHODS

2.1 Patients and MR acquisition

This retrospective study was approved by our institute, and the informed consent was waived. In this study, all cases underwent PSA testing for suspected PCa from May 2015 to March 2018 were enrolled. Cases who met the following criteria were included: (a)PSA levels of 4-10ng/ml; (b) MR imaging performed less than two weeks before prostate biopsy; (c) MR images required on the same 3-T scanner; (d) DWI images are available. Cases who underwent hormone therapy or chemotherapy were excluded. A total of 140 patients were included and were randomly divided into the training set (n = 93) and the test set (n = 47) at a ratio of 2:1. The training set was used to build the predictive model for PCa, the performance of which was first assessed on the training set and then validated on the test set. The workflow of this study was shown in **Fig. 1**.

The patient characteristics on the training and test sets were shown in **Table 1**. PI-RADSv2 score for each patient was evaluated by a senior radiologist (Reader 1) who have twelve-year experience in interpretation of MRI of the prostate. All clinic-radiological factors showed no significant difference ($p > 0.4$) between the training and test sets, which justified their use as training and test sets.

The acquisition parameters of DWI data were as following: b values = 0 and 800s/mm^2; repetition time/echo time = 3000 ms/56 ms; slice thickness = 6 mm; and acquisition matrix = 128*128. The regions of interest (ROIs) were delineated manually on the slice where the area of the suspected PC was maximal on DWI images using itk-SNAP (www.itksnap.org). When delineating the ROIs on the axial sections by another radiologist (Reader 2) with five-year experience in interpretation of MRI of the prostate, polygon tool in itk-SNAP was used and the delineation process was under the supervision of Reader 1.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 18 Mar 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Table 1. The clinical and radiological characteristics of patients in the training and test sets

Characteristics	Training set	Test set	*P*-value
Age (years)	64.03 [10.09]	63.19 [7.39]	0.616
PSA levels (ng/ml)	6.91 [1.64]	7.10 [1.58]	0.520
PI-RADSv2 score			0.441
1-2/ 3/ 4-5	35/ 9/ 49	13/ 4/ 30	
Prostate cancer			0.959
Positive/ Negative	40/ 53	20/ 27	

Note: PSA, Prostate specific antigen; PI-RADSv2, Prostate Imaging Reporting and Data System version 2; Age and PI-RADSv2 score were shown as mean [standard deviance]. Pearson chi-square test and Student t test were used.

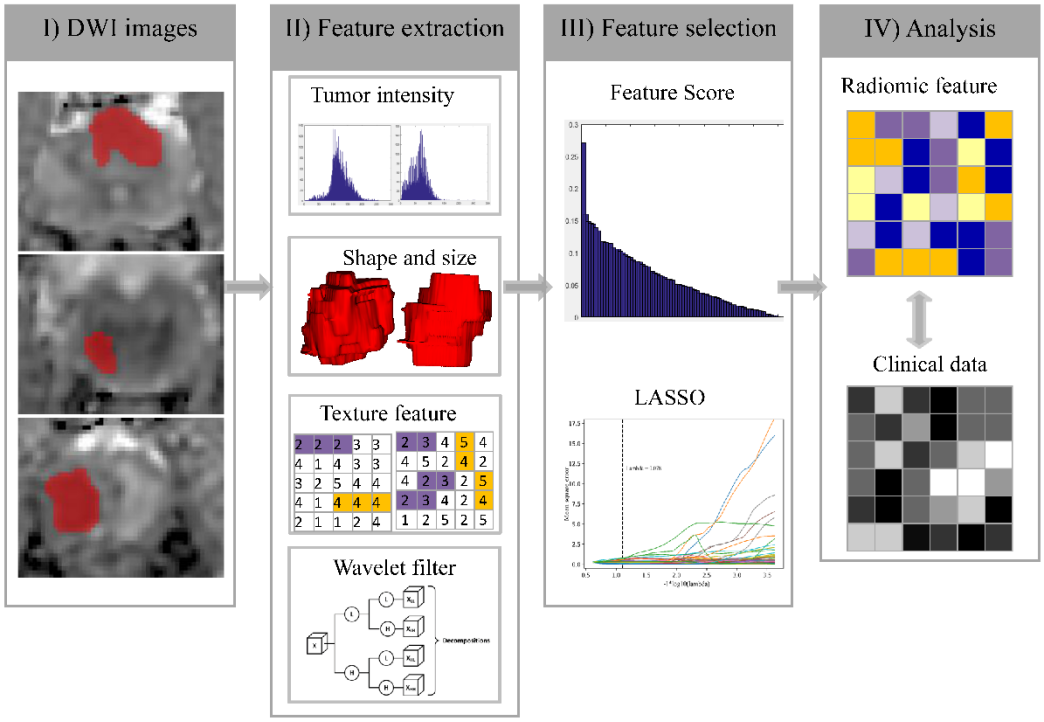

Figure 1. The workflow of this study.

2.2 Radiomic features extraction and selection

Package Pyradiomics version 2.0.1 [13] was used to extract radiomic features and 700 radiomic features were extracted automatically based on the delineated ROI for each case. These features could be included into five groups: (I) image intensity features (n = 18), (II) shape and size features (n = 12), (III) texture features (n = 68), and (IV) wavelet-based features (n = 344), and (V) Laplace of Gaussian (LoG) filter-based features (n = 258). Before feature selection, radiomic features were normalized using a z-score method. The feature selection was conducted on the training set. First, features with low variance (less than 0.1) were removed because these features contain relatively little information (i.e. The entropy of these features are small) and are of little predictive value [14, 15]. After that, two sample t test was performed to select features with potential predictive value, and features with p value larger than 0.05 were removed. Next, Redundant features were then removed based on the Pearson

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 18 Mar 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

correlation coefficient [16]. Features with Pearson correlation coefficients > 0.9 were grouped, and from each group, features with the smallest p value (i.e. the strongest predictive value) were selected. Finally, least absolute shrinkage and selection operator (LASSO) [17] was used to selected the final predictive features. LASSO is one of the well-known feature selection method which selects features by shrinking the coefficients of irrelevant features to zero. In this study, leave-one-out cross-validation was performed to select the best hyperparameter λ using mean square error criteria. The optimal subset of radiomic features selected on the training set were also applied on the test set for the following analysis.

2.3 Development and validation of Radiomics model

Random forest classification algorithm [18, 19] was used to build the PCa predictive model on the training set. The parameters of this algorithm were chosen based on the maximal out-of-bag score. The performance of the predictive model was first assessed on the training set and then validated on the test set using area under the curve (AUC) and receiver operation characteristics (ROC) curve [20]. Furthermore, classification accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were also used to assess the performance of the predictive model. Notably, specificity and NPV were very important. Relatively high specificity indicated the low probability of misclassifying nonPCa patient as PCa, which might avoid overdiagnosis and reduce unnecessary prostate biopsy; Relatively high NPV indicated the low probability of misclassifying PCa patients as nonPCa, which might avoid under-diagnosis.

3. RESULTS

Seven radiomic features with potential predictive value were selected using LASSO with LOOCV. Mean square errors on each cross-validation fold and the optimal regulation parameter were shown in **Fig. 2**. Based on these features, the Radiomics model yielded AUC values of 0.900 and 0.844 on the training and test sets. PI-RADSv2 score yielded AUC values of 0.773 and 0.690 on the training and test sets. DeLong test [21] showed that radiomics model performed significantly better than the PI-RADSv2 on the both training and test sets ($p < 0.001$ and $p < 0.05$). Accuracy, sensitivity, specificity, NPV, and PPV values for radiomics model and PI-RADSv2 were also displayed in **Table 2**. specificity and NPV values for radiomics model and PI-RADSv2 were 0.815 vs 0.481 and 0.765 vs 0.786. The ROC curves of the Radiomics model and PI-RADSv2 were shown in **Figure 3**.

Table 2. Performances of PI-RADSv2 and Radiomics model

Model	Performance	AUC	Accuracy	Sensitivity	Specificity	NPV	PPV
PI-RADSv2	Training set	0.773	0.774	0.850	0.717	0.864	0.694
	Test set	0.690	0.617	**0.800**	0.481	0.765	0.533
Radiomics	Training set	0.900	0.871	0.800	0.925	0.860	0.889
	Test set	**0.844**	**0.766**	0.700	**0.815**	**0.786**	**0.737**

Note: PI-RADSv2, Prostate Imaging Reporting and Data System version 2; AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value; The best performances on the test set were indicated as bold font.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 18 Mar 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

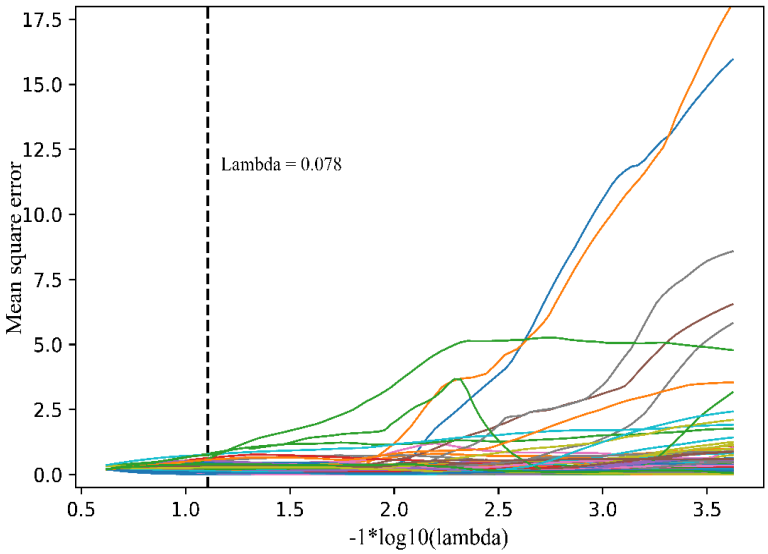

Figure 2. Mean square errors on each cross-validation fold. The optimal value of the regulation parameter 0.078 were determined based on LASSO with LOOCV.

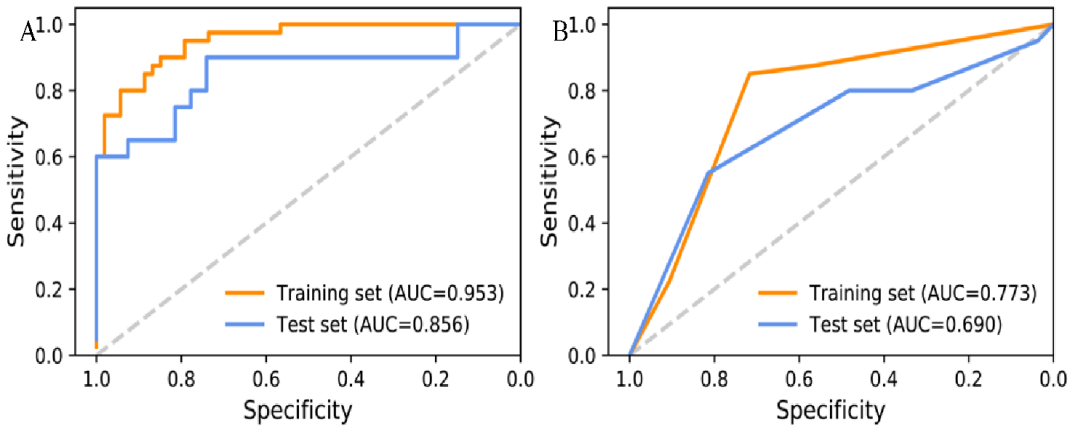

Figure 3. The ROC curves (A) Radiomics model and (B) PI-RADSv2 on the training and test sets.

New or breakthrough work to be presented

In this study, we applied radiomics analysis of DWI data to reduce the over-diagnosis of PCa with PSA concentrations of 4-10 ng/ml. Compared with PI-RADSv2, the Radiomics model performed significantly better (AUC: 0.900 vs 0.773 and 0.844 vs 0.690 on the training and test sets). It might indicate that radiomics analysis of DWI data might help to predict the probability of suspected PCa patients being positive PCa before biopsy and to improve the personalized management. Furthermore, the specificity of Radiomics model was higher than PI-RADSv2, which might indicate that radiomics model had a lower probability of misclassifying nonPCa patients as PCa and might reduce the over-diagnosis and avoid unnecessary biopsies.

CONCLUSION

In this study, we evaluated that whether radiomics analysis of DWI data could reduce the over-diagnosis of PCa patients with PSA levels of 4-10 ng/ml.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 18 Mar 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

[1] Siegel, R. L., Miller K. D., Jemal A., “Cancer statistics, 2018,” Ca. Cancer J. Clin. Papers 68(1), 7-30 (2018).
[2] Attard G., Parker C., Eeles R. A., Schröder, F., Tomlins S. A., Tannock I., et al., “Prostate cancer,” Lancet. Papers 387, 70-82 (2016).
[3] Murphy L., Prencipe M., Gallagher W. M., Watson R. M., “Commercialized biomarkers: new horizons in prostate cancer diagnostics,” Expert Rev. Mol. Diagn. Papers 15(4), 491-503 (2015).
[4] Carroll P. H., Mohler J. L., “NCCN guidelines updates: prostate cancer and prostate cancer early detection,” J. Natl. Compr. Canc. Netw. Papers 16(5S), 620-623 (2018).
[5] Catalona W. J., Richie J. P., deKernion J. B., et al, “Comparison of prostate specific antigen concentration versus prostate specific antigen density in the early detection of prostate cancer: receiver operating characteristic curves,” J. Urol. Papers 152(6), 2031–2036 (1994).
[6] Weinreb J. C., Barentsz J. O., Choyke P. L., Cornud F., Haider M.A., Macura K.J., et al, “PI-RADS prostate imaging—reporting and data system: 2015 Version 2,” Eur. Urol. Papers 69(1), 16-40 (2016).
[7] Wang R., Wang J., Gao G., Hu J., Jiang Y., Zhao Z., et al, “Prebiopsy mp-MRI can help to improve the predictive performance in prostate cancer: a prospective study in 1478 consecutive patients,” Clin. Cancer Res. Papers 23(14), 3692-3699 (2017).
[8] Choi M. H., Lee Y. J., Jung S. E., Rha S. E., Byun J. Y., “Prebiopsy biparametric MRI: differences of PI-RADS version 2 in patients with different PSA levels,” Clin. Radiol. Papers S0009-9260(18), 30184-3 (2018).
[9] Zhang Y., Zeng N., Zhu Y. C., Huang Y. X. R., Guo Q., Tian Y., “Development and internal validation of PI-RADs v2-based model for clinically significant prostate cancer,” World J. Surg. Oncol. Papers 16(1), 102 (2018).
[10] Xu N., Wu Y. P., Chen D. N., Ke Z. B., Cai. H., Wei H., et al, “Can Prostate Imaging Reporting and Data System Version 2 reduce unnecessary prostate biopsies in men with PSA levels of 4-10 ng/ml?”, J. Cancer Res. Clin. Oncol. Papers 144(5), 987-995 (2018).
[11] Lambin P., Rios-Velazquez E., Leijenaar R., Carvalho S., van Stiphout R. G., Granton P., et al, “Radiomics: extracting more information from medical images using adcanced feature analysis,” Eur. J. Cancer Papers 48(4), 441-446 (2012).
[12] Kumar V., Gu Y., Basu S., Berglund A., Eschrich S. A., Schabath M. B., et al, “Radiomics: the process and the challenges,” Magn. Reson. Imaging. Papers 30(9), 1234-1248 (2012).
[13] van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., et al, “Computational Radiomics System to Decode the Radiographic Phenotype”, Cancer Res. Papers 77(21), e104–e107 (2017).
[14] Wang H. F., Yoon S. W., “A machine learning model for medical image recognition using texture-based features”, Proc. IIE Annual Conference Institute of Industrial and Systems Engineers (IISE), 1655-1660 (2017).
[15] Huang Q. H., Zhang F., Li X. L., “A new breast tumor ultrasonography CAD system based on decision tree and BI-RADS features”, World Wide Web. Papers 21(6), 1491-1504 (2018).
[16] Wiedermann W., Hagmann M., “Asymmetric properties of the Pearson Correlation Coefficient: Correlation as the negative association between linear regression residuals”, Commun. Stat-Theor M. Papers 45(21):6263-6283 (2016).
[17] Tibshirani R, “Regression shrinkage and selection via the LASSO”, J. R. Statist. Soc. B Papers 58(1):267-288 (1996).
[18] Breiman L, “Random forests”, Mach. Learn. Papers 45(1):5-32 (2001).
[19] Mahesh P., “Random forest classifier for remote sensing classification”, Int. J. Remote Sens. Papers 26(1):217-222 (2005).
[20] Hanley J. A., McNeil B. J., “The meaning and use of the Area under a Receiver Operating Characteristic (ROC) curve”, Radiology Papers 143(1):26-39 (1982).
[21] DeLong E. R., DeLong D. M., Clarke-Pearson D. L., “Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach”, Biometrics Papers 44(3):837-845 (1988).

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 18 Mar 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use