Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

ADC Map Radiomics for Non-Invasive Detection of BRCA1/2 Mutation Status in Prostate Cancer: A Machine Learning-Based Radiogenomics Study

AILA DE MENEZES FERREIRA

ailamf@usp.br

Universidade de São Paulo

Julio Cesar Nather Junior
Hospital Universitário da Universidade de São Paulo

Lucianna Fonseca Barreto
Universidade de São Paulo

Fernando Chahud
Universidade de São Paulo

Daniel Lopes da Cunha
Universidade Federal do Piauí

Larissa Grazielle Souza Ribeiro
Hospital de Câncer de Barretos

Fernanda Coeli Lacchini
Universidade de São Paulo

Leandro Machado Colli
Hospital Universitário da Universidade de São Paulo

Valdair Francisco Muglia
Universidade de São Paulo

Research Article

Keywords: Prostate cancer, BRCA mutation, radiomics, apparent diffusion coefficient, machine learning, radiogenomics, multiparametric MRI

Posted Date: August 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10594837/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

Conventional multiparametric magnetic resonance imaging (mpMRI) parameters have limited ability to identify prostate cancer (PCa) patients harboring BRCA1/2 mutations. We investigated whether radiomic features extracted from apparent diffusion coefficient (ADC) maps can non-invasively discriminate BRCA-mutated from doubly negative (BRCA−) PCa patients.

Materials and Methods

This prospective cohort study with retrospective imaging analysis included 107 PCa patients (58 BRCA-positive [BRCA+], confirmed by germline and/or somatic testing; 49 doubly negative controls [BRCA−]) at a quaternary referral center (2021–2024). Patients with a negative result in one genetic pathway who were not tested in the other were excluded. Conventional mpMRI parameters (lesion size, mean ADC, zonal location, PI-RADS category, prostate volume) were compared between groups. Radiomic features were extracted from ADC maps using PyRadiomics (107 features; 33 retained after morphological robustness filtering). Twenty-five machine-learning algorithm families were evaluated with stratified k-fold cross-validation. The primary endpoint was area under the precision-recall curve (PR-AUC); secondary endpoint was ROC-AUC.

Results

No significant differences were found in conventional mpMRI parameters between BRCA + and BRCA− groups (mean ADC: 0.631 ± 0.153 vs. 0.675 ± 0.177 × 10^{-3} mm^2/s, p = 0.19; lesion size, PI-RADS, zonal location: all p > 0.10). For radiomic analysis (n = 98; 51 BRCA+, 47 BRCA−), the best overall model was CatBoost applied to robustness-filtered radiomic features (PR-AUC = 0.6936; ROC-AUC = 0.6612). Restricting analysis to lesions ≥ 500 voxels (n = 31) improved performance (PR-AUC = 0.778; ROC-AUC = 0.615).

Conclusion

Conventional mpMRI parameters cannot distinguish BRCA-mutated from non-mutated PCa. ADC map radiomics combined with machine learning demonstrated intermediate yet promising discriminatory performance, suggesting that quantitative imaging features may capture biologically relevant information associated with BRCA mutational status. Larger multicenter validation studies are warranted.

Clinical trial number: not applicable.

INTRODUCTION

Prostate cancer (PCa) is the second most frequently diagnosed malignancy in men worldwide, with approximately 1.4 million new cases and 375,000 deaths annually [1]. BRCA1 and BRCA2 germline and somatic mutations occur in approximately 5–12% of localized PCa and up to 25% of metastatic castration-resistant disease, conferring greater genomic instability, higher metastatic potential, and significantly worse overall survival [2, 3].

Early identification of BRCA mutation carriers is clinically critical: it informs eligibility for poly(ADP-ribose) polymerase (PARP) inhibitor therapy, guides intensified surveillance initiation—recommended from age 40 in carriers—and enables cascade germline testing of first-degree relatives [4, 5]. Current detection relies on invasive germline blood sampling and/or tissue-based somatic sequencing, both resource-intensive modalities that are not universally accessible and are temporally decoupled from imaging evaluation.

Multiparametric MRI (mpMRI) of the prostate, particularly the apparent diffusion coefficient (ADC) map derived from diffusion-weighted imaging (DWI), provides quantitative information about tumor microarchitecture that correlates with cellularity and histological aggressiveness [6]. Radiomics—the automated high-throughput extraction of quantitative imaging features—has demonstrated capacity to capture intratumoral heterogeneity reflecting underlying molecular and genomic characteristics in multiple cancer types [7, 8].

In PCa, radiomic analyses have been applied to Gleason/ISUP grade stratification, PI-RADS scoring enhancement, and biochemical recurrence prediction [9–12]. However, radiogenomic studies linking radiomic phenotypes specifically to BRCA1/2 mutational status in PCa remain scarce. We hypothesized that BRCA dysfunction—impairing homologous recombination, promoting genomic instability, and altering tumor microarchitecture—would generate ADC-map texture patterns distinguishable from wild-type tumors.

The aims of this study were: (1) to determine whether conventional mpMRI parameters discriminate BRCA-mutated from doubly negative PCa patients; and (2) to assess the discriminatory performance of machine-learning models trained on ADC map radiomic features for this task using a rigorously defined doubly negative control group.

MATERIALS AND METHODS

Study design and patient selection

This prospective cohort study with retrospective imaging analysis was conducted at the Hospital das Clínicas, Ribeirão Preto Medical School, University of São Paulo (HCFMRP-USP), a quaternary referral center, from January 2021 to July 2024. Clinical and genetic data were collected prospectively as part of routine oncological care and institutional research protocols; radiomic feature extraction and machine-learning modeling were performed retrospectively on the prospectively acquired MRI examinations. The

study received ethics committee approval (Protocol No. 4.629.838; CAAE: 40856720.8.0000.5440). Written informed consent was obtained from all participants.

Per institutional urology protocol, men with elevated PSA (> 2.5 ng/mL) and/or abnormal digital rectal examination were referred for prostate mpMRI and, when indicated, histological evaluation. The institutional Radiology and Hospital Information Systems were queried for the period January 2021–July 2024, returning 1,587 records of patients with clinical suspicion of PCa who had undergone both prostate mpMRI and histopathological evaluation. After exclusion of 833 records with benign histology or ISUP grade 1, 317 with ISUP grade 2, and 95 duplicates, 342 patients with PCa (ISUP ≥ 3) remained. Subsequent exclusion criteria were applied: (a) unsatisfactory mpMRI quality limiting interpretation (n = 32); (b) interval between MRI and histological confirmation > 6 months (n = 26); (c) mpMRI performed after a therapeutic intervention (n = 25); (d) no visible lesion or lesion < 5 mm (n = 46), yielding 213 patients. An additional 25 patients with histopathologically confirmed prostate cancer and available BRCA genetic testing from two independent prospective cohorts were incorporated, yielding a final study population of 238 patients and improving cohort representativeness.

Patient groups and critical exclusion criterion

To ensure unambiguous BRCA classification, the present analysis included exclusively: (a) BRCA+ patients—those with a confirmed pathogenic variant in BRCA1 or BRCA2 in at least one genetic pathway (germline, somatic, or both); and (b) BRCA− controls—those with documented negative results in both germline and somatic pathways (doubly negative). The 131 patients who had a negative result in only one pathway and were not tested in the other were deliberately excluded, as a negative result in one pathway does not exclude a mutation in the other, and their inclusion in the control group would risk label contamination. This selection yielded 58 BRCA + and 49 BRCA− patients (n = 107 for conventional MRI analysis) (Fig. 1).

BRCA mutation testing

Germline testing: Peripheral blood DNA was extracted using the QIAsymphony DSP DNA Mini Kit (Qiagen, Hilden, Germany) and amplified by multiple displacement amplification. Genotyping used the Illumina Global Screening Array chip processed on an iScan platform (Illumina, San Diego, CA, USA), detecting single-nucleotide polymorphisms across BRCA1/2.

Somatic testing: DNA was extracted from prostatectomy or TRUS-biopsy specimens (Maxwell system, Promega, Madison, WI, USA). Whole-exome sequencing was performed on the Illumina HiSeq 4000 platform (minimum coverage 15× in ≥ 80% of targets). Variant calling used GATK UnifiedGenotyper; functional impact was assessed by PolyPhen-2, SIFT, PROVEAN, MutationAssessor, and MutationTaster. Variants were classified against established BRCA databases.

mpMRI protocol

All examinations were performed on 1.5T or 3.0T scanners (Philips Achieva, Best, Netherlands) using a 16-channel body coil. The standardized protocol included axial T1-weighted, multiplanar T2-weighted (T2W), DWI (b-values: 0, 250, 500, 1000, 1500 s/mm²), ADC map reconstruction, and dynamic contrast-enhanced (DCE) sequences after intravenous gadopentetate dimeglumine (0.1 mmol/kg). Examinations were performed and interpreted before any therapeutic intervention.

Conventional MRI parameters

Five parameters were assessed for each patient's index lesion by a radiologist with 10 years of prostate MRI experience: (a) maximum lesion diameter (mm) on DWI; (b) mean ADC value ($\times 10^{-3}$ mm²/s)—single measurement for lesions ≤ 1 cm, mean of three measurements for lesions > 1 cm; (c) dominant lesion zonal location; (d) PI-RADS v2.1 category; and (e) prostate volume (cm³). A second radiologist with 20 years of abdominal imaging experience independently evaluated a random 20% sample to assess reproducibility.

Radiomic analysis pipeline

ADC maps were exported in DICOM format to 3D Slicer (v5.5.0; slicer.org). The index lesion was manually segmented slice-by-slice on axial ADC maps, generating a three-dimensional volume of interest (VOI) (Fig. 2). Feature extraction used the open-source PyRadiomics library (Python; van Griethuysen et al., 2017) with SimpleITK preprocessing, generating 107 radiomic features per patient encompassing first-order statistics (n = 18), shape descriptors (n = 14), gray-level co-occurrence matrix (GLCM, n = 24), gray-level run-length matrix (GLRLM, n = 16), gray-level size zone matrix (GLSZM, n = 16), gray-level dependence matrix (GLDM, n = 14), and neighboring gray-tone difference matrix (NGTDM, n = 5). Additionally, 14 quantitative ADC and morphological descriptors were computed (mean, median, percentiles, IQR, skewness, kurtosis, entropy, volume, surface area, sphericity).

Robustness filtering: Each VOI was perturbed by discrete erosion and dilation to simulate segmentation variability. Features were retained if they achieved a concordance coefficient ≥ 0.80 across original and perturbed masks. This reduced the feature set from 107 to 33 stable descriptors (the 'robustness-filtered radiomic block'). Nine patients (7 BRCA+, 2 BRCA−) were excluded from radiomic analysis due to low ADC map signal-to-noise ratio, yielding a radiomic cohort of 98 patients (51 BRCA+, 47 BRCA−).

Machine-learning modeling strategy

Three predictor blocks were defined: (1) 14 quantitative ADC and morphological descriptors; (2) 33 robustness-filtered radiomic features; (3) combined block (n = 47). Stratified k-fold cross-validation (5 external folds, 3 internal folds for hyperparameter tuning) was applied throughout. Three successive modeling phases were conducted: Phase 1 (benchmark)—logistic regression and histogram-based

gradient boosting; Phase 2—12 algorithm families (including random forests, extremely randomized trees, support vector machines, XGBoost, CatBoost, k-nearest neighbors, Gaussian Naïve Bayes, and multilayer perceptron); Phase 3 (broad search)—25 algorithm families with systematic grid-search hyperparameter tuning, adding penalized regressions, ridge classifier, polynomial/radial SVM variants, online linear learners, passive-aggressive classifier, perceptron, linear/quadratic discriminant analysis, AdaBoost, bagging with trees, and Gaussian process.

Statistical analysis

Baseline clinical, laboratory, and pathological characteristics were compared between the BRCA-positive and doubly negative control groups using the Mann-Whitney U test for continuous variables and the chi-squared or Fisher's exact test for categorical variables, as appropriate. Univariable and multivariable logistic regression analyses were performed to identify independent predictors of BRCA mutation status; variables with $p < 0.20$ on univariable analysis were included in the multivariable model. Results are reported as odds ratios (OR) with 95% confidence intervals (95% CI). Statistical significance was set at $p < 0.05$.

For the radiomic analysis, volumetric segmentation of the dominant lesion was performed on ADC maps using 3D Slicer (v5.5.0). Radiomic features were extracted using PyRadiomics (Python), with image loading and pre-processing handled by SimpleITK and numerical manipulation performed with NumPy. The extraction encompassed first-order statistics, shape descriptors, and texture matrices (GLCM, GLRLM, GLSZM, GLDM, and NGTDM), yielding 107 features per patient. Feature stability was assessed through morphological perturbation (erosion and dilation of the region of interest), and only features with an intraclass correlation coefficient ≥ 0.80 were retained, reducing the set to 33 robust descriptors.

Machine-learning models were evaluated using stratified nested cross-validation with five outer folds and two to three inner folds for hyperparameter tuning; 25 algorithm families were compared across three predictor blocks (direct ADC/shape quantitative descriptors, robustness-filtered radiomic features, and their combination). Probabilistic calibration using a sigmoid function was applied within the cross-validation loop in the initial benchmark phase. The primary performance metric was the precision-recall area under the curve (PR-AUC), selected for its informativeness in positive-class identification. Secondary metrics included ROC-AUC, Brier score, sensitivity, specificity, balanced accuracy, Matthews correlation coefficient, positive predictive value, negative predictive value, and F1-score. Operating thresholds were defined on the training fold targeting 85% sensitivity and subsequently applied to the corresponding test fold. All analyses were performed in Python using scikit-learn and PyRadiomics.

English translation and language editing assistance were provided using Claude (Anthropic), a large language model (LLM). The content was reviewed and verified by the authors.

RESULTS

Patient characteristics

The primary analysis cohort comprised 107 patients (BRCA+: n = 58; BRCA−: n = 49). The only statistically significant difference between groups was age: BRCA+ patients were older (mean 72.4 ± 7.9 years) than BRCA− controls (68.9 ± 9.3 years; p = 0.038). PSA, PSA density (dPSA), and ISUP grade distribution did not differ significantly (all p > 0.10; Table 1).

Table 1
Clinical, laboratory, and pathological characteristics of BRCA + and BRCA− groups (n = 107)

Parameter	BRCA− (n = 49)	BRCA+ (n = 58)	p value
Age, years (mean ± SD)	68.9 ± 9.3	72.4 ± 7.9	0.038*
PSA, median (IQR), ng/mL	11.4 (7.2–29.0)	11.1 (6.0–55.9)	0.69
PSA density, median (IQR), ng/mL/cm³	0.36 (0.16–0.70)	0.30 (0.15–0.74)	0.53
ISUP grade, n (%)			0.12
Grade 1	2 (4.1%)	4 (6.9%)	
Grade 2	3 (6.1%)	11 (19.0%)	
Grade 3	19 (38.8%)	12 (20.7%)	
Grade 4	9 (18.4%)	14 (24.1%)	
Grade 5	16 (32.6%)	17 (29.3%)	
**Statistically significant (p < 0.05). SD standard deviation; IQR interquartile range.*			

Conventional MRI parameters

No significant differences were detected between BRCA + and BRCA− groups for dominant lesion size (21.7 ± 16.5 vs. 19.4 ± 11.5 mm; p = 0.42), prostate volume (48.9 ± 36.7 vs. 39.7 ± 15.8 cm³; p = 0.11), zonal location (p = 0.47), mean ADC (0.631 ± 0.153 vs. 0.675 ± 0.177 × 10^{-3} mm²/s; p = 0.19) (Fig. 3), or PI-RADS category (p = 0.91; Table 2, Fig. 4).

Table 2
Conventional mpMRI parameters for BRCA − and BRCA+ groups

Parameter	BRCA− (n = 49)	BRCA+ (n = 58)	p value
Dominant lesion size, mm (mean ± SD)	19.4 ± 11.5	21.7 ± 16.5	0.42
Prostate volume, cm^3 (mean ± SD)	39.7 ± 15.8	48.9 ± 36.7	0.11
Mean ADC, $\times 10^{-3}$ mm^2/s (mean ± SD)	0.675 ± 0.177	0.631 ± 0.153	0.19
Zonal location, n (%)			0.47
Peripheral zone	35 (71.4%)	42 (72.4%)	
Transition zone	3 (6.1%)	1 (1.7%)	
Both zones	11 (22.5%)	15 (25.9%)	
PI-RADS category, n (%)			0.91
PI-RADS 3	5 (10.2%)	5 (8.6%)	
PI-RADS 4	18 (36.7%)	20 (34.6%)	
PI-RADS 5	26 (53.1%)	33 (56.8%)	

mpMRI multiparametric magnetic resonance imaging; ADC apparent diffusion coefficient; SD standard deviation.

Logistic regression confirmed that none of the conventional MRI or clinical parameters independently predicted BRCA mutational status. In multivariate analysis, age (OR 1.039, 95%CI 0.987–1.094, p = 0.13), prostate volume (OR 1.009, 95%CI 0.991–1.028, p = 0.29), and mean ADC (OR 0.990, 95%CI 0.996–1.009, p = 0.18) all failed to reach significance (Table 3).

Table 3
Logistic regression analysis for prediction of BRCA mutational status

Variable	Odds Ratio	95% CI	p value
Univariate analysis			
Age	1.049	1.001–1.100	0.04*
PSA	1.000	0.998–1.002	0.71
PSA density	0.986	0.907–1.073	0.75
ISUP grade	0.79	0.55–1.13	0.20
Lesion size	1.011	0.984–1.039	0.42
Prostate volume	1.013	0.996–1.030	0.12
PI-RADS	1.111	0.275–4.470	0.88
Mean ADC	0.998	0.995–1.000	0.19
Multivariate analysis			
Age	1.039	0.987–1.094	0.13
Prostate volume	1.009	0.991–1.028	0.29
Mean ADC	0.990	0.996–1.009	0.18

**Significant in univariate analysis but not multivariate. CI confidence interval; ADC apparent diffusion coefficient.*

Radiomic feature extraction

Feature extraction was successfully completed in all 98 patients (0% failure rate). After robustness filtering, 33 of 107 radiomic features were retained. The cohort showed near-equal class distribution (52.0% BRCA+), avoiding class imbalance. Lesion size distribution by voxel count is shown in Table 4; 31 patients (31.6%) had VOIs ≥ 500 voxels.

Table 4
Distribution of segmented VOI sizes and BRCA status

VOI size (voxel threshold)	Total (n)	BRCA+ (n)	BRCA− (n)	%
All lesions	98	51	47	100%
≥ 100 voxels	78	40	38	79.6%
≥ 300 voxels	45	26	19	45.9%
≥ 500 voxels	31	18	13	31.6%

VOI volume of interest.

Machine-learning model performance

Table 5 summarizes progressive performance gains across the three modeling phases (full cohort, n = 98). Phase 1 (benchmark) yielded a maximum PR-AUC of 0.565 and ROC-AUC of 0.450 (logistic regression, combined block). Phase 2 (12-family expansion) elevated performance to PR-AUC 0.689 and ROC-AUC 0.638 (CatBoost, quantitative ADC block). Phase 3 (25-family broad search) achieved the best overall result: CatBoost applied to the robustness-filtered radiomic-only block (PR-AUC = 0.6936; ROC-AUC = 0.6612; Brier score = 0.260; Table 5, Table 6, Fig. 5).

Table 5
Progressive performance across three modeling phases (all lesions, n = 98)

Modeling phase	Best configuration	PR-AUC	ROC-AUC	Brier
Phase 1: Benchmark (2 algorithms)	Logistic regression + combined block	0.565	0.450	0.268
Phase 2: Expanded (12 algorithms)	CatBoost + ADC-quantitative block	0.689	0.638	0.299
Phase 3: Broad search (25 algorithms)†	CatBoost + radiomic-only block	0.694	0.661	0.260

Best overall result. PR-AUC area under precision-recall curve; ROC-AUC area under receiver operating characteristic curve.

Table 6
Top 10 model combinations – Phase 3 broad search (all lesions, n = 98)

Rank	Feature block	Algorithm	PR-AUC	ROC-AUC
1†	Radiomic-only (33 robust features)	CatBoost	0.694	0.661
2	ADC-quantitative (14 features)	Elastic net LR	0.693	0.609
3	ADC-quantitative (14 features)	Ridge classifier	0.689	0.611
4	Combined block (47 features)	k-nearest neighbors	0.683	0.635
5	ADC-quantitative (14 features)	L2-penalized LR	0.675	0.597
6	ADC-quantitative (14 features)	L1-penalized LR	0.673	0.596
7	ADC-quantitative (14 features)	Gaussian Naïve Bayes	0.667	0.593
8	Combined block (47 features)	Gaussian process	0.667	0.598
9	ADC-quantitative (14 features)	Linear discriminant analysis	0.663	0.588
10	ADC-quantitative (14 features)	Quadratic discriminant analysis	0.659	0.544

Best overall result of the study. LR logistic regression.

When analysis was restricted to lesions ≥ 500 voxels (n = 31), performance improved substantially (random forest, ADC-only block: PR-AUC = 0.778; ROC-AUC = 0.615; Table 7, Fig. 6). This analysis confirmed that small VOIs contribute disproportionate noise to matrix-based texture descriptors.

Table 7
Best model performance by VOI size threshold (sensitivity analysis)

VOI threshold	n	Best model configuration	PR-AUC	ROC-AUC	Brier
All lesions	98	CatBoost + radiomic-only	0.694	0.661	0.260
≥ 100 voxels	78	Extra-trees + radiomic block	0.615	0.568	0.247
≥ 300 voxels	45	Extra-trees + radiomic block	0.693	0.567	0.259
≥ 500 voxels	31	Random forest + ADC-only block	0.778	0.615	0.239

VOI volume of interest; PR-AUC precision-recall area under curve; ROC-AUC receiver operating characteristic area under curve.

DISCUSSION

Our study investigated the ability of ADC map radiomic features combined with machine learning to discriminate BRCA1/2-mutated from doubly negative PCa patients. The principal findings were: (1)

conventional mpMRI parameters—including mean ADC, lesion size, PI-RADS category, and zonal location—failed to discriminate BRCA+ from BRCA− patients; (2) radiomic analysis yielded intermediate but methodologically consistent discriminatory performance (best ROC-AUC = 0.661; PR-AUC = 0.694), with CatBoost applied to robustness-filtered features outperforming all other configurations; and (3) performance improved substantially when analyses were restricted to lesions with ≥ 500 voxels, highlighting the critical influence of VOI size on radiomic reliability.

The inability of conventional MRI parameters to differentiate BRCA-mutated from non-mutated PCa is biologically plausible. Although BRCA1/2 mutations promote genomic instability and more aggressive phenotypes [2, 3], the resulting macroscopic alterations—cell density, necrosis, extracellular matrix reorganization—may not be sufficiently pronounced to generate distinct conventional imaging signatures. This is consistent with prior reports demonstrating that PSA, histological grade, and clinical stage are equally insufficient to identify BRCA mutation carriers [13, 14]. Our logistic regression confirmed this: even age, which showed marginal univariate significance ($p = 0.038$), lost significance in multivariate analysis, indicating confounding rather than true independent prediction.

The radiogenomic hypothesis of this study rests on a mechanistic link between BRCA1/2 dysfunction and ADC-measurable microarchitectural changes. Loss of homologous recombination capability leads to accumulation of double-strand DNA breaks, chromosomal rearrangements, and oncogenic pathway activation [15], which may modify nuclear-to-cytoplasmic ratios, intracellular water content, and extracellular matrix composition—all of which influence water molecule diffusion and therefore ADC-derived texture patterns. Descriptors capturing spatial heterogeneity (GLCM entropy, contrast, homogeneity; GLRLM and GLSZM features) would theoretically reflect these microarchitectural differences better than scalar ADC values [16, 17].

The superiority of the radiomic-only block over the combined block merits discussion. In moderately-sized cohorts ($n = 98$), adding 14 quantitative ADC descriptors to 33 radiomic features increases dimensionality without proportionally increasing discriminative information, as first-order statistical radiomic features already implicitly encode mean ADC behavior. This 'dimensionality penalty' is well-recognized in radiomics literature and underscores the importance of feature selection and robustness filtering [18]. The CatBoost algorithm's superiority aligns with its strength in handling ordered categorical interactions and capturing complex non-linear patterns in moderate-dimensional feature spaces [19].

The strong effect of VOI size deserves emphasis for future study design. Radiomic texture features based on spatial matrices (GLCM, GLRLM, GLSZM) require sufficient voxel-to-voxel relationships for statistical stability. Small VOIs (< 100–300 voxels) generate unreliable matrix estimations, introducing noise that degrades classifier performance [20]. Our data confirm this: restricting to ≥ 500 voxels increased PR-AUC from 0.638 to 0.778—a clinically meaningful improvement—despite halving the sample size.

The study of Ogbonnaya et al. [21] represents one of the most directly comparable works in PCa radiogenomics, demonstrating correlations between biparametric MRI radiomic features and hypoxia-

related gene expression (ANGPTL4, VEGFA, P4HA1) in localized PCa with AUCs of 0.68–0.77—consistent with our findings. Other radiomic studies in PCa report ROC-AUCs of 0.65–0.85 for molecular target discrimination with comparable sample sizes [9, 12], supporting that our intermediate performance level (ROC-AUC 0.661) is appropriate for this exploratory radiogenomic task.

This study has several limitations. The relatively modest sample size and single-center design may limit the generalizability of the findings. In addition, external validation was not available and will be necessary to confirm the robustness and reproducibility of the proposed radiomic signatures. Nevertheless, the study was conducted in a well-characterized cohort and provides preliminary evidence supporting the feasibility of non-invasive BRCA phenotyping using ADC-based radiomics.

The strengths of this study include the use of a rigorously defined doubly negative control group, comprehensive genetic characterization, successful radiomic feature extraction in all eligible cases, and a robust machine-learning framework incorporating reproducibility-based feature selection. Furthermore, the observed influence of lesion volume on model performance provides a clinically relevant methodological insight for future radiogenomic studies in prostate cancer.

Taken together, these findings suggest that ADC-based radiomics may capture biologically relevant information associated with BRCA mutational status that is not reflected by conventional MRI parameters, supporting the continued development of imaging-based biomarkers for precision oncology in prostate cancer.

CONCLUSION

Conventional mpMRI parameters are insufficient to discriminate BRCA1/2-mutated from non-mutated prostate cancer. In contrast, ADC map radiomics combined with CatBoost machine learning demonstrated promising discriminatory performance (PR-AUC = 0.694; ROC-AUC = 0.661), suggesting that quantitative imaging features may capture biologically relevant information associated with BRCA mutational status that is not reflected by conventional imaging parameters. Performance improved substantially in larger lesions (PR-AUC = 0.778 for VOIs ≥ 500 voxels). These findings justify future multicenter, prospectively designed studies with standardized MRI protocols, external validation, and multiparametric radiomic integration toward non-invasive BRCA phenotyping in prostate cancer.

Declarations

Funding: This study was supported by the São Paulo Research Foundation (FAPESP). The funders had no role in study design, data collection, analysis, interpretation, or manuscript writing.

Competing Interests: The authors declare no competing interests.

Ethics Approval: This study was approved by the Institutional Ethics Committee of HCFMRP-USP (Protocol No. 4.629.838; CAAE: 40856720.8.0000.5440) and conducted in accordance with the

principles of the Declaration of Helsinki.

Informed Consent: Written informed consent was obtained from all participants prior to enrollment.

Availability of Data and Materials: The dataset(s) supporting the conclusions of this article is/are not publicly available due to institutional data governance requirements and patient privacy protections mandated by the Institutional Ethics Committee of HCFMRP-USP, but are available from the corresponding author on reasonable request.

Acknowledgments:

The authors thank the patients who participated in this study and the staff of the HCFMRP-USP Department of Medical Imaging for technical support.

Author Contributions (CRediT taxonomy):

Aila de Menezes Ferreira: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing

Julio Cesar Nather Junior: Formal analysis, Investigation, Writing – review & editing

Lucianna Fonseca Barreto: Data curation, Writing – review & editing

Fernando Chahud: Data curation, Writing – review & editing

Daniel Lopes da Cunha: Writing – review & editing

Larissa Grazielle Souza Ribeiro: Writing – review & editing

Fernanda Coeli Lacchini: Writing – review & editing

Leandro Machado Colli: Resources (BRCA genetic analysis), Writing – review & editing

Valdair Francisco Muglia: Conceptualization, Supervision, Writing – review & editing

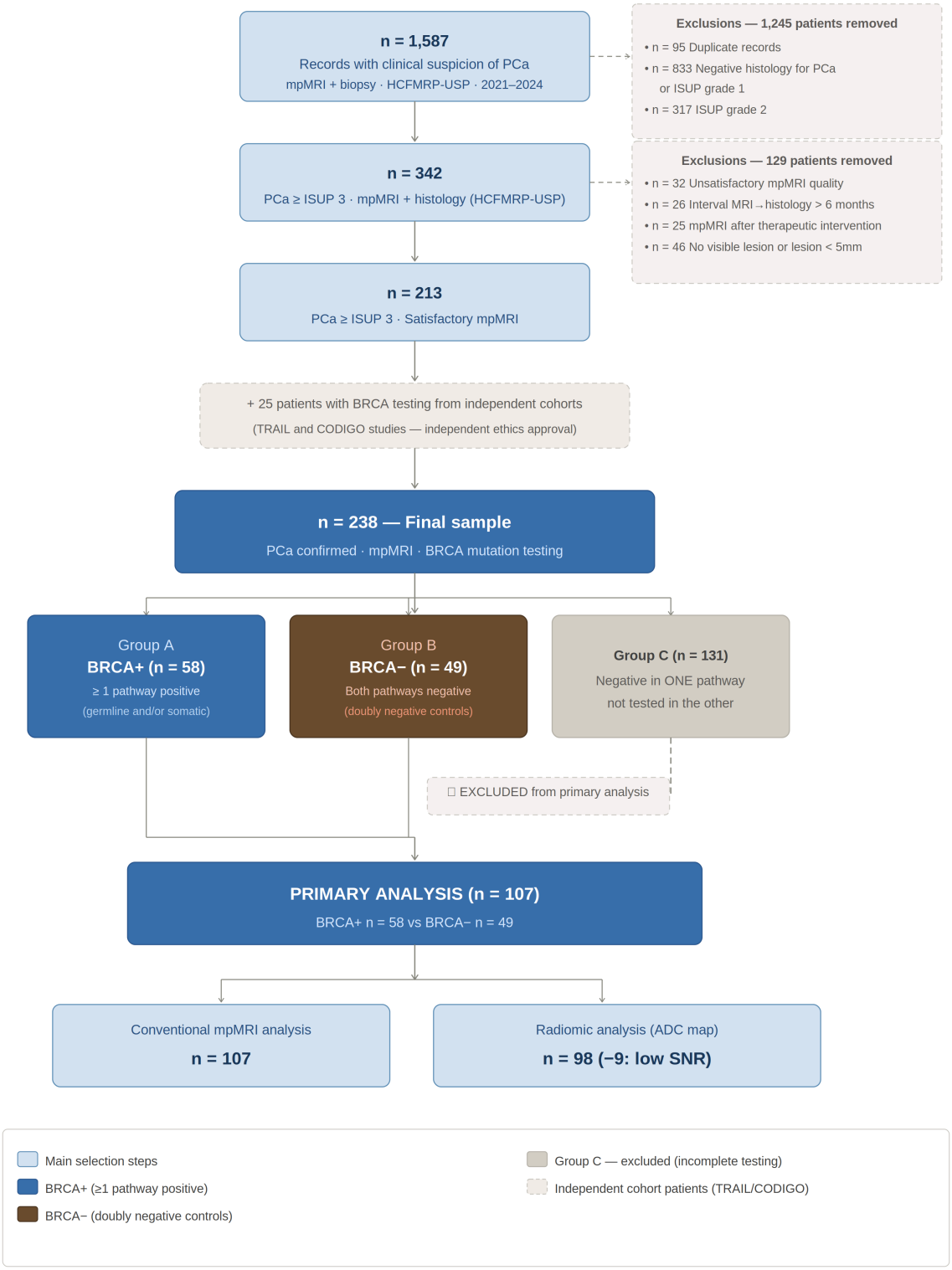
References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024;74(1):12–49.
2. Pritchard CC, Mateo J, Walsh MF, et al. Inherited DNA-repair gene mutations in men with metastatic prostate cancer. N Engl J Med. 2016;375(5):443–53.
3. Castro E, Goh C, Olmos D, et al. Germline BRCA mutations are associated with higher risk of nodal involvement, distant metastasis, and poor survival outcomes in prostate cancer. J Clin Oncol. 2013;31(14):1748–57.

4. Parker C, Gillessen S, Heidenreich A, Horwich A. Cancer of the prostate: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2020;31(9):1119–34.
5. Shah S, Chen B. Testing for inherited susceptibility to prostate cancer: BRCA1/2 and beyond. Int J Mol Sci. 2021;22(23):12628.
6. Donati OF, Mazaheri Y, Afaq A, et al. Prostate cancer aggressiveness: assessment with whole-lesion histogram analysis of the apparent diffusion coefficient. Radiology. 2014;271(1):143–52.
7. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749–62.
8. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278(2):563–77.
9. Wibmer A, Hricak H, Gondo T, et al. Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. Eur Radiol. 2015;25(10):2840–50.
10. Nketiah G, Elschot M, Kim E, et al. T2-weighted MRI-derived textural features reflect prostate cancer aggressiveness: preliminary results. Eur Radiol. 2017;27(7):3050–9.
11. Chaddad A, Niazi T, Probst S, et al. Predicting Gleason score of prostate cancer patients using radiomic analysis. Front Oncol. 2018;8:630.
12. Min X, Li M, Dong D, et al. Multi-parametric MRI-based radiomics signature for discriminating between clinically significant and insignificant prostate cancer: cross-validation of a machine learning method. Eur J Radiol. 2019;115:16–21.
13. Gallagher DJ, Gaudet MM, Pal P, et al. Germline BRCA mutations denote a clinicopathologic subset of prostate cancer. Clin Cancer Res. 2010;16(7):2115–21.
14. Barbieri CE, Bangma CH, Bhatt DL, et al. The mutational landscape of prostate cancer. Eur Urol. 2013;64(4):567–76.
15. Zimmer K, Kocher F, Bodner J, et al. Targeting BRCA and DNA damage repair genes in GI cancers: pathophysiology and clinical perspectives. Front Oncol. 2021;11:662055.
16. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern. 1973;SMC–3(6):610–21.
17. 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):328–38.
18. van Timmeren JE, Cester D, Tanadini-Lang S, et al. Radiomics in medical imaging—'how-to' guide and critical reflection. Insights Imaging. 2020;11(1):91.
19. Spohn SKB, Bettermann AS, Bruns P, et al. Radiomics in prostate cancer imaging for a personalized treatment approach—current aspects of methodology and a systematic review on validated studies. Theranostics. 2021;11(16):8027–42.

20. Baessler B, Weiss K, Dos Santos DP. Robustness and reproducibility of radiomics in magnetic resonance imaging: a phantom study. Invest Radiol. 2019;54(4):221–8.
21. Ogbonnaya CN, Sala E, Higgins GS, et al. Radiogenomics reveals correlation between quantitative texture radiomic features of biparametric MRI and hypoxia-related gene expression in men with localised prostate cancer. J Clin Med. 2023;12(8):2605.

Figures



PCa prostate cancer; mpMRI multiparametric MRI; HCFMRP-USP Hospital das Clinicas da FMRP-USP; BRCA breast cancer susceptibility gene; SNR signal-to-noise ratio.

Figure 1

Patient selection flowchart. Starting from 1,587 records with clinical suspicion of prostate cancer (PSA > 2.5 ng/mL and/or abnormal digital rectal examination, with available MRI and histopathology), the selection process yielded 107 patients for the primary analysis: 58 BRCA+ (confirmed pathogenic variant in ≥ 1 genetic pathway) and 49 doubly negative controls (BRCA− in both germline and somatic pathways). Patients with a negative result in one pathway and not tested in the other (n = 131) were

excluded. mpMRI multiparametric magnetic resonance imaging; ISUP International Society of Urological Pathology.

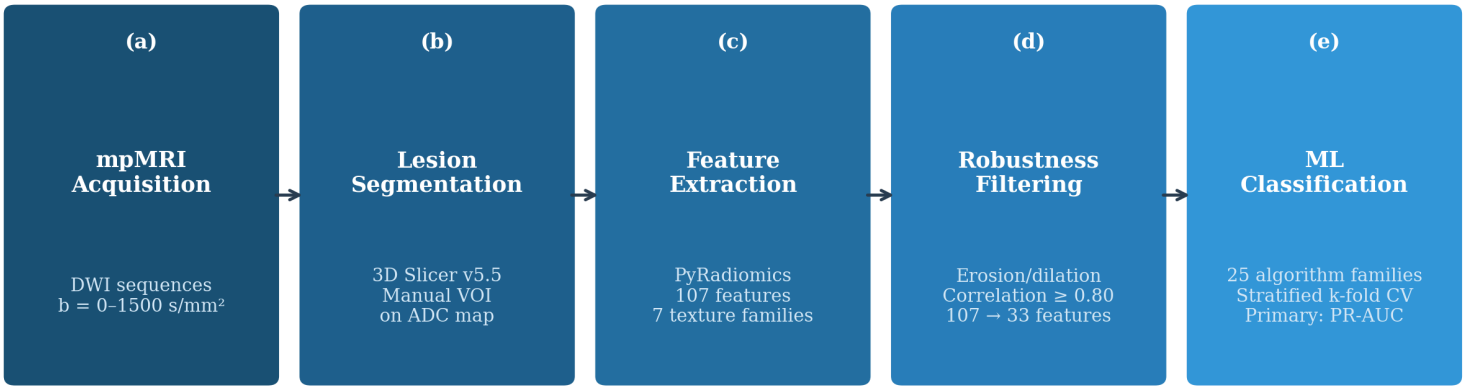


Figure 2

Radiomic pipeline applied in this study. (a) ADC map acquisition from mpMRI DWI sequences (b = 0, 250, 500, 1000, 1500 s/mm²); (b) manual 3D segmentation of index lesion on axial ADC maps in 3D Slicer; (c) feature extraction with PyRadiomics (107 original features from 7 texture families); (d) morphological robustness filtering retaining 33 stable descriptors; (e) machine-learning classification using stratified k-fold cross-validation.

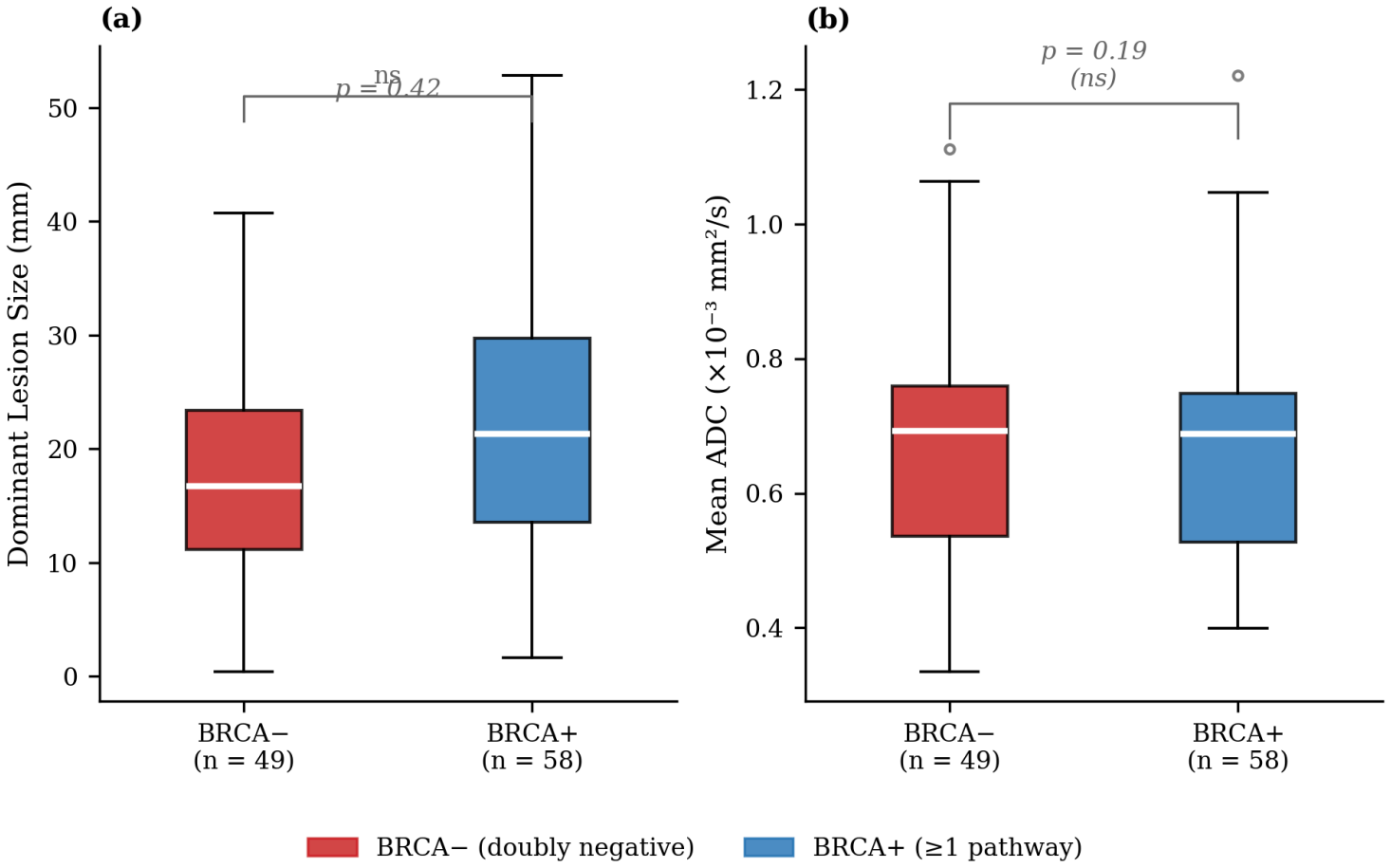


Figure 3

Box plots showing distribution of conventional mpMRI parameters between BRCA+ and BRCA− groups. (a) Dominant lesion size (mm); (b) mean ADC value (× 10^{-3} mm^2/s). No statistically significant differences were observed (p = 0.42 and p = 0.19, respectively). Horizontal lines inside boxes indicate medians; boxes represent IQR; whiskers extend to 1.5 × IQR.

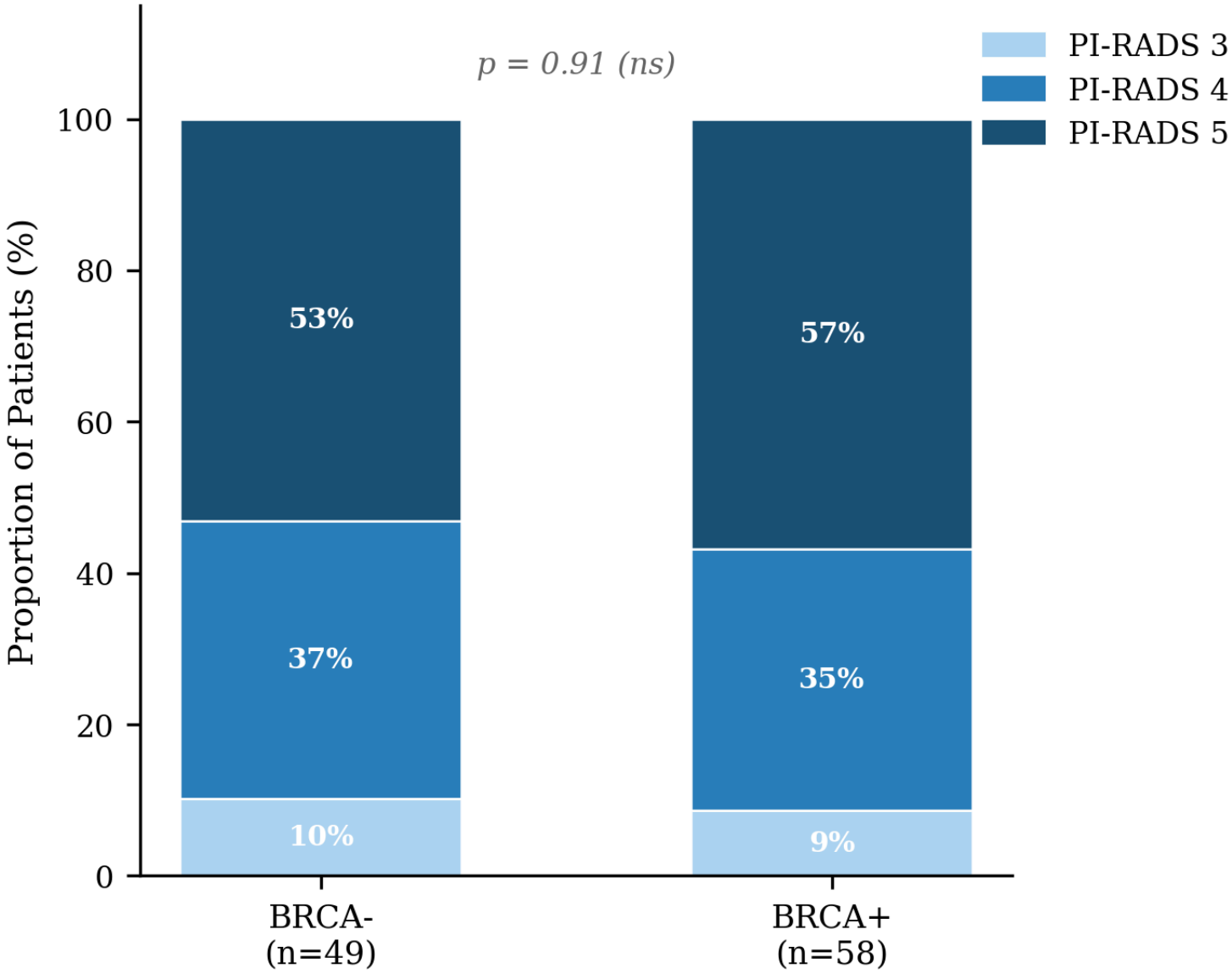


Figure 4

Stacked bar chart showing PI-RADS category distribution for BRCA+ and BRCA− groups. No significant difference was observed (p = 0.91). PI-RADS Prostate Imaging Reporting and Data System.

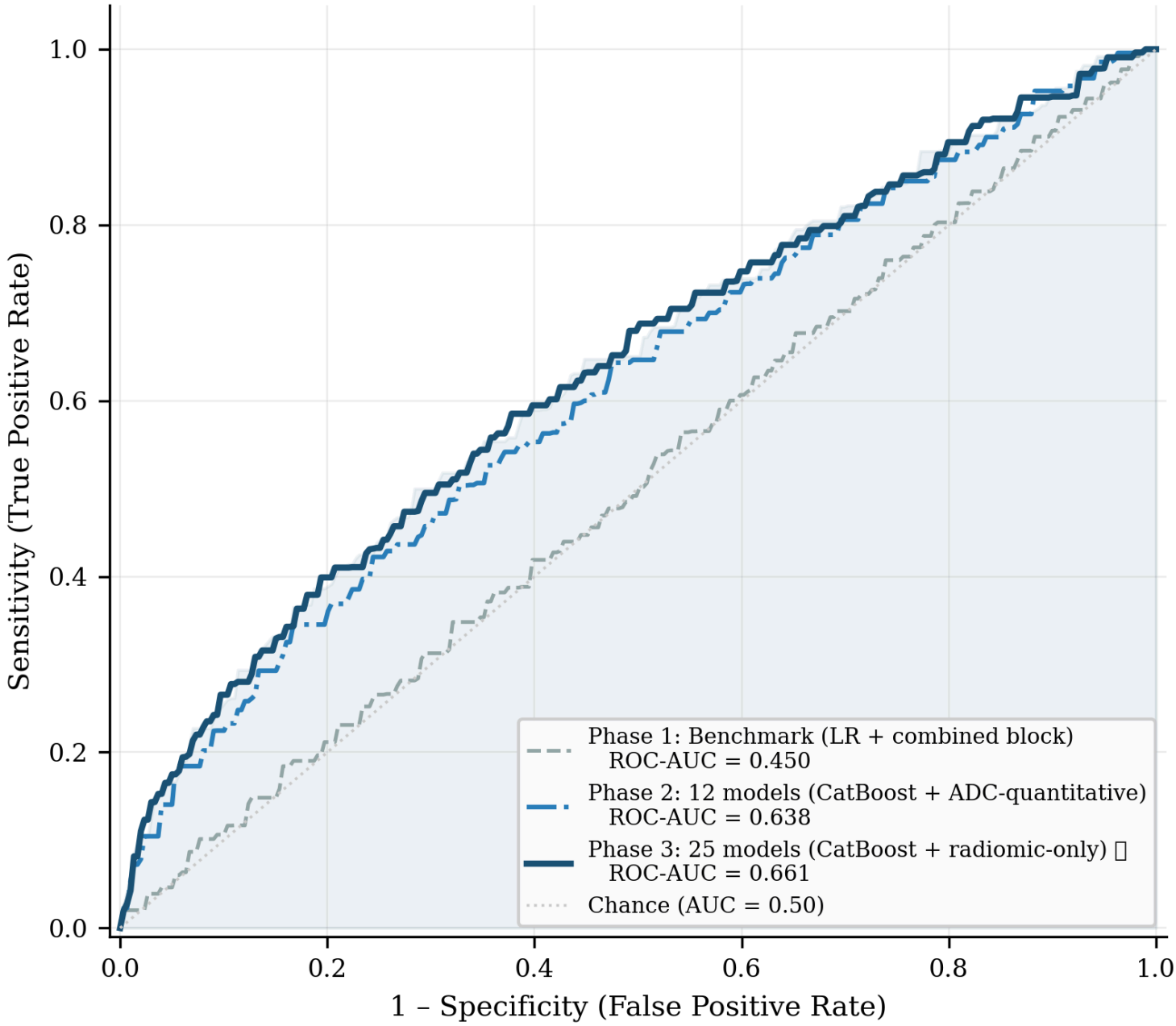


Figure 5

ROC curves for the best-performing models across three modeling phases (full radiomic cohort, n = 98). Phase 3 CatBoost with robustness-filtered radiomic features (ROC-AUC = 0.661, solid line) represents the overall best result. The diagonal dashed line represents chance-level discrimination (ROC-AUC = 0.50). ROC receiver operating characteristic; AUC area under the curve.

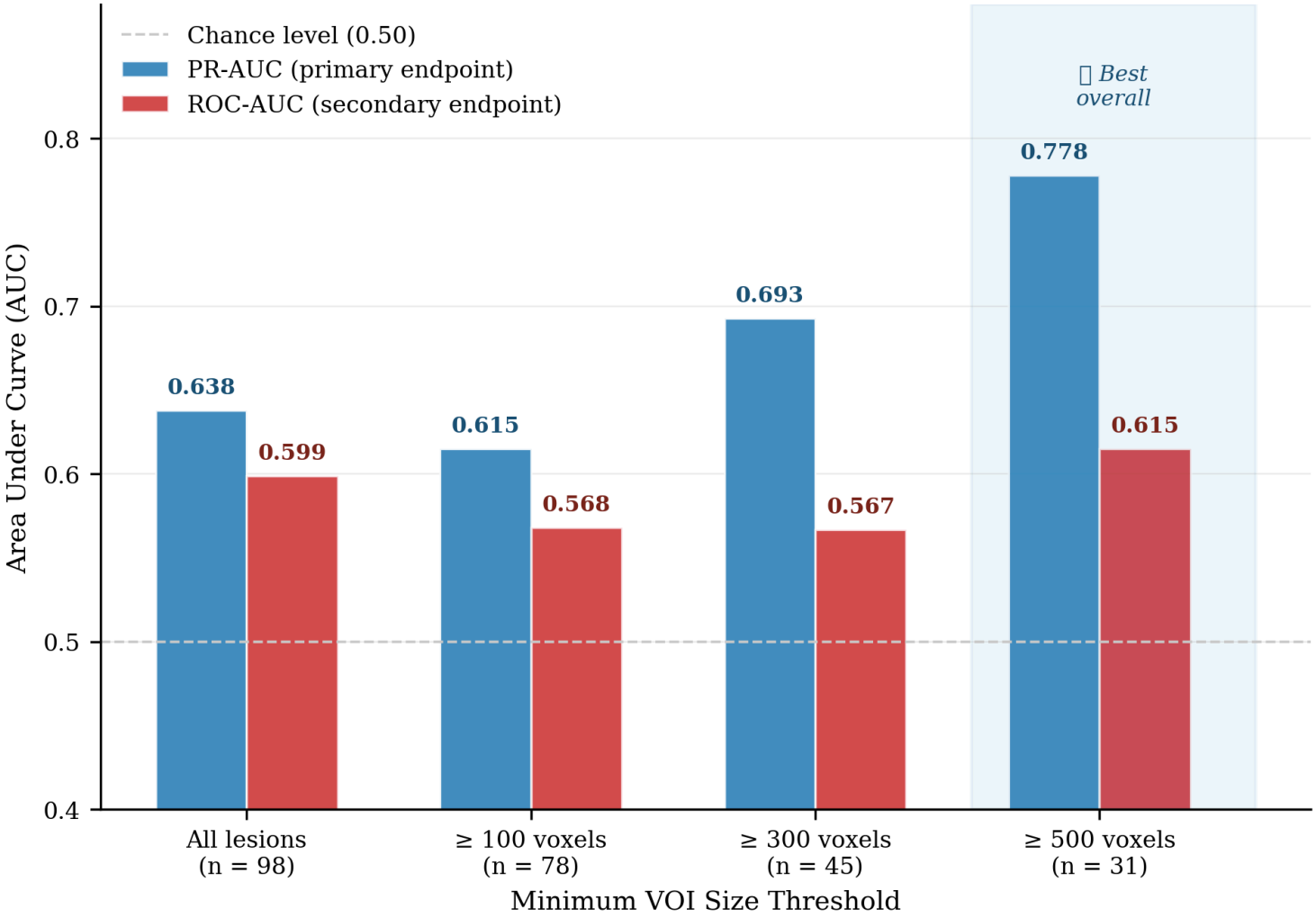


Figure 6

Effect of volume-of-interest (VOI) size on radiomic model performance. PR-AUC (primary endpoint, solid bars) and ROC-AUC (secondary endpoint, hatched bars) as a function of minimum VOI voxel threshold, demonstrating progressive performance improvement with larger lesion volumes. Error bars represent 95% confidence intervals from cross-validation.