

# Application of Artificial Intelligence/ Machine Vision & Learning for the Development of a Live Single-cell Phenotypic Biomarker Test to Predict Prostate Cancer Tumor Aggressiveness

Jonathan S. Varsanik, PhD,<sup>1\*</sup> Michael S. Manak, PhD,<sup>1\*</sup> Matthew J. Whitfield, PhD,<sup>1\*</sup>  
Brad J. Hogan, PhD,<sup>1</sup> Wendell R. Su, MEng,<sup>1</sup> CJ Jiang, PhD,<sup>1</sup> Grannum R. Sant, MD, FRCS, FACS,<sup>2</sup>  
David M. Albala, MD,<sup>3</sup> Ashok C. Chander, PhD<sup>1\*</sup>

<sup>1</sup>Cellanyx, Beverly, MA; <sup>2</sup>Tufts University School of Medicine, Boston, MA; <sup>3</sup>Crouse Hospital, Syracuse, NY

To assess the usefulness and applications of machine vision (MV) and machine learning (ML) techniques that have been used to develop a single cell-based phenotypic (live and fixed biomarkers) platform that correlates with tumor biological aggressiveness and risk stratification, 100 fresh prostate samples were acquired, and areas of prostate cancer were determined by post-surgery pathology reports logged by an independent pathologist. The prostate samples were dissociated into single-cell suspensions in the presence of an extracellular matrix formulation. These samples were analyzed via live-cell microscopy. Dynamic and fixed phenotypic biomarkers per cell were quantified using objective MV software and ML algorithms. The predictive nature of the ML algorithms was developed in two stages. First, random forest (RF) algorithms were developed using 70% of the samples. The developed algorithms were then tested for their predictive performance using the blinded test dataset that contained 30% of the samples in the second stage. Based on the ROC (receiver operating characteristic) curve analysis, thresholds were set to maximize both sensitivity and specificity. We determined the sensitivity and specificity of the assay by comparing the algorithm-generated predictions with adverse pathologic features in the radical prostatectomy (RP) specimens. Using MV and ML algorithms, the biomarkers predictive of adverse pathology at RP were ranked and a prostate cancer patient risk stratification test was developed that distinguishes patients based on surgical adverse pathology features. The ability to identify and track large numbers of individual cells over the length of the microscopy

---

\*Contributed equally.

experimental monitoring cycles, in an automated way, created a large biomarker dataset of primary biomarkers. This biomarker dataset was then interrogated with ML algorithms used to correlate with post-surgical adverse pathology findings. Algorithms were generated that predicted adverse pathology with  $>0.85$  sensitivity and specificity and an AUC (area under the curve) of  $>0.85$ . Phenotypic biomarkers provide cellular and molecular details that are informative for predicting post-surgical adverse pathologies when considering tumor biopsy samples. Artificial intelligence ML-based approaches for cancer risk stratification are emerging as important and powerful tools to compliment current measures of risk stratification. These techniques have capabilities to address tumor heterogeneity and the molecular complexity of prostate cancer. Specifically, the phenotypic test is a novel example of leveraging biomarkers and advances in MV and ML for developing a powerful prognostic and risk-stratification tool for prostate cancer patients.

---

[*Rev Urol.* 2020;22(4):159–167]

© 2021 MedReviews®, LLC

#### KEY WORDS

Artificial intelligence • Phenotypic biomarkers • Machine vision • Machine learning  
• Prostate cancer

**P**rostate cancer (PCa) is the most prevalent male malignancy and the third most common cause of death worldwide.<sup>1</sup> There is an expanding effort in the applications of computer science via artificial intelligence (AI) to improve disease diagnosis, management, and the development of effective therapies in clinical health science.<sup>2–4</sup> With an estimated 191,000 men expected to be diagnosed with prostate cancer in the United States in 2021, the application of artificial intelligence methods such as machine vision (MV) and machine learning (ML) to develop a phenotypic cancer-cell signature for risk stratification hold promise for better patient care.<sup>1</sup>

Currently, assessment of prostate cancer aggressiveness relies on formalin fixed–paraffin embedded (FFPE) techniques such as Gleason score (GS), Prostate

Cancer Grading Group scores (PCGG), or tissue genomic analyses.<sup>5–7</sup> These techniques are fraught with limitations including tumor multifocality, heterogeneity, and tissue under-sampling with needle-biopsy techniques. The current predictors and correlates of tumor aggressiveness and risk assessment do not include live-cell phenotypic and biological attributes of the individual cancer cells (eg, actin dynamics, cell motility, subcellular protein localization, etc).<sup>8</sup> Integrating such a live-cell phenotypic signature that includes molecular, cellular, and functional dynamic behavior approaches may improve the clinical risk stratification (RS) in men with PCa.<sup>9–11</sup>

We recently reported on the successful culture of primary PCa cells obtained immediately after radical prostatectomy (RP).<sup>12</sup> Additionally, machine vision and

learning (MVL) techniques have been used to develop a single cell-based phenotypic (live and fixed biomarkers) platform that correlates with tumor biological aggressiveness and RS.

This article addresses the MVL components and applications to the development of a new single-cell phenotypic biomarker signature in PCa. MVL approaches are being applied and/or are in development in medicine for screening, diagnosis, prognosis, treatment selection, and monitoring. Other applications of MVL in PCa include methods to stratify formalin-fixed prostate tissue into stroma, benign/normal tissue, or prostate cancerous tissue,<sup>3</sup> and analysis of radiographic multi-parametric MRI images to improve the performance of Prostate Imaging Reporting and Data System (PI-RADS).<sup>13</sup>

Herein, we detail characteristics of our systematically optimized MVL approaches to development of a live-cell phenotypic biomarker platform for PCa. The increases in computer processing power, data storage (cloud), and communications infrastructure over the past two decades have enabled MVL techniques (AI) and algorithms to effectively process and train massive image and biomarker datasets for clinical medicine diagnostics and imaging applications.

MV technology is a valuable, if not a necessary, prerequisite to ensure accurate extraction of all significant dynamic cellular features and uniformity across biomarker quantifications that are used to train and test ML algorithms. Further, to ensure ML algorithms are not biased and do not overfit data during the learning process, ML algorithms are built to objectively rank biomarkers in a blinded fashion. To confirm biomarkers are objectively ranked, we present data that show a reduction of ML predictive performance dependent on higher-ranked biomarkers inclusion in the analysis.

## Materials and Methods

### *Sample Collection and Preparation*

Fresh prostate samples (100) were acquired from 10 sites in the United States (5 university hospitals, 3 large urology group practices, and 2 bio-banks) with institutional review board (IRB) approval. Areas of PCa were determined by post-surgery pathology reports logged by an independent pathologist. The prostate samples were dissociated into single-cell suspensions in the presence of an extracellular matrix (ECM) formulation containing collagen and fibronectin and allowed to acclimate to the conditions prior to imaging like conditions previously described.<sup>12</sup>

### *Microscopy and Image Collection*

A Nikon TE-2000 system with a 20× Differential Interference Contrast DIC objective and prism set-up for live-cell imaging and 40× objective for fluorescence imaging were used to capture all images. Cellular behavior was monitored every 3 minutes for 1 hour and every 3 seconds for 1.5 minutes. All cells were individually tracked every 5 minutes for 2 hours. Thousands of images were collected for each sample to define live-cell morphological dynamic characteristics and molecular details.

### *Immunofluorescence*

Fixed-cell analysis was performed by staining with mouse pFAK antibody (BD Biosciences, San Jose, CA) for activated focal adhesion staining, rabbit integrin-linked kinase (ILK) antibody (Abcam, Cambridge, UK), and rat alpha-tubulin antibody (Sigma-Aldrich, St. Louis, MO). Secondary alexa-fluor 488 anti-mouse, 647 anti-rabbit, and 594 anti-rat antibodies (Life Technologies, Carlsbad, CA) were used to stain the pFAK, ILK, and alpha-tubulin primary antibodies.<sup>9,10</sup>

### *Machine Vision Algorithms*

Image analysis was described previously.<sup>9,10</sup> Cells were identified using both center of mass and grayscale values and each given an individual tracking number and followed throughout the analysis both spatially and temporally. Cellular dynamics were measured over 4 hours and measurement of the tortuosity or curves in the cell surface was monitored 4 hours. Retrograde force velocity (RFV) of the cellular membrane was monitored via kymographs, which were scored for the slope as the membrane flow velocity. Each individual focal adhesion measurement

was monitored by scoring the highest grayscale values around the perimeter of the cell membrane above background for total size per focal adhesion, distance from the cell perimeter, and total number of focal adhesions per cell.

### *Machine Learning Algorithms*

Random forest (RF) decision tree-based ML algorithms were used to correlate biomarker measurements with surgical adverse pathologies. The RF algorithms are trained collections of decision trees with random selection to average the outcome, reduce variance, reduce over-fitting, and minimize bias. Individual cells were analyzed and thereafter the whole population of analyzed cells was used to create a sample score. The predictive nature of the ML algorithms is developed in two stages. The first stage involves the development of the RF algorithm with the index dataset using 70% of the samples. The developed algorithm is then tested for its predictive performance with the test dataset that contains 30% of the samples in the second stage. Based on ROC (receiver operating characteristic) curve analysis, thresholds were set to maximize both sensitivity and specificity. We determined the sensitivity and specificity of the assay by comparing the algorithm-generated predictions with adverse pathologic features in the radical prostatectomy specimens.

## Results

To assess the dynamic biomarkers in individually cultured primary PCa cells, we developed MV algorithms for monitoring cellular phenotypes with both spatial and temporal resolution (Figure 1). Simply, cells are exposed to an optimized ECM environment<sup>12</sup> that allows rapid single-cell culturing and biomarker recognition to distinguish healthy

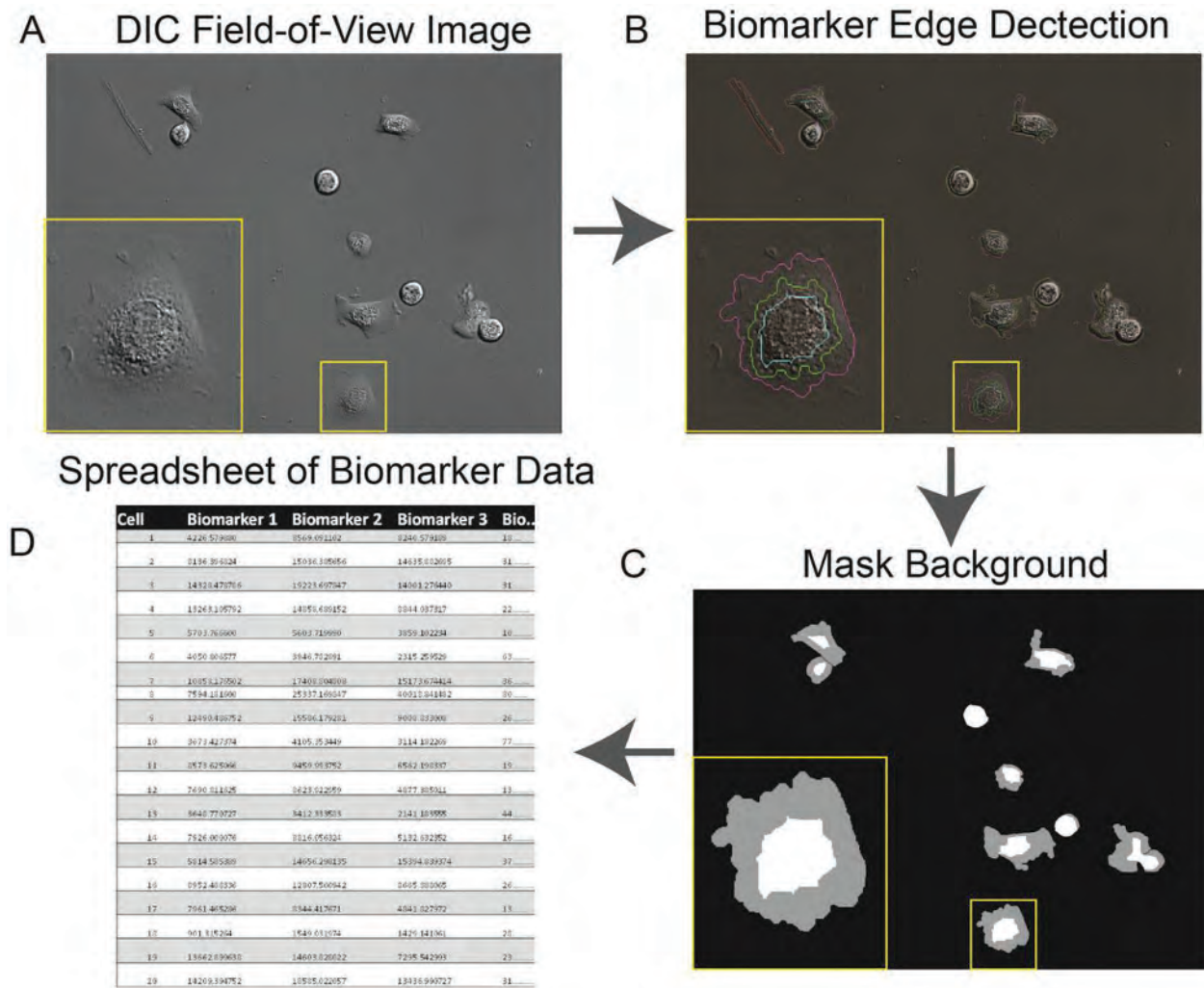


Figure 1. Schematic of the machine vision process for live-cell phenotypic biomarker measurements. Images are initially captured using time-lapse differential interference contrast (DIC) microscopy. Machine vision software is then utilized to identify boundary regions of biomarkers using edge detection techniques. Next, the background is minimized, and layers are reduced and flattened. Finally, single-cell biomarker measurements are recorded in a spreadsheet for downstream analysis with nanometer-second resolution.

cells from suspected cancer cells. The ability to identify and track large numbers of individual cells over the length of the microscopy experimental monitoring cycles, in an automated way, created a large biomarker dataset of primary biomarkers. This biomarker dataset was then interrogated with ML algorithms (Figure 2) used to correlate with post-surgical adverse pathology findings (Table 1).

Although there are many possible ML algorithm methods, based on the needs of physicians making decisions about a given patient's

PCa disease state, we chose to produce outputs that were binary and interpretable in nature versus a low-, medium-, or high-risk classification with zones of uncertainty. RF ML algorithms give the option of being a classifier that can create binary outputs based on setting a threshold to allow for the highest sensitivity and specificity (Figure 2[B]). Further, RF algorithms use multiple decision trees with many nodes that can be separated across decision trees and used to find the best possible correlative outcomes. From the high degree

of correlation based on thresholds set for maximizing sensitivity and specificity, strong prediction can be made above current nomogram-style decision classification systems (Table 2, Table 3, Figure 3). This ML technique of RF correlation classification provides a stronger predictive power, relative to current clinical nomograms used by physicians to make treatment decisions (Table 2, Table 3) following biopsy diagnosis of prostate cancer.

Robustness of the ML algorithms predictions for biomarker correlation with post-surgical adverse



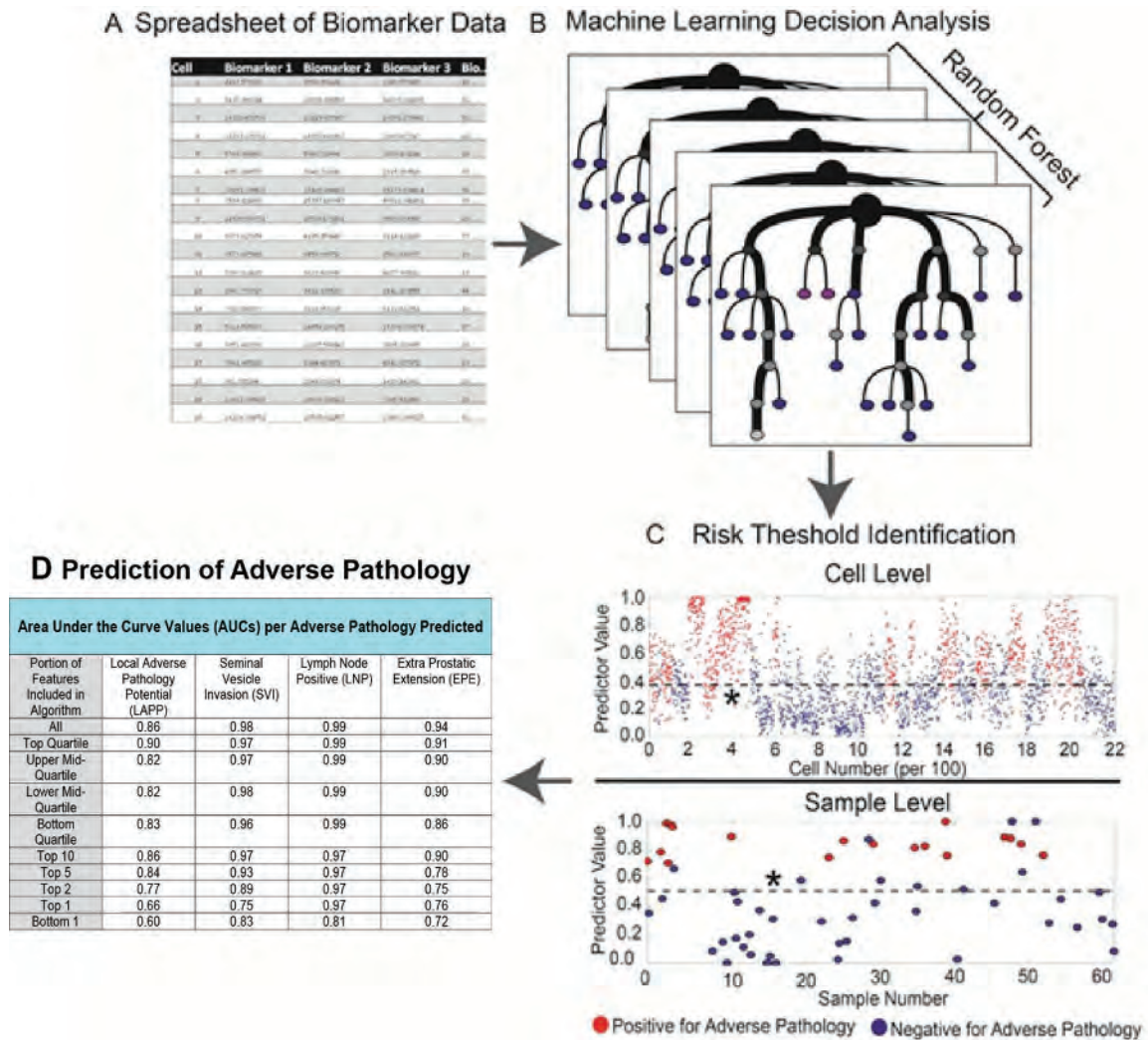


Figure 2. Schematic of machine learning data flow employed to analyze live-cell phenotypic biomarker measurements. (A) 100s of primary novel static and dynamic biomarker measurements with single-cell resolution are quantified and organized in a database such that quantified biomarkers may be input into machine learning software for analysis. (B) Machine learning-based, objective, random forest techniques are utilized to rank biomarkers based on their respective predictive power to maximize stratification of single cells for their given local and metastatic aggressive potential. (C) Single-cell stratification is then translated/transformed to a sample/patient level prediction by objective thresholding based on the ensemble of single-cell predictions. (D) Sample/patient predictions are then assessed against unblinded clinical reports for accuracy as measured by receiver operating characteristic–area under the curve.

**TABLE 1**

**Specific Adverse Pathologies Currently Predicted by Live-cell Phenotypic Biomarker Machine Vision Measurements and Machine Learning Analysis**

**Associated With Local Disease**

- Seminal vesicle invasion (SVI)
- Positive surgical margins (PSM)
- Extra-prostatic extension (EPE)

**Associated With Metastatic Disease**

- Perineural invasion (PNI)
- Lymph node positive (LNP)
- Lymph-vascular invasion (LVI)

Each specific adverse pathology is categorized given its relative correlation with local tumor growth or metastatic disease in prostate cancer.

**TABLE 2****Examples and Reported Performance of Currently Applied Nomograms in Prostate Cancer Prognosis and Risk Stratification**

Nomograms Used in Prostate Cancer	AUC Value for Associated Prediction
Partin Tables	Stage prediction T1c AUC = 0.57, T2a AUC = 0.51 <sup>14</sup>
Kattan and D'Amico Nomogram	Prostate cancer at biopsy PSA AUC = 0.69, fPSA AUC = 0.77 <sup>15</sup>
Memorial Sloan Kettering Nomogram	2-year risk of 70% recurrence AUC = 0.78 <sup>16</sup>
Fox Chase Cancer Center Nomograms	Prostate cancer at biopsy cancer AUC = 0.74, high grade AUC = 0.77 <sup>17</sup>
UCSF-CAPRA Nomogram	Metastatic progression at radical prostatectomy AUC = 0.79 <sup>18</sup>

AUC, area under the curve; fPSA, free PSA; PSA, prostate-specific antigen.

Classic multidimensional analysis utilizing qualitative, semi-quantitative, and quantitative patient data to diagnose, predict patient risk, patient outcomes, and guide treatment decisions rely on nomograms. Several nomograms have been developed to address various aspects of patient treatment decisions based on tumor measurements.

**TABLE 3****Machine Learning Analysis Utilizing Live Primary Cell Phenotypic Biomarkers Yields Predictive Power > 0.85 AUC Dependent on Objectively Ranked Biomarkers**

Portion of Features Included in Algorithm	Local Adverse Pathology Potential (LAPP)	Seminal Vesicle Invasion (SVI)	Lymph-node Positive (LNP)	Extra-prostatic Extension (EPE)
All	0.86	0.98	0.99	0.94
Top quartile	0.90	0.97	0.99	0.91
Upper mid-quartile	0.82	0.97	0.99	0.90
Lower mid-quartile	0.82	0.98	0.99	0.90
Bottom quartile	0.83	0.96	0.99	0.86
Top 10	0.86	0.97	0.97	0.90
Top 5	0.84	0.93	0.97	0.78
Top 2	0.77	0.89	0.97	0.75
Top 1	0.66	0.75	0.97	0.76
Bottom 1	0.60	0.83	0.81	0.72

AUC, area under the curve.

As different biomarkers are utilized in the prediction of specific adverse pathologies, the performance of predictions, as measured by AUC, reduces as objectively, lower-ranked/subordinate biomarkers are employed.

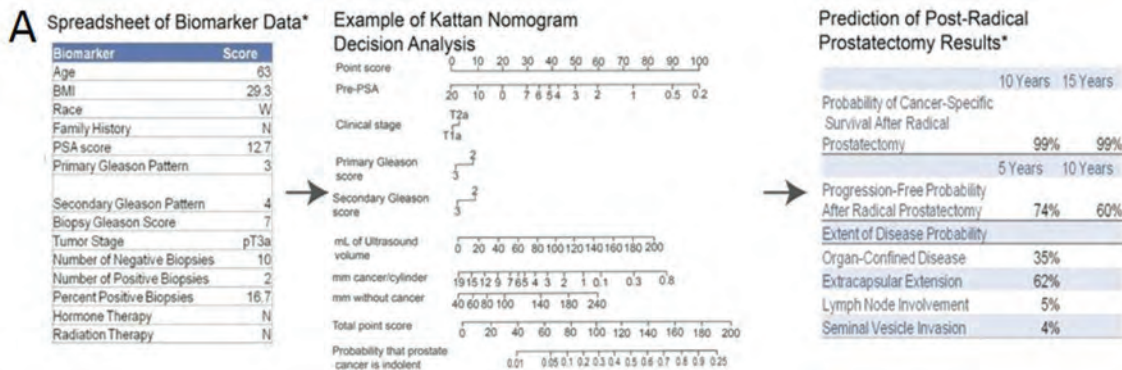


Figure 3. Nomogram-based analysis. Nomograms are the current and primary prediction tools designed to help patients and their physicians understand the nature of a given patient’s prostate cancer. Nomograms are utilized, clinically, in various ways including assessing risk based on specific characteristics of a patient and their disease, as well as predicting the likely outcomes of a given patient’s treatment.

pathologies was assessed by using different fractions of the biomarker features and based on sensitivity, specificity, and AUC values using ROC curve analysis (Figure 4, Table 2). For example, each post-surgical adverse pathology prediction was tested by only including

the features from the top 10 or top 5 or to the top 2 or 1 with the top 10 generally having the best performance relative to the removal of 5 or more top influencing features (Figure 4, Table 2). The ~600 features included to make predictions are only slightly better than

including just the top 10 biomarkers demonstrating that over-fitting from including “too many variables” was not an issue based on the nature of the RF algorithms and their classification scheme.

New biomarkers are being developed in precision oncology. With

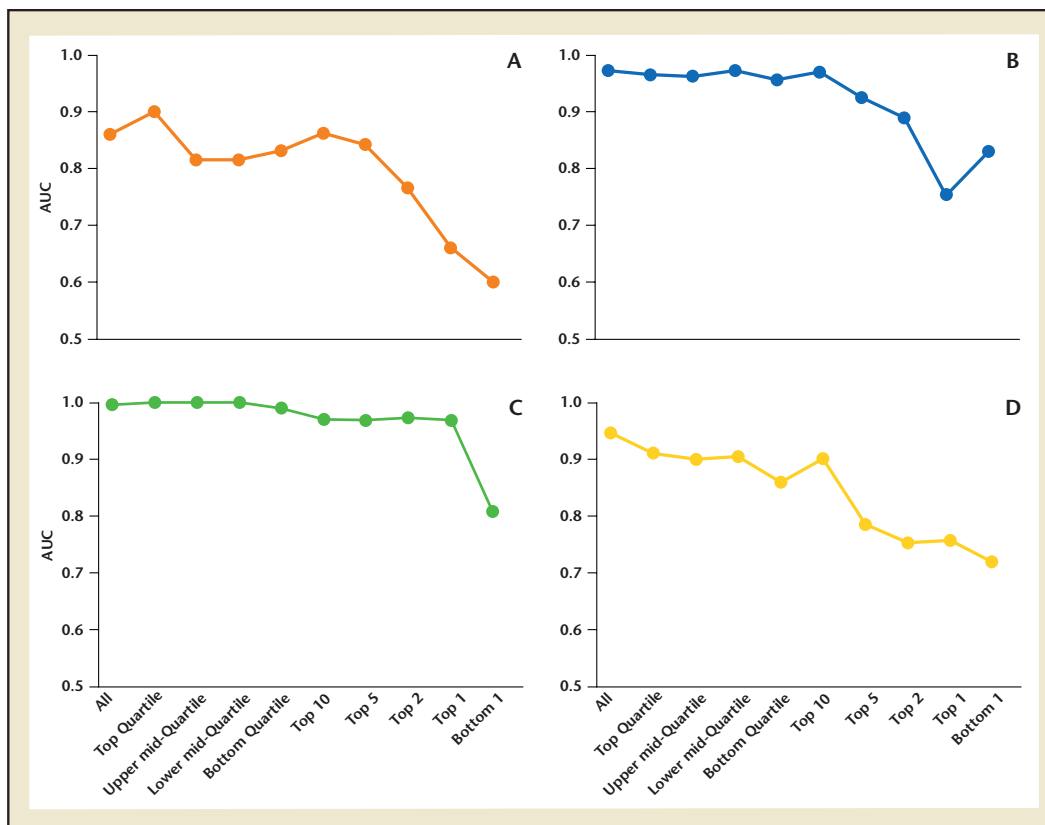


Figure 4. Predictive performance is reduced when, machine learning–based, objectively ranked, “top” biomarkers are excluded in the analysis. The predictive power of machine learning–derived algorithms is dependent on the number of objectively ranked biomarkers utilized to derive a given prediction. Graphs showing area under the curve (AUC) values plotted for each subset of features included in the algorithm show the effects on performance. (A) Local adverse pathology potential (LAPP). (B) Seminal vesicle invasion (SVI). (C) Lymph node positive (LNP). (D) Extra-prostatic extension (EPE).

the inclusion of MV and advances in culturing technology, new types of analyses are possible that extend beyond traditional FFPE-based analysis of tissue. Additional applications of ML to precision oncology and treatment of PCa include digital pathology computer-aided biomarker detection and big-data techniques such as proteomics, genomics, and data mining of electronic health records. A novel application of ML to PCa risk assessment has been reported utilizing the live primary cell phenotypic (LPCP) biomarker test.<sup>9</sup> The LPCP test utilizes the ability to rapidly culture live-biopsy cells then measure, via MV, both dynamic and fixed cellular and molecular phenotypic biomarkers, such as cell motility, protein dynamics, and protein modification states. The LPCP test then leverages a rich phenotypic data set to generate ML-derived statistical algorithms that output clinical scores capable of risk stratifying patients based on several adverse pathology features in radical prostatectomy specimens, for example, seminal vesicle invasion (SVI), lymph-node positive (LNP), and extra-prostatic extension (EPE).

### Discussion

RS of newly diagnosed PCa is currently based on needle biopsy,

tumor volume, and GS or Gleason Grade Group classification. The precision of this RS is challenged by tumor heterogeneity and multifocality and the fact that the Gleason scoring is based on the glandular pattern of the PCa rather than the individual cellular features of the prostatic epithelial cancer cells.

Genomic tissue biomarkers have been developed in recent years to improve RS of PCa and they correlate with adverse pathology at RP and clinical outcomes. The phenotypic platform<sup>9</sup> developed with MVL has shown a high correlation with adverse pathology (SVI, LNP, EPE, etc) with better performance than what is reported for genomic tissue biomarkers. It is important to ensure all biomarker data inputs into ML algorithms are as relevant, quantitative, and objective as possible (Figure 2).

MV automatically, accurately, and precisely identifies patterns and signals within a digitized image to quantify features such as biomarker staining intensity, individual cell shape, and dynamic measurements such as cell movement (division, adhesion) or protein dynamics.

An important consideration in any application of ML to data analysis is over-fitting. To prevent over-fitting, care has been taken to carefully select an optimal ML methodology (RF), the features

(biomarkers) included, the division of training and validation sets (70/30 split), and the methods of evaluation (sensitivity and specificity).

Additional optimization of the current ML algorithms will include testing deep-learning algorithms such as a convolutional neural network (CNN), a specialized type of deep neural network model.

### Conclusions

Current tissue genomic biomarkers used for RS in PCa are based on an a priori selected number of pre-selected genes. The ML algorithm (Figure 2) developed with the live-cell phenotypic biomarker platform is constructed from a large number of phenotypic covariates that have biological and/or clinical significance and putting them together in a statistical model for determining probability (Table 2) of a pathologic outcome.

Clinical nomograms exhibit numerous limitations such as subjective biomarker inputs and suboptimal predictive performance. Artificial intelligence ML-based approaches for cancer RS are emerging as important and powerful tools to compliment current measures of RS, techniques with capabilities to address tumor heterogeneity, and molecular complexity of PCa.

### MAIN POINTS

- Artificial intelligence in the form of machine vision and machine learning are promising technologies that can be applied to improve prostate cancer risk stratification.
- Phenotypic biomarkers provide cellular and molecular details that are informative for predicting post-surgical adverse pathologies when considering tumor biopsy samples.
- Live-cell imaging provides added information regarding cellular and molecular dynamic behavior leading to improved prostate cancer risk stratification.
- The live primary cell phenotypic (LPCP) test has practical clinical utility value applicable to prostate cancer risk stratification.



Specifically, the phenotypic test is a novel example of leveraging biomarkers and advances in MV and ML for developing a powerful prognostic and RS tool for PCa patients. Via the identification of a diverse and objectively quantifiable biomarker set, ML techniques can be utilized to risk stratify patients by being trained on well-established clinical criteria and endpoints toward enhanced patient RS and precision treatment selection. The live single-cell phenotypic platform has the promise of a better measure of tumor biology and aggressiveness compared with Gleason scoring and genomic tissue biomarkers. ■

*For this study involving human patient samples, informed consent was obtained from all individual participants included in the study and under the guidance of IRB approval.*

*This research was funded by Cellanix.*

*Jonathan S. Varsanik, Michael S. Manak, Matthew J. Whitfield, Brad J. Hogan, Wendell R. Su, Grannum R. Sant, David M. Albala, and Ashok C. Chander have stock options in Cellanix.*

## References

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin.* 2014;64:9–29.
2. IBM and Quest Diagnostics Launch Watson-Powered Genomic Sequencing Service to Help Physicians Bring Precision Cancer Treatments to Patients Nationwide. <http://newsroom.questdiagnostics.com/2016-10-18-IBM-and-Quest-Diagnostics-Launch-Watson-Powered-Genomic-Sequencing-Service-to-Help-Physicians-Bring-Precision-Cancer-Treatments-to-Patients-Nationwide>. Published October 18, 2018. Accessed November 30, 2020.
3. Gertych A, Ing N, Ma Z, et al. Machine learning approaches to analyze histological images of tissues from radical prostatectomies. *Comput Med Imaging Graph.* 2015;46:197–208.
4. Li R, Dong X, Ma C, Liu L. Computational identification of surrogate genes for prostate cancer phases using machine learning and molecular network analysis. *Theor Biol Med Model.* 2014;11:37.
5. Gonzalzo ML, Bastian PJ, Mangold LA, et al. Relationship between primary Gleason pattern on needle biopsy and clinicopathologic outcomes among men with Gleason score 7 adenocarcinoma of the prostate. *Urology.* 2006;67:115–119.
6. He Y, Gu J, Strom S, et al. The prostate cancer susceptibility variant rs2735839 near KLK3 gene is associated with aggressive prostate cancer and can stratify gleason score 7 patients. *Clin Cancer Res.* 2014;20:5133–5139.
7. Little J, Wilson B, Carter R, et al. Multigene panels in prostate cancer risk assessment: a systematic review. *Genet Med.* 2016;18:535–544.
8. Sant GR, Knopf KB, Albala DM. Live-single-cell phenotypic cancer biomarkers-future role in precision oncology? *NPJ Precis Oncol.* 2017;1:21.
9. Albala D, Manak MS, Varsanik JS, et al. Clinical proof-of-concept of a novel platform utilizing biopsy-derived live single cells, phenotypic biomarkers, and machine learning toward a precision risk stratification test for prostate cancer grade groups 1 and 2 (Gleason 3+3 and 3+4). *Urology.* 2019;124:198–206.
10. Manak MS, Varsanik JS, Hogan BJ, et al. Live-cell phenotypic-biomarker microfluidic assay for the risk stratification of cancer patients via machine learning. *Nat Biomed Eng.* 2018;2:761–772.
11. Kim SJ, Wang C, Zhao B, et al. Deep transfer learning-based hologram classification for molecular diagnostics. *Sci Rep.* 2018;8:17003.
12. Chander AC, Manak MS, Varsanik JS, et al. Rapid and short-term extracellular matrix-mediated in vitro culturing of tumor and nontumor human primary prostate cells from fresh radical prostatectomy tissue. *Urology.* 2017;105:91–100.
13. Wang J, Wu CJ, Bao ML, et al. Machine learning-based analysis of MR radiomics can help to improve the diagnostic performance of PI-RADS v2 in clinically relevant prostate cancer. *Eur Radiol.* 2017;27:4082–4090.
14. Boyce S, Fan Y, Watson RW, Murphy TB. Evaluation of prediction models for the staging of prostate cancer. *BMC Med Inform Decis Mak.* 2013;13:126.
15. Karakiewicz PI, Benayoun S, Kattan MW, et al. Development and validation of a nomogram predicting the outcome of prostate biopsy based on patient age, digital rectal examination and serum prostate specific antigen. *J Urol.* 2005;173:1930–1934.
16. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol.* 2015;16:e173–e180.
17. Chang AJ, Autio KA, Roach M 3rd, Scher HI. High-risk prostate cancer-classification and therapy. *Nat Rev Clin Oncol.* 2014;11:308–323.
18. Reese AC, Cowan JE, Brajtford JS, et al. The quantitative Gleason score improves prostate cancer risk assessment. *Cancer.* 2012;118:6046–6054.