Computationally Derived Image Signature of Stromal Morphology Is Prognostic of Prostate Cancer Recurrence Following Prostatectomy in African American Patients

Hersh K. Bhargava[1], Patrick Leo[2], Robin Elliott[3], Andrew Janowczyk[2], Jon Whitney[2], Sanjay Gupta[4,5], Pingfu Fu[6], Kosj Yamoah[7], Francesca Khani[8], Brian D. Robinson[8], Timothy R. Rebbeck[9], Michael Feldman[10], Priti Lal[10], and Anant Madabhushi[2,5]

ABSTRACT

Purpose: Between 30%–40% of patients with prostate cancer experience disease recurrence following radical prostatectomy. Existing clinical models for recurrence risk prediction do not account for population-based variation in the tumor phenotype, despite recent evidence suggesting the presence of a unique, more aggressive prostate cancer phenotype in African American (AA) patients. We investigated the capacity of digitally measured, population-specific phenotypes of the intratumoral stroma to create improved models for prediction of recurrence following radical prostatectomy.

Experimental Design: This study included 334 radical prostatectomy patients subdivided into training (V_T, $n = 127$), validation 1 (V_1, $n = 62$), and validation 2 (V_2, $n = 145$). Hematoxylin and eosin–stained slides from resected prostates were digitized, and 242 quantitative descriptors of the intratumoral stroma were calculated using a computational algorithm. Machine learning and elastic net Cox regression models were constructed using V_T to predict biochemical recurrence-free survival based on these features. Performance of these models was assessed using V_1 and V_2, both overall and in population-specific cohorts.

Results: An AA-specific, automated stromal signature, AAstro, was prognostic of recurrence risk in both independent validation datasets [$V_{1,AA}$: AUC = 0.87, HR = 4.71 (95% confidence interval (CI), 1.65–13.4), $P = 0.003$; $V_{2,AA}$: AUC = 0.77, HR = 5.7 (95% CI, 1.48–21.90), $P = 0.01$]. AAstro outperformed clinical standard Kattan and CAPRA-S nomograms, and the underlying stromal descriptors were strongly associated with IHC measurements of specific tumor biomarker expression levels.

Conclusions: Our results suggest that considering population-specific information and stromal morphology has the potential to substantially improve accuracy of prognosis and risk stratification in AA patients with prostate cancer.

Introduction

Prostate cancer has the highest incidence of any cancer among males in the United States (1). Surgical resection of the prostate (radical prostatectomy) is prescribed as a curative therapy for approximately 75,000 newly diagnosed patients each year (2), while 30%–40% of patients experience biochemical recurrence (BCR) following radical prostatectomy (3, 4). Clinical decisions about the prescription of adjuvant therapy are made based on estimates of the probability of prostate cancer recurrence following surgery (5).

Increasing evidence (1) suggests that African Americans (AA) have a higher likelihood of being diagnosed with prostate cancer and may experience more aggressive disease. Compared with Caucasian American (CA) males, AA men have a 1.76-fold higher lifetime probability of developing prostate cancer, and a 2.20-fold greater chance of disease-related death (1). Recent investigations at the genomic, epigenomic, transcriptomic, and proteomic levels have suggested significant differences in the biology of AA versus CA tumors (6–9). Despite these findings, race is not considered by current BCR prognosis tools.

Numerous studies have demonstrated a role for the stroma in the pathogenesis of a number of cancers. These studies have indicated changes in stromal cell phenotypes, alterations in extracellular matrix (ECM) composition, and the presence of biomarkers similar to those observed during wound repair (10–12). A growing number of artificial intelligence–driven digital pathology studies have employed quantitative histomorphometry (QH) to precisely analyze tumor structure from scanned images (13). A recent study (14) found that quantitative histomorphometric features from tumor-adjacent benign regions were prognostic of prostate cancer biochemical recurrence, and that combining these tumor-adjacent features with tumor-specific features resulted in increased prognostic accuracy with respect to cancer recurrence. Despite these findings, stromal morphology is not explicitly considered in characterization of prostate cancer.

The objectives of this work were to evaluate whether (i) population-specific, quantitative structural information describing the prostate

[1]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, California. [2]Department of Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio. [3]Department of Pathology, Case Western Reserve University, Cleveland, Ohio. [4]Department of Urology, Case Western Reserve University, Cleveland, Ohio. [5]Louis Stokes Cleveland Veterans Administration Medical Center, Cleveland, Ohio. [6]Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, Ohio. [7]Moffitt Cancer Center & Research Institute and Department of Radiation Oncology, University of South Florida, Tampa, Florida. [8]Departments of Pathology and Laboratory Medicine and Urology, Weill Cornell Medicine, New York, New York. [9]T.H. Chan School of Public Health and Dana Farber Cancer Institute, Harvard University, Boston, Massachusetts. [10]Department of Pathology, University of Pennsylvania, Philadelphia, Pennsylvania.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Current address for H.K. Bhargava: Biophysics Graduate Program, University of California, San Francisco, San Francisco, California.

Corresponding Author: Anant Madabhushi, Case Western Reserve University, 2071 Martin Luther King Jr. Drive, Cleveland, OH 44106. Phone: 216-368-8519; Fax: 216-368-4969; E-mail: axm788@case.edu

Clin Cancer Res 2020;26:1915–23

doi: 10.1158/1078-0432.CCR-19-2659

Translational Relevance

Mounting evidence suggests that there are molecular and phenotypic differences in prostate cancer biology in African Americans (AA) and Caucasian Americans (CA). However, extant models for predicting outcome in prostate cancer (e.g., Kattan, Sharait, Swanson, and CAPRA-S nomograms) do not consider population-specific differences in the disease phenotype. Most of these models were largely trained on non-AA patients, and hence may not be appropriate for predicting disease aggressiveness in AA patients. In this work, computer vision and machine learning techniques were employed to evaluate the capacity of quantitative measurements of the intratumoral stroma computed from digitized H&E-stained slides from radical prostatectomy specimens to predict recurrence risk following surgery. Significant differences in stromal morphology were observed in samples derived from AA versus CA patients, and an AA-specific prognostic model was found to significantly outperform population-agnostic models for AA patients. In addition, stromal phenotypes identified as prognostic in AA patients correlated with IHC measured tumor biomarkers. The results of this study suggest that taking into account stromal morphology and population-specific differences could lead to more accurate risk stratification for AA patients with prostate cancer post radical prostatectomy.

cancer intratumoral stroma could be used to create prognostic models for biochemical recurrence-free survival (BRFS) following radical prostatectomy and (ii) prognostic descriptors of stromal morphology were associated with markers of tumor biology. An automated computational approach was used to calculate 242 quantitative metrics of stromal morphology. A subset of these metrics was identified as prognostic of BRFS in three subdivisions of the 127-patient training cohort—AA only ($V_{T,AA}$), CA only ($V_{T,CA}$), and AA + CA ($V_{T,AA+CA}$). Machine learning (ML) and elastic net Cox (ENC) models were trained using the stromal image features to predict BCR likelihood following radical prostatectomy. These models were validated using two independent holdout datasets V_1 ($n = 64$) and V_2 ($n = 145$). The performance of these models was compared with current clinical standards for postoperative risk prognosis, namely the Kattan and CAPRA-S nomograms. The models created in this work were found to outperform existing nomograms in both validation datasets. Finally, IHC stains for 10 biomarkers were obtained for a 95-patient AA subset, and expression levels were tested for correlation with stromal morphology descriptors. Twenty-six biomarker–image feature pairs had a significant association, three of which included stromal features prognostic of BRFS. These results suggest that there are morphologic variations specific to AA tumors that are associated with prostate cancer aggressiveness, and that quantitative characterization of tissue morphology that accounts for these variations can improve BRFS estimates.

Materials and Methods

Datasets and sample preparation

A total of 334 patients with prostate cancer who underwent radical prostatectomy were identified for inclusion in this study. A CONSORT-style flow diagram describing the handling of patient information is provided in Supplementary Fig. S2. The surgical procedures and sample preparation were conducted at three different institutions. Resected prostates were fixed overnight in formalin, serially sectioned, and entirely submitted in quadrants. Hematoxylin and eosin–stained (H&E) slides were prepared using the formalin-fixed, paraffin-embedded tissue. Each case was reviewed by a genitourinary pathologist to select a single representative slide. The slides were then scanned using a whole-slide scanner. The images were reviewed by a pathologist, and a single representative cancerous region was digitally annotated in each image. This annotated region was used for the quantitative histomorphometry experiments in this study. An example annotation is shown in **Fig. 1A**.

The patients were divided into three cohorts: V_T (training, $n = 127$), V_1 (validation 1, $n = 62$), and V_2 (validation 2, $n = 145$). V_T was used for feature discovery and model training, and V_1 and V_2 were used for independent holdout validation of prognostic model performance. Dataset V_2 was obtained for additional validation following the experiments conducted using V_T and V_1. All datasets were approximately class balanced between AA and CA patients. Patient race was self-reported. V_T and V_1 were also approximately class balanced with respect to patients who experienced BCR versus those who did not. Selected demographic, clinical, and pathologic features of the entire patient dataset are provided in **Table 1**. Patient features for each dataset studied are presented in Supplementary Table S8.

V_T ($n = 127$) and V_1 ($n = 62$) consisted of samples collected at the Hospital of the University of Pennsylvania (Philadelphia, PA). The slides corresponding to these samples were scanned at 40× magnification using an Aperio Scanscope Whole-Slide Scanner (Leica) at the Department of Pathology, University of Pennsylvania (Philadelphia, PA). V_2 consisted of samples collected at University Hospitals Cleveland Medical Center (UHCMC; $n = 70$) and at New York Presbyterian Weill Cornell Medical Center (NYP, New York, NY; $n = 75$). Samples prepared at UHCMC were scanned at 40× magnification on a Zeiss Axio Scan.Z1 Slide Scanning Microscope (Zeiss). The samples prepared at NYP were scanned at 40× magnification using an Aperio Scanscope Whole-Slide Scanner (Leica) at the Department of Pathology and Laboratory Medicine, Weill Cornell Medicine (New York, NY). An example of a digitized H&E slide image is shown in **Fig. 1A** and **B**.

All slides and patient records were gathered in accordance with U.S. Common Rule guidelines in protocols approved by the institutional review board at the respective institution. The need for written consent from participants was waived because of the use of retrospective data.

Nuclear and stromal detection and segmentation

Nuclei and stroma were segmented using a previously developed deep learning method based on convolutional neural networks (15). The outputs of the deep learning models were confidence maps that represented the probability that each pixel in the image was part of a nucleus and that it belonged to the stroma (**Fig. 1C** and **D**). A confidence threshold was determined by inspection of image data from V_T and applied in conjunction with maximum and minimum size thresholds to convert the confidence map to a binary matrix. Closed shapes corresponding to nuclear or stromal borders were traced from the binary maps to yield the final matrix of boundary coordinates.

Image feature calculation

Using the boundary coordinates of stromal nuclei and of the stromal compartment (the output of the segmentation process), 242 quantitative histomorphometric (QH) image features were calculated for each patient. The extracted features included metrics derived from the

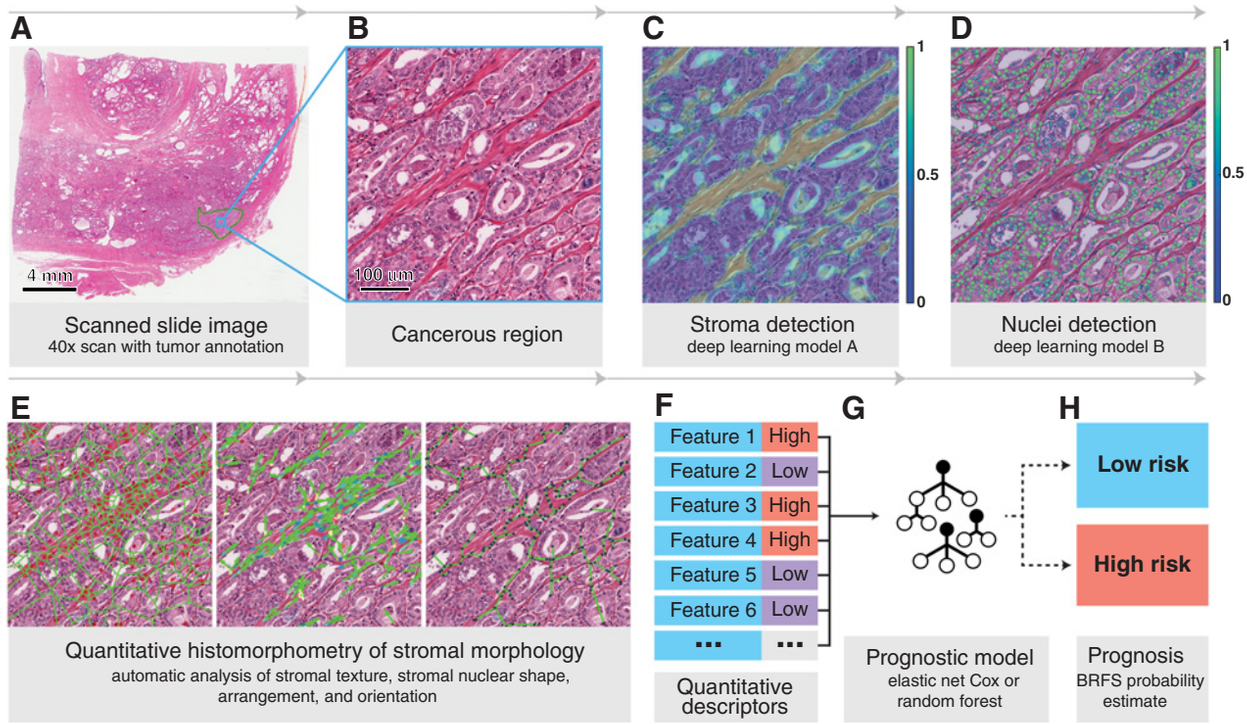

Figure 1.

Dataset preparation, analysis, and prognostic model construction. Paraffin-embedded, resected prostate glands were sectioned using a microtome. H&E slides were then prepared and digitally scanned (**A**). For each slide, a single representative cancerous region was annotated on the digital image by a pathologist (green line in a, magnified view in **B**). Stroma and nuclei were then segmented from the region of interest by deep learning models, yielding class probability maps (**C**) and (**D**). These probability maps were thresholded and used to compute stromal image features (**E** and **F**). Stromal morphology descriptors were used to train prognostic models (**G**), which estimate biochemical recurrence risk score (**H**).

stromal texture, the global and local connectivity graphs of stromal nuclei, nuclear centroids, and descriptors of nuclear shape and orientation (feature classes further described in Supplementary Table S2).

Table 1. Summary of clinicopathologic features of the whole dataset.

Variables	Subvariables	Fraction (%) or mean (STD)
Number of patients	Total	334
	Training	127 (38%)
	Validation	207 (62%)
Age in years		59.54 (7.15)
Race	AA	170 (51%)
	CA	164 (49%)
Preoperative PSA (ng/mL)	Recurrence	10.95 (11.37)
	Nonrecurrence	7.92 (11.52)
Pathologic Gleason grade sum	6 or less	109 (33%)
	7	197 (59%)
	8 or greater	28 (8%)
Pathologic stage	T2x	170 (51%)
	T3x	94 (28%)
	Data not available (either T2x or T3x)	70 (21%)
Time to event (days)	Recurrence	1,063 (1120)
	Last follow-up (nonrecurrence)	1,717 (1411)

Visualizations of selected stromal image feature calculations are shown in Supplementary Figure 1A–P for patients from each race and BCR status combination.

IHC

We obtained IHC stain intensity H-scores for 10 biomarkers relevant to prostate cancer progression for a total of 76 patients in $V_{T,AA}$ and $V_{1,AA}$ combined. The biomarkers investigated were TMPRSS2-ERG fusion, PTEN, PMSA, Racemase (AMACR), C-MYC, AR, Ki-67, P-53, nuclear RB, and cytoplasmic RB. The mean IHC H-scores and number of patients with nonzero scores for each marker are provided in Supplementary Table S9.

Statistical methods and definitions

BRFS was measured from the date of surgery to the date of BCR, which was defined as at least two PSA test results greater than 0.2 ng/mL. Patients who did not experience BCR were right censored at the date of last follow-up in survival analyses.

Within each racial subset of V_T ($V_{T,AA}$, $V_{T,CA}$, and $V_{T,AA+CA}$), all possible combinations of features were tested for correlation by calculating the Pearson correlation coefficient (PCC). To remove redundant features (i.e., features that were almost linearly dependent upon one another), the feature with smaller absolute β value in a Cox model for BRFS was removed from pairs of features that had PCC greater than 0.90.

ENC models were built using a modified version of Glmnet for MATLAB (16).

All statistical tests were two-sided, and were performed with the significance level set at 0.05. All statistical and feature analyses were conducted using MATLAB (The Mathworks, Inc.) and Python 3 (Python Software Foundation, https://www.python.org/).

Experiment 1: identification of stromal nuclear features prognostic of BCR

The independent prognostic capability of each stromal image feature with respect to BRFS was assessed in V_T using univariable Cox proportional hazards regression to determine which features might be suitable for risk prognosis model construction. The classes of features tested are described in Supplementary Table S5, and all the features and their median values for each training subset are provided in Supplementary Table S6. Cox proportional hazards regression models were fit to each stromal feature within each population subset ($V_{T,AA}$, $V_{T,CA}$, and $V_{T,AA+CA}$) to assess the prognostic power of each feature within each racial group.

Experiment 2: BCR prognosis model construction using stromal image features

Random forest and ENC models were constructed to estimate BCR risk using stromal image feature values. These models take a patient's vector of image feature values as input, and output an estimate of the recurrence risk for the patient.

Random forest (RF) classifiers were tested with input numbers of features between 1 and 25 for optimal performance averaged over 10 iterations of 3-fold cross-validation in V_T. The hyperparameters that produced the highest statistically significant Cox proportional hazards regression HR were identified for each training cohort ($V_{T,AA}$, $V_{T,CA}$, and $V_{T,AA+CA}$). Following model parameter optimization using the training set, models were trained on the entire training set, locked down, and tested against V_1 and V_2. For each validation experiment, AUC values were calculated, Kaplan–Meier survival curves constructed, and univariable Cox proportional hazards regression applied to determine HR and *P* value.

Elastic net-penalized Cox proportional hazards regression was implemented to estimate BRFS time based on the quantitative histomorphometry features. These models were fit to the survival data. To determine the optimal risk score threshold for stratifying high-recurrence risk from low-recurrence risk patients, we used the following algorithm: (i) risk scores were calculated for each patient in the training set using the ENC model. (ii) Risk scores between the 20th and 80th percentiles were retained as candidate thresholds. (iii) Each candidate was tested as a threshold in the training set and log-rank *P* values and HRs were calculated. (iv) Candidate thresholds with statistically significant performance were retained, and the value corresponding to the largest HR was selected. Following parameter optimization and threshold determination, model parameters were locked down and the models were tested against the holdout validation sets V_1 and V_2. Analysis of validation set performance was performed using the same method as described for the ML models.

To assess the performance of the AA-specific, automated stromal signature (AAstro) model relative to the clinical gold standard, two postoperative recurrence risk prognosis nomograms were implemented: CAPRA-S and Kattan (17, 18). The Kattan nomogram was implemented with the most up-to-date model parameters retrieved from the Memorial Sloan-Kettering Cancer Center website (accessed May 2019), and both nomograms were implemented with t = 5-year prediction targets. Classification models were created from the nomograms by thresholding the nomogram output probabilities at a recurrence risk probability value of 0.5. In addition, Kaplan–Meier and Cox proportional hazards analyses were performed on the output of these nomogram-based classifiers for each cohort to evaluate the differences in outcomes for the predicted low-risk and high-risk classes.

Experiment 3: comparison of AAstro with clinical variables and nomograms

To determine whether AAstroENC was independent of clinical variables, multivariable Cox proportional hazards models were fit using the model score as well as clinical and pathologic variables. To assess the performance of the AAstro model relative to the clinical gold standard, two postoperative recurrence risk prognosis nomograms were implemented, CAPRA-S and Kattan (17, 18). We also performed experiments to evaluate the performance of these nomograms for CA patients.

Experiment 4: association of stromal morphology descriptors with biomarker expression levels

To determine whether any stromal image features were associated with the expression level of tumor biomarkers, PCC values and associated *P* values were calculated for each biomarker–stromal feature pair. Statistically significant ($P < 0.05$) associations with PCC absolute value greater than 0.4 were reported.

Results

Clinicopathologic features of the patient datasets

Clinical and pathologic features of the patient dataset are provided in **Table 1**. The mean time to BCR was approximately 3 years, and the mean to last follow-up overall was approximately 5 years. Patients in all cohorts were about 60 years of age at the time of surgery. No statistically significant differences in the distributions of clinical or demographic features were found between the training and validation datasets by Wilcoxon rank sum test.

Experiment 1: descriptors of stromal morphology are associated with biochemical recurrence

In dataset V_T overall ($V_{T,AA+CA}$), 15 uncorrelated (PCC < 0.90) stromal image features were prognostic of BRFS based on Cox proportional hazards regression analysis. These features included quantitative descriptors of tissue texture, nuclear shape, and nuclear arrangement. In the AA-only training cohort ($V_{T,AA}$), six features were significantly prognostic, and in the CA training cohort ($V_{T,CA}$), 22 were identified. Notably, while there were fewer prognostic features in $V_{T,AA}$, the AA-specific features exhibited more dramatic hazard ratios than features identified in the other cohorts. In addition, while significant features in $V_{T,CA}$ and $V_{T,AA+CA}$ consisted of a mixture of feature types (descriptors of shape, orientation, arrangement, and texture), those discovered in $V_{T,AA}$ were exclusively shape and texture descriptors. All features with significantly differing distributions between BCR and non-BCR patients of each racial cohort are presented in Supplementary Table S6, with median values for each group.

Experiment 2: a stromal morphologic signature (AAstro) is prognostic of biochemical recurrence in AA patients

No ML or ENC model with any training set ($V_{T,AA}$, $V_{T,CA}$, or $V_{T,AA+CA}$) was prognostic of disease recurrence for CA or AA+CA cohorts in more than one validation set (Supplementary Table S1). However, random forest and ENC models trained on AA patients were prognostic of recurrence risk for AA patients in both validation datasets. For these patients, the ENC model fit using survival data

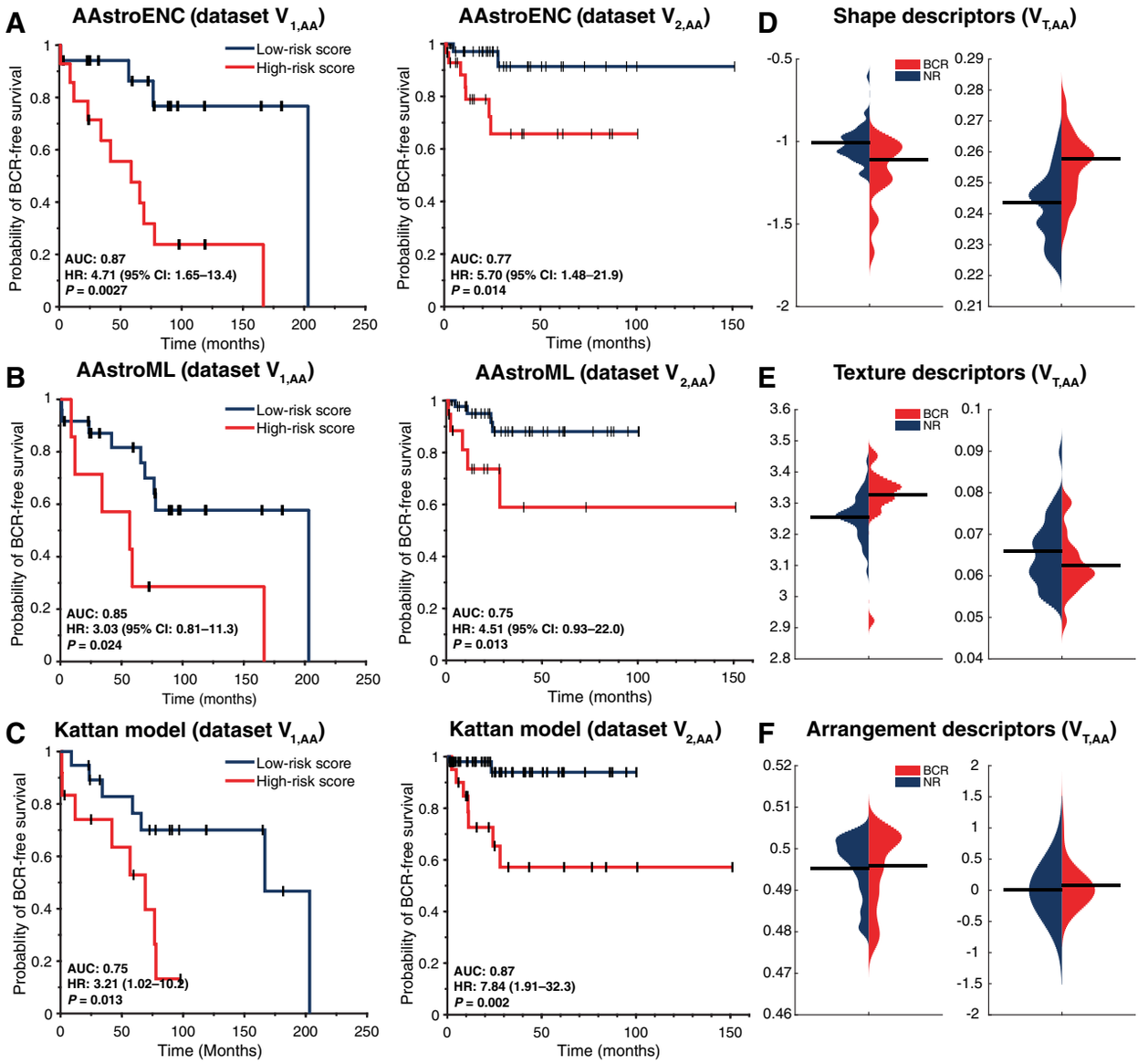

Figure 2.

Classifier performance in validation datasets and distributions of features prognostic for AA patients. Kaplan–Meier survival curve estimates of predicted low- versus high-BCR risk groups in $V_{1,AA}$ and $V_{2,AA}$ for AAstroENC (**A**), AAstroML (**B**), and Kattan nomogram (**C**). Distributions in $V_{T,AA}$ of QH descriptors of stromal nuclear shape (min/max Fourier descriptor 4 and mean fractal dimension; **D**), texture (Haralick mean information measure 1 and mean contrast inverse moment; **E**), and nuclear arrangement (COrE mean tensor correlation and subgraph number of isolated nodes; **F**) used by AAstro models.

outperformed the ML models trained on binary BCR/non-BCR labels. The results of classification experiments for each training and validation cohort combination are presented in Supplementary Table S1, and results for AA patients are displayed in **Fig. 2A** and **B**.

The best performing RF model for AA patients, AAstroML, was trained on the top six stromal image features identified by univariable HR. These features were all descriptors of stromal nuclear shape and stromal texture (specific features listed in Supplementary Table S3). AAstroML achieved an AUC of 0.85 and HR of 3.03 [95% confidence interval (CI), 0.812–11.3; $P = 0.024$] in validation on $V_{1,AA}$, and an AUC of 0.75 and HR of 4.51 (95% CI, 0.925–22; $P = 0.013$) on $V_{2,AA}$.

The best performing ENC model for AA patients, AAstroENC, was trained on 10 descriptors of stromal morphology (listed in Supplementary Table S3). The features selected were similar to those selected by AAstroML, but also included two descriptors of nuclear arrangement. AAstroENC achieved an AUC of 0.87 and HR of 4.71 (95% CI, 1.65–13.4; $P = 0.0027$) in $V_{1,AA}$, and AUC of 0.77 and HR of 5.7 (95% CI, 1.48–21.90; $P = 0.014$) in $V_{2,AA}$, outperforming AAstroML in both datasets.

Distributions of selected features used by the AAstro models are visualized in **Fig. 2D–F**. Kaplan–Meier curves estimating the relative survival of the estimated high- versus low-recurrence risk groups for each classifier in each validation set are shown in **Fig. 2A** and **B**.

Table 2. Multivariable Cox proportional hazards modeling of AAstroENC model scores adjusting for clinical and pathologic variables.

Parameter	Dataset $V_{1,AA}$ ($n = 31$) HR (95% CI)	*P*	Dataset $V_{2,AA}$ ($n = 93$) HR (95% CI)	*P*
AAstroENC score	**4.62 (1.31–16.23)**	**0.017**	**2.58 (1.38–4.83)**	**0.0029**
Age at the time of surgery	0.27 (0.091–0.82)	0.02	0.74 (0.32–1.69)	0.47
Gleason score ≤6	Reference	1.00	Reference	1.00
Gleason score = 7	1.89 (0.86–4.13)	0.11	0.68 (0.19–2.48)	0.57
Gleason score ≥8	1.09 (0.42–2.84)	0.85	1.07 (0.31–3.78)	0.91
Preoperative PSA value	2.60 (1.11–6.10)	0.027	1.75 (0.82–3.74)	0.14
Presence of SVI	**3.27 (1.21–8.86)**	**0.019**	**2.53 (1.20–5.33)**	**0.014**
Presence of ECE	0.82 (0–infinity)	1.00	0.94 (0.23–3.76)	0.93
Positive surgical margins	0.77 (0.30–1.92))	0.57	1.18 (0.61–2.31)	0.62
Pathologic stage pT2x	Reference	1.00	Reference	1.00
Pathologic stage pT3x	0.62 (0–infinity)	1.00	1.36 (0.81–2.28)	1

Note: Predictors with significant *P* values in bold.

Experiment 3: AAstro signature is independent of clinical variables and outperforms clinical nomograms for AA patients

In a multivariable Cox regression analysis, AAstroENC was prognostic in both $V_{1,AA}$ and $V_{2,AA}$ independent of clinical and pathologic variables used in existing risk estimation models (**Table 2**). The only other variable retaining a significant HR in both validation datasets when adjusting for other variables and AAstroENC score was presence of seminal vesicle invasion (SVI).

AAstroENC and AAstroML outperformed clinical models in both validation datasets. The only nomogram prognostic in an AA validation cohort was the Kattan nomogram in $V_{1,AA}$. The HR achieved by the Kattan nomogram was lower than that of AAstroENC. Performance of models created using CAPRA-S and Kattan nomograms for AA patients is provided in Supplementary Table S2, and equivalent modeling conducted for CA patients in Supplementary Table S3.

Experiment 4: association of stromal morphology with tumor biomarker expression

Twenty-nine pairs of IHC-derived biomarkers and stromal morphology descriptors were identified as having significant association with PCC > 0.4 or PCC < −0.4. Three of these pairs included stromal morphology descriptors that were prognostic of BRFS in $V_{T,AA}$. These three pairs were between stromal nuclear shape mean fractal dimension with cytoplasmic RB (PCC = 0.606, $P = 0.0005$), stromal texture Haralick mean information measure 1 with TMPRSS2-ERG fusion protein (PCC = −0.447, $P = 0.033$), and stromal nuclear shape mean fractal dimension with AR expression (PCC = 0.41, $P = 4.12e\text{-}4$). PTEN was found to be associated with 15 image features computed from the stroma. All of these descriptors were measurements of nuclear shape. PTEN was the marker associated with the most stromal image features, being statistically correlated with 15 features. All of these descriptors were measurements of nuclear shape. The QH feature-biomarker pairing with the highest absolute PCC was the association of PTEN with mean Fourier descriptor 4 of stromal nuclear shape (PCC = −0.623, $P = 7.56e\text{-}3$).

The biomarkers studied are outlined in **Fig. 3A**, and selected pairs with significant correlation are highlighted in **Fig. 3B**. Scatter plots depicting the association of prognostic stromal image features with biomarker expression values are presented in **Fig. 3C–E**. The full list of pairs identified is presented in Supplementary Table S4.

Discussion

The goal of this work was to evaluate quantitative descriptors of stromal morphology and population-specific tuning for post-radical prostatectomy prognosis. We calculated quantitative descriptors of stromal morphology from routine H&E slides, and found that stromal morphology differs between AA and CA patients. Models created using our stromal morphologic signature were found to be prognostic for AA patients in two independent holdout validation datasets. These models, AAstroML and AAstroENC, predicted risk independently of routine clinical variables, and outperformed current clinical nomograms in the AA validation set patients. Multiple stromal image features were also associated with the expression level of tumor biomarkers measured using IHC.

In our analysis, the stromal features most prognostic of risk in all three cohorts (AA, CA, and AA+CA) were descriptors of stromal texture, stromal nuclear shape, and stromal nuclear arrangement. These features specifically included Fourier and invariant descriptors, which measure the fundamental shape of the nuclear boundary. The features selected by the model also included measures of nuclear arrangement, which describe the relative nuclear spatial density, and stromal texture. The distributions of these features in each cohort appear to indicate that high-risk tumors have relatively higher levels of intratumoral heterogeneity in stromal nuclear shape and arrangement and stromal texture. Differences in stromal morphologic phenotypes between AA and CA tumors are supported by other studies. Kinseth and colleagues (19) found that the majority of genes with significant expression differences between CA and AA men were associated with the tumor-adjacent stroma. A number of these differentially expressed genes were found to be involved with cell organization and structure, including ECM regulation, cellular adhesion, and cytoskeleton maintenance.

The prognostic utility of stromal features observed in this study adds to a growing base of knowledge implicating tissue regions not traditionally examined by pathologists as harboring prognostic cues. In a study involving more than 6,000 image features derived from breast cancer images, Beck and colleagues (20) identified stromal image features as being more strongly prognostic of survival compared with tumor epithelial features. Similarly, Lee and colleagues (14) showed that computer-extracted descriptors of nuclear morphology, derived from benign, tumor-adjacent regions were strongly associated with the likelihood of biochemical recurrence post-surgery. Collectively, these findings reinforce the importance of interrogating patterns within the stroma and tumor-adjacent regions on histopathology.

The variations in cell and tissue phenotype measured by our computer vision approach are the result of molecular pathways invisible on H&E slides. While the precise molecular etiology of prostate cancer is not fully understood, we hypothesized that the image features used by our AAstro models might be associated with the expression of biomarkers important to prostate cancer pathogenesis. To investigate this hypothesis, we obtained IHC stains against 10 well-characterized prostate cancer protein biomarkers (**Fig. 3A**). These biomarkers included proteins involved in cellular metabolism, tumor suppression, androgen signaling, and transcriptional regulation. We tested the expression level of these biomarkers for association with our stromal morphology descriptors.

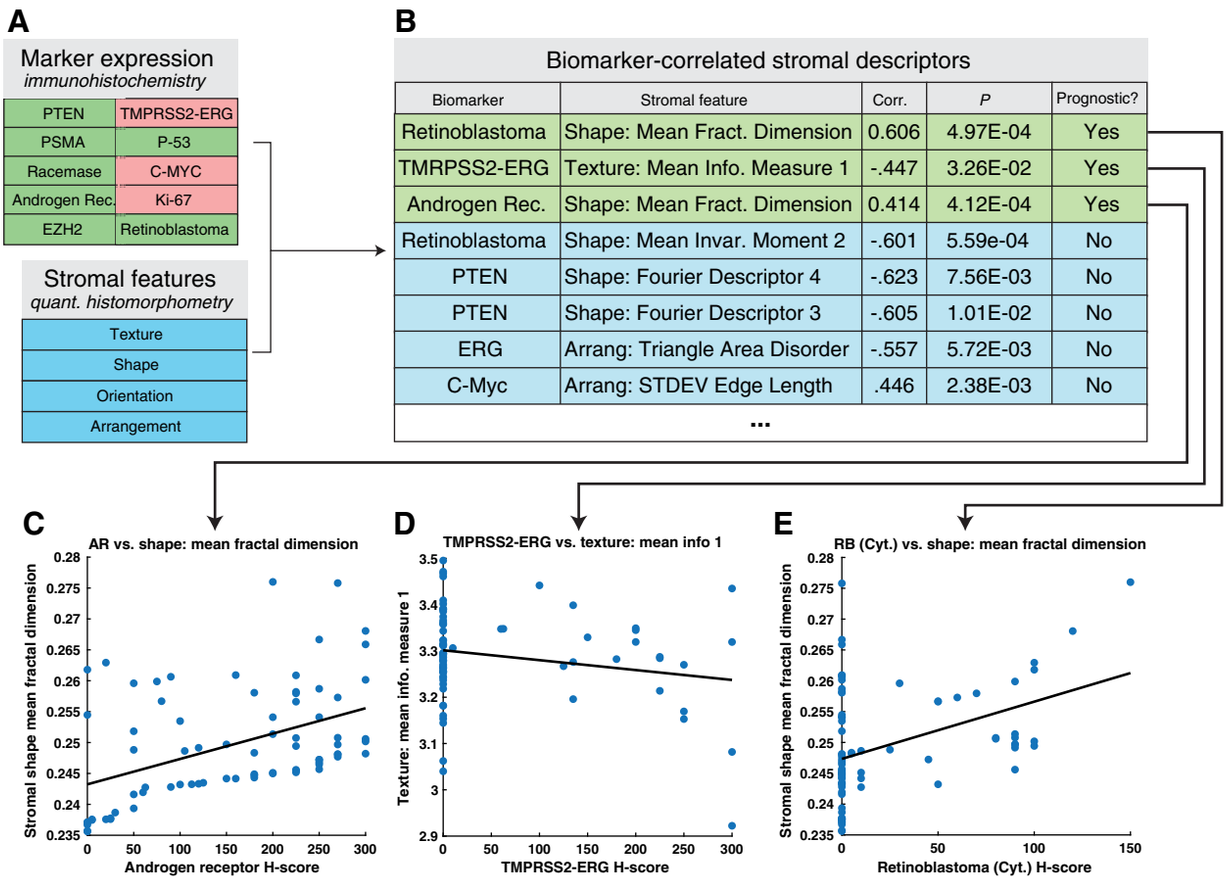

Figure 3.

Association of stromal morphology descriptors with biomarker expression levels. **A,** Expression levels of selected prostate cancer tumor biomarkers were measured using IHC. These values were tested for association with stromal image features calculated from H&E-stained images. Selected pairings of biomarkers and stromal image features with significant correlation are shown in **B**, with prognostic stromal features highlighted in green. Scatter plot visualizations of correlation between prognostic stromal features and biomarker expression levels are shown in **C–E**.

This experiment revealed a strong association between stromal nuclear shape (mean fractal dimension) and RB, a tumor suppressor protein. Loss of RB is associated with a transition to incurable prostate cancer (21). RB has been shown to control androgen receptor (AR) expression: depletion of RB induces dysregulation of AR activity, which is associated with therapeutic bypass and tumor progression (21). An association was also identified between stromal texture (Haralick mean information measure 1) and TMPRSS2-ERG gene fusion, a molecular aberration frequently observed in prostate cancer. TMPRSS2-ERG expression level was negatively associated with recurrence-free survival (22). TMPRSS-ERG expression has also been shown to be associated with specific stromal biomarkers (23). Associations were also identified between stromal nuclear shape (mean fractal dimension) and the AR, a key element of the androgen signaling system. Interestingly, AR signaling in the stroma has been shown to influence tumor behavior: loss of stromal AR is related with substantially increased risk (24). Finally, we identified an association between 15 image features of stromal nuclear shape and expression of PTEN, a tumor suppressor protein. Loss of PTEN is the most common genetic change in prostate cancer, and is associated with more aggressive disease and castration resistance (25). Interestingly, loss of PTEN is associated with specific adverse histologic features, including intraductal carcinoma, cribriform Gleason pattern 4, and stromogenic carcinoma (26).

These preliminary experiments support the existence of a connection between stromal architecture and tumor biochemistry. Notably, previous work on these biomarkers has focused on the prostate cancer epithelium, and our results suggest potential roles for these biomarkers in the intratumoral stroma. It is also possible that stroma–epithelium interactions, which have been implicated in tumorigenesis, mediate both the morphologic and biochemical changes we observed.

This study did have some limitations. First, the size of the training dataset was limited, especially given its division into AA and CA subgroups. It is possible that with a larger training dataset, models prognostic for CA patients could have been trained. In addition, some patients in the study had relatively short follow-up durations. In addition, BCR prognosis models trained on CA patients failed to effectively stratify BCR risk in CA men. It is possible that the poor model performance on the CA cohort was in part due to increased diversity and/or genetic heterogeneity within this group (27, 28).

In spite of these limitations, this study is the first to show the role of stromal features in prostate cancer BCR likelihood estimation, and is the first QH study to assess population-specific differences in prostate cancer. In addition, the findings of this study provide some biological

insight into differences in prostate cancer morphology between AA and CA patients. Future work will entail reproducing and validating these findings in larger cohorts and studying stromal features in the context of features from additional tissue compartments.

Disclosure of Potential Conflicts of Interest

A. Madabhushi reports receiving other commercial research support from Philips, holds ownership interest (including patents) in Elucid Bioimaging and Inspirata Inc., and was the primary inventor on two patents "Tumor+Adjacent Benign Signature (TABS) For Quantitative Histomorphometry," Anant Madabhushi, George Lee, Sahirzeeshan Ali, United States Serial Number (USSN): 9,424,460, August 23, 2016 and "High-throughput adaptive sampling for whole-slide histopathology image analysis," Anant Madabhushi, Angel Cruz, Fabio Gonzalez, United States Serial Number (USSN): 10,235,755, March 19, 2019, both owned by Case Western Reserve University, and is an advisory board member/unpaid consultant for Inspirata Inc. No potential conflicts of interest were disclosed by the other authors.

Disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the United States Department of Veterans Affairs, the Department of Defense, the National Science Foundation, or the United States Government.

Authors' Contributions

Conception and design: H.K. Bhargava, P. Leo, A. Janowczyk, T.R. Rebbeck, P. Lal, A. Madabhushi

Development of methodology: H.K. Bhargava, P. Leo, A. Janowczyk, J. Whitney, A. Madabhushi

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H.K. Bhargava, P. Leo, R. Elliott, K. Yamoah, F. Khani, B.D. Robinson, M. Feldman

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): H.K. Bhargava, P. Leo, R. Elliott, A. Janowczyk, P. Fu, K. Yamoah, T.R. Rebbeck, M. Feldman, A. Madabhushi

Writing, review, and/or revision of the manuscript: H.K. Bhargava, P. Leo, R. Elliott, A. Janowczyk, J. Whitney, S. Gupta, P. Fu, K. Yamoah, F. Khani, B.D. Robinson, T.R. Rebbeck, P. Lal, A. Madabhushi

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H.K. Bhargava, S. Gupta, A. Madabhushi

Study supervision: A. Madabhushi

Acknowledgments

We thank the members of the Case Western Reserve University Center for Computational Imaging and Personalized Diagnostics for their support and helpful discussions. This work was supported by the NCI of the NIH (1U24CA199374-01, R01CA202752-01A1, R01CA208236-01A1, R01CA216579-01A1, R01CA220581-01A1, and 1U01 CA239055-01); National Center for Research Resources (1 C06 RR12463-01); Veterans Association Merit Review Award the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service (IBX004121A); the Department of Defense Prostate Cancer Idea Development Award (W81XWH-15-1-0558); the Department of Defense Lung Cancer Investigator-Initiated Translational Research Award (W81XWH-18-1-0440); the Department of Defense Peer Reviewed Cancer Research Program (W81XWH-16-1-0329); the Department of Defense Prostate Cancer Disparity Award (W81XWH-19-1-0720); the Hartwell Foundation T32 grant to J. Whitney; the Case Western Reserve University Nephrology Training Grant (5T32DK007470 to A. Janowczyk); the National Science Foundation Graduate Research Fellowship Program (CON501692 to P. Leo); the Ohio Third Frontier Technology Validation Fund; and the Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering and the Clinical and Translational Science Award Program at Case Western Reserve University.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 13, 2019; revised December 17, 2019; accepted January 30, 2020; published first March 5, 2020.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69:7–34.
2. Burkhardt JH, Litwin MS, Rose CM, Correa RJ, Sunshine JH, Hogan C, et al. Comparing the costs of radiation therapy and radical prostatectomy for the initial treatment of early-stage prostate cancer. J Clin Oncol 2002;20:2869–75.
3. Zincke H, Oesterling JE, Blute ML, Bergstralh EJ, Myers RP, Barrett DM. Long-term (15 years) results after radical prostatectomy for clinically localized (stage T2c or lower) prostate cancer. J Urol 1994;152:1850–7.
4. Freedland SJ, Humphreys EB, Mangold LA, Eisenberger M, Dorey FJ, Walsh PC, et al. Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy. JAMA 2005;294:433–9.
5. Freedland SJ, Rumble RB, Finelli A, Chen RC, Slovin S, Stein MN, et al. Adjuvant and salvage radiotherapy after prostatectomy: American Society of Clinical Oncology Clinical Practice Guideline Endorsement. J Clin Orthod 2014;32:3892–8.
6. Karakas C, Wang C, Deng F, Huang H, Wang D, Lee P. Molecular mechanisms involving prostate cancer racial disparity. Am J Clin Exp Urol 2017;5:34–48.
7. Bonilla C, Hooker S, Mason T, Bock CH, Kittles RA. Prostate cancer susceptibility loci identified on chromosome 12 in African Americans. PLoS One 2011;6:e16044.
8. Rose AE, Satagopan JM, Oddoux C, Zhou Q, Xu R, Olshen AB, et al. Copy number and gene expression differences between African American and Caucasian American prostate cancer. J Transl Med 2010;8:70.
9. Elliott B, Zackery DL, Eaton VA, Jones RT, Abebe F, Ragin CC, et al. Ethnic differences in TGFβ signaling pathway may contribute to prostate cancer health disparity. Carcinogenesis 2018;39:546–55.
10. Tyekucheva S, Bowden M, Bango C, Giunchi F, Huang Y, Zhou C, et al. Stromal and epithelial transcriptional map of initiation progression and metastatic potential of human prostate cancer. Nat Commun 2017;8:420.
11. Tomas D, Spajić B, Milosević M, Demirović A, Marusić Z, Kruslin B. Intensity of stromal changes predicts biochemical recurrence-free survival in prostatic carcinoma. Scand J Urol Nephrol 2010;44:284–90.
12. Condon MS. The role of the stromal microenvironment in prostate cancer. Semin Cancer Biol 2005;15:132–7.
13. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. Nat Rev Clin Oncol 2019;16:703–15.
14. Lee G, Veltri RW, Zhu G, Ali S, Epstein JI, Madabhushi A. Nuclear shape and architecture in benign fields predict biochemical recurrence in prostate cancer patients following radical prostatectomy: preliminary findings. Eur Urol Focus 2017;3:457–66.
15. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inform 2016;7:29.
16. Qian J, Hastie T, Friedman J, Tibshirani R, Simon N. Glmnet for Matlab. 2013. Available from: http://www.stanford.edu/~hastie/glmnet_matlab/.
17. Cooperberg MR, Hilton JF, Carroll PR. The CAPRA-S score: a straightforward tool for improved prediction of outcomes after radical prostatectomy. Cancer 2011;117:5039–46.
18. Kattan MW, Eastham JA, Stapleton AM, Wheeler TM, Scardino PT. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. J Natl Cancer Inst 1998;90:766–71.
19. Kinseth MA, Jia Z, Rahmatpanah F, Sawyers A, Sutton M, Wang-Rodriguez J, et al. Expression differences between African American and Caucasian prostate cancer tissue reveals that stroma is the site of aggressive changes. Int J Cancer 2014;134:81–91.
20. Beck AH, Sangoi AR, Leung S, Marinelli RJ, Nielsen TO, van de Vijver MJ, et al. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. Sci Transl Med 2011;3:108ra113.

21. Sharma A, Yeow W-S, Ertel A, Coleman I, Clegg N, Thangavel C, et al. The retinoblastoma tumor suppressor controls androgen signaling and human prostate cancer progression. J Clin Invest 2010;120:4478–92.
22. Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia 2008;10:177–88.
23. Hägglöf C, Hammarsten P, Strömvall K, Egevad L, Josefsson A, Stattin P, et al. TMPRSS2-ERG expression predicts prostate cancer survival and associates with stromal biomarkers. PLoS One 2014;9:e86824.
24. Leach DA, Buchanan G. Stromal androgen receptor in prostate cancer development and progression. Cancers 2017;9:pii: E10.
25. Jamaspishvili T, Berman DM, Ross AE, Scher HI, De Marzo AM, Squire JA, et al. Clinical implications of PTEN loss in prostate cancer. Nat Rev Urol 2018;15:222–34.
26. Shah RB, Shore KT, Yoon J, Mendrinos S, McKenney JK, Tian W. PTEN loss in prostatic adenocarcinoma correlates with specific adverse histologic features (intraductal carcinoma, cribriform Gleason pattern 4 and stromogenic carcinoma). Prostate 2019;79:1267–73.
27. Ioannidis JPA, Ntzani EE, Trikalinos TA. "Racial" differences in genetic effects for complex diseases. Nat Genet 2004;36:1312–8.
28. Perez AD, Hirschman C. The changing racial and ethnic composition of the US population: emerging American identities. Popul Dev Rev 2009;35:1–51.