Multicentre Evaluation of Artificial Intelligence Risk Classification for Detection of Clinically Significant Prostate Cancer on Biparametric MRI

ELECTRONIC SUPPLEMENTARY MATERIAL

SUPPLEMENTARY METHODS

Data were obtained from three independent cohorts: Site A (Germany; n = 254), Site B (The Netherlands; n = 29), and Site C (United States; n = 504). Portions of the Site A and Site B cohorts have been described previously [1, 2]. None of the imaging or pathology data included in the present study were used during development or training of the evaluated AI system.

All data were collected as part of routine clinical care and subsequently curated for retrospective analysis.

1. Ethical Concerns

The study adhered to the Declaration of Helsinki and complied with applicable data protection regulations, including the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA) in the United States. Reporting follows the STARD 2015 and CLAIM guidelines.

For the Site A cohort, institutional review board approval was granted by the Berlin Medical Association (reference Eth-28/22), and informed consent was obtained at inclusion. For the Site B cohort, approval was obtained from the St. Antonius Hospital Institutional Review Board (reference W19.165; 28 August 2019), with waiver of informed consent. For the RadNet cohort, institutional approval was obtained with waiver of informed consent due to retrospective data use.

2. Data Availability Statement

The imaging, pathology, and associated clinical data analyzed in this study contain sensitive patient-level health information and are therefore not publicly available, even in anonymized form, due to institutional, ethical, and data protection restrictions. Access to de-identified data may be considered upon reasonable request to the corresponding author, subject to approval by the relevant institutional review boards, data custodians, and data-sharing agreements. The AI software evaluated in this study is commercially available. Analysis code generated for this study may be made available upon reasonable request, where permissible and without compromising patient confidentiality or proprietary software restrictions.

3. Population

Across cohorts, eligible participants were men aged 18 years or older who underwent prostate MRI for suspicion of prostate cancer followed by biopsy or prostatectomy with available Gleason Grade Group (GG). Prior negative biopsy was not a study-wide exclusion criterion, and such men were included; prior-biopsy history was not harmonized across cohorts. Indications for biopsy were determined by referring urologists and included elevated prostate-specific antigen (PSA), abnormal digital rectal examination findings, or suspicious MRI findings (typically PI-RADS 3–5). All imaging and pathology data were clinically acquired.

Exclusion criteria common to all cohorts included missing histopathologic results, incomplete required MRI sequences, prior radical prostatectomy, major pelvic surgery, insufficient image quality, or missing metadata required for analysis.

Site A

The Site A dataset was acquired consecutively between January and August 2022 across 27 institutions (six referral centers with multidisciplinary tumor board experience and 21 private practices). Eligible patients were men with suspected prostate cancer scheduled for transperineal prostate biopsy. Exclusion criteria included absence of pre-biopsy multiparametric MRI, inadequate image quality, or absence of a documented PI-RADS score.

Site B

The Site B cohort included patients from Site B (SAZ) and Canisius Wilhelmina Hospital (CWZ). Patients underwent 3-T multiparametric MRI within 12 months prior to radical prostatectomy between September 2014 and July 2021 (SAZ) or June 2016 and December 2021 (CWZ). MRI examinations from Radboud University Medical Center were included as an external imaging source. Exclusion criteria included severe image artifacts, absence of a visible primary tumor on MRI, incomplete pathological data, or pathological stage T2 disease with positive surgical margins at the lesion side.

Site C

The RadNet cohort was a convenience sample of MRI examinations performed at 13 centers in California. Patients underwent 1.16-T to 3-T MRI with sequences sufficient for PI-RADS assessment. Exclusion criteria included insufficient image quality (e.g., motion artifacts, endorectal coil use, inadequate field of view), prior prostate cancer treatment at the time of imaging, significant post-biopsy hemorrhage or other confounding findings, inflammatory conditions, or uninterpretable examinations.

4. Image Acquisition

Across cohorts, multiparametric MRI was performed using scanners with field strengths ranging from 1.16 to 3 T from multiple vendors. Protocols included axial T2-weighted imaging, diffusion-weighted imaging (DWI) with high b-values (minimum ≥800 s/mm²), and corresponding apparent diffusion coefficient (ADC) maps. Dynamic contrast-enhanced (DCE) imaging was acquired in accordance with PI-RADS v2.1 recommendations where applicable. Detailed acquisition parameters per cohort are provided in Supplementary Table S1.

Site A examinations met minimum PI-RADS technical standards, including axial T1-weighted, axial and orthogonal T2-weighted, DWI with low and high b-values (0–100 and 800–1800 s/mm²), ADC maps, and DCE imaging with gadolinium-based contrast.

Site B imaging was performed on Siemens Magnetom Skyra and Philips Ingenia 3-T systems using torso array coils. T2-weighted and DWI sequences with ADC maps were included.

RadNet imaging included axial T2-weighted and diffusion-weighted sequences with ADC maps. Additional orthogonal T2-weighted and DCE sequences were available in selected cases.

5. Radiology Reporting

Across cohorts, MRI examinations were interpreted according to the Prostate Imaging Reporting and Data System (PI-RADS) version 2 or 2.1 by genitourinary radiologists. All reporting radiologists had substantial prior experience in prostate MRI interpretation (typically >1000 prostate MRI examinations interpreted in clinical practice). Radiologists were blinded to the AI outputs, and the evaluated AI system was not used during clinical reporting.

Lesions were assigned PI-RADS scores on the standard five-point ordinal scale. Where local staging information was available, MRI-based staging was recorded at the lesion level. Lesions classified as organ-confined (radiologic T2) with uncertain extraprostatic extension (EPE) were considered non-EPE, consistent with the original radiology report. In cohorts with prostatectomy reference standard, histopathologic correlation between MRI-visible lesions and surgical pathology was extracted from structured pathology reports, including laterality and anatomic location when EPE was present.

Site A

In the Site A cohort, radiologists reported findings according to PI-RADS v2.1 and documented up to four lesions per examination, including an index lesion. PI-RADS 3 lesions were considered suspicious in clinical practice. No computer-aided detection or diagnosis tools were used by the reporting radiologists.

Site B

In the Site B cohort, reporting was performed according to PI-RADS v2 by experienced uro-radiologists with extensive prostate MRI case volumes. MRI-based local staging was dichotomized at the lesion level, and lesion-to-pathology correlation was established using structured surgical pathology reports following radical prostatectomy.

Site C

In the Site C cohort, two reporting paradigms were present. A subset of examinations underwent a structured 2+1 adjudication procedure. In this approach, two independent expert radiologists—each having interpreted more than 100 prostate MRI examinations in the preceding year—assigned PI-RADS scores independently. A third senior adjudicating radiologist, with experience exceeding 250 prostate MRI examinations annually, subsequently reviewed each case with access to the initial PI-RADS scores and resolved discrepancies or modified scores when appropriate. This adjudicated PI-RADS score served as the reference radiologic assessment for that subset.

The remaining Site C examinations were interpreted in routine clinical practice by experienced radiologists according to PI-RADS standards.

6. Biopsy Planning and Reporting

Across cohorts, histopathologic evaluation was performed according to International Society of Urological Pathology (ISUP) modified Gleason grading criteria. For each biopsy core, primary and secondary Gleason patterns were recorded, and an ISUP Grade Group (GG) was assigned. When multiple cores were obtained from a lesion or patient, the highest GG was used to define lesion-level and patient-level clinically significant prostate cancer (csPCa; GG ≥2), consistent with contemporary diagnostic practice.

Site A

In the Site A cohort, biopsy planning and execution were performed by an experienced urologist (K.G.) with experience exceeding 3000 MRI/ultrasound fusion-guided transperineal biopsies and active participation in multidisciplinary tumor board discussions. Planning was conducted one day prior to biopsy.

First, multiparametric MRI was reviewed by the urologist using a licensed computer-aided detection (CAD) tool while blinded to the official radiology report. CAD-identified targets were exported for planning. Subsequently, the official radiology report was reviewed, and PI-RADS–reported lesions were contoured manually on MRI without additional radiologic input. All targets were integrated into the fusion biopsy system, and biopsy planning was completed solely by the urologist.

Biopsies were performed via a transperineal MRI/ultrasound fusion-guided approach. Both targeted cores and systematic sampling were obtained according to institutional protocol. In the overlapping source cohort, the median number of systematic cores was 4 (IQR, 2–5), and the median total number of histologic cores was 10. Histopathologic grading followed ISUP standards.

Site B

In the Site B cohort, MRI-targeted biopsy grades were incorporated where available. In cases in which targeted biopsies were not performed, systematic biopsy grades from the ipsilateral side of the MRI-suspected lesion were used. For patients undergoing radical prostatectomy, surgical pathology served as the reference standard, with lesion-level correlation derived from structured pathology reports. Histopathologic grading adhered to ISUP criteria. Biopsy route, core counts, and cognitive versus software-assisted fusion information were not consistently available in the retrospective source records.

Site C

For Site C, biopsy results were obtained from routine clinical records. Detailed biopsy route, core counts, and cognitive versus software-assisted fusion information were not consistently available.

7. Lesion Level Selection

For lesion-level analyses, lesion records from the clinical database (n = 3745) and AI output database (n = 2387) were matched using a hierarchical approach based on three anatomical identifiers documented in pathology reports: prostate zone (peripheral, transition, or central), laterality (right or left), and axial slice location (apex, mid-gland, or base). Matched lesions were retained to create a set of fully paired lesions. After verification of key integrity and removal of duplicate keys (none identified), 1507 lesions had matching identifiers in both datasets and constituted the paired lesion cohort. Lesions without a match in the other dataset (2238 clinical-only; 880 AI-only) were not included in lesion-level performance analyses. Among the 1507 paired lesions, 339 had missing required fields for histopathological truthing, leaving 1168 complete paired lesions for lesion-level diagnostic performance evaluation. These exclusions represented method-specific detections and incomplete truthing rather than failed patient linkage.

8. Lesion Matching Algorithm

The lesion-matching algorithm was used for 214 Site A and Site C cases in which AI lesions had not already been linked to radiology and pathology outcomes through expert-adjudication or pathology-truthing procedures.

First, a lesion-localization model assigned anatomic locations to AI-generated segmentations using a probabilistic atlas constructed from 378 labeled lesions with existing MRI/pathology annotations. Second, the resulting location probabilities were used to match AI lesions to report-derived anatomic locations using a cutoff selected from the annotated dataset. The six steps are detailed below and summarized in Supplementary Methods Table 1.

Step	Purpose
1. Prostate-centered normalization	Standardize MRI geometry across patients to enable spatial comparison despite differences in acquisition matrix, voxel spacing, and field of view.
2. Anatomical probability atlas construction	Learn the relationship between spatial prostate coordinates and anatomical region descriptors.
3. AI lesion localization	Assign anatomical descriptors to AI-detected lesions using the probability atlas.
4. Report-based descriptor extraction	Convert MRI and pathology report locations into structured anatomical descriptors.
5. Lesion matching	Determine whether pathology-confirmed lesions corresponded to radiologist-reported PI-RADS lesions or AI-detected lesions.
6. Visual review and adjudication	Confirm anatomical plausibility and resolve uncertain automated matches.

Cohort for Lesion Localization Modeling

Lesion localization modeling used cases with structured anatomical lesion annotations, MRI, and segmentation data. Across the full cohort, 787 studies were available; 763 had T2-weighted geometry data, and 753 were successfully normalized for spatial analysis.

Manual lesion segmentations and pathology-positive region masks linked to anatomical location descriptors were available for 378 labeled lesions. These lesions formed the dataset for construction and validation of the anatomical probability model. Across these lesions, 38 unique anatomical region descriptors were identified. Most lesions were assigned to a single anatomical location, although a minority was associated with multiple anatomical descriptors (range, 1–12), reflecting complex or extended lesion involvement.

Baseline MRI Geometry

Substantial variability in acquisition geometry was observed across the original examinations. Among 763 T2-weighted MRI studies, 117 distinct matrix configurations were present. The median matrix size was 512 × 512 × 27 voxels (minimum, 192 × 192 × 20; maximum, 1024 × 1024 × 80).

Median in-plane voxel spacing was 0.469 mm (IQR, 0.43–0.469 mm; range, 0.195–0.833 mm) in orthogonal in-plane directions. Through-plane spacing was markedly larger, with a median of 3.0 mm (IQR, 3.0–3.0 mm; range, 1.2–5.19 mm), reflecting typical anisotropic slice acquisition in clinical prostate MRI. The median in-plane field of view was 220 mm × 220 mm (IQR, 200–240 mm; range, 140–260 mm), with craniocaudal coverage of 80.5 mm (IQR, 75–87 mm; range, 60–135 mm). This heterogeneity precluded direct voxel-wise spatial comparison across patients.

Prostate-Centered Spatial Normalization

To enable voxel-level aggregation and probabilistic modeling, all MRI examinations and associated segmentations were transformed into a standardized prostate-centered coordinate system.

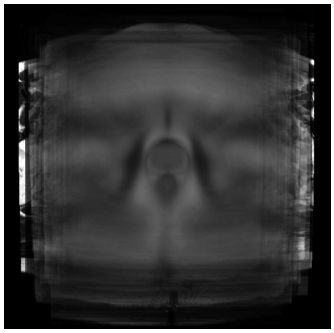
All images were first reoriented to a consistent anatomical axis convention. The prostate segmentation was used to compute the gland centroid in physical coordinates and to estimate prostate extent along each principal axis. The cohort median prostate dimensions were defined as the target reference size.

For each study, a scaling transformation was applied in physical space about the prostate centroid, independently along each axis, such that the gland dimensions matched the cohort median. Following scaling, a translation was applied so that all prostate centroids coincided at a common origin, thereby establishing a prostate-centered coordinate framework.

A single output grid was then constructed using the cohort median voxel spacing. All images were resampled onto this grid using linear interpolation for T2-weighted intensity images and nearest-neighbor interpolation for binary segmentation masks to preserve label integrity.

After normalization, all 753 processed studies shared identical geometry: matrix size 781 × 867 × 85 voxels with voxel spacing 0.469 × 0.469 × 3.0 mm, corresponding to a field of view of 366 × 406 × 255 mm. All lesion masks used for probabilistic modeling conformed to this standardized spatial configuration. Supplementary Methods Figure 1 shows the median intensity of all T2 scans in this space.

The resulting composite image illustrates the anatomical alignment achieved after normalization, with consistent prostate positioning and orientation across patients despite substantial heterogeneity in original acquisition geometry. This standardized coordinate framework enabled voxel-level aggregation and construction of the probabilistic anatomical atlas used for automated lesion localization and matching.



Supplementary Methods Figure 1. Prostate-Centered Spatial Normalization Framework

Voxel-wise median T2-weighted MRI intensity across all spatially normalized examinations displayed in prostate-centered coordinate space. All MRI volumes were reoriented, scaled about the prostate centroid to match cohort median gland dimensions, translated to a common origin, and resampled onto a unified output grid (781 × 867 × 85 voxels; voxel spacing 0.469 × 0.469 × 3.0 mm).

Lesion Volume Characteristics

Lesion volumes were calculated for the 378 labeled masks using voxel dimensions derived from image affine matrices. Median lesion volume was 1.34×10^3 mm³ (IQR, 381–3.36×10^3 mm³), with a range from 0 to 3.89×10^4 mm³, indicating substantial heterogeneity in lesion size within the modeling cohort.

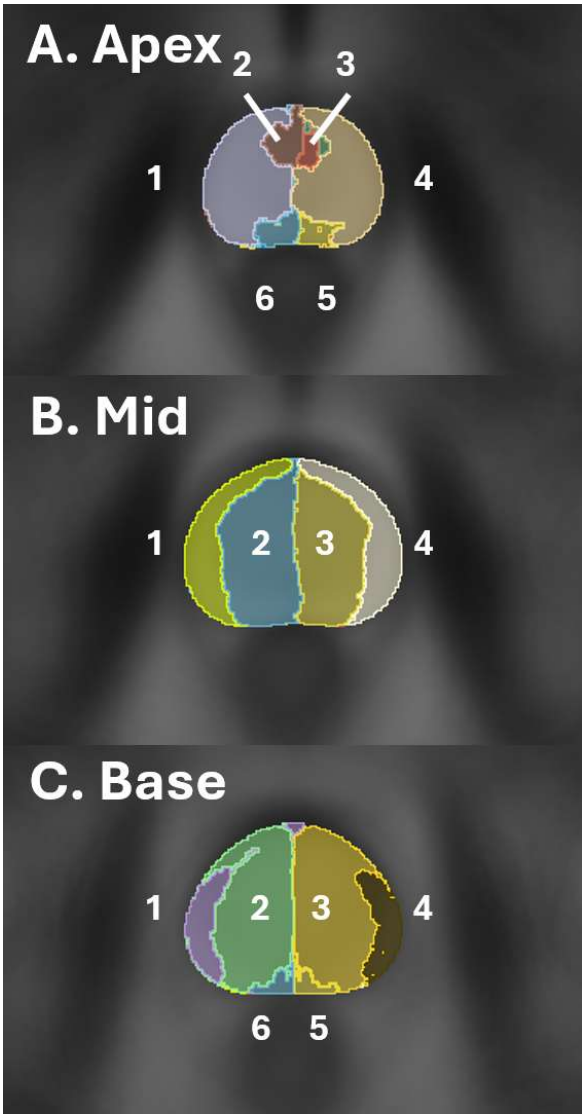
Construction of the Voxel-Wise Anatomical Probability Model

A voxel-wise probabilistic anatomical model was constructed using the 378 spatially normalized lesion masks and their associated anatomical descriptors. For each of the 38 anatomical regions, voxel-level counts were accumulated across all lesions labeled with that region. Laplace smoothing was applied to generate stable conditional probability estimates and avoid zero-probability voxels.

This process yielded a four-dimensional probability tensor representing, for each voxel within the prostate mask, the empirical probability of belonging to each anatomical region. A dominant anatomical labelmap was generated by assigning, at each voxel, the region with the highest posterior probability.

To reduce potential lateral asymmetry resulting from unequal distribution of left- and right-sided annotations, region probabilities were mirrored across the prostate midline on a slice-by-slice basis and averaged. This symmetry constraint improved spatial coherence without altering overall anatomical topology. Minor morphological smoothing was subsequently applied within the prostate mask to reduce voxel-level noise at region interfaces.

The resulting probabilistic atlas provides a continuous spatial mapping between prostate coordinates and clinically interpretable anatomical regions (e.g., apex, mid gland, base; medial and lateral subdivisions).



Supplementary Methods Figure 2. Voxel-Wise Dominant Anatomical Regions Derived from the Probabilistic Atlas

Dominant anatomical regions from the data-driven probabilistic atlas are displayed in prostate-centered space and overlaid on the cohort-average T2-weighted MRI volume. The background image represents the voxel-wise mean intensity of all spatially normalized examinations. At each voxel within the prostate mask, the anatomical region with the highest posterior probability is shown.

Three representative craniocaudal levels are depicted: (A) apex, (B) mid-gland, and (C) base. Numbered labels indicate the dominant anatomical descriptors at each level.

(A) Apex: 1 = Right Lateral Apex; 2 = Right Apex; 3 = Left Apex; 4 = Left Lateral Apex; 5 = Left Mid; 6 = Right Mid.
(B) Mid-gland: 1 = Right Lateral Mid; 2 = Right Mid; 3 = Left Mid; 4 = Left Lateral Mid.
(C) Base: 1 = Right Lateral Base; 2 = Right Base; 3 = Left Base; 4 = Left Lateral Base; 5 = Left Mid; 6 = Right Mid.

This atlas illustrates the spatial distribution of the 38 anatomical region descriptors used for automated lesion localization and hierarchical matching.

Automated Lesion Localization

For a new lesion segmentation in prostate-centered space, anatomical localization was determined by averaging voxel-wise region probabilities across all lesion voxels contained within the prostate mask. Regions were ranked according to mean probability, and the highest-ranking anatomical descriptors were returned as predicted lesion locations.

This approach incorporates the full spatial extent of the lesion rather than relying solely on lesion centroid position and is robust to irregular lesion shapes.

Internal Validation

Internal validation of the anatomical localization method was performed using ten-fold cross-validation across the 378 labeled lesions. The dataset was randomly partitioned into ten folds of approximately equal size (37–38 lesions per fold). For each iteration, the probabilistic atlas was evaluated on the held-out fold.

Performance was assessed using a hit@5 metric [3], defined as correct if at least one of the top five predicted anatomical regions matched any of the ground-truth anatomical descriptors assigned to the lesion. Across folds, hit@5 ranged from 0.737 to 0.921. The overall hit@5 across all lesions was 0.841, with a mean fold performance of 0.841 (standard deviation, 0.050), indicating consistent performance across resampled subsets.

Automated Reading of MRI and Pathology Reports

Lesion location descriptors were extracted from MRI and pathology reports using the Claude Sonnet large language model (model ID: anthropic.claude-sonnet-4-5-20250929-v1:0; Anthropic): laterality (right/left), craniocaudal level (apex/mid-gland/base or equivalent terminology), prostate zone (peripheral/transition/central), and clock-face position (X o'clock).

If prostate zone and/or craniocaudal level were not explicitly documented in the pathology report, corresponding location information was obtained from the MRI report when clearly stated. Automated extraction was quality checked by M.G.P. for unexpected and missing values.

Lesion Matching

Lesion matching was performed within each patient to determine whether a pathology-referenced lesion corresponded to (i) a radiologist-reported PI-RADS lesion or (ii) an AI-detected lesion localized using the statistical anatomical model.

For each pathology-confirmed lesion, structured anatomical descriptors were extracted, including laterality (left/right), craniocaudal level (apex, mid-gland, base), and prostate zone (peripheral, transition, central). When available, clock-face position was also recorded. These descriptors defined the ground-truth anatomical target region for that lesion.

Radiologist-reported PI-RADS lesions were matched by direct comparison of anatomical descriptors documented in the MRI report. A correspondence was accepted if the PI-RADS lesion shared concordant laterality and craniocaudal level with the pathology descriptor. When prostate zone information was available in both sources, concordance was additionally required. If multiple PI-RADS lesions satisfied these criteria within a patient, clock-face position was used to resolve ambiguity by selecting the lesion with the closest angular correspondence.

For AI-detected lesions, anatomical localization was derived from the probabilistic atlas constructed in prostate-centered space. For each AI segmentation, voxel-wise region probabilities were averaged across the lesion mask to determine the most likely anatomical region. The highest-probability region defined the AI lesion's anatomical descriptor. A pathology lesion was considered matched to an AI lesion if the dominant AI-derived region was concordant with the pathology descriptors for laterality and craniocaudal level, with zone concordance required when available. In cases with multiple candidate AI lesions satisfying descriptor concordance, the lesion whose centroid angle most closely approximated the pathology-reported clock-face position was selected.

Visual Inspection

If no radiologist-reported lesion or AI-detected lesion satisfied the required anatomical concordance criteria, the pathology lesion was classified as unmatched. All automated correspondences were subsequently visually reviewed to confirm anatomical plausibility, with final adjudication performed by an experienced genitourinary radiologist in cases of uncertainty.

This approach ensured that lesion matching was based on explicit anatomical agreement between pathology-defined location and imaging-derived descriptors, using direct descriptor comparison for radiologist-reported lesions and atlas-based localization for AI-detected lesions.

All matches produced by the automated matching procedure were visually reviewed by M.G.P. using imaging data and lesion localization summaries. In cases of uncertainty, an expert radiologist (F.G.) was consulted to confirm or revise the correspondence.

9. Data Processing

All data were pseudoanonymized at source. Analyses were performed in Python (pandas, scikit-learn, statsmodels, matplotlib, seaborn), with code version-controlled and reproducibility ensured via configuration files and seed control. Secondary and exploratory analyses were interpreted descriptively, and no adjustment for multiple comparisons was applied.

10. Statistical Analysis

The linear regression model described in the Association Between Continuous AI Risk Score and Tumor Grade subsection of the manuscript used heteroskedasticity-robust (HC3) standard errors. Findings were consistent with Kendall's tau (τ=0.237, $p = 1.19\times10^{-12}$). Higher GG was associated with higher AI risk scores ($\beta = 0.051$ per GG increase, 95% CI 0.038–0.064; $p < 0.001$).

SUPPLEMENTARY REFERENCES

1. van den Berg I, Soeterik TFW, van der Hoeven EJRJ et al (2023) The development and external validation of artificial intelligence-driven MRI-based models to improve prediction of lesion-specific extraprostatic extension in patients with prostate cancer. Cancers (Basel) 15:5452.

2. Guenzel K, Baumgaertner GL, Padhani AR et al (2024) Diagnostic utility of artificial intelligence-assisted transperineal biopsy planning in prostate cancer suspected men: a prospective cohort study. Eur Urol Focus 10:833–842.

3. Koren Y, Bell R, Volinsky C (2009) Matrix factorization techniques for recommender systems. Computer 42:30–37.

SUPPLEMENTARY TABLES

Supplementary Table S1. Scanner and Pulse Sequence Information for Each Scan Center.

Reported values represent the range observed across examinations at each site. Abbreviations: Manufacturer: Si = Siemens; GE = General Electric; Ph = Philips; Hi = Hitachi; To = Toshiba. TR = repetition time; TE = echo time; ADC = apparent diffusion coefficient; T = Tesla. Parameters marked with an asterisk (*) denote isotropic in-plane resolutions.

Property	All sites	Site A	Site B	Site C
Field strength (T)	1.16–3	1.5–3	3	1.16–3
Manufacturer	Si=320, GE=293, Ph=98, Hi=8, To=1	Si=238, Ph=15, To=1	Si=24, Ph=5	GE=292, Ph=78, Si=50, Hi=8
Number of scanner models	37	20	2	19
Structural MRI				
TR (ms)	1200–9270	1200–9130	4052.47–9270	2061–8770
TE (ms)	80–207	90–135	92–120	80–207
Flip angle (°)	90–180	90–180	90–160	90–160
In-plane pixel spacing (mm)*	0.19–0.83	0.22–0.83	0.46–0.6	0.20–0.74
Slice thickness (mm)*	1.2–3.5	1.2–3.5	2.5–3	1.5–3.5
Matrix size*	192–1024	192–896	320–432	256–1024
Diffusion MRI				
b value (s/mm²)	800–1800	800–1800	800–1400	800–1500
TR (ms)	2018–10173	3000–7200	3000–6200	2018–10173
TE (ms)	57–111.322	57–111.322	57–87.012	60.684–91
Diffusion MRI & ADC				
In-plane pixel spacing (mm)*	0.56–2.30	0.56–1.96	1.00–2.30	0.66–1.97
Slice thickness (mm)	2.5–5	3–5	2.5–3	3–4
Matrix (rows×cols)*	96–332	96–320	122–160	102–332

Supplementary Table S2. Distribution of AI Risk Categories and PI-RADS Scores According to Gleason Grade Group.

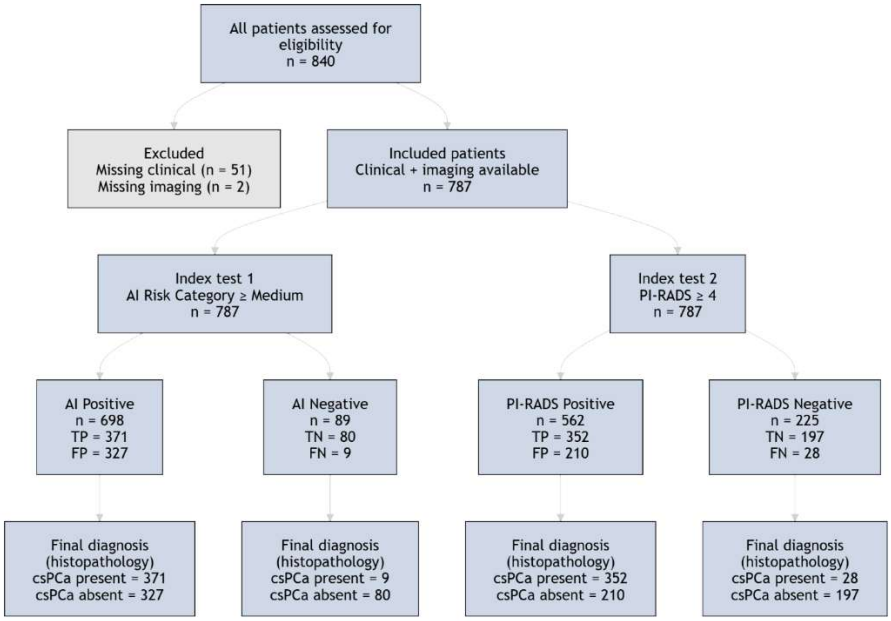
Counts and percentages of patients classified as non–clinically significant prostate cancer (non-csPCa; Gleason score <7, GG <2) and clinically significant prostate cancer (csPCa; Gleason score ≥7, GG ≥2), stratified by AI Risk Category (Low, Medium, High) and radiologist-assigned PI-RADS score (3–5). AI-category percentages are calculated within each histopathologic group. PI-RADS percentages are calculated within patients assigned PI-RADS 3–5 in each histopathologic group; patients with no index lesion (PI-RADS 2) are excluded from the PI-RADS percentage denominator.

Gleason Grade	AI Risk Category	N	Percent
non-csPCa (Gleason<7)	high	91	22.4
non-csPCa (Gleason<7)	low	80	19.7
non-csPCa (Gleason<7)	medium	236	58.0
csPCa (Gleason≥7)	high	258	67.9
csPCa (Gleason≥7)	low	9	2.4
csPCa (Gleason≥7)	medium	113	29.7

Gleason Grade	PI-RADS	N	Percent
non-csPCa (Gleason<7)	3	73	25.8
non-csPCa (Gleason<7)	4	164	58.0
non-csPCa (Gleason<7)	5	46	16.3
csPCa (Gleason≥7)	3	23	6.1
csPCa (Gleason≥7)	4	193	51.5
csPCa (Gleason≥7)	5	159	42.4

SUPPLEMENTARY FIGURES

Supplementary Figure S1. Patient-level STARD Flow Diagram.



Supplementary Figure S2. Lesion-level STARD Flow Diagram.