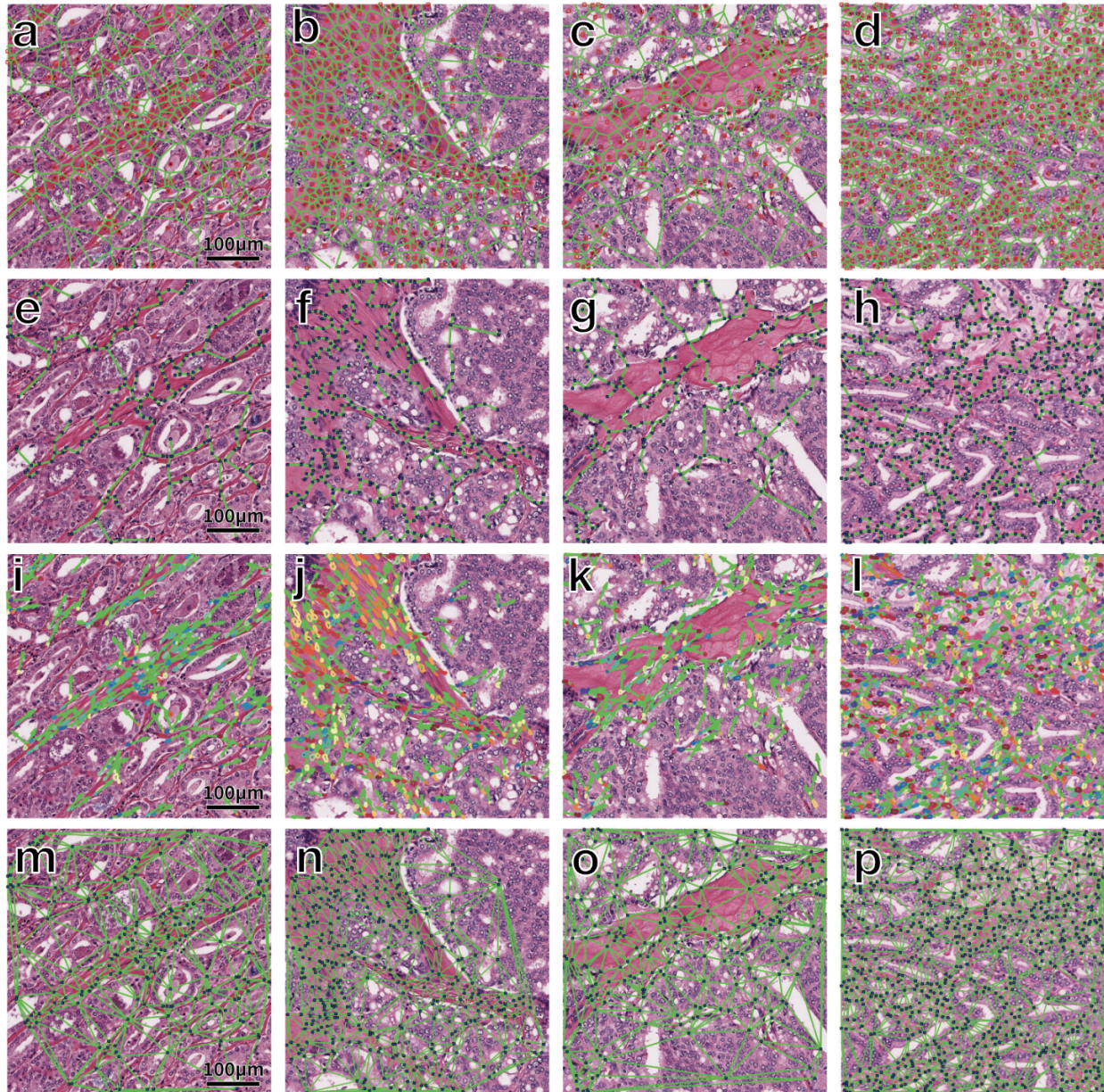


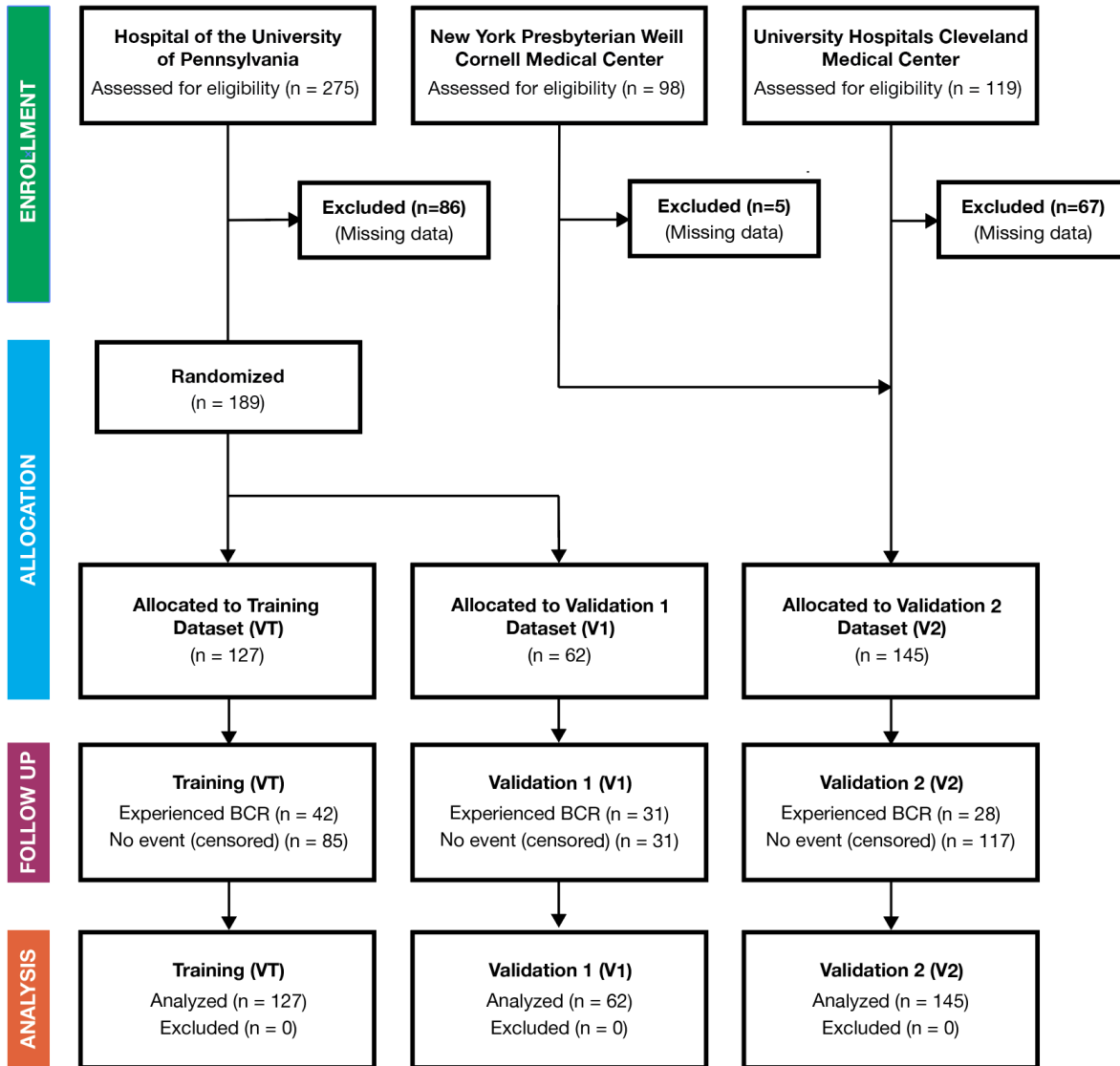
Supplementary Information

Supplementary Figure S1: Comparison of stromal histology between BCR and non-BCR patients in different populations.

Voronoi Diagrams (a-d), minimal spanning trees (e-h), directionality colormaps (i-l), and global graphs (m-p) for stromal nuclei in AA-BCR+ (leftmost column), AA-BCR- (second column from left), CA-BCR+ (third column from left), and CA-BCR- (fourth column from left) patient samples.



Supplementary Figure S2: CONSORT-style flow diagram depicting patient categorization and processing in the study.



Supplementary Table S1: Results of ML and ENC classification experiments.

For each training and validation cohort combination, the results of Random Forest and Elastic Net Cox models are reported. Statistics reported are Area Under the Receiver Operating Characteristic Curve (AUC), Hazard Ratio (HR) between predicted high and low risk groups, 95% Confidence Interval (95% CI) for the HR, and p value (p). Statistically significant results are indicated in bold.

	T _{AA} Training		T _{CA} Training		T _{AA+CA} Training	
	Random Forest	Elastic Net Cox	Random Forest	Elastic Net Cox	Random Forest	Elastic Net Cox
V1_{AA}	AUC: 0.85 HR: 3.03 (95% CI: 0.812 - 11.3) p: 0.024	AUC: 0.87 HR: 4.71 (95% CI: 1.65 - 13.4) p: 0.0027	AUC: 0.39 HR: 1.62 (0.569 - 4.63) p: 0.62	AUC: 0.48 HR: 1.74 (0.58 - 5.23) p: 0.323	AUC: 0.78 HR: 1.96 (0.65 - 5.85) p: 0.14	AUC: 0.55 HR: 0.83 (0.09 - 7.50) p: 0.86

V2_{AA}	AUC: 0.75 HR: 4.51 (95% CI: 0.925 - 22) p: 0.013	AUC: 0.77 HR: 5.70 (95% CI: 1.48 - 21.9) p: 0.014	AUC: 0.60 HR: 1.15 (0.296 - 4.47) p: 0.84	AUC: 0.56 HR: 0.31 (0.0099 - 9.74) p: 0.242	AUC: 0.68 HR: 4.56 (1.1 - 18.9) p: 0.018	AUC: 0.56 HR: 0.61 (0.05 - 7.78) p: 0.63
V1_{CA}	AUC: 0.33 HR: 0.466 (0.106 - 2.05) p: 0.45	AUC: 0.50 HR: 0.921 (0.338 - 2.51) p: 0.87	AUC: 0.56 HR: 0.987 (0.32 - 3.05) p: 0.98	AUC: 0.47 HR: 0.54 (0.17 - 1.69) p: 0.22	AUC: 0.49 HR: 1.01 (0.29 - 3.55) p: 0.99	AUC: 0.40 HR: 0.87 (0.29 - 2.58) p: 0.79
V2_{CA}	AUC: 0.44 HR: 0.546 (95% CI: 0.183 - 1.62) p: 0.33	AUC: 0.49 HR: 1.74 (0.56 - 5.38) p: 0.28	AUC: 0.44 HR: 0.832 (0.299 - 2.32) p: 0.72	AUC: 0.49 HR: Infinite p: 0.224	AUC: 0.45 HR: 0.47 (0.16 - 1.37) p: 0.22	AUC: 0.46 HR: 1.47 (0.40 - 5.43) p: 0.60
V1_{AA+C} A	AUC: 0.66 HR: 1.57 (0.56 - 4.43) p: 0.31	AUC: 0.68 HR: 1.84 (0.90 - 3.78) p: 0.082	AUC: 0.55 HR: 0.744 (0.368 - 1.51) p: 0.40	AUC: 0.46 HR: 0.84 (0.38 - 1.89) p: 0.64	AUC: 0.58 HR: 1.32 (0.53 - 3.30) p: 0.51	AUC: 0.46 HR: 0.61 (0.21 - 1.75) p: 0.26
V2_{AA+C} A	AUC: 0.71 HR: 2.10 (0.83 - 5.32) p: 0.21	AUC: 0.61 HR: 2.37 (0.99 - 5.69) p = 0.028	AUC: 0.52 HR: 1.19 (0.593 - 2.38) p: 0.63	AUC: 0.54 HR: 1.67 (0.34 - 8.21) p: 0.56	AUC: 0.53 HR: 1.14 (0.62 - 2.08) p: 0.68	AUC: 0.49 HR: 1.06 (0.325 - 3.47) p: 0.92

Supplementary Table S2: Performance of CAPRA-S and Kattan Classifiers on AA validation datasets. Statistics reported are Area Under the Receiver Operating Characteristic Curve (AUC), Hazard Ratio (HR) between predicted high and low risk groups, 95% Confidence Interval (95% CI) for the HR, and p value (p).

Model	V_{1,AA} Performance	V_{2,AA} Performance
CAPRA-S Nomogram	AUC: 0.74 HR: 2.53 (0.847 - 7.53) p: 0.059	AUC: 0.70 HR: 2.99 (0.603 - 14.8) p: 0.086
Kattan Nomogram	AUC: 0.75 HR: 2.67 (0.826 - 8.65) p: 0.045	AUC: 0.62 HR: 2.20 (0.409 - 11.8) p: 0.252
AAstroENC Classifier	AUC: 0.87 HR: 4.71 (1.65 - 13.4) p: 0.0027	AUC: 0.77 HR: 5.70 (1.48 - 21.9) p: 0.014

Supplementary Table S3: Performance of CAPRA-S and Kattan Classifiers on CA datasets. Statistics reported are Area Under the Receiver Operating Characteristic Curve (AUC), Hazard Ratio (HR) between predicted high and low risk groups, 95% Confidence Interval (95% CI) for the HR, and p value (p).

Model	V _{T,CA} Performance	V _{1,CA} Performance
CAPRA-S Nomogram	AUC: 0.90 HR: 2.86 (1.15 - 7.13) p: 0.0066	AUC: 0.92 HR: 8.19 (1.17 - 57.20) p: 3.91e-7
Kattan Nomogram	AUC: 0.82 HR: 3.18 (1.06 - 9.49) p: 0.0035	AUC: 0.92 HR: 8.19 (1.17 - 57.20) p: 3.91e-7

Supplementary Table S3: Features selected by the AAstro models.

Model	Features
AAstroML	Shape: Min / Max Fourier Descriptor 4 Shape: Mean Fractal Dimension Shape: Median Fractal Dimension Haralick: Mean Information Measure 1 Shape: Std. Deviation Variance of Distance Shape: Min / Max Distance Ratio
AAstroENC	Shape: Mean Distance Ratio Shape: Mean Fractal Dimension Shape: Std. Deviation Perimeter Ratio Shape: Median Fractal Dimension Shape: Min / Max Perimeter Ratio Shape: Min / Max Fourier Descriptor 4 CGT: Mean Tensor Correlation Sub-Graph: Number Isolated Nodes Haralick: Mean Contrast Inverse Moment Haralick: Mean Intensity Average

Supplementary Table S4. Association of stromal image features with tumor biomarkers. Pearson's correlation coefficient (PCC) and associated p value are reported. Features prognostic of BRFS for AA patients are indicated in bold.

Biomarker	Feature Name	PCC	p-value
PTEN	Shape:Mean Fourier Descriptor 4	-0.623	7.56E-03
RB (cyt.)	Shape:Mean Fractal Dimension	0.606	4.97E-04
PTEN	Shape:Std. Deviation Fourier Descriptor 3	-0.605	1.01E-02
RB (cyt.)	Shape:Mean Invariant Moment 2	-0.601	5.59E-04
PTEN	Shape:Std. Deviation Fourier Descriptor 4	-0.598	1.12E-02
PTEN	Shape:Std. Deviation Fourier Descriptor 8	-0.586	1.34E-02

PTEN	Shape:Std. Deviation Fourier Descriptor 2	-0.585	1.35E-02
PTEN	Shape:Std. Deviation Fourier Descriptor 10	-0.583	1.40E-02
PTEN	Shape:Mean Fourier Descriptor 10	0.580	1.46E-02
PTEN	Shape:Std. Deviation Fourier Descriptor 9	-0.576	1.54E-02
PTEN	Shape:Mean Fourier Descriptor 7	0.571	1.66E-02
PTEN	Shape:Mean Fourier Descriptor 8	0.565	1.81E-02
ERG	Delaunay:Triangle Area Disorder	-0.557	5.72E-03
PTEN	Shape:Mean Fourier Descriptor 1	0.552	2.15E-02
PTEN	Shape:Std. Deviation Fourier Descriptor 6	-0.549	2.24E-02
PTEN	Shape:Std. Deviation Fourier Descriptor 1	-0.547	2.29E-02
PTEN	Shape:Mean Distance Ratio	-0.536	2.65E-02
PTEN	Shape:Std. Deviation Fourier Descriptor 5	-0.535	2.70E-02
ERG	Voronoi:Area Disorder	-0.454	2.96E-02
ERG	Haralick:Mean Information Measure 1	-0.447	3.26E-02
C-MYC	Sub-Graph:Std. Deviation Edge Length	0.446	2.38E-03
RB (nuc.)	Haralick:Mean Intensity Average	-0.439	1.06E-02
ERG	Shape:Min / Max Fourier Descriptor 8	-0.438	3.64E-02
C-MYC	Sub-Graph:Kurtosis Edge Length	-0.422	4.37E-03
C-MYC	Sub-Graph:Mean Edge Length	-0.419	4.68E-03
AR	Shape:Mean Fractal Dimension	0.414	4.12E-04
RB (cyt.)	Arch:Disorder of Nearest Neighbors in A 10 Pixel Radius	-0.412	2.64E-02
AR	Shape:Mean Distance Ratio	-0.410	4.72E-04
RB (nuc.)	Arch:Std. Deviation Nearest Neighbors in A 20 Pixel Radius	0.404	1.97E-02

Supplementary Table S5: Overview of 242 features extracted.

Feature class	Derived attributes
---------------	--------------------

Voronoi Tessellation (12 features)	Number of nodes, number of edges, area, chord parameters
Delaunay Triangulation (8 features)	Side lengths, triangle geometry
Minimum Spanning Tree (4 features)	Number of nodes, edge length, degree, number of neighbors
Local Nuclear Cluster Graph (26 features)	Structure of clusters, patterning of clusters within graph
Nuclear Shape (100 features)	Nuclear area, perimeter, Fourier descriptors, invariant moments
Cell Orientation Entropy (39 features)	Contrast energy, Contrast inverse moment, Contrast average, Contrast variance, Contrast entropy, Intensity average, Intensity variance, Intensity entropy, Entropy, Energy, Correlation, 2 measures of information
Sub-Graph Features (26 features)	Number of nodes, number of edges, eccentricity, clustering coefficients
Texture Features (26 features)	Texture, edges, gradients, spots, and homogeneity of the image. Entropy, variance, and energy are calculated from the co-occurrence matrix describing how often pixels of various intensities are found near pixels of another intensity.

Supplementary Table S6: Significantly Differing Features based on Cox Proportional Hazards Regression

Cohort	Feature Name	P Value	Haz. Ratio	BCR Mean	NR Mean	% Diff.
AA	Shape: Min / Max Fourier Descriptor 4	1.577E-05	4.334E-01	-1.128	-1.026	9.409
AA	Shape: Mean Fractal Dimension	2.683E-05	3.041	2.580E-01	2.438E-01	5.684
AA	Shape: Median Fractal Dimension	1.294E-04	2.131	2.309E-01	2.164E-01	6.526
AA	Haralick: Mean Information Measure 1	1.289E-02	1.950	3.346	3.261	2.581
AA	Shape: Std. Deviation Variance Of Distance	1.456E-02	5.760E-01	1.774E-02	1.802E-02	-1.592
AA	Shape: Min / Max Distance Ratio	4.507E-02	1.619	5.936E-01	5.955E-01	-.03137
CA	Delaunay: Triangle Area Std. Deviation	5.777E-05	2.030	3.489E+01	2.145E+01	47.74
CA	CGT: Std. Deviation Tensor Contrast Energy	3.370E-04	1.845	4.631E-01	6.099E-02	153.5
CA	Sub-Graph: Number Isolated Nodes	7.278E-04	2.401	3.333E-01	1.538E-01	73.68

CA	CGT: Std. Deviation Tensor Information Measure 1	7.896E-04	1.703	5.529E-03	1.094E-03	133.9
CA	MST: Edge Length Disorder	1.081E-03	2.280	4.453E-01	4.166E-01	6.655
CA	Delaunay: Side Length Disorder	1.097E-03	2.081	4.658E-01	4.462E-01	4.290
CA	MST: Edge Length Std. Deviation	3.902E-03	1.550	2.323	1.809	24.90
CA	Shape: Mean Fourier Descriptor 2	7.965E-03	5.440E-01	1.668E-05	3.271E-05	-64.91
CA	CGT: Range Tensor Information Measure 1	8.150E-03	1.639	6.677E-02	2.066E-02	105.5
CA	Voronoi: Chord Minimum / Maximum	1.080E-02	5.201E-01	5.451E-02	5.853E-02	-7.112
CA	Arch: Disorder Of Distance To 3 Nearest Neighbors	1.115E-02	1.827	3.622E-01	3.372E-01	7.151
CA	Shape: Mean Invariant Moment 3	1.307E-02	5.627E-01	3.623E-04	3.965E-04	-9.023
CA	Delaunay: Side Length Minimum / Maximum	1.535E-02	5.467E-01	6.983E-02	8.020E-02	-13.82
CA	Voronoi: Perimeter Minimum / Maximum	1.760E-02	5.664E-01	3.282E-02	3.487E-02	-6.059
CA	Shape: Min / Max Perimeter Ratio	2.067E-02	1.707	2.228E-01	2.151E-01	3.549
CA	Shape: Mean Invariant Moment 7	2.131E-02	6.458E-01	-6.373E-10	9.628E-12	206.1
CA	Arch: Density Of Polygons	2.625E-02	4.829E-01	4.553E-02	6.173E-02	-30.21E
CA	Shape: Mean Area Ratio	2.754E-02	1.705	3.552E-01	3.454E-01	2.792
CA	Sub-Graph: Std. Deviation Edge Length	3.601E-02	1.450	6.416	6.304	1.752
CA	Voronoi: Perimeter Disorder	3.674E-02	1.641	4.578E-01	4.474E-01	2.296
CA	Shape: Mean Invariant Moment 1	4.198E-02	6.477E-01	1.028E-01	1.052E-01	-2.293
CA	Haralick: Std. Deviation Information Measure 2	4.338E-02	1.446	1.207E-01	1.099E-01	9.369
AA+CA	Sub-Graph: Number Isolated Nodes	3.026E-05	1.941	2.143E-01	7.059E-02	101.0
AA+CA	Shape: Min / Max Fourier Descriptor 4	1.232E-03	6.268E-01	-1.122	-1.050	6.61
AA+CA	Delaunay: Triangle Area Std. Deviation	1.755E-03	1.524	3.095E+01	2.263E+01	31.0
AA+CA	Delaunay: Triangle Area Disorder	5.609E-03	1.622	6.611E-01	6.435E-01	2.69
AA+CA	Shape: Mean Invariant Moment 4	6.037E-03	6.019E-01	1.653E-05	1.817E-05	-9.44
AA+CA	Shape: Mean Fractal Dimension	1.092E-02	1.633	2.490E-01	2.432E-01	2.36
AA+CA	CGT: Std. Deviation Tensor Contrast Entropy	1.171E-02	1.354	9.053E-03	2.386E-03	117.0
AA+CA	Shape: Mean Invariant Moment 1	1.175E-02	6.752E-01	1.027E-01	1.046E-01	-1.83
AA+CA	Sub-Graph: Number Central Nodes	1.645E-02	1.281	1.095	1.012	7.92
AA+CA	Shape: Median Fourier Descriptor 4	1.653E-02	7.686E-01	-1.562E-06	-8.570E-08	179.0
AA+CA	Haralick: Mean Correlation	2.063E-02	6.941E-01	-1.683E-01	-1.500E-01	11.5
AA+CA	CGT: Range Tensor Energy	2.358E-02	1.293	3.512E-02	9.824E-03	113.2

AA+CA	Shape: Median Fractal Dimension	2.558E-02	1.461	2.219E-01	2.165E-01	2.48
AA+CA	CGT: Std. Deviation Tensor Intensity Variance	3.151E-02	1.279	6.353E-02	1.866E-02	109.2
AA+CA	Shape: Mean Distance Ratio	3.407E-02	1.427	5.649E-01	5.558E-01	1.62

Supplementary Table S8: Clinicopathologic features by dataset

<i>Variable</i>	<i>Subvariables</i>	<i>Dataset V_T</i>	<i>Dataset V₁</i>	<i>Dataset V₂</i>
<i>Number of Patients</i>		127	62	145
<i>Age (years)</i>		58.68 (7.03)	60.08 (7.02)	60.2 (7.66)
<i>Race</i>	<i>African-American</i> <i>Caucasian-American</i>	64 (50.4%) 63 (49.6%)	31 (50.0%) 31 (50.0%)	93 (64.1%) 52 (35.9%)
<i>Preop. PSA (ng/mL)</i>	<i>Recurrence</i> <i>Non-Recurrence</i>	12.11 (13.12) 9.54 (12.97)	10.50 (13.d14) 5.79 (13.01)	10.07 (5.43) 5.98 (5.28)
<i>Gleason Grade Group</i>	<i>6 or less</i> <i>Equal to 7</i> <i>8 or greater</i>	46 (36.2%) 69 (54.3%) 12 (9.4%)	29 (46.8%) 26 (41.9%) 7 (11.3%)	44 (35.3%) 91 (62.8%) 10 (6.9%)
<i>Pathological Stage</i>	<i>T2x</i> <i>T3x</i> <i>Either T2x or T3x</i> <i>(data unavailable)</i>	73 (57.5%) 52 (40.9%) 2 (1.6%)	35 (56.5%) 27 (43.5%) 0 (0%)	59 (40.2%) 16 (11.0%) 60 (41.4%)
<i>Time to Event (days)</i>	<i>Recurrence</i> <i>Last follow-up (non-recurrence)</i>	1019.95 (1419.36) 1728.67 (1455.13)	1237.16 (1416.10) 1859.26 (1474.48)	783.74 (1296.84) 1796.0 (1287.31)

Supplementary Table S9: Number of patients with nonzero values for each biomarker, number of patients assayed, and mean IHC H-Score for each biomarker in V1_{AA}.

Biomarker	Number Nonzero	Number of Cases	Mean H-Score
ERG	23	76	56.81
PTEN	17	76	32.38
PMSA	61	76	89.61
RACEMASE	71	76	174.24
C-MYC	44	76	37.05
AR	69	76	155.99

Ki-67	81	95	1.55
P-53	67	76	89.79
RB-Nuc	33	76	22.11
RB-Cyt	29	76	24.87