ACCSCIENCE PUBLISHING

Eurasian Journal of Medicine and Oncology

ORIGINAL RESEARCH ARTICLE

Incremental value of needle-trajectory radiomics using publicly available three-dimensional transrectal ultrasound data for detecting core-level clinically significant prostate cancer

Yongzhe Zheng[†], Lixin Wang[†], Guangming Jin*

Department of Ultrasound Medicine, Affiliated Hospital of Yanbian University (Yanbian Hospital), Yanji, Jilin, China

Abstract

Introduction: Transrectal ultrasound (TRUS) radiomics has shown potential for prostate cancer detection, but previous studies have primarily extracted features from lesions or broad anatomical regions. The diagnostic value of ultrasound features surrounding individual biopsy needle trajectories remains unclear.

Objective: To determine whether radiomic features extracted from regions surrounding tracked needle trajectories in three-dimensional (3D) TRUS improve core-level discrimination of clinically significant prostate cancer (csPCa) beyond clinical and anatomical location information.

Methods: We retrospectively analyzed the public Prostate-MRI-US-Biopsy cohort, including 1,129 patients with 17,559 cores. csPCa was defined as a Gleason score of ≥7. Cylindrical regions (5-mm radius, with 3-mm and 7-mm sensitivity analyses) were centered on trajectory paths, and 93 PyRadiomics features were extracted. Three L1-penalized logistic regression models (clinical/location baseline, trajectory-radiomics, and combined) were evaluated using nested patient-grouped cross-validation and patient-cluster bootstrap resampling.

Results: After quality control, 16,259 cores (1,561 csPCa) were analyzed. No radiomic feature was selected in any outer fold of the combined model. Areas under the receiver operating characteristic curve (AUCs) for baseline, trajectory-radiomics, and combined models were 0.812, 0.543, and 0.819, respectively. The combined-minus-baseline AUC difference was 0.0065 ($p < 0.001$), attributable to fold-specific regularization rather than radiomic information, and the combined model had a higher (worse) Brier score than the baseline model (0.196 vs. 0.189).

Conclusion: Tracked coordinates enabled region construction, but conventional trajectory radiomics provided no attributable or clinically meaningful incremental value over clinical/location data. Independent validation is required before clinical application.

Keywords: Prostate cancer; Transrectal ultrasound; Radiomics; Biopsy; Needle

[†]*These authors contributed equally to this work.*

***Corresponding author:**
Guangming Jin
(jgm920@sina.com)

Citation: Zheng Y, Wang L, Jin G. Incremental value of needle-trajectory radiomics using publicly available three-dimensional transrectal ultrasound data for detecting core-level clinically significant prostate cancer. *Eurasian J Med Oncol.* doi: 10.36922/EJMO026350451

Received: August 27, 2026

Revised: October 1, 2026

Accepted: October 22, 2026

Published online: November 16, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Transrectal ultrasound (TRUS) remains a practical imaging platform for prostate biopsy

guidance, yet conventional B-mode images often provide limited direct conspicuity for cancer. Three-dimensional (3D) tracking changes the information available during biopsy: each pathology result can be linked to a sampled path rather than merely to a prostate sector. Tracked systems have been used to record and revisit biopsy locations and to support magnetic resonance imaging-ultrasound fusion workflows, but tracking accuracy alone does not establish that the surrounding ultrasound signal contains useful diagnostic information.[1-3]

Quantitative ultrasound studies have approached prostate cancer localization from several directions. Multiparametric ultrasound approaches combine B-mode imaging with shear-wave elastography, contrast-enhanced ultrasound, or both[4-6] while recent work in high-frequency micro-ultrasound has used core labels for representation learning and classification.[7,8] Conventional TRUS radiomics has also been investigated, although most studies define lesions or broad anatomic regions rather than the exact path associated with an individual biopsy specimen.[9] These tasks, modalities, and reference standards are related but not interchangeable.

The public Prostate-MRI-US-Biopsy collection contains 3D TRUS volumes, core-level pathology, and tracked tip and base coordinates.[10] It therefore permits a direct test of a narrow question: can the procedural geometry be reused to define a radiomic sampling region, and does that region contribute information beyond prostate-specific antigen (PSA), prostate volume, PSA density, and the recorded core location? A negative answer would still be informative because coordinate availability does not guarantee adequate B-mode contrast, registration accuracy, or tissue support along the full trajectory.

We evaluated the incremental value of trajectory-centered radiomics for core-level clinically significant prostate cancer (csPCa). The primary analysis used a 5-mm cylindrical region and compared a combined model with a clinical/location baseline. Secondary analyses evaluated a stand-alone radiomics model and the effects of using 3- and 7-mm cylinder radii. All data partitioning, preprocessing, model selection, and uncertainty estimation respected patient clustering.

2. Data and methods

2.1. Study design and public data source

This retrospective modeling study used the deidentified Prostate-MRI-US-Biopsy collection from The Cancer Imaging Archive.[10] The unit of prediction was the biopsy core, whereas the patient was the unit of data partitioning and resampling. The source manifest retained one biopsy-linked ultrasound series per patient, choosing the series with the greatest number of linked biopsy rows when more than one candidate was available. Eligible series were required to have at least eight biopsy cores with complete ultrasound tip and base coordinates in ultrasound space and an image digest matching the official record. After series selection and eligibility screening, the study manifest comprised 1,129 patients and 17,559 biopsy cores before trajectory-level quality control.

The public biopsy table did not provide variables suitable for sex- or gender-based analyses; neither sex nor gender was evaluated as a predictor or subgroup. Other demographic variables such as age, race, and body mass index were unavailable or incomplete in the public manifest and were therefore not included. This limitation is discussed further below.

2.2. Reference standard and candidate predictors

Core-level biopsy pathology served as the reference standard. We defined csPCa as a primary plus secondary Gleason score of at least 7, including Gleason 3+4 disease; every other core was negative for this endpoint. The baseline candidate predictors were serum PSA, prostate volume, PSA density calculated as PSA divided by prostate volume, and the recorded core-location label. No patient-level outcome was substituted for a core label. The tracked trajectory and radiomic measurements were not used to derive the reference standard.

2.3. Image verification and coordinate mapping

Each selected Digital Imaging and Communications in Medicine (DICOM) file was downloaded from the public collection and accepted only after its pixel data had been decoded or, for uncompressed transfer syntaxes, checked against the expected pixel-data length. The full download audit was completed for all 1,129 selected patients without an unresolved file error. Isotropic voxel spacing was read from private DICOM element (1129,1016). Because the reconstructed volumes did not provide a usable patient-space origin for the tracked coordinates, we placed the image origin at the geometric center of the voxel array and documented this as an assumption.

Tracked ultrasound coordinates were transformed to the image coordinate convention in two steps: ultrasound (x,y,z) was mapped to 3D Slicer right-anterior-superior coordinates (y,z,x), then to left-posterior-superior coordinates (−y,−z,x). Each continuous tip-to-base trajectory was sampled at 201 equally spaced points. A core was excluded if less than 80% of this path lay within the reconstructed image, if the trajectory length was less than 2 mm, or if its endpoints duplicated an earlier core with the

same recorded location within 0.5 mm. This pathology-blind duplicate rule removed one core; an audit found no near-duplicate pair with discordant pathology.

Before full-cohort feature extraction, we conducted a technical pilot involving 20 patients and visually inspected image overlays to assess trajectory orientation and intersection with tissue regions. The 20 patients were purposively selected from the 1,129-patient manifest to cover the major acquisition software versions, a range of prostate volumes, and cores from different prostate regions. Two readers inspected orthogonal overlays. Alignment was considered acceptable if the projected tip-to-base path remained inside the reconstructed ultrasound field, the cylindrical region intersected positive-valued ultrasound support, and the path direction was anatomically concordant with the expected base-to-apex course. No independent patient-space geometric reference was available; therefore, trajectory localization accuracy could not be quantified against a gold standard. The trajectory-containment and signal-support thresholds specified in Sections 2.3 and 2.4 were used as quantitative proxies for acceptable localization. This limitation is revisited in Section 4.

2.4. Trajectory regions and image preprocessing

For each core, we created a capped cylindrical region around the line segment joining its tracked tip and base. The prespecified primary cylinder radius was 5 mm. Three- and 7-mm cylinders were generated independently for sensitivity analysis. The geometric mask was intersected with positive-valued ultrasound support, so padding and voxels outside the reconstructed acoustic field did not enter feature extraction. Regions with fewer than 100 supported voxels or less than 50% supported volume were excluded.

An illustrative 5-mm trajectory region in three orthogonal planes is shown in Figure 1. Intensity normalization was performed separately for each 3D volume. Positive-valued voxels were centered by the volume-specific mean and divided by the population standard deviation, after which values were clipped to the interval from −3 to 3; voxels outside ultrasound support remained zero. Images and masks were resampled to 1-mm isotropic voxel spacing using B-spline interpolation for image intensities and nearest-neighbor interpolation for masks. Gray levels were discretized with a fixed bin width of 0.25. The parameter file set a five-voxel padding distance, 3D extraction, a minimum mask size of 64 voxels, and no additional image filter. The bin width was fixed before modeling in accordance with the Image Biomarker Standardisation Initiative (IBSI) reporting principles; no post hoc bin-width sensitivity analysis was performed.

2.5. Radiomic feature extraction

PyRadiomics 3.1.0 (https://pyradiomics.readthedocs.io/en/latest/) was used to calculate original-image features according to fixed settings.[11] The feature set comprised first-order, gray-level co-occurrence matrix, gray-level run-length matrix, gray-level size-zone matrix, gray-level dependence matrix, and neighboring gray-tone difference matrix features. Shape features were excluded because cylinder shape was imposed by the algorithm and clipping would encode image support rather than tumor morphology. The final feature set comprised 93 radiomic features per eligible biopsy core. Preprocessing, discretization, software version, and feature families were fixed before full-cohort modeling and were reported in accordance with IBSI and CheckList for EvaluAtion of Radiomics research (CLEAR) principles.[12,13]

2.6. Prediction models and internal validation

We compared three logistic regression models: a clinical/location baseline, a trajectory-radiomics model, and a combined model. Categorical core location was imputed with the most frequent training value and one-hot encoded. Continuous variables and radiomic features were median-imputed and standardized; zero-variance radiomic features were removed. Each model used an L1 penalty, balanced class weights, and inverse patient core-count sample weights so that patients with many biopsy cores did not dominate estimation. The regularization parameter C was selected from {0.003, 0.01, 0.03, 0.1, 0.3, 1, 3}.

Model evaluation used five outer patient-stratified folds. Within each outer training set, four patient-stratified folds selected C by area under the receiver operating characteristic curve (AUC). Imputation, encoding, variance filtering, scaling, and penalized estimation were repeated within the training data for each split. No core from an outer test patient was available during tuning or fitting. The resulting outer-fold predictions were concatenated to form one out-of-fold probability for every analyzed core. This was internal validation within a single public collection, not external validation.

L1 logistic regression was chosen a priori because it is transparent, reproducible, compatible with patient-clustered resampling, and permits direct auditing of retained predictors. More flexible classifiers might capture nonlinear interactions but would not resolve the primary question, which was whether handcrafted trajectory radiomics contributes incremental information beyond clinical/location variables.

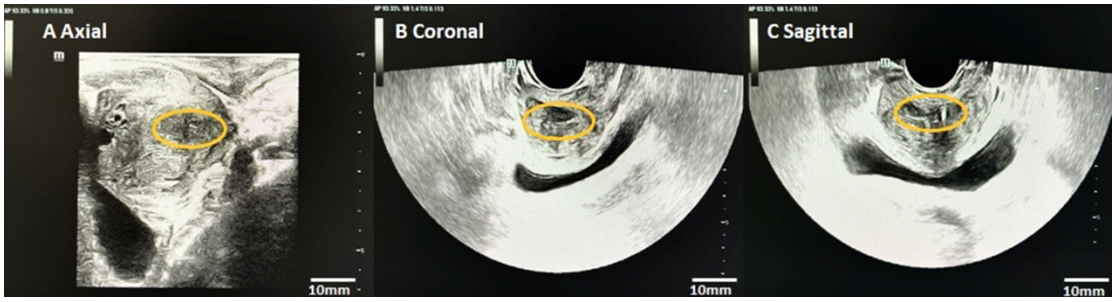


Figure 1. Construction of a trajectory-centered radiomic region. (A) Axial. (B) Coronal. (C) Sagittal. Note: Orthogonal three-dimensional transrectal ultrasound sections from one deidentified public examination show a 5-mm region centered on the tracked tip-to-base biopsy path. Yellow contours delineate the supported radiomic region after intersection of the cylindrical mask with positive-valued ultrasound voxels. Each panel shows a 65 × 65 mm field of view, with a 10-mm scale bar. The image illustrates geometry and quality control, not model performance.

2.7. Statistical analysis

The primary estimand was the paired difference in core-level AUC between the combined 5-mm model and the clinical/location baseline, interpreted together with the radiomic coefficient retention audit. Secondary comparisons evaluated trajectory radiomics against the baseline and the combined model against trajectory radiomics. We also calculated average precision, Brier score, calibration intercept and slope, and exploratory decision curves over threshold probabilities from 0.02 to 0.50.[14,15] AUC and average-precision confidence intervals and paired AUC-difference confidence intervals were obtained from 2,000 bootstrap samples drawn with replacement at the patient level. These intervals were conditional on the realized out-of-fold predictions and did not repeat fold assignment or model fitting. Two-sided bootstrap *p*-values were adjusted across the three model comparisons with the Benjamini–Hochberg procedure.

The radius analysis was restricted to cores with successful extraction at 3, 5, and 7 mm. Separate trajectory-radiomics models were trained for each radius with the same patient-grouped nested validation. The 5-mm predictions were compared pairwise with the 3- and 7-mm predictions, with Benjamini–Hochberg adjustment across those two comparisons. Feature robustness across radii was summarized by paired Spearman correlation for each feature. Analyses followed the prespecified hierarchy and reporting concepts in TRIPOD+AI and CLEAR.[16]

No formal sample-size calculation was performed because the analysis included all eligible patients in the public collection. Modeling used Python 3.9.13 (https://www.python.org) and scikit-learn 1.5.2 (http://scikit-learn.org); the complete package versions and random seeds are retained with the analysis materials.

3. Results

3.1. Cohort flow and extraction quality

All 1,129 selected DICOM series passed pixel-data verification, and trajectory records were generated for all 17,559 biopsy cores without a patient-level extraction failure (Figure 2).

At the primary 5-mm radius, 16,259 cores (92.6%) from all patients met geometry and signal-support criteria; 1,300 cores were excluded. Exclusion reasons could overlap and most often included less than 80% trajectory containment (856 cores) or less than 50% ultrasound signal support within the geometric region (800 cores). The modeled set contained 1,561 csPCa cores from 454 patients and 93 finite radiomic features per eligible core (Table 1).

3.2. Primary model comparison

The clinical/location baseline had an out-of-fold AUC of 0.812 (95% cluster-bootstrap confidence interval [CI], 0.794–0.827) and average precision of 0.345 (95% CI, 0.301–0.399) (Figure 3 and Tables 2 and 3). The trajectory-radiomics model showed weak discrimination with an AUC of 0.543 (95% CI, 0.521–0.565) and average precision of 0.106 (95% CI, 0.094–0.119). The combined model had an AUC of 0.819 (95% CI, 0.802–0.834) and average precision of 0.352 (95% CI, 0.308–0.406).

The paired combined-minus-baseline AUC difference was 0.0065 (95% CI, 0.0031–0.0100; Benjamini–Hochberg-adjusted $p < 0.001$). This numerical difference did not represent retained radiomic information. In all five outer folds, the combined L1 model selected the same four baseline predictors and zero radiomic features. Baseline and combined models selected the same regularization value in folds 1–3; in folds 4–5, the combined model selected stronger shrinkage (C = 0.003 vs. 0.01). Their within-

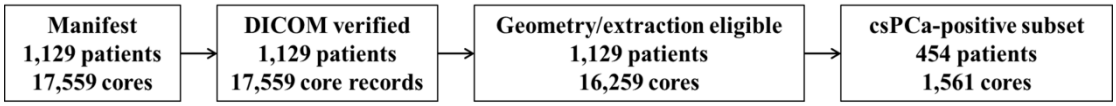


Figure 2. Cohort and biopsy-core flow. Note: One biopsy-linked three-dimensional transrectal ultrasound series was selected for each of 1,129 patients. All selected DICOM files passed pixel-data verification, corresponding to 17,559 biopsy-core records. At the prespecified 5-mm radius, 16,259 of 17,559 biopsy cores met trajectory-containment, signal-support, and extraction criteria and entered the primary analysis; 1,300 cores were excluded, with overlapping reasons most commonly including < 80% trajectory containment or < 50% ultrasound signal support. The final modeled set comprised 1,561 csPCa cores from 454 patients.
Abbreviations: csPCa: Clinically significant prostate cancer; DICOM: Digital Imaging and Communications in Medicine.

fold AUC differences ranged from 0 to 0.00064, showing that the pooled difference arose from fold-specific fitting and probability scaling rather than an added radiomic predictor. Refitted models reproduced the saved out-of-fold probabilities to within 1.2×10^{-16}. The trajectory-only AUC was 0.269 lower than the baseline (95% CI, −0.294 to −0.244; adjusted $p < 0.001$).

Table 1. Cohort and extraction characteristics

Characteristic	Value
Manifest patients	1,129
Manifest biopsy cores	17,559
Analyzed patients at 5 mm	1,129
Analyzed biopsy cores at 5 mm	16,259
Patients with at least one csPCa core	454
csPCa biopsy cores	1,561
PSA, ng/mL, median (IQR)	6.71 (4.60–10.00)
Prostate volume, cm^3, median (IQR)	47.07 (34.26–67.50)
Biopsy cores per patient, median (IQR)	15.0 (12.0–17.0)
Ultrasound software version 2.3.0.25, patients	72
Ultrasound software version 2.3.0.81, patients	220
Ultrasound software version 2.3.2.3, patients	28
Ultrasound software version 2.3.3.13, patients	73
Ultrasound software version 3.1.1.8, patients	11
Ultrasound software version 3.1.2.9, patients	215
Ultrasound software version 3.3.0.45, patients	4
Ultrasound software version 3.3.1.35, patients	86
Ultrasound software version 3.4.0.28, patients	420

Note: Values are counts unless otherwise indicated.
Abbreviations: csPCa: Clinically significant prostate cancer; IQR: Interquartile range; PSA: Prostatespecific antigen.

3.3. Calibration and exploratory decision analysis

The baseline, trajectory-only, and combined Brier scores were 0.189, 0.250, and 0.196, respectively. Calibration intercepts were −2.259, −2.243, and −2.286, with slopes of 1.046, 0.255, and 1.204. Thus, the class-balanced model outputs substantially overestimated observed core-level risk, and the trajectory-only predictions also showed marked overfitting. In exploratory decision analysis, none of the models had positive net benefit above a threshold probability of approximately 0.10; the combined model did not show a consistent advantage over the baseline. These findings indicate that discrimination and calibration should be interpreted separately, and the raw probabilities are not suitable for clinical decision-making without recalibration and external validation.

3.4. Radius sensitivity analysis

The paired radius analysis included 16,087 cores from 1,129 patients, including 1,543 csPCa cores. The trajectory-radiomics models yielded AUCs of 0.536 (95% CI, 0.516–0.557), 0.547 (95% CI, 0.526–0.567), and 0.541 (95% CI, 0.519–0.564) for cylinder radii of 3, 5, and 7 mm, respectively. The AUC difference between the 5- and 3-mm models was 0.011 (95% CI, −0.003 to 0.024; adjusted $p = 0.226$), whereas the difference between the 5- and 7-mm models was 0.006 (95% CI, −0.009 to 0.022; adjusted $p = 0.403$). Median feature-wise Spearman correlations were 0.924 for 3 versus 5 mm, 0.954 for 5 versus 7 mm, and 0.828 for 3 versus 7 mm (Figure 4 and Table 4–6).

4. Discussion

This public-data study established that tracked biopsy coordinates can be converted into quality-controlled radiomic regions at scale, but it did not establish incremental diagnostic value from the surrounding conventional 3D TRUS signal. No radiomic feature survived penalization in any outer fold. The small pooled AUC difference of 0.0065 occurred because the two pipelines selected different shrinkage in two folds; within-fold AUC differences were at most 0.00064. Together with the worse Brier score and absence of improved net benefit, this shows that the statistically detectable pipeline-level difference was neither attributable to radiomics nor clinically meaningful.

Weak stand-alone discrimination is plausible given this setting. Conventional B-mode TRUS provides limited

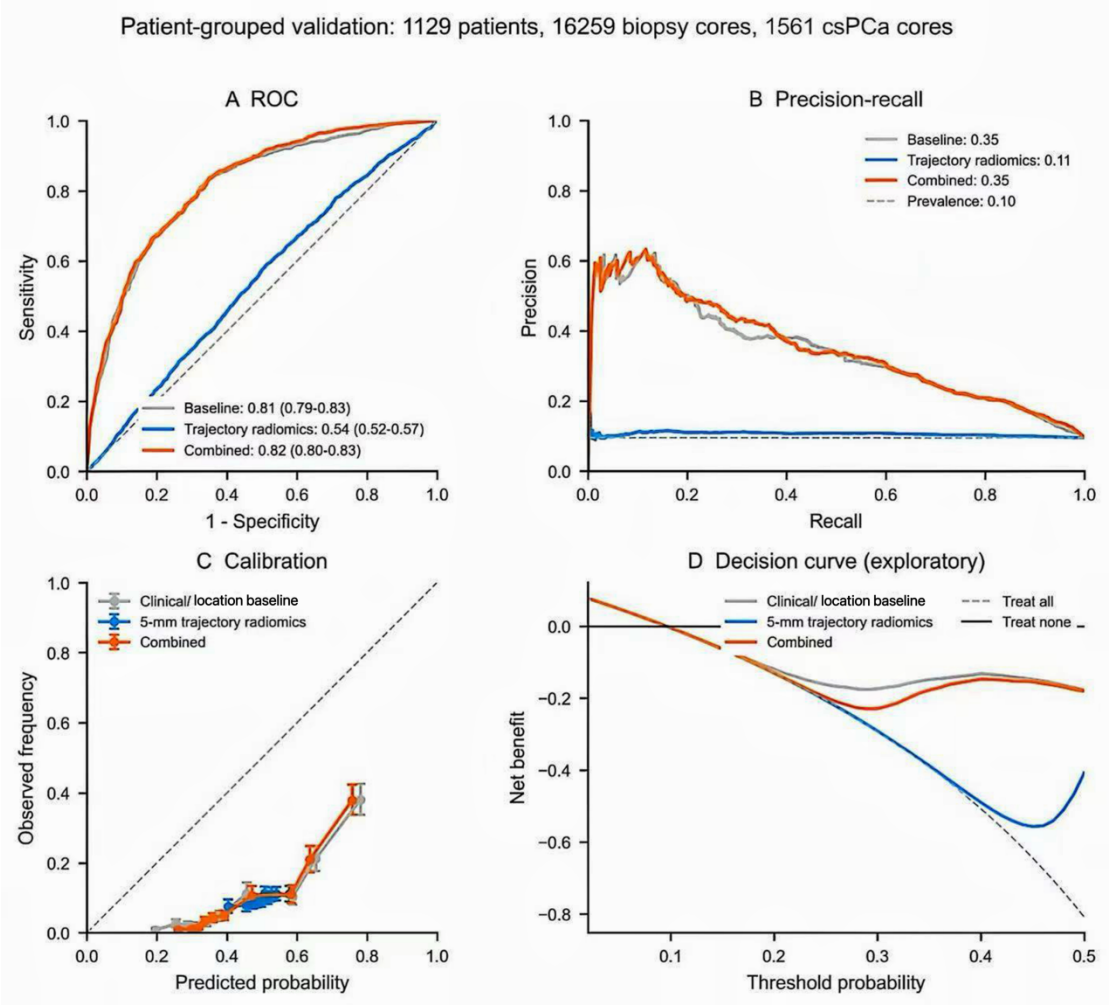


Figure 3. Out-of-fold model performance in the primary 5-mm analysis. Patient-grouped validation included 1,129 patients, 16,259 biopsy cores, and 1,561 clinically significant prostate cancer (csPCa) cores. (A) Receiver operating characteristic curves. (B) Precision–recall curves. (C) Calibration plots. (D) Exploratory decision curves. The panels compare the clinical/location baseline, 5-mm trajectory-radiomics, and combined models. Out-of-fold predictions were generated using nested patient-grouped cross-validation. Confidence intervals were estimated using patient-cluster bootstrap resampling. The calibration results indicate that the raw probabilities generated by the class-balanced models should not be interpreted as estimates of calibrated clinical risk. Decision-curve results should be interpreted as exploratory.

Table 2. Out-of-fold model performance

Model	AUC	AUC, 95% CI	Average precision	Avg precision, 95% CI	Brier score	Calibration intercept	Calibration slope
Clinical/location baseline	0.812	0.794–0.827	0.345	0.301–0.399	0.189	−2.259	1.046
5-mm trajectory radiomics	0.543	0.521–0.565	0.106	0.094–0.119	0.250	−2.243	0.255
Combined	0.819	0.802–0.834	0.352	0.308–0.406	0.196	−2.286	1.204

Note: CIs were obtained from 2,000 patient-cluster bootstrap samples.
Abbreviations: AUC: Area under the receiver operating characteristic curve; CI: Confidence interval.

Table 3. Paired model comparisons

Comparison	AUC difference	95% CI	Bootstrap *p*-value	BH-adjusted *p*-value
Combined minus clinical/location baseline	0.0065	0.0031–0.0100	<0.001	<0.001
Trajectory radiomics minus clinical/location baseline	−0.269	−0.294 to −0.244	<0.001	<0.001
Combined minus trajectory radiomics	0.276	0.251–0.303	<0.001	<0.001

Note: Differences were calculated from paired out-of-fold predictions. *p*-values used patient-cluster bootstrap and BH correction across three model comparisons.
Abbreviations: AUC: Area under the receiver operating characteristic curve; BH: Benjamini–Hochberg; CI: Confidence interval.

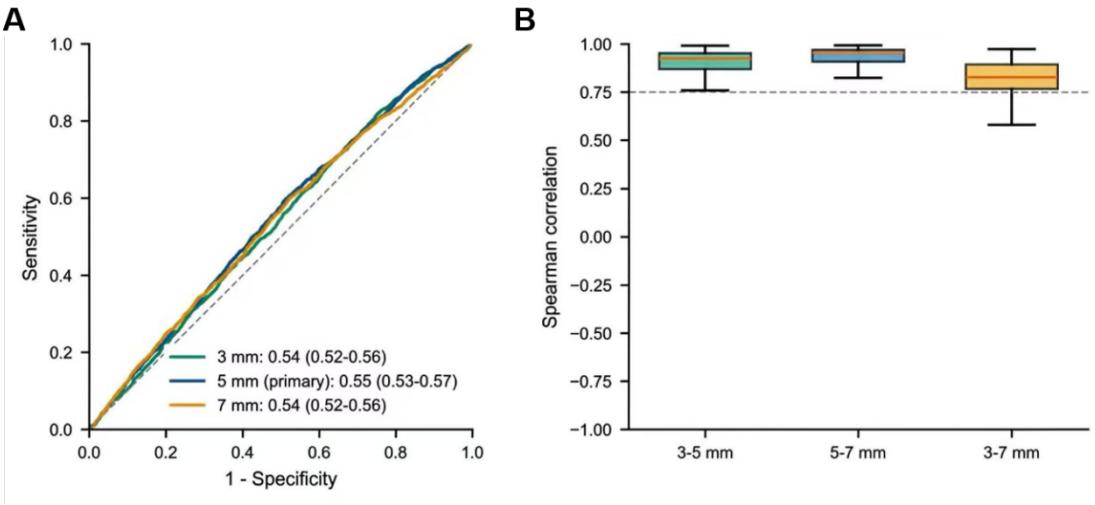


Figure 4. Prespecified radius sensitivity analysis. The paired analysis included 1,129 patients and 16,087 biopsy cores, of which 1,543 were clinically significant prostate cancer cores. (A) Receiver operating characteristic curves comparing trajectory-radiomics models using biopsy cores eligible for analysis at three radii (3, 5, and 7 mm). The areas under the curve were 0.536, 0.547, and 0.541 for the 3-, 5-, and 7-mm models, respectively. (B) Box plots showing feature-wise Spearman correlation coefficients between radiomic features extracted at each pair of cylinder radii (3 vs. 5 mm, 5 vs. 7 mm, and 3 vs. 7 mm). The dashed horizontal line indicates a correlation coefficient of 0.75.

Table 4. Prespecified radius model performance

Radius	AUC	AUC, 95% CI	Average precision	Brier score	Calibration intercept	Calibration slope
3 mm	0.536	0.516–0.557	0.104	0.249	−2.245	0.761
5 mm	0.547	0.526–0.567	0.108	0.248	−2.242	0.733
7 mm	0.541	0.519–0.564	0.107	0.246	−2.234	0.206

Note: All radius models used the same patients, paired biopsy cores, outer folds, and inner tuning design.
Abbreviations: AUC: Area under the receiver operating characteristic curve; CI: Confidence interval.

tumor contrast, while a biopsy trajectory samples a narrow path whose pathology may reflect only part of the imaged tissue. Coordinate uncertainty, incomplete acoustic support, and the imposed cylindrical geometry may further dilute local texture signals. The radius analysis provided an additional check on whether these findings depended on one arbitrary cylinder width. Stand-alone trajectory discrimination remained weak at every radius; the paired AUC differences were 0.011 for 5 versus 3 mm and 0.006 for 5 versus 7 mm. Although many feature values were correlated between adjacent radii, the median 3-versus-7-mm correlation was 0.828. Changing the cylinder width

Table 5. Paired radius comparisons

Comparison	AUC difference	95% CI	Bootstrap *p*-value	BH-adjusted *p*-value
5 mm minus 3 mm	0.011	−0.003 to 0.024	0.113	0.226
5 mm minus 7 mm	0.006	−0.009 to 0.022	0.403	0.403

Note: *p*-values used patient-cluster bootstrap and BH correction across the two prespecified radius comparisons.
Abbreviations: AUC: Area under the receiver operating characteristic curve; BH: Benjamini–Hochberg; CI: Confidence interval.

therefore altered some measurements without revealing a radius at which conventional trajectory radiomics became diagnostically informative.

Table 6. Cross-radius feature stability

Radius pair	Median Spearman ρ	Features with ρ ≥ 0.75
3–5 mm	0.924	87
5–7 mm	0.954	88
3–7 mm	0.828	72

Note: Spearman correlations were calculated feature-wise across paired cores eligible at all three radii.

Earlier prostate ultrasound studies have reported useful localization or classification with multiparametric B-mode, elastography, contrast-enhanced ultrasound, or high-frequency micro-ultrasound. In particular, transrectal shear-wave elastography has been evaluated for the differential diagnosis of benign and malignant prostate lesions, with promising but not uniformly reproducible results.[17] Those results do not directly contradict the present findings: the image formation, tissue contrast, labeling unit, and available clinical covariates differ. Our analysis instead isolates a narrower question about exact tracked paths in conventional public 3D TRUS. The approach was technically feasible but diagnostically limited. More recent work using deep learning on TRUS video and multiparametric ultrasound has reported higher AUCs for csPCa detection, but those models typically learn from whole-gland or lesion-level representations rather than from the specific trajectory geometry tested here, and they require substantially larger training cohorts and external validation.[18-23]

4.1. Clinical positioning and workflow implications

The core-level reference standard is methodologically appropriate for the question addressed, but it does not imply direct clinical applicability. Biopsy cores are obtained after the decision to biopsy, and their trajectories are shaped by systematic sampling, targeted sampling, operator technique, and tracking-system geometry.[24-26] A model that predicts csPCa from a trajectory region therefore cannot guide the initial decision to biopsy, nor can it replace pre-biopsy risk assessment. At most, such a model might support retrospective quality assurance by identifying whether biopsy trajectories passed through potentially informative tissue. If real-time tracking and radiomic inference become sufficiently mature, future models might also help determine whether additional samples should be obtained along or near an existing trajectory after the decision to biopsy has been made. The present negative result does not support any of these uses clinically; it only demonstrates methodological feasibility of constructing trajectory-centered regions from public tracked coordinates.

4.2. Discrimination is not clinical usefulness

The baseline and combined models achieved AUCs of 0.812 and 0.819, which indicate modest-to-good ranking ability. However, calibration intercepts were markedly negative, calibration slopes were not consistently close to 1, Brier scores were worse for the combined model than for the baseline, and exploratory decision curves showed no positive net benefit above a threshold probability of approximately 0.10. These findings separate discrimination from clinical usefulness. The AUC should not be interpreted as evidence that the model is ready for clinical decision-making. Class-balanced fitting and inverse patient core-count weighting were used to avoid domination by patients with many cores, but they also produced probabilities that overestimated observed core-level risk. Recalibration methods such as Platt scaling, isotonic regression, or prior correction could be considered in future work, but they would need to be fitted within independent data and would not by themselves establish clinical utility.

4.3. Technical assumptions, registration, and batch effects

The coordinate transform required a centered-volume origin assumption because the public reconstructed

volumes did not provide a usable patient-space origin. A 20-patient technical pilot was used for visual quality control, but no independent quantitative localization reference was available. Small spatial errors could substantially affect radiomic measurements around a narrow biopsy trajectory. Additional sources of registration error include tracking-system error, tissue deformation, needle bending, patient movement, and the fact that pathology represents a specimen rather than an exact in vivo point. These issues set a ceiling on what trajectory-centered radiomics can achieve in this dataset.

Acquisition software varied across the collection (Table 1), and software version may affect gain, dynamic range, speckle characteristics, and texture features. Because the public manifest lacks a center identifier and software versions may be confounded with collection batches, we did not perform formal batch-effect correction. This is a limitation; future studies should consider harmonization methods and external validation across scanners and software versions. Other demographic variables such as age, race, and BMI were unavailable or incomplete, so their potential influence on prostate volume and baseline clinical models could not be evaluated.

4.4. Modeling choices and future directions

L1 logistic regression was chosen as the sole primary model because it is transparent, reproducible, and allows direct auditing of retained predictors. The zero-radiomic-feature result across all outer folds is therefore informative: it is not an artifact of a black-box model. More flexible classifiers or representation-learning architectures might capture nonlinear texture patterns, but they would require larger labeled datasets, careful architecture selection, and external validation; in the present dataset, they would be at high risk of overfitting and would not answer the prespecified question about handcrafted trajectory radiomics. Similarly, the fixed bin width of 0.25 was prespecified according to IBSI principles, but bin-width sensitivity was not evaluated. Future work should test discretization, harmonization, and alternative model families.

The Brier score for the trajectory-only model (0.250) should not be attributed to a single regularization failure. Rather, it reflects a combination of weak underlying signal, AUC-oriented inner-loop tuning rather than calibration-oriented tuning, class-balanced weighting, and patient-level sample weighting. These choices were appropriate for the primary discrimination question but produced poorly calibrated probabilities. The small combined-minus-baseline AUC difference of 0.0065 is reported for transparency because it demonstrates that even a nominal pipeline-level difference was not due to retained radiomic information. Discarding the radiomic pipeline entirely would obscure this audit.

The broader radiomics literature has repeatedly emphasized that methodological rigor, transparent reporting, and external validation are essential before clinical translation.[27-34] Prostate ultrasound radiomics specifically has been criticized for heterogeneity in region definition, reference standards, and validation strategies, which limits cross-study comparability.[35-38] Our negative result contributes to this literature by showing that even when procedural geometry is reused in a controlled and auditable pipeline, conventional handcrafted texture features may still lack incremental value. This aligns with the growing recognition that deep representation learning, multimodal fusion, and larger multi-institutional cohorts are likely necessary to extract clinically useful signal from ultrasound data.[39-42]

4.5. Limitations

Several limitations warrant emphasis. This was internal validation within one public collection, without a geographic or temporal external cohort, reader study, or prospective workflow evaluation. Biopsy pathology is an imperfect reference standard and is affected by sampling and registration error. The coordinate transform required a centered-volume origin assumption that was checked qualitatively but not against an independent geometric reference, and acquisition software varied across the collection. The bootstrap intervals were conditional on one set of out-of-fold predictions and did not incorporate variability from repeated partitioning and refitting. Class-balanced fitting produced poorly calibrated raw probabilities; consequently, the decision analysis is evidence of a limitation, not clinical utility. The baseline model included a core-location category labeled "TARGET OR PRIOR POSITIVE," which may encode information beyond anatomical location. These constraints set a clear ceiling: the study evaluates regions produced by this procedural-geometry assumption and their internal discrimination, not deployment or all possible trajectory-centered ultrasound representations.

5. Conclusion

Tracked biopsy coordinates enabled the reproducible construction of trajectory-centered radiomic regions across this public 3D TRUS cohort. Handcrafted trajectory radiomics did not provide attributable incremental value over PSA, prostate volume, PSA density, and core-location information, and the resulting probabilities were not suitable for clinical use without independent validation and recalibration. Future work should prioritize external validation, harmonization across acquisition platforms,

and representation-learning approaches that can leverage larger, multi-institutional datasets.

Acknowledgments

The authors acknowledge The Cancer Imaging Archive and the contributors of the Prostate-MRI-US-Biopsy collection for making the deidentified source data publicly available.

Funding

This research was funded by the Natural Science Foundation of Jilin Province (grant no. YDZJ202201ZYTS204).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Guangming Jin
Formal analysis: Lixin Wang
Investigation: Lixin Wang
Methodology: Yongzhe Zheng
Writing–original draft: Yongzhe Zheng
Writing–review & editing: Guangming Jin

Ethical approval and consent to participate

This study was a secondary analysis of a deidentified public dataset (Prostate-MRI-US-Biopsy, The Cancer Imaging Archive) and did not involve direct patient contact, intervention, or access to protected health information. The Institutional Review Board of Yanbian University Hospital reviewed the study protocol and determined that this research did not constitute human subjects research, thereby granting an exemption from full ethical review (approval number: [IRB-2025-110]). Informed consent was not required for this retrospective analysis because the study used fully anonymized, publicly available data. All original data collection and consent procedures were carried out by the source dataset contributors in accordance with their institutional guidelines.

Consent for publication

This manuscript contains no identifiable personal data, images, or clinical details of any individual. Separate consent for publication is therefore not applicable.

Availability of data

The source images, biopsy coordinates, and pathology annotations are available from The Cancer Imaging Archive as the Prostate-MRI-US-Biopsy collection (doi:10.7937/tcia.2020.a61ioc1a). The geometry, quality-control, feature-extraction, modeling, and audit code are provided in the supplementary reproducibility package. The core-level feature matrices can be regenerated from the public source data with the supplied scripts.

Further disclosure

During the preparation of this work, the authors used AI-assisted tools to organize references, assist code development, and improve language readability. After using these tools, the authors reviewed and edited the content as needed, independently verified the analyses and citations, and take full responsibility for the final publication.

References

1. Natarajan S, Marks LS, Margolis DJA, *et al*. Clinical application of a 3D ultrasound-guided prostate biopsy system. *Urol Oncol*. 2011;29(3):334-342.

   doi: 10.1016/j.urolonc.2011.02.014

2. Baumann M, Mozer P, Daanen V, Troccaz J. Prostate biopsy tracking with deformation estimation. *Med Image Anal*. 2012;16(3):562-576.

   doi: 10.1016/j.media.2011.01.008

3. Sonn GA, Natarajan S, Margolis DJA, *et al*. Targeted biopsy in the detection of prostate cancer using an office based magnetic resonance ultrasound fusion device. *J Urol*. 2013;189(1):86-92.

   doi: 10.1016/j.juro.2012.08.095

4. Wildeboer RR, Mannaerts CK, van Sloun RJG, *et al*. Automated multiparametric localization of prostate cancer based on B-mode, shear-wave elastography, and contrast-enhanced ultrasound radiomics. *Eur Radiol*. 2020;30(2):806-815.

   doi: 10.1007/s00330-019-06436-w

5. Chen P, Turco S, Wang Y, *et al*. Can 3D multiparametric ultrasound imaging predict prostate biopsy outcome? *Ultrasound Med Biol*. 2024;50(8):1194-1202.

   doi: 10.1016/j.ultrasmedbio.2024.04.007

6. Sun Y, Fang J, Shi Y, *et al*. Machine learning based on radiomics features combing B-mode transrectal ultrasound and contrast-enhanced ultrasound to improve peripheral zone prostate cancer detection. *Abdom Radiol (NY)*. 2024;49(1):141-150.

   doi: 10.1007/s00261-023-04050-5

7. Wilson PFR, Gilany M, Jamzad A, *et al*. Self-supervised learning with limited labeled data for prostate cancer detection in high-frequency ultrasound. *IEEE Trans Ultrason Ferroelectr Freq Control*. 2023;70(9):1073-1083.

   doi: 10.1109/TUFFC.2023.3297840

8. Gilany M, Wilson P, Perera-Ortega A, *et al*. TRUSformer: improving prostate cancer detection from micro-ultrasound using attention and self-supervision. *Int J Comput Assist*

Radiol Surg. 2023;18(7):1193-1200.

doi: 10.1007/s11548-023-02949-4

9. Huang Y, Qian H, Zheng Y, Song H, Liu X. A radiomics model based on transrectal ultrasound for predicting prostate cancer. *Med Ultrason*. 2024;26(2):138-146.

doi: 10.11152/mu-4329

10. Natarajan S, Priester A, Margolis D, Huang J, Marks L. Prostate MRI and Ultrasound With Pathology and Coordinates of Tracked Biopsy (Prostate-MRI-US-Biopsy) (version 2) [data set]. The Cancer Imaging Archive. 2020.

doi: 10.7937/TCIA.2020.A61IOC1A

11. van Griethuysen JJM, Fedorov A, Parmar C, *et al*. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77(21):e104-e107.

doi: 10.1158/0008-5472.CAN-17-0339

12. Zwanenburg A, Vallières M, Abdalah MA, *et al*. The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. 2020;295(2):328-338.

doi: 10.1148/radiol.2020191145

13. Kocak B, Baessler B, Bakas S, *et al*. CheckList for EvaluAtion of Radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMII. *Insights Imaging*. 2023;14:75.

doi: 10.1186/s13244-023-01415-8

14. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230.

doi: 10.1186/s12916-019-1466-7

15. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26(6):565-574.

doi: 10.1177/0272989X06295361

16. Collins GS, Moons KGM, Dhiman P, *et al*. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378.

doi: 10.1136/bmj-2023-078378

17. Li H, Yu WL, Wu ML. Jing zichang jianqiebo tanxingchengxiang dui cunzai xia niaolu zhengzhuang de qianliexian bingbian liang-exing de jianbie zhenduan jiazhi ji qi yingxiang yinsu fenxi [Value of transrectal shear-wave elastography in differential diagnosis of benign and malignant prostate lesions with lower urinary tract symptoms and factors related to severity of such symptoms]. *Chin J Med Ultrasound (Electronic Edition)*. 2023;20(11):1193-1198. [In Chinese].

doi: 10.3877/cma.j.issn.2096-1537.2023.11.014

18. Zhang Q, Zhou C, Chen Y, Luo Y. Artificial intelligence for prostate cancer detection and risk stratification using transrectal ultrasound: a narrative review. *Gland Surg*. 2026;15(2):52.

doi: 10.21037/gs-2025-aw-504

19. Lou W, Chen P, Wu C, *et al*. Prostate cancer classification using 3D deep learning and ultrasound video clips: a multicenter study. *Front Oncol*. 2025;15:1582035.

doi: 10.3389/fonc.2025.1582035

20. Liu X, Zhang ZX, Zheng B, Xu M, Cao XY, Huang HM. A retrospective study on predicting clinically significant prostate cancer via a bi-parametric ultrasound-based deep learning radiomics model. *Front Oncol*. 2025;15:1538854.

doi: 10.3389/fonc.2025.1538854

21. Sun YK, Zhou BY, Miao Y, *et al*. Three-dimensional convolutional neural network model to identify clinically significant prostate cancer in transrectal ultrasound videos: a prospective, multi-institutional, diagnostic study. *EClinicalMedicine*. 2023;60:102027.

doi: 10.1016/j.eclinm.2023.102027

22. Qi XY, Wang K, Feng BJ, *et al*. Comparison of machine learning models based on multi-parametric magnetic resonance imaging and ultrasound videos for the prediction of prostate cancer. *Front Oncol*. 2023;13:1157949.

doi: 10.3389/fonc.2023.1157949

23. Miao KH, Miao JH, Finkelstein M, Chatterjee A, Oto A. Next-generation advances in prostate cancer imaging and artificial intelligence applications. *J Imaging*. 2025;11(11):390.

doi: 10.3390/jimaging11110390

24. Pepe P, Pepe L, Fiorentino V, Curduman M, Fraggetta F. Multiparametric MRI targeted prostate biopsy: when omit systematic biopsy? *Arch Ital Urol Androl*. 2024;96(4):12992.

doi: 10.4081/aiua.2024.12992

25. Drost FJH, Osses DF, Nieboer D, *et al*. Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. *Cochrane Database Syst Rev*. 2019;4(4):CD012663.

doi: 10.1002/14651858.CD012663.pub2

26. Kasivisvanathan V, Rannikko AS, Borghi M, *et al*. MRI-targeted or standard biopsy for prostate-cancer diagnosis. *N Engl J Med*. 2018;378(19):1767-1777.

doi: 10.1056/NEJMoa1801993

27. Lambin P, Leijenaar RTH, Deist TM, *et al*. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749-762.

doi: 10.1038/nrclinonc.2017.141

28. Park JE, Kim D, Kim HS, *et al*. Quality of science and reporting of radiomics in oncologic studies: room for

improvement according to radiomics quality score and TRIPOD statement. *Eur Radiol*. 2020;30(1):523-536.

doi: 10.1007/s00330-019-06360-z

29. Kocak B, Durmaz ES, Erdim C, Ates E, Kaya OK, Kilickesmez O. Radiomics of renal masses: systematic review of reproducibility and validation strategies. *AJR Am J Roentgenol*. 2020;214(1):129-136.

doi: 10.2214/AJR.19.21709

30. Park JE, Kim HS, Kim D, *et al*. A systematic review reporting quality of radiomics research in neuro-oncology: toward clinical utility and quality improvement using high-dimensional imaging features. *BMC Cancer*. 2020;20(1):29.

doi: 10.1186/s12885-019-6504-5

31. Kocak B, Akinci D'Antonoli T, Mercaldo N, *et al*. METhodological RadiomICs Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMII. *Insights Imaging*. 2024;15(1):8.

doi: 10.1186/s13244-023-01572-w

32. Cuocolo R, Cipullo MB, Stanzione A, *et al*. Machine learning for the identification of clinically significant prostate cancer on MRI: a meta-analysis. *Eur Radiol*. 2020;30(12):6877-6887.

doi: 10.1007/s00330-020-07027-w

33. Stanzione A, Gambardella M, Cuocolo R, Ponsiglione A, Romeo V, Imbriaco M. Prostate MRI radiomics: a systematic review and radiomic quality score assessment. *Eur J Radiol*. 2020;129:109095.

doi: 10.1016/j.ejrad.2020.109095

34. Woźnicki P, Westhoff N, Huber T, *et al*. Multiparametric MRI for prostate cancer characterization: combined use of radiomics model with PI-RADS and clinical parameters. *Cancers (Basel)*. 2020;12(7):1767.

doi: 10.3390/cancers12071767

35. Brembilla G, Giganti F, Sidhu H, *et al*. Diagnostic accuracy of abbreviated bi-parametric MRI (a-bpMRI) for prostate cancer detection and screening: a multi-reader study. *Diagnostics (Basel)*. 2022;12(2):231.

doi: 10.3390/diagnostics12020231

36. Castillo TJM, Starmans MPA, Arif M, *et al*. A multi-center, multi-vendor study to evaluate the generalizability of a radiomics model for classifying prostate cancer: high grade vs low grade. *Diagnostics (Basel)*. 2021;11(2):369.

doi: 10.3390/diagnostics11020369

37. Smith CP, Czarniecki M, Mehralivand S, *et al*. Radiomics and radiogenomics of prostate cancer. *Abdom Radiol (NY)*. 2019;44(6):2021-2029.

doi: 10.1007/s00261-018-1660-7

38. Chaddad A, Kucharczyk MJ, Niazi T. Multimodal radiomic features for the predicting Gleason score of prostate cancer. *Cancers (Basel)*. 2018;10(8):249.

doi: 10.3390/cancers10080249

39. Rouvière O, Puech P, Renard-Penna R, *et al*. Use of prostate systematic and targeted biopsy on the basis of multiparametric MRI in biopsy-naive patients (MRI-FIRST): a prospective, multicentre, paired diagnostic study. *Lancet Oncol*. 2019;20(1):100-109.

doi: 10.1016/S1470-2045(18)30569-2

40. Ahdoot M, Wilbur AR, Reese SE, *et al*. MRI-targeted, systematic, and combined biopsy for prostate cancer diagnosis. *N Engl J Med*. 2020;382(10):917-928.

doi: 10.1056/NEJMoa1910038

41. Turkbey B, Rosenkrantz AB, Haider MA, *et al*. Prostate Imaging Reporting and Data System Version 2.1: 2019 update of Prostate Imaging Reporting and Data System Version 2. *Eur Urol*. 2019;76(3):340-351.

doi: 10.1016/j.eururo.2019.02.033

42. Padhani AR, Barentsz J, Villeirs G, *et al*. PI-RADS Steering Committee: the PI-RADS multiparametric MRI and MRI-directed biopsy pathway. *Radiology*. 2019;292(2):464-474.

doi: 10.1148/radiol.2019182946