


















ORIGINAL ARTICLE

Open Access



Combined radiomics, PI-RADS, and clinical model improve significant prostate cancer prediction and guide biopsy decision

Andreu Antolin^{1,2*} , Richard Mast³ , Nuria Roson⁴ , Javier Arce¹ , Ramon Almodovar¹ , Roger Cortada² , Anna Alberti¹ , Berta Miro⁵ , Olga Mendez⁶ , Almudena Maceda⁷ , Esther Serrano⁷ , Carmen Pietro-de-la-Lastra⁸ , Ana Jimenez-Pastor⁸ , Anna Nogué-Infante⁸ , Manuel Escobar⁹ , Enrique Trilla^{2,6,10}  and Juan Morote^{2,6,11} 

Abstract

Objectives The aim of this study was to develop and validate an MRI radiomics-based predictive model to discriminate significant prostate cancer (sPCa), compare it with PI-RADS, and determine whether incorporating PI-RADS and other clinical variables improves clinical performance.

Materials and methods A retrospective observational study was conducted using a cohort of 1497 MRI cases from 1395 men to develop the models. For each case, the index-lesion PI-RADS score, systematic \pm targeted biopsy results, and six additional clinical variables were collected. Prostate biopsy samples served as the reference standard, defining sPCa as Gleason Grade ≥ 7 . Handcrafted radiomic features were extracted from automatically segmented prostate glands. Four machine learning models were developed: (1) Radiomics, (2) PI-RADS, (3) PI-RADS + Radiomics, and (4) PI-RADS + Radiomics + Clinical Variables. Model performance and comparisons were evaluated using the area under the curve (AUC), while clinical utility was assessed through the decision curve analysis plot, Clinical Utility plot, and the number of avoided biopsies.

Results The radiomics model did not perform significantly better than PI-RADS in the validation cohort (AUC 0.838 vs. 0.833, $p = 0.874$). The combination of radiomics, PI-RADS, and clinical variables achieved the highest performance, with an AUC of 0.891 (95% CI: 0.853–0.930), significantly outperforming the other models ($p < 0.05$). It also showed the highest specificity (29.41%) and biopsy avoidance rate (18.15%), although the differences were not statistically significant ($p = 0.313$).

Conclusions Incorporating radiomics and clinical variables into PI-RADS enhances its ability to discriminate sPCa, potentially decreasing false positives and unnecessary biopsies.

Critical relevance statement The incorporation of clinical variables and radiomics into PI-RADS enhances its ability to predict significant prostate cancer, helping mitigate some of PI-RADS's current limitations, such as a significant false-positive rate, and might help reduce unnecessary biopsies.

Enrique Trilla and Juan Morote contributed equally to this work.

*Correspondence:

Andreu Antolin

antolin.andreu@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

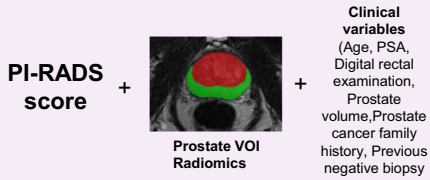
Key Points

- PI-RADS limitations result in overdiagnosis of indolent prostatic lesions and unnecessary biopsies.
- Radiomics and clinical variables enhance PI-RADS ability to detect significant prostate cancer.
- Combined clinical-radiological models reduce false positives and help avoid unnecessary biopsies.


Keywords Prostatic neoplasms, Magnetic resonance imaging, Radiomics, Models (Statistical)

Graphical Abstract

Combined radiomics, PI-RADS and clinical model improves significant prostate cancer prediction and guides biopsy decision

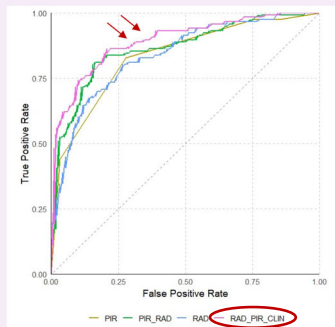



The addition of clinical variables and radiomic features into PI-RADS (RAD_PIR_CLIN) improves significant prostate cancer prediction and might help reduce the number of false positives and unnecessary prostate biopsies



	PIR	RAD	PIR-RAD	PIR-RAD-CLIN
Avoided biopsies, %	15.18	13.86	15.84	18.15
Undetected sPCa, %	2.59	2.59	2.59	2.59

PIR = PI-RADS model; PIR_RAD = radiomic combined with PI-RADS model; RAD = radiomic model; RAD_PIR_CLIN = combination of PI-RADS, radiomics and clinical variables model; sPCa = significant prostate cancer





Insights Imaging (2026) Antolin A, Mast R, Roson N et al;
DOI: 10.1186/s13244-026-02295-4

Introduction

Prostate cancer (PCa) is the second most common malignancy and the fifth leading cause of cancer death in men [1]. Based on the International Society of Urological Pathology (ISUP) Gleason Grading [2], PCa is classified as indolent PCa (iPCa; ISUP 1/Gleason 6) or significant (sPCa; ISUP >1/Gleason ≥ 7). iPCa is usually managed with active surveillance, while sPCa requires treatment due to a poorer prognosis.

Early detection of sPCa relies on multiparametric or biparametric prostate magnetic resonance imaging (mpMRI/bpMRI) in men with a serum prostate-specific antigen (PSA) > 3.0 ng/mL and/or abnormal digital rectal examination (DRE). Suspicious lesions are identified through the Prostate Imaging-Reporting and Data System (PI-RADS), currently in version 2.1, which provides a semiquantitative risk assessment of sPCa [3]. PI-RADS ≥ 3 lesions undergo targeted biopsies, complemented with a systematic prostate biopsy [4].

PI-RADS has limitations [5], including significant inter-reader variability [6, 7]. Its positive predictive value for PI-RADS 3 is 20%, leading to unnecessary biopsies and overdiagnosis of iPCa [8]. New predictive biomarkers are needed to reduce false positives [9].

Radiomics have shown promising results, but published studies have methodological limitations in data selection, model development, and validation [10–15]. Therefore, new radiomic models should be developed, compared with PI-RADS, and integrated with clinical variables to evaluate their added predictive value.

This study aims to: (1) develop and validate a radiomics-based predictive model for sPCa using MRI prostate segmentations in men with clinical suspicion of PCa; (2) compare it with PI-RADS, and develop/validate a combined model; (3) build and validate a multivariable model incorporating PI-RADS, radiomic features, and clinical variables, to assess improvements over previous models

Materials and methods

Study design

An observational retrospective study was conducted to develop, validate and compare different predictive models (see Model Development subsection) according to the study aims, following the Standards for the Reporting of Studies of Diagnostic Accuracy (STARD) [16].

Inclusion criteria were: (1) men with suspected PCa defined as serum PSA > 3 ng/mL and/or abnormal DRE; (2) prostate bpMRI or mpMRI reported following PI-RADS v2 [17] or v2.1 [3]; (3) prostate biopsy performed according to the standard protocol, consisting of targeted biopsy of suspicious lesions (PI-RADS \geq 3) complemented by systematic biopsies, or systematic biopsies alone in PI-RADS < 3 but high clinical suspicion of sPCa; (4) availability of clinical variables (see Clinical Variables subsection).

Exclusion criteria were: (1) prior diagnosis of PCa and/or men in active surveillance; (2) prostate biopsies before MRI; (3) prostate biopsies performed later than 3 months after MRI; (4) missing information regarding MRI, clinical variables or prostate biopsies; (5) MRI artifacts that prevent an accurate reading.

1497 MRIs from 1395 men with suspected PCa were consecutively selected according to the inclusion/exclusion criteria from November 2015 to November 2022 at Vall Hebron Hospital, Barcelona, Spain. MRIs were reviewed by an abdominal radiologist with 5 years of experience, supervised by a senior radiologist with > 10 years of experience in prostate imaging.

The study was conducted in accordance with Checklist for Artificial Intelligence in Medical Imaging (CLAIM) guidelines [18].

Ethical issues

The study was approved by the institutional Ethics Committee (PR(AG)02/2021). Data were collected and pseudonymized in accordance with ethical committee approval, using a unique code per case. The Ethics Committee granted a waiver of written informed consent, as participants were verbally informed at the time of specimen collection about the research use of their data. A Data Protection Impact Assessment was completed and reviewed by the committee.

Image acquisition protocol and image pre-processing

MRI images were acquired according to PI-RADS and European Society of Urogenital Radiology guidelines [3, 17, 19]. Studies were performed in 1.5-T (Siemens MAGNETOM *Symphony*, Siemens Healthineers) and 3-T (Siemens MAGNETOM *Trio*, Siemens Healthineers) scanners. Biparametric studies consisted of T2, diffusion-weighted imaging (DWI) and apparent diffusion

coefficient (ADC) sequences, while multiparametric studies included an additional dynamic contrast-enhanced (DCE) sequence. Between late 2018 and early 2019, mpMRI was progressively replaced by bpMRI for early diagnosis of PCa due to better cost-efficiency, absence of intravenous contrast, and comparable performance [20, 21]. Acquisition parameters are provided in Supplementary Tables S1 and S2. A phased array body coil was used in all.

Axial T2 and DWI sequences were selected, while DCE sequences were excluded since there is no clear evidence that quantitative features extracted from this sequence have added value in radiomic models [10]. Images were analyzed with QP-Prostate[®] (Quibim S.L.), which is a medical device software (SaMD) with CE (*Conformité Européenne*), UKCA (United Kingdom Conformity Assessed) and FDA 510k (Food and Drug Administration) certification. This software automatically segments peripheral and transitional (including the central zone) zones and seminal vesicles. It also provides the required image pre-processing steps prior to feature extraction to ensure reproducibility across patients (see Supplementary Material) and generates a new ADC map and DWI b-1400 sequence.

Extraction of radiomic features

Handcrafted radiomic features were extracted using the QP-Insights[®] platform (Quibim S.L.) from the peripheral and transitional volume of interest (VOI) masks from T2 and generated DWI b1400/ADC maps. Seminal vesicle masks were discarded since prostate carcinoma is not generated in this region. 1379 radiomic features were obtained from each mask (see Supplementary Material). Therefore, 8274 characteristics were obtained per case, corresponding to 1379 features extracted from each VOI across the three sequences.

Clinical variables

Seven clinical variables were collected according to the Barcelona Predictive Model of sPCa [22], which were obtained from the patient's medical records. These were: age (years), serum PSA (ng/mL), prostate volume (mL, based on MRI), PI-RADS index lesion score (1–5), DRE (abnormal or normal), previous negative biopsy (yes/no) and previous family history of prostate cancer (yes/no). Clinical variables were preprocessed in accordance with the requirements of the algorithms used for model development. There were no missing values in the dataset.

Reference standard

The presence or absence of sPCa in the prostatic biopsy samples was used as the outcome variable for this binary classification task. It was defined as ISUP > 1 (Gleason

score ≥ 7), while non-sPCa included ISUP 1 (Gleason score 6) and benign lesions [2].

Prostate biopsy protocol

Men with ≥ 1 suspicious MRI lesion (PI-RADS ≥ 3) underwent 2 to 4-core targeted biopsy per suspicious lesion, complemented with 12-core systematic biopsy. PI-RADS < 3 cases underwent only 12-core systematic biopsy. Targeted biopsies were performed using cognitive fusion with a transrectal ultrasound (TRUS) approach between 2015 and 2020, and software fusion with transperineal ultrasound (TPUS) from 2021 onward. All procedures were carried out by a single urologist with > 10 years of experience in prostate biopsy.

Tissue samples were examined by a dedicated genitourinary pathologist with > 15 years of experience.

Model development

Four classification models were developed to assess the aim of the study: (1) PI-RADS model (PIR); (2) radiomics model (RAD); (3) radiomics + PI-RADS model (PIR_RAD); and (4) radiomics + PI-RADS + clinical variables model (RAD_PIR_CLIN). Model development is detailed in Supplementary Material.

The final models were also analyzed using SHAP (SHapley Additive exPlanations) to assess the contribution of individual features to the predictions [23], thereby enhancing model interpretability and explainability.

Statistical analysis

Quantitative variables were summarized using mean and standard deviation (SD), as well as median and interquartile range (IQR). Categorical variables were reported as absolute and relative frequencies. Mean comparisons between partitions were performed using the independent samples of Student's t-test when normality was met, or the Wilcoxon rank-sum test otherwise. Statistical differences in categorical variables between distributions were assessed using the chi-squared test. Discrimination of the four predictive models was evaluated based on the area under the Receiver Operating Characteristic (ROC) curve (AUC) with 95% confidence intervals (CIs). Additionally, the following metrics were calculated using a fixed probability threshold of 0.5: (1) accuracy; (2) balanced accuracy; (3) precision; and (4) F1-score. AUCs were compared using DeLong's test. Decision curve analysis (DCA) was performed to evaluate net benefit, and clinical effectiveness was assessed through the clinical utility curve (CUC), representing the rates of avoided prostate biopsies and the undetected rate of sPCa according to the continuous probability threshold.

To ensure comparability, all models were evaluated at a fixed sensitivity of 97.4%, corresponding to the PI-

RADS ≥ 3 threshold. For each model, the predicted probability threshold was adjusted accordingly, and biopsy avoidance was defined as the proportion of patients without sPCa whose predicted probability fell below this threshold. To statistically compare model performance against the PI-RADS baseline in terms of specificity and biopsy avoidance, McNemar's test was applied. The analysis was restricted to patients without sPCa to assess whether the specificity improvements observed with the alternative models were statistically significant.

All statistical tests were two-sided, and a p -value < 0.05 was considered indicative of statistical significance.

Statistical analysis was done with R software version 4.5.0 (R Foundation for Statistical Computing).

Results

Clinical characteristics of the cohorts

1497 MRI cases from 1395 men with suspected PCa were consecutively selected according to the inclusion/exclusion criteria. The cohort was randomly split into a training cohort (80%) and a test cohort (20%). The clinical characteristics of the overall cohort and each partition are summarized in Table 1. The prevalence of sPCa was 39.2%. PSA had an SD of 130.4 ng/mL since the minimum and maximum values were 0.24 and 4414.4 ng/mL, respectively. 77.7% were biopsy naive, while 22.3% had a previous negative prostate biopsy. No statistically significant differences were observed between the training and test cohorts in any of the assessed variables, including sPCa detection rate and PI-RADS score of the index lesion.

The CLAIM checklist is depicted in Supplementary Material Table S3.

Model performance in discriminating sPCa

Random Forest was the algorithm that yielded the best overall performance among all classifiers tested. The AUC values for each Random Forest-based model, as well as the PI-RADS logistic regression model, are summarized in Table 2, which also includes p -values from pairwise comparisons between each model and the subsequent one. The ROC curve of each model is plotted in Fig. 1.

The PIR model achieved an AUC of 0.833 (95% CI: 0.788–0.878), while the RAD model reached an AUC of 0.838 (95% CI: 0.792–0.884), with no statistically significant difference between them in sPCa detection ($p = 0.874$). The combination of PI-RADS and radiomic features (PIR_RAD) showed a higher AUC of 0.863 (95% CI: 0.818–0.907) but was not significantly superior over both individual models ($p = 0.05$ and $p = 0.874$ for PIR and RAD, respectively). Finally, the addition of clinical variables to the combined PI-RADS and radiomics model (RAD_PIR_CLIN) resulted in an AUC of 0.891 (95% CI: 0.853–0.930), which was

Table 1 Clinical characteristics of the entire cohort and the respective training and test partitions

Clinical characteristics	Entire cohort	Training cohort	Test cohort	<i>p</i> -value
Number of cases	1497	1194	303	-
Mean age, years (SD)	68.5 (8.4)	68.5 (8.3)	68.5 (8.7)	0.719
Median age, years [IQR]	69.1 (62.8, 74.1)	69 (62.8, 74.2)	69.1 (63, 74)	
Mean serum PSA, ng/mL (SD)	13.7 (116.6)	14.7 (130.4)	9.6 (14.9)	0.382
Median serum PSA, ng/mL (IQR)	6.1 (4.3, 9.8)	6.1 (4.3, 9.9)	5.8 (4.3, 9.5)	
Abnormal DRE, <i>n</i> (%)	313 (20.8)	256 (21.4)	56 (18.5)	0.292
PCa family history, <i>n</i> (%)	103 (6.9)	86 (7.2)	17 (5.6)	0.395
Previous negative prostate biopsy, <i>n</i> (%)	334 (22.3)	268 (22.4)	66 (21.8)	0.865
Mean prostate volume, mL (SD)	61.7 (31.5)	61.7 (31.2)	61.5 (32.6)	0.617
Median prostate volume, mL (IQR)	55 (40, 76)	55 (40, 76)	54 (40, 76)	
PI-RADS score of index lesion, <i>n</i> (%)				0.293
1	131 (8.8)	100 (8.4)	31 (10.2)	
2	61 (4.1)	43 (3.6)	18 (5.9)	
3	526 (35.1)	420 (35.2)	106 (35)	
4	477 (31.9)	386 (32.3)	91 (30)	
5	302 (20.2)	245 (20.5)	57 (18.8)	
sPCa detection rate, <i>n</i> (%)	577 (39.2)	471 (39.4)	116 (38.3)	0.761

Data of quantitative variables is represented in mean and standard deviation (SD) at the upper part of the row, as well as median and interquartile range (IQR) at the bottom. Data of categorical variables is represented in absolute frequency (*n*) and relative frequency in %

DRE digital rectal examination, IQR interquartile range, *n* absolute frequency, PCa prostate cancer, PI-RADS Prostate Imaging-Reporting and Data System, PSA prostate-specific antigen, SD standard deviation, sPCa significant prostate cancer

Table 2 Models' performance with pairwise comparisons between each model and the subsequent one

Classification model	AUC (95%, CI)	<i>p</i> -value	Holm adjusted <i>p</i> -value
PI-RADS (PIR)	0.833 (0.788–0.878)	-	-
Radiomics (RAD)	0.838 (0.792–0.884)	0.874	0.874
PI-RADS–Radiomics (PIR_RAD)	0.863 (0.818–0.907)	0.236	0.472
Radiomics–PI-RADS–Clinical Variables (RAD_PIR_CLIN)	0.891 (0.853–0.930)	0.003	0.012

AUC area under the curve, CI confidence interval

significantly higher than the PIR, RAD and PIR_RAD models ($p < 0.05$ for each comparison).

The accuracy, balanced accuracy, precision, and F1-score for RAD, PIR_RAD and RAD_PIR_CLIN, considering a threshold of 0.5, are presented in Table S7.

Net benefit

The DCA of the four models is plotted in Fig. 2. The PIR_RAD model and the addition of clinical variables (PIR_RAD_CLIN) showed the highest net clinical benefit in the widest range of threshold probabilities (approximately 30–60%). The PIR and RAD model showed a lower net clinical benefit.

Clinical efficacy

The sensitivity of all models was fixed to match that of the PI-RADS ≥ 3 reference threshold, which in this dataset was

97.40%. This ensures a fair comparison in terms of specificity and biopsy avoidance, consistent with current clinical practice that recommends targeted biopsy for lesions scored PI-RADS ≥ 3 . Table 3 summarizes the resulting diagnostic metrics for each model based on this supposition. The PIR model achieved a specificity of 24.60% and a biopsy avoidance rate of 15.18%. The RAD model showed slightly lower specificity (22.46%) and biopsy avoidance (13.86%), with no statistically significant difference compared to PIR ($p = 0.711$). The combined PIR_RAD model improved the specificity (25.67%) and biopsy avoidance (15.84%) over both standalone models but did not achieve statistical significance in comparison to PIR ($p = 0.896$). The addition of clinical variables further improved specificity and biopsy avoidance in comparison to the previous models (29.41% and 18.15%), but the difference was not statistically significant ($p = 0.313$). Importantly, the proportion of

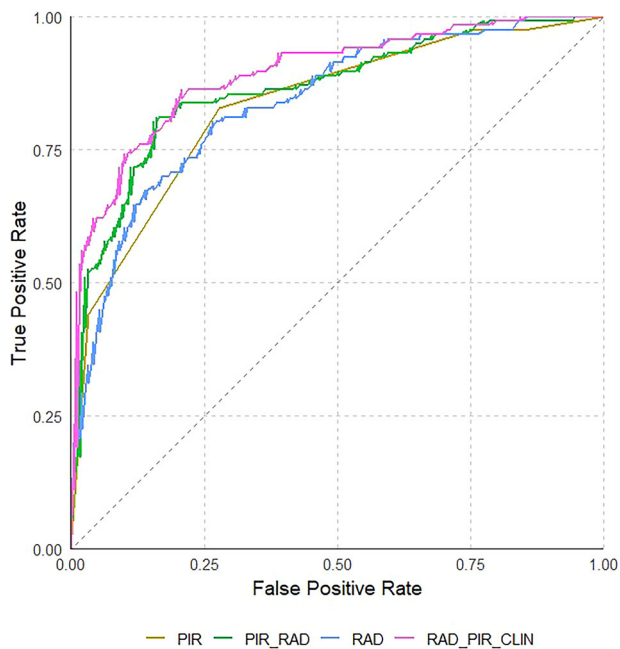


Fig. 1 Receiver operating characteristics (ROC) curves showing the discrimination ability for significant prostate cancer (sPCa) of the PI-RADS model (PIR), radiomics model (RAD), PI-RADS + radiomics model (PIR_RAD), and radiomics + PI-RADS + clinical variables model (RAD_PIR_CLIN)

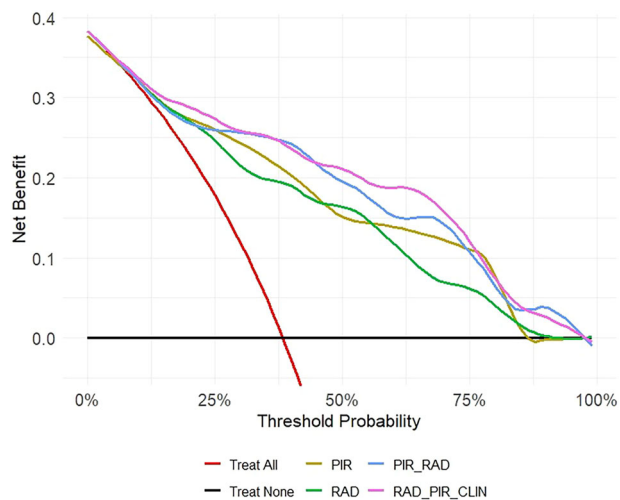


Fig. 2 Decision curve analysis (DCA) showing the net benefit of the PI-RADS model (PIR), radiomic model (RAD), PI-RADS + radiomic model (PIR_RAD), and the radiomic + PI-RADS + clinical variables model (RAD_PIR_CLIN) compared to performing biopsy on all or none men with prostate cancer suspicion

undetected sPCa remained constant across all models (2.59%) due to the fixed sensitivity.

The CUC for the four models is presented in Fig. 3, showing the balance between the proportion of biopsies

Table 3 Comparison of model clinical efficacy at fixed sensitivity, following recommendations to avoid biopsies for PI-RADS < 3

	PIR	RAD	PIR-RAD	PIR-RAD-CLIN
Threshold, %	10.90	13.82	8.50	10.45
True positive, <i>n</i>	113	113	113	113
False negatives, <i>n</i>	3	3	3	3
True negative, <i>n</i>	46	42	48	55
False positives, <i>n</i>	141	145	139	132
Sensitivity, %	97.41	97.41	97.41	97.41
Specificity, %	24.60	22.46	25.67	29.41
Youden Index, %	22.01	19.87	23.08	26.83
PPV, %	44.49	43.80	44.84	46.12
NPV, %	93.89	93.33	94.11	94.83
Avoided biopsies, %	15.18	13.86	15.84	18.15
Undetected sPCa, %	2.59	2.59	2.59	2.59
<i>p</i> -value	-	0.711	0.896	0.314

NPV negative predictive value, PIR PI-RADS model, PIR_RAD radiomic combined with PI-RADS model, PPV positive predictive value, RAD radiomic model, RAD_PIR_CLIN combination of PI-RADS, radiomics and clinical variables model, sPCa significant prostate cancer

that could be avoided and the proportion of undetected sPCa across a range of threshold probabilities for performing targeted biopsy. Like the DCA results, the PIR_RAD model and the addition of clinical variables (RAD_PIR_CLIN) allowed for a greater number of biopsies to be safely avoided while maintaining similar or lower rates of undetected sPCa compared to the other two models.

Explainability

SHAP analysis was applied to the RAD_PIR_CLIN model, which combined 25 variables to predict the presence or absence of sPCa (Fig. 4).

PI-RADS score was the most influential predictor in the model, followed by patient age. Higher values of these two variables were associated with an increased risk of sPCa. Prostate volume showed an inverse correlation, with larger volumes associated with a lower chance of sPCa. Radiomic features showed variable effects depending on their values. Those derived from DWI b-1400 and ADC sequences showed greater and more consistent positive impact on model predictions of sPCa.

False positives were associated with higher PI-RADS scores, suggesting that imaging-based suspicion is important in model prediction, in agreement with the SHAP analysis. This dependence suggests that model performance may be influenced by reader expertise, potentially performing better in experienced centers.

False negatives were more frequent in patients with larger prostate volumes, which may reflect reduced lesion

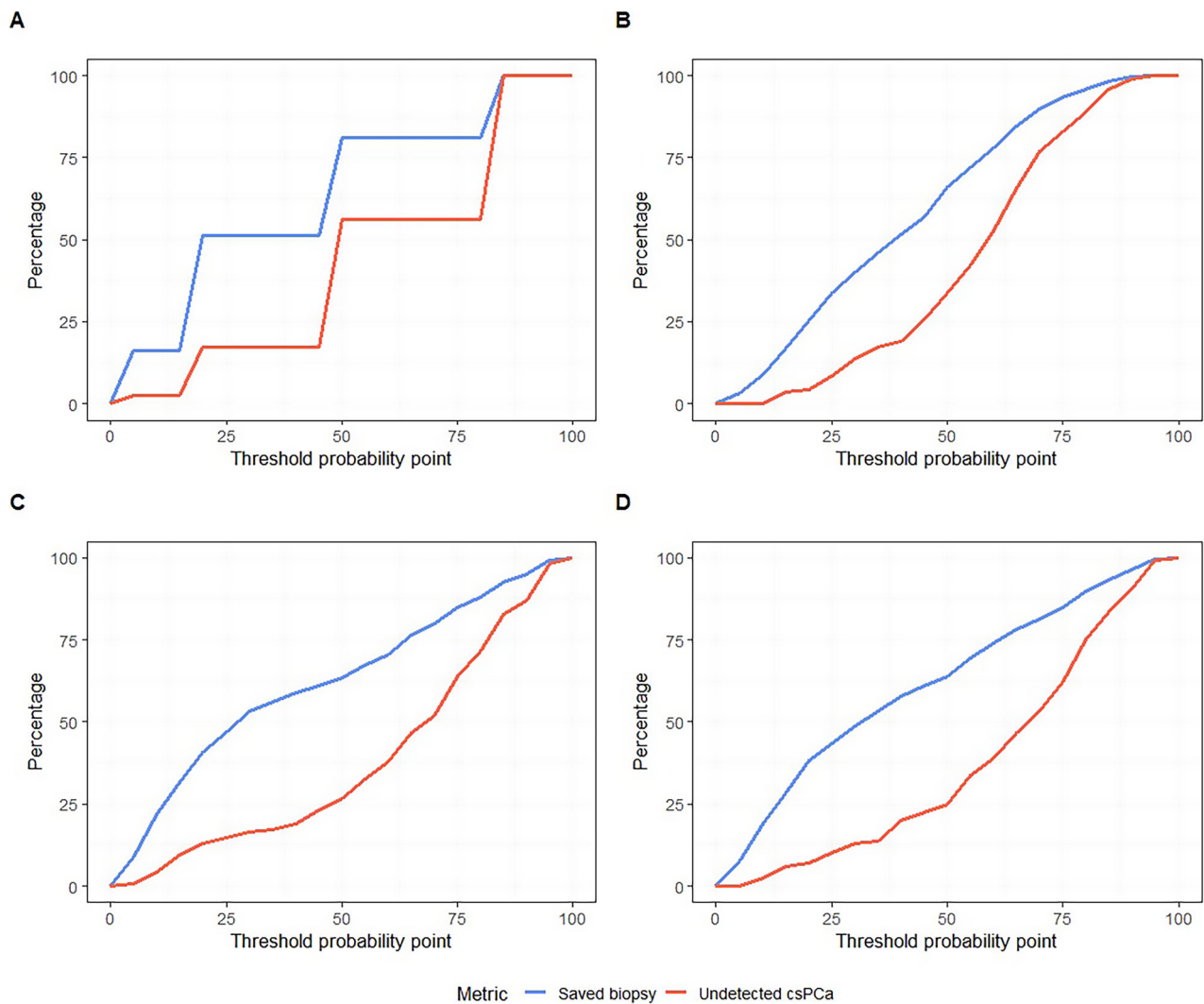


Fig. 3 Clinical utility curve (CUC) showing the proportion of biopsies that could be avoided and the proportion of undetected sPCa across a range of threshold probabilities for performing targeted biopsy, in the PI-RADS model (A), radiomic model (B), PI-RADS + radiomic model (C), and PI-RADS + radiomic + clinical variables (D)

conspicuity in larger glands. This could affect PI-RADS assessment and potentially influence radiomic feature extraction.

Discussion

A handcrafted radiomic model based on prostate segmentations was developed to discriminate sPCa in bpMRI from men with suspected PCa. The radiomic and PI-RADS models achieved moderate and comparable performance in discriminating sPCa. Combining PI-RADS and radiomic features yielded slightly higher performance but was not significantly superior to either model alone. Adding clinical variables further improved discrimination and achieved statistical significance when compared to the other models.

Our results are consistent with the published literature, which shows considerable variability, as noted in previous systematic reviews and meta-analyses [10–12, 15]. Methodological limitations, including small cohorts, non-uniform reference standards, manual segmentation variability, and the lack of multimodal models, contribute to this variability and hinder the true potential of radiomics. In this study, we mitigated some of these limitations by using a large cohort, a clear biopsy-based reference standard, automatic prostate segmentation, and multimodal models integrating clinical features to improve sPCa discrimination.

Clinical implementation of radiomic models is partially halted by the lack of evidence comparing these models



Fig. 4 SHapley Additive exPlanations (SHAP) summary plot showing the top 25 features contributing to significant prostate cancer (sPCa) prediction in the PI-RADS + radiomics + clinical variables model (RAD_PIR_CLIN)

with PI-RADS [15]. In our study, no significant differences were observed between the radiomic and PI-RADS models, consistent with previous reviews [15].

However, differences might emerge when radiologist experience is considered. Bao et al [24] reported better performance in junior radiologists who used artificial intelligence (AI), but not clearly in senior/expert readers. Similarly, Hamm et al [25] found improved performance among nonexperts, particularly in PI-RADS 3 lesions. Other studies reported no differences when compared to expert radiologists [26]. In contrast, a recent multicenter study including over 9000 patients and more than 60 radiologists (with a median of 7 years of experience) reported significantly better performance of the AI model compared to radiologists, although the model incorporated clinical metadata [27]. However, in a multidisciplinary setting where radiologists had access to full medical records, the AI model was neither significantly superior nor inferior to radiologists [27]. Prospective studies are needed to clarify the potential of radiomic models.

The combination of radiomic features and PI-RADS yielded better but non-significant discriminating capacity

than the models alone, consistent with available literature [15, 28]. The addition of clinical variables significantly improved performance in comparison to the rest of the models. This might support the integration of these models as a complementary tool to aid the radiologist, which might help overcome some of the limitations of PI-RADS v2.1 [5] and inherent inter-reader variability in MRI interpretation [6, 7].

Age, DRE, PSA, prostate volume, and biopsy status (biopsy-naïve vs. previously biopsied) are among the clinical variables commonly incorporated into sPCa risk models [22, 29–31]. Prior studies have shown that adding these variables to PI-RADS enhances the detection of sPCa and helps reduce unnecessary biopsies [22]. The combined PI-RADS and radiomics model, as well as the inclusion of clinical variables, demonstrated a higher net clinical benefit across a wide range of threshold probabilities compared to standalone models, in line with previous findings [24, 32]. This improvement correlates with a slightly higher specificity and biopsy avoidance observed in the combined models, when evaluated at a fixed sensitivity. However, the differences were not statistically significant. A recent study reported that the

integration of deep radiomics, PI-RADS, and clinical variables improved patient risk stratification prior to biopsy and led to a reduction in unnecessary procedures [32]. Further studies with larger cohorts or alternative methodological approaches are needed to fully assess this potential, as overdiagnosis of non-sPCa lesions remains a challenge [9].

There were several limitations in this study. First, a retrospective single-center and single-vendor cohort was used, which might limit model generalizability, although some studies reported no significant differences based on magnetic field strength or MRI vendor [33, 34]. Second, PI-RADS scoring was based on both versions 2.0 and 2.1, which may introduce bias. However, it remains unclear whether these differences have a significant clinical impact [35, 36]. Third, deep radiomics and lesion-mask feature extraction were not tested. Studies using whole-prostate segmentations on large datasets have shown comparable results to those based on lesion-specific segmentations [27, 32]. These studies were based on deep radiomics, so the use of handcrafted radiomics might limit radiomics' performance. This is consistent with past reviews that favored deep learning techniques [12]. We believe deep radiomics models will continually outperform handcrafted ones with increasingly larger datasets. Lastly, no external validation was performed due to the difficulty in obtaining the clinical variables in external cohorts.

In conclusion, the addition of radiomics and clinical variables to PI-RADS improves its discriminatory capacity for detecting sPCa, potentially reducing false positives and unnecessary biopsies in men with suspected PCa. However, further methodological refinement and prospective validation studies are necessary to fully establish the clinical impact and support the integration of such models into routine clinical practice.

Abbreviations

ADC	Apparent diffusion coefficient
AI	Artificial intelligence
AUC	Area under the curve
bpMRI	Biparametric prostate MRI
CI	Confidence intervals
CUC	Clinical utility curve
DCA	Decision curve analysis
DCE	Dynamic contrast-enhanced
DRE	Digital rectal examination
DWI	Diffusion-weighted imaging
iPCa	Indolent prostate cancer
IQR	Interquartile range
ISUP	International Society of Urological Pathology
mpMRI	Multiparametric prostate MRI
MRI	Magnetic resonance imaging
PCa	Prostate cancer
PIR	PI-RADS model
PIR_RAD	Radiomics + PI-RADS model
PI-RADS	Prostate Imaging-Reporting and Data System

PSA	Prostate-specific antigen
RAD	Radiomics model
RAD_PIR_	Radiomics + PI-RADS + Clinical Variables model
CLIN	
SD	Standard deviation
SHAP	SHapley Additive exPlanations
sPCa	Significant prostate cancer
VOI	Volume of interest

Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1186/s13244-026-02295-4>.

ELECTRONIC SUPPLEMENTARY MATERIAL

Acknowledgements

The final cohort included part of the patients that were used for creating a clinical predictive model for significant prostate cancer published in "Morote J, Borque-Fernando A, Triquell M, et al (2022) The Barcelona Predictive Model of Clinically Significant Prostate Cancer. *Cancers (Basel)* 14(6):1589. <https://doi.org/10.3390/cancers14061589>", in which some of the authors already participated. The new cohort includes some of those patients, as well as others recruited afterward. This new cohort was used to improve the predictive model by incorporating radiomics.

Author contributions

Study conception and design: A.A., J.M. Acquisition of data: A.A., J.M., B.M., O.M., R.M., J.A., R.A., R.C., A. Alberti, A.M., E.S. Analysis and interpretation of data: A.A., J.M., B.M., C.P.L., A.J.P., A.N.I. Drafting of manuscript: A.A., J.M. Critical revision: all authors.

Funding

This research was funded by the European project FLUTE, which has received funding from HORIZON-HLTH2022-IND-13 action under the Horizon Europe Framework with grant agreement Nr.101095382. This research was funded by the *Ministerio de Asuntos Económicos y Transformación Digital* (SP) (MIA.2021.M02.0005) and the *Instituto de Salud Carlos III* (SP) through the project "PI20/01666" (co-funded by European Regional Development Fund "A way to make Europe"). It was also partially funded by the *Instituto de Salud Carlos III* through the project "PI23/01310" to Olga Mendez. HORIZON-HLTH2022-IND-13 (European Project FLUTE) enabled the acquisition, extraction, and preprocessing of the data used in this study. PI20/01666 supported the development, implementation, and evaluation of the computational models presented in this article.

Data availability

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

Declarations

Ethics approval and consent to participate

The study was approved by the institutional Ethics Committee (PR(AG)02/2021). Data were collected and pseudonymized in accordance with the ethical committee's approval, with a single code per case. Data Protection Impact Assessment was completed and reviewed by the committee.

Consent for publication

The Ethics Committee granted a waiver of written informed consent, as participants were verbally informed at the time of specimen collection about the research use of their data.

Competing interests

C.P.L., A.J.P., and A.N.I. are employed by Quibim S.L. A.J.P. is a member of the Scientific Editorial Board of *Insights into Imaging* and, as such, did not participate in the selection or review processes for this article.

Author details

¹Department of Radiology, Institut de Diagnòstic per la Imatge (IDI), Hospital Universitari Vall d'Hebron Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain. ²Department of Surgery, Universitat Autònoma de Barcelona (UAB) Plaça Cívica, 08193 Bellaterra, Barcelona, Spain. ³Department of Radiology, Hospital Universitari Vall d'Hebron Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain. ⁴Head of Radiology Department, Hospital Universitari Vall d'Hebron Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain. ⁵Statistics and Bioinformatics Department, Vall d'Hebron Research Institute (VHIR) Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain. ⁶Group of Biomedical Research in Urology, Vall Hebron Research Institute (VHIR) Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain. ⁷Vall d'Hebron Research Institute (VHIR) Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain. ⁸Quantitative Imaging Biomarkers in Medicine, Quibim S.L. Avenida Aragón 30 13th floor, office I–J, 46021 Valencia, Spain. ⁹Clinical Director of Radiology Department, Hospital Universitari Vall d'Hebron Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain. ¹⁰Head of Department of Urology, Hospital Universitari Vall d'Hebron Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain. ¹¹Department of Urology, Hospital Universitari Vall d'Hebron Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain.

Received: 4 December 2025 Revised: 14 April 2026 Accepted: 19 April 2026
Published online: 25 April 2026

References

- Bray F, Laversanne M, Sung H et al (2024) Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 74:229–263
- Epstein JI, Egevad L, Amin MB et al (2016) The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol* 40:244–252
- Turkbey B, Rosenkrantz AB, Haider MA et al (2019) Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. *Eur Urol* 76:340–351
- Cornford P, van den Bergh RCN, Briers E et al (2024) EAU-EANM-ESTRO-ESUR-ISUP-SIOG guidelines on prostate cancer—2024 update. Part I: screening, diagnosis, and local treatment with curative intent. *Eur Urol* 86:148–163
- Puryško AS, Baroni RH, Giganti F et al (2021) PI-RADS Version 2.1: a critical review, from the AJR Special Series on Radiology Reporting and Data Systems. *AJR Am J Roentgenol* 216:20–32
- Bhayana R, O'Shea A, Anderson MA et al (2021) PI-RADS Versions 2 and 2.1: interobserver agreement and diagnostic performance in peripheral and transition zone lesions among six radiologists. *AJR Am J Roentgenol* 217:141–151
- Smith CP, Harmon SA, Barrett T et al (2019) Intra- and interreader reproducibility of PI-RADS v2: a multireader study. *J Magn Reson Imaging* 49:1694–1703
- Oerther B, Engel H, Bamberg F, Sigle A, Gratzke C, Benndorf M (2022) Cancer detection rates of the PI-RADSV2.1 assessment categories: systematic review and meta-analysis on lesion level and patient level. *Prostate Cancer Prostatic Dis* 25:256–263
- Osses DF, Roobol MJ, Schoots IG (2019) Prediction medicine: biomarkers, risk calculators and magnetic resonance imaging as risk stratification tools in prostate cancer diagnosis. *Int J Mol Sci* 20:1637
- Castillo TJM, Arif M, Niessen WJ, Schoots IG, Veenland JF (2020) Automated classification of significant prostate cancer on MRI: a systematic review on the performance of machine learning applications. *Cancers (Basel)* 12:1606
- Cuocolo R, Cipullo MB, Stanzione A et al (2020) Machine learning for the identification of clinically significant prostate cancer on MRI: a meta-analysis. *Eur Radiol* 30:6877–6887
- Sushentsev N, Moreira Da Silva N, Yeung M et al (2022) Comparative performance of fully-automated and semi-automated artificial intelligence methods for the detection of clinically significant prostate cancer on MRI: a systematic review. *Insights Imaging* 13:59
- Syer T, Mehta P, Antonelli M et al (2021) Artificial intelligence compared to radiologists for the initial diagnosis of prostate cancer on magnetic resonance imaging: a systematic review and recommendations for future studies. *Cancers (Basel)* 13:3318
- Twilt JJ, van Leeuwen KG, Huisman HJ, Fütterer JJ, de Rooij M (2021) Artificial intelligence based algorithms for prostate cancer classification and detection on magnetic resonance imaging: a narrative review. *Diagnostics (Basel)* 11:959
- Antolin A, Roson N, Mast R et al (2024) The Role of Radiomics in the Prediction of Clinically Significant Prostate Cancer in the PI-RADS v2 and v2.1 Era: A Systematic Review. *Cancers (Basel)* 16:2951
- Cohen JF, Korevaar DA, Altman DG et al (2016) STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 6:e012799
- Weinreb JC, Barentsz JO, Choyke PL et al (2016) PI-RADS Prostate Imaging-Reporting and Data System: 2015, Version 2. *Eur Urol* 69:16–40
- Tejani AS, Klontzas ME, Gatti AA et al (2024) Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. *Radiol Artif Intell* 6:e240300
- de Rooij M, Israël B, Tummers M et al (2020) ESUR/ESUI consensus statements on multi-parametric MRI for the detection of clinically significant prostate cancer: quality requirements for image acquisition, interpretation and radiologists' training. *Eur Radiol* 30:5404–5416
- Ng ABCD, Asif A, Agarwal R et al (2025) Biparametric vs multiparametric MRI for prostate cancer diagnosis: the PRIME diagnostic clinical trial. *JAMA* 334:1170–1179
- Salinas-Miranda E, Birosh A, McInnes MDF et al (2025) NPV of biparametric and multiparametric prostate MRI: a comparative systematic review and meta-analysis. *AJR Am J Roentgenol* 224:e2432328
- Morote J, Borque-Fernando A, Triquell M et al (2022) The Barcelona predictive model of clinically significant prostate cancer. *Cancers (Basel)* 14:1589
- Anjum M, Khan K, Ahmad W, Ahmad A, Amin MN, Nafees A (2022) New SHapley Additive ExPlanations (SHAP) approach to evaluate the raw materials interactions of steel-fiber-reinforced concrete. *Materials (Basel)* 15:6261
- Bao J, Qiao X, Song Y et al (2024) Prediction of clinically significant prostate cancer using radiomics models in real-world clinical practice: a retrospective multicenter study. *Insights Imaging* 15:68
- Hamm CA, Baumgärtner GL, Biessmann F et al (2023) Interactive explainable deep learning model informs prostate cancer diagnosis at MRI. *Radiology* 307:e222276
- Lin Y, Yilmaz EC, Belue MJ et al (2024) Evaluation of a cascaded deep learning-based algorithm for prostate lesion detection at biparametric MRI. *Radiology* 311:e230750
- Saha A, Bosma JS, Twilt JJ et al (2024) Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study. *Lancet Oncol* 25:879–887
- Cai JC, Nakai H, Kuanar S et al (2024) Fully automated deep learning model to detect clinically significant prostate cancer at MRI. *Radiology* 312:e232635
- Radtke JP, Wiesenfarth M, Kesck C et al (2017) Combined clinical parameters and multiparametric magnetic resonance imaging for advanced risk modeling of prostate cancer—patient-tailored risk stratification can reduce unnecessary biopsies. *Eur Urol* 72:888–896
- van Leeuwen PJ, Hayen A, Thompson JE et al (2017) A multiparametric magnetic resonance imaging-based risk model to determine the risk of significant prostate cancer prior to biopsy. *BJU Int* 120:774–781
- Alberts AR, Roobol MJ, Verbeek JFM et al (2019) Prediction of high-grade prostate cancer following multiparametric magnetic resonance imaging: improving the Rotterdam European Randomized Study of Screening for Prostate cancer Risk Calculators. *Eur Urol* 75:310–318
- Schrader A, Netzer N, Hielscher T et al (2024) Prostate cancer risk assessment and avoidance of prostate biopsies using fully automatic deep learning in prostate MRI: comparison to PI-RADS and integration with clinical data in nomograms. *Eur Radiol* 34:7909–7920
- Transin S, Souchon R, Goninard-Melodelima C et al (2019) Computer-aided diagnosis system for characterizing ISUP grade \geq 2 prostate cancers

- at multiparametric MRI: a cross-vendor evaluation. *Diagn Interv Imaging* 100:801–811
34. Sun Z, Wu P, Cui Y et al (2023) Deep-learning models for detection and localization of visible clinically significant prostate cancer on multiparametric MRI. *J Magn Reson Imaging* 58:1067–1081
 35. Byun J, Park KJ, Kim MH, Kim JK (2020) Direct comparison of PI-RADS Version 2 and 2.1 in transition zone lesions for detection of prostate cancer: preliminary experience. *J Magn Reson Imaging* 52:577–586
 36. Rudolph MM, Baur ADJ, Cash H et al (2020) Diagnostic performance of PI-RADS version 2.1 compared to version 2.0 for detection of peripheral and transition zone prostate cancer. *Sci Rep* 10:15982

Publisher's Note

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