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)

David Albala[1], Michael S. Manak[1], Jonathan S. Varsanik[1], Hani H. Rashid, Vladimir Mouraviev, Stephen M. Zappala, Ene Ette, Naveen Kella, Kimberly M. Rieger-Christ, Grannum R. Sant, and Ashok C. Chander

OBJECTIVE To examine the ability of a novel live primary-cell phenotypic (LPCP) test to predict postsurgical adverse pathology (P-SAP) features and risk stratify patients based on SAP features in a blinded study utilizing radical prostatectomy (RP) surgical specimens.

METHODS Two hundred fifty-one men undergoing RP were enrolled in a prospective, multicenter (10), and proof-of-concept study in the United States. Fresh prostate samples were taken from known areas of cancer in the operating room immediately after RP. Samples were shipped and tested at a central laboratory. Utilizing the LPCP test, a suite of phenotypic biomarkers was analyzed and quantified using objective machine vision software. Biomarkers were objectively ranked via machine learning-derived statistical algorithms (MLDSA) to predict postsurgical adverse pathological features. Sensitivity and specificity were determined by comparing blinded predictions and unblinded RP surgical pathology reports, training MLDSAs on 70% of biopsy cells and testing MLDSAs on the remaining 30% of biopsy cells across the tested patient population.

RESULTS The LPCP test predicted adverse pathologies post-RP with area under the curve (AUC) via receiver operating characteristics analysis of greater than 0.80 and distinguished between Prostate Cancer Grade Groups 1, 2, and 3/Gleason Scores 3 + 3, 3 + 4, and 4 + 3. Further, LPCP derived-biomarker scores predicted Gleason pattern, stage, and adverse pathology with high precision—AUCs>0.80.

CONCLUSION Using MLDSA-derived phenotypic biomarker scores, the LPCP test successfully risk stratified Prostate Cancer Grade Groups 1, 2, and 3 (Gleason 3 + 3 and 7) into distinct subgroups predicted to have surgical adverse pathologies or not with high performance (>0.85 AUC). UROLOGY 124: 198–206, 2019. © 2018 Elsevier Inc.

[1] *Contributed equally to this work.*

Conflict of Interest Statement: *The authors declare the following competing financial interests: DMA: Genomic Health, speaker; Myriad Genetics, speaker; Cellanyx Diagnostics, stock options; Applied Medical, stock options. MSM: Cellanyx Diagnostics, stock options. JSV: Cellanyx Diagnostics, stock options. HHR: Cellanyx Diagnostics, consultant stock options. VM: Cellanyx Diagnostics, stock options; 3D Biopsy, consultant. SMZ: Cellanyx Diagnostics, stock options. EE: Cellanyx Diagnostics, stock options. NK declares no conflict of interest. KMRC declares no conflict of interest. GRS: Cellanyx Diagnostics, stock options. ACC: Cellanyx Diagnostics, stock options. The following patents cover aspects of the technology presented in this article: WO2012048269A2 and WO2016138041A3.*

From the Department of Urology, Crouse Hospital, Syracuse, NY; the Associated Medical Professionals of New York, Syracuse, NY; the Cellanyx Diagnostics, Beverly, MA; the University of Rochester Medical Center School of Medicine and Dentistry, Rochester, NY; the Urology Department, Central Florida Cancer Institute, Davenport, FL; the Department of Urology, Tufts University School of Medicine, Boston, MA; the Andover Urology, Andover, MA; the Anoixis Corporation, Natick, MA; the The Urology Place, San Antonio, TX; and the Lahey Hospital and Medical Center, Burlington, MA

Address correspondence to: Ashok C. Chander, Ph.D., Cellanyx Diagnostics, 100 Cummings Center, Suite 451C, Beverly, MA 19015. E-mail: ashok@cellanyx.com

Submitted: January 23, 2018, accepted (with revisions): June 14, 2018

https://doi.org/10.1016/j.urology.2018.09.032
0090-4295

Historically, Gleason score (GS) 7 is the most common Prostate cancer (PCa) diagnosis next to GS 6 (49% of cases); in a recent study of >1800 PCa patients, 39% were diagnosed as GS 7.[1] A study of 320 GS 7 patients published in 2006 found 75% had the less aggressive form and were classified as (3 + 4) instead of (4 + 3) postprostatectomy.[2] Improved risk stratification (RS) of GS 6 and 7 by differentiating indolent from aggressive GS 6 and GS 7 cancers may reduce rates of over-treatment.[3] This observation is underscored by the new guidelines that the International Society of Urological Pathology recently issued in 2014.[4] The Prostate Cancer Grading Group (PCGG) 1-5 scoring system defines group 1 as GS 6, grade group 2 as GS 3 + 4 = 7, grade group 3 as GS 4 + 3 = 7, grade group 4 as GS 8, and grade group 5 as GS 9-10. While the new PCGGs have demonstrated the ability to more accurately grade stratify than the current Gleason system,[5] the ability to predict postsurgical adverse pathology (P-SAP) features with precision (>0.85% receiver operator characteristic: area under the curve [AUC]), before surgery, for low-risk and indolent prostate cancers (Gleason 3 + 3/PCGG 1 and Gleason 3 + 4/PCGG 2), remains the "holy grail" of prostate risk-stratification. Toward that aim, and given that P-SAP features remain the "gold-standard" of clinical diagnosis, prediction of P-SAP features from radical prostatectomy (RP) specimens has been the basis of validation of a number of new biopsy-based immunohistochemistry, molecular, and genomic biomarker tests[6-13] that aim to ameliorate the problem of PCa patient RS.

Currently, the primary means to predict P-SAP features and risk stratify patients to guide definitive treatment decisions are the GS, the PCGG scores, as well as genomic tests. Importantly, current methods utilize fixed-biopsy tissue that allows for single time point biomarker measurements. Further, the GS and PCGGs are both based on architectural glandular patterns of the tumor rather than individual cellular characteristics[14] and do not include live-cell phenotypic and biological attributes of individual cancer cells (eg, actin dynamics, cell motility, subcellular protein localization, etc.). Integrating information from a complimentary live-cell prognostic test that leverages innovations in molecular, cellular, and functional dynamic behavioral approaches may improve RS of PCa. Recently, we described a novel patient RS platform capable of predicting P-SAP features—STRAT-AP—[15] that leverages the ability to rapidly culture prostate biopsy cells to measure dynamic and fixed molecular and cellular phenotypic biomarkers with the potential of predicting SAP features. This platform stratified low- and intermediate-risk patients with high sensitivity and specificity in a cohort of ~60 men.[15] The current study (Fig. 1) attempts to validate and demonstrate clinical proof-of-concept of a live primary-cell phenotypic (LPCP) test based on the STRAT-AP technology (Fig. 2) in a sample set of 251 patients by predicting specific and clinically actionable adverse pathology features as well as risk stratifying men based on predicted post-RP GS/PCGG and number of P-SAP features.

In the following study (Fig. 1) we present data that supports the clinical proof-of-concept of the LPCP test (Fig. 1D-F) and its ability to apply novel machine vision to quantify LPCP biomarkers (Fig. 2) and machine learning (Fig. 3) to: (1) predict specific SAP features, (Fig. 3) (2) generate clinical scores capable of risk stratifying patients based on PCGG and GS and number of P-SAP features (Fig. 4) as well as (3) predict biochemical recurrence (BCR) (Supplemental Figure 5), (4) predict stage, (Supplemental Figure 7), and (5) predict predominant postsurgical GS/PCGG, (Supplemental Figure 7) with AUCs>0.80.

MATERIALS AND METHODS

Ten sites (5 university hospitals, 3 large urology group practices, and 2 bio-banks provided fresh prostate samples after Institutional Review Board approval (Fig. 1A). Samples were obtained from the fresh RP specimens via needle core biopsy or excision from (1) known areas of prostate cancer and (2) well away from the know cancer area, as determined by visual inspection of the prostatectomy sample in the operating room and prior diagnostic biopsy results. Areas of prostate cancer were determined by diagnostic biopsy results and confirmed by a board certified pathologist. Samples presented with the following postsurgical GS percentage distribution of the total sample size in the cohort: PCGG 1/Gleason 3 + 3 (13.87%), PCGG 2/Gleason 7 (3 + 4) (49.16%), PCGG 3/Gleason 7 (4 + 3) (25.21%), PCGG 4/Gleason 8 (5.04%), PCGG 5/Gleason 9 (5.04%), PCGG 5/Gleason 10 (1.68%) and tumor stage of pT2a (7.14%), PT2b (0.84%), pT2c (53.78%), pT3a (14.71%), pT3b (8.4%), pT3c (0.42%) (Fig. 1 B-C).

Cell culturing methods are described in reference[16] and in Supplemental information.

Phenotypic biomarkers were measured using differential interference contrast and fluorescence microscopy and described in reference[15] and Supplemental information. Machine vision based measurement of phenotypic biomarkers is described in reference[15] and Supplemental information.

Statistical analysis algorithms (Fig. 3) were used to analyze the data generated from the machine vision image analysis algorithms. Data were independently validated by an external third party statistical analysis group - Anoixis, LLC (Natick, MA).

Machine learning-derived algorithms then ranked individual biomarkers in an unbiased/nonsubjective manner on their ability to predict a given group-of or specific adverse pathology state(s) to determine the exact formal machine learning-derived statistical algorithm (MLDSA) used to for clinical predictions. Machine learning-derived algorithm-mediated biomarker analysis is described in reference[15] and Supplemental information.

RESULTS

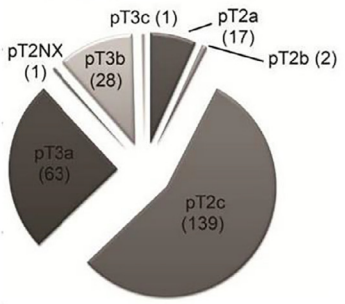
Patient Characteristics, RP Analysis, and Reported Postsurgical Adverse Pathology Features

Two hundred fifty-one RP samples were collected in a blinded fashion with Institutional Review Board approval

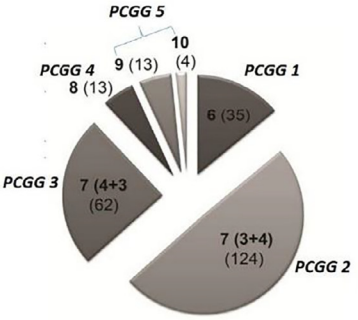
A

Gleason Score	PCGG	Stage	Number of Cases	%	%
6	1	pT2a	6	2.4	13.9
		pT2b	1	0.4	
		pT2c	25	10.0	
		pT3a	3	1.2	
7(3+4)	2	pT2a	9	3.6	49.4
		pT2b	1	0.4	
		pT2c	87	34.7	
		pT2NX	1	0.4	
		pT3a	23	9.2	
		pT3b	3	1.2	
7(4+3)	3	pT2a	2	0.8	24.7
		pT2c	23	9.2	
		pT3a	26	10.4	
		pT3b	11	4.4	
8	4	pT2c	2	0.8	5.2
		pT3a	5	2.0	
		pT3b	5	2.0	
		pT3c	1	0.4	
9	5	pT2c	1	0.4	5.2
		pT3a	4	1.6	
		pT3b	8	3.2	
10	5	pT2c	1	0.4	1.6
		pT3a	2	0.8	
		pT3b	1	0.4	
Total			251	100	100

B Tumor Stage Distribution (Stage (Total number))



C Gleason Score & *Prostate Cancer Grading Group* Distribution (Total number)



IRB Approved Collection Sites

- University Hospitals
- Large Urology Groups
- Biobanks

IRB Accepted Patient Inclusion Criteria

- Male, All Races
- 40-80 years old
- Patients with Gleason Score 6+
- Patients with known clinical outcomes
- Patients not receiving chemotherapy for cancer

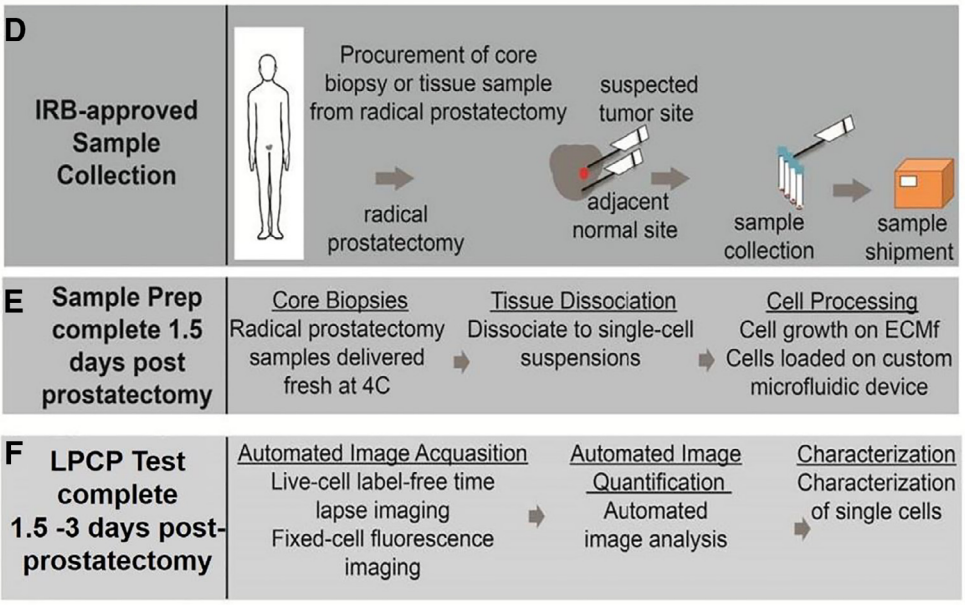
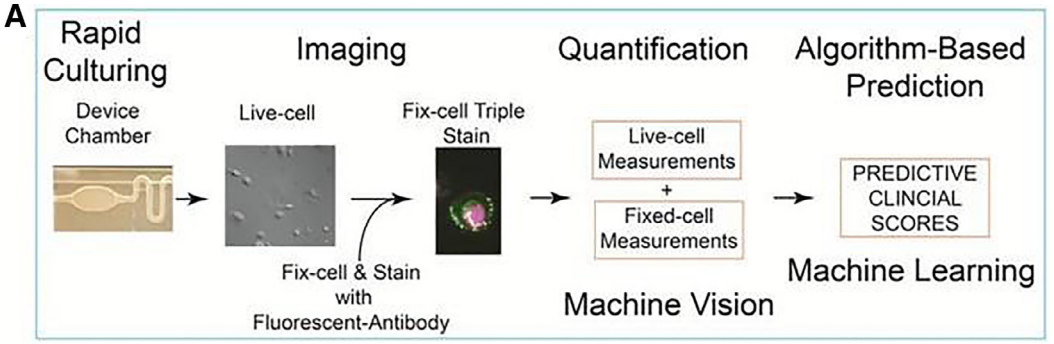
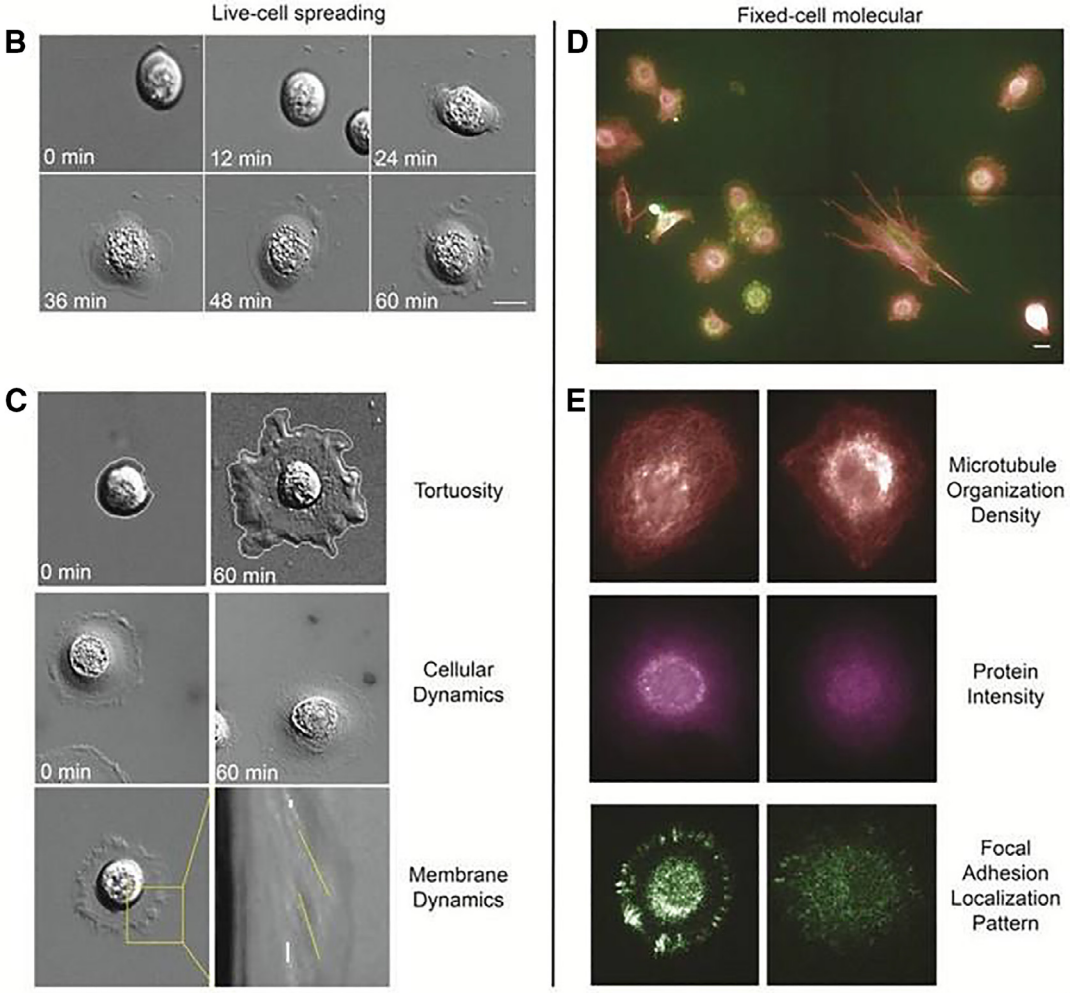


Figure 1. Patient sample postsurgical adverse pathology distribution: Gleason score, Prostate Cancer Grade Group, stage as well as sample selection criteria and live primary-cell phenotypic (LPCP) test biomarker measurement and analysis work flow. (**A**) The cohort distribution is shown for unblinded postsurgical Gleason Score and stage of patient samples. (**B**) Cohort distribution based on tumor stage. (**C**) Cohort distribution based on Gleason score and PCGG. (**D**) Sample collection and sample transport to central laboratory workflow. (**E**) Flow diagram highlighting processing of live-primary samples and sample. (**F**) Phenotypic, (cellular and molecular) biomarkers are measured on live and subsequently fixed patient samples. The flow diagram outlines the process of the LPCP test's biomarker measurement, algorithmic analysis, and generation of predictive measurements. Machine vision image analysis measure biomarkers from the images collected via live- and subsequently fixed-primary samples. Machine learning derived statistical algorithms are employed to characterize cellular biomarkers with single-cell resolution. Abbreviations: PCGG, prostate cancer grade group. (Color version available online.)





F

Representative Biomarker Average and Standard Deviation for n = 237 samples		
Biomarker Feature	**Mean**	**+/- Standard Deviation**
Spreading Velocity	87.00 (nm/sec)	207.71 (nm/sec)
Tortuosity	1.87 (AU)	0.46 (AU)
Cellular Dynamics	11.11 (nm/sec)	36.83 (nm/sec)
Membrane Dynamics	0.04 (nm/sec)	0.01 (nm/sec)
Protein Intensity	0.02 (AU)	0.01 (AU)
Focal Adhesions	28.73	22.70
Focal Adhesion Distance	1.14 (nm)	2.33 (nm)

Figure 2. The LPCP test utilizes an automated machine vision process for biomarker measurements to rapidly assess primary prostate tissue, allowing machine learning algorithms to objectively analyze biomarker measurements. **(A)** The LPCP test process of rapidly culturing primary prostate tissue within 72 hours, imaging and quantification of both live and fixed cells, matching live and fixed cell measurements to be input into machine learning algorithms. **(B)** Live-cell biomarkers are measured via Differential Interference Contrast (DIC) time-course imaging of primary prostate cells rapidly derived from radical prostatectomy samples. **(C)** Examples of biomarkers measured during live cell analysis: tortuosity, cellular dynamics such as motility, and membrane dynamics that provide information about actin dynamics. Two different phenotypes are presented to show the possible range of phenotypes for tortuosity and cellular dynamics. **(D)** A representative field of view for label-based fixed-biomarker cell analysis. **(E)** Example of fixed-biomarker measurements: microtubule organization, protein intensity and modification state, and focal adhesion localization pattern. Two different phenotypes are shown for each example of fixed cell biomarkers. **(F)** Automated quantification by machine vision algorithms and statistics on cell population for example biomarkers. Abbreviations: DIC, differential interference contrast; LPCP, live primary-cell phenotypic. (Color version available online.)

and exhibited the following distribution for GS/PCGG: 14%—GS 3 + 3/PCGG 1, 49%—Gleason 3 + 4/PCGG 2, 25%—GS 4 + 3/PCGG 3, 5%—GS 8/PCGG 4, and 7%—GS 9 (GS10/PCGG 5). Patients were males between the ages of 40 and 80 years old of any race that did not have history of infectious disease or were previously treated with chemotherapy (Fig. 1A-C).

Of the 251 samples, only 238 samples were successfully cultured as 13 samples exhibited yeast or bacterial infection during collection or transport and were unable to be cultured. A total of 237 samples were successfully cultured and analyzed after a machine vision-based biomarker quantification error occurred during the analysis process of one sample.

LPCP BIOMARKER MEASUREMENT USING MACHINE VISION AND ANALYSIS VIA MACHINE LEARNING DERIVED ALGORITHMS

Based on the STRAT-AP technology, developed by Cellanyx Diagnostics (Beverly, MA),[15] the LPCP test analyzes primary biopsy tissue biomarkers using both live-cell and fixed-cell imaging techniques to generate both dynamic and static phenotypic biomarker measurements that can be input into machine learning algorithms to predict P-SAP features as well as other clinical end points such as BCR (Fig. 2A). Each patient's live cells are imaged to assess dynamic cellular and molecular biomarkers including, but not limited to, adhesion, tortuosity, cell motility, membrane dynamics, and actin dynamics (Fig. 2B and C)—indirect measures of cell-extra-cellular matrix (ECM) interactions, cellular mechanics, and signal pathway activation responsible for motility and membrane dynamics. Fixed-biomarkers are subsequently analyzed utilizing immunofluorescence to assess specific protein localization and modification states enabling the interrogation of specific signaling pathways (Fig. 2D and E).

Machine vision software track individual cells and correlate both live-cell and fixed-cell LPCP biomarker measurements for each corresponding cell, with single cell resolution, to achieve the quantification of cellular and molecular phenotypic primary and aggregate biomarkers for each cell (Fig. 2F). The LPCP machine learning algorithms, based on random forest, decision tree analysis[15] are trained to develop MLDSAs that can predict if a patient is positive or negative for a given P-SAP feature (Fig. 3). MLDSAs then generate threshold values that determine if a patient is positive or negative for a specific or group of P-SAP features allowing prediction performance to be assessed.

Clinical Workflow and Prediction of Specific Postsurgical Adverse Pathology Features, Postsurgical Stage, Postsurgical GS/PCGG, and BCR

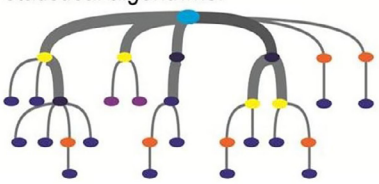
The LPCP test was designed as a laboratory-developed test that could be run in any CLIA laboratory. Specifically, patient sample characteristics and acquisition were optimized to fit seamlessly in the prostate biopsy workflow of the urologist. Figure 3A summarizes the current surgical workflow for histopathological analysis and postsurgical GSPCGG reporting compared to the LPCP test workflow (Fig. 3B).

Specific postsurgical adverse pathologies (ie, EP, positive surgical margins [PSM], and seminal vesicle invasion) were predicted and compared to unblinded postsurgical report findings. The performance (AUCs > 0.80) of these predictions is summarized in Figure 3C and ROC curves are presented in Supplemental Figure 3. Specifically, prediction of P-SAP features such as PSM, extra-prostatic extension (EPE), seminal vesicle invasion, and lymph node positive exhibited AUCs>0.85. Importantly, when field samples containing nontumor or field, tissue was analyzed on MLDSAs trained on tumor/cancer tissue the MLDSAs demonstrated suboptimal AUCs of 0.57-0.74 for specific P-SAP features (PSM, EPE, Peri-neural invasion (PNI), lymph node positive, and lympho-vascular invasion (LVI)) confirming that the cancer trained MLDSAs are specific to tumor tissue and not applicable to nontumor cells (Supplemental Figure 3). Interestingly, the LPCP test was also able to predict BCR (Supplemental Figure 5) as well as postsurgical stage and postsurgical GS/PCGG with AUCs>0.80 (Supplemental Figure 7). Similar to the prediction of specific P-SAP features, the tumor-trained MLDSAs were specific to tumor-samples and performed suboptimally when applied to nontumor/field tissue supporting the idea that tumor-based MLDSAs are specific to tumor samples and not random sampling of the prostate (Supplemental Figure 7).

A Biomarkers are automatically measured for each cell imaged from a patient sample.

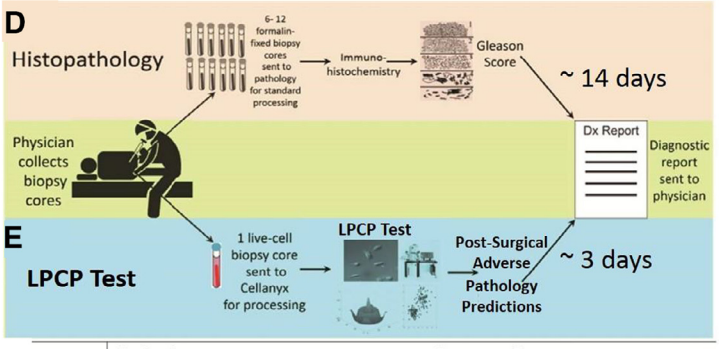
	Patient				
Cell ID	Marker 1	Marker 2	Marker 3	Marker 4	..
1	7353.5	509.42	2.9523	30744	..
2	40526	1576.3	5.2008	22409	..
3	7063.6	578.66	4.0195	75730	..
4	18066	1263	7.2896	2.12E+05	..
5	9470.8	488.55	2.0773	2.09E+05	..

B Biomarker measurements are analyzed through decision tree statistical algorithms.



C **Scores are generated by Machine Learning Derived Statistical Algorithms (MLDSAs) from biomarker analysis useful for predicting post-surgical adverse pathology features:**

- **Seminal Vesicle Invasion (SVI)**
- **Positive Surgical Margins (PSM)**
- **Extra-Prostatic Extension (EPE)**
- **Perineural Invasion (PNI)**
- **Lympho-Vascular Invasion (LVI)**
- **Lymph Node Positive (LNP)**
- **T3 – stage (T3)**
- **> Gleason Pattern 4 (PCGG 3-5)**
- **SVI & LNP & T3 & PCGG 3-5**



F

Predicted Adverse Pathology	Number of Samples per Adverse Pathology	Sensitivity	Specificity	AUC	True Positive for Adverse Pathology	True Negative for Adverse Pathology	Predicted Positive	Predicted Negative
Positive surgical margin	236	0.81	0.83	0.87	69	167	56	139
Seminal vesicle invasion	228	0.90	0.79	0.87	29	199	26	157
Extra-prostatic extension	221	0.83	0.86	0.89	90	131	75	113
Perineural invasion	202	0.87	0.90	0.94	163	39	142	35
Lymph node positive	189	0.92	0.97	0.94	12	177	11	172
Lymph vascular invasion	223	0.99	0.80	0.90	21	202	21	161
T3	237	0.85	0.77	0.82	88	149	75	115
PCGG 3-5	237	0.76	0.86	0.85	86	151	65	130
SVI & LNP & T3 & PCGG 3-5	185	0.84	0.77	0.84	115	122	97	94
SVI or LNP or T3 or PCGG 3-5	237	0.99	0.92	0.96	5	179	5	164

Figure 3. Quantified biomarkers are input into random forest decision trees to generate Machine Learning Derived Statistical Algorithms (MLDSAs) that subsequently generate predictions of postsurgical adverse pathology features. **(A)** Biomarker measurements, quantified from machine vision analysis are recorded in a database that is input into machine learning algorithms to train, and subsequently test, Machine Learning Derived Statistical Algorithms (MLDSAs). **(B)** MLDSAs are based on random forest decision tree analysis. **(C)** MLDSAs predict generate predictions of clinically relevant postsurgical adverse pathology features. Clinical workflow of Live-Primary Cell Phenotypic test to standard histology/Gleason Score/Prostate Cancer Grading Group workflow and predictive performance for post-surgical adverse pathology features. **(D)** Clinical workflow comparison of histopathology (Gleason/PCGG) analysis with the LPCP test. **(E)** The LPCP test requires only one biopsy compared to 12 and takes 72 hours to return results after being processed in a central laboratory via high-content microscopy, machine vision and machine learning-derived statistical algorithms (MLDSAs). **(F)** Statistical performance table summarizing the LPCP test's ability to predict specific postsurgical adverse pathology features. Abbreviations: LPCP, live-primary cell phenotypic; PCGG, prostate cancer grade group. (Color version available online.)

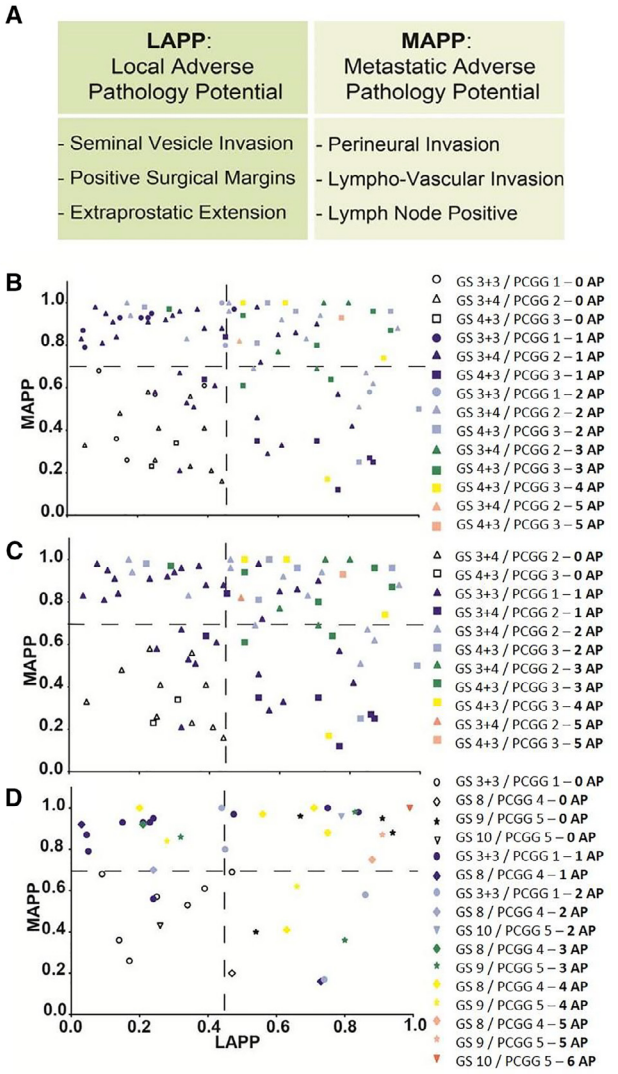


Figure 4. MLDSAs of the LPCP test generate predictive clinical scores LAPP and MAPP toward enhanced patient risk stratification of Gleason 3 + 3/PCGG 1 and Gleason 3 + 4/PCGG 2 patients based on number of postsurgical adverse pathology features. Patient risk stratification of prostate cancers has been demonstrated with the LPCP test in an early clinical trial. Each point represents the ensemble behavior of the cell population of a single patient sample. Cells score as a distinct population with high MAPP and LAPP when derived from samples with high number of adverse pathologies. Intermediate number of pathology samples cluster into different areas than those of higher or lower reported adverse pathologies. The lines that cross at 0.45 on the x-axis and 0.69 on the y-axis represent the thresholds of LAPP and MAPP, respectively. **(A)** Definition of Local Adverse Pathology Predictor (LAPP) and Metastatic Adverse Pathology Predictor (MAPP) scores **(B)** 100 samples were randomly chosen such that the Gleason score and adverse pathology distributions are the same as the complete data set. MAPP and LAPP scatter plot comparing Gleason 3/PCGG 1 samples to Gleason 7/PCGG 1 and 2 samples shows how these Gleason samples can be risk stratified using the phenotypic biomarker based scoring metrics. **(C)** LAPP and MAPP scatter plot comparing Gleason 3 + 4/PCGG 1 samples to Gleason 4 + 3/PCGG 2 samples shows how these patient samples can be risk stratified using the LPCP test scoring metrics — LAPP and MAPP. **(C)** LAPP and MAPP scatter plot comparing Gleason 3 + 3/PCGG 1 samples to Gleason 8/PCGG 4 or higher samples confirms that these patient samples can be risk stratified using the LPCP based scoring metrics of LAPP and MAPP. Importantly, samples with no adverse pathologies cluster below the threshold scores in the lower left quadrant regardless of Gleason score. Samples with more adverse pathologies cluster toward the right and above the threshold score. In each case the level of detail provided by the LPCP's scoring metrics supplements the Gleason score/PCGG. Abbreviations: AP, post surgical adverse pathology; LPCPs, live primary-cell phenotypic; MLDSAs, machine learning derived statistical algorithms; PCGG, prostate cancer grade group. (Color version available online.)

Predictive Performance of MAPP, and LAPP Scores (Adverse Pathology Scores) and Enhanced RS of GS 3 + 3/PCGG 1 and GS 3 + 4/PCGG 2 Patients Based on LAPP and MAPP Scores

Toward enhanced patient risk-stratification, P-SAP scores were developed and defined as: Local Adverse Pathology Potential (LAPP) score (prediction of EPE, seminal vesicle invasion, and positive surgical margins), and Metastatic Adverse Pathology Potential (MAPP) score (prediction of only perineural invasion, lympho-vascular invasion, and lymph node involvement). Both the LAPP and MAPP score incorporate features that are used in clinical practice. Given such a definition, the data support that LAPP and MAPP enable a clinically meaningful identification of normal vs malignant tissue and enhanced risk-stratification of indolent and aggressive patients within GS 3 + 3/PCGG 1 and GS 3 + 4/PCGG 2. To demonstrate RS based on adverse pathologies at the time of RP, a plot of LAPP vs MAPP (Fig. 4) of a random selection of 100 patient data points (representative set of data of full data set [Supplemental Figure 4]) demonstrates the ability of the diagnostic platform to stratify samples based on the number of postsurgical adverse pathological features noted on the unblended, respective RP surgical pathology reports (Fig. 4 and Supplemental Figure 4). The plots highlight how the thresholds produced by the MLDSAs can be used to risk stratify patients that have similar postsurgical GSs/PCGGs yet end up having different associated postsurgical adverse pathologies for example, distinguishing a GS 3 + 3/PCGG 1 patient with SAP features from a GS 3 + 3/PCGG 1 patient without P-SAP features.

DISCUSSION

In this study we utilized a first-in-class LPCP test[15] (Fig. 1), to measure a suite of novel phenotypic and dynamic biomarkers in live prostate cells with single cell resolution. This LPCP test was assessed for its ability to (1) predict specific P-SAP features and (2) risk stratify GS

3 + 3/PCGG1 and GS 3 + 4/PCGG 2 patients for adverse pathology to demonstrate the clinical proof-of-concept of a prostate cancer RS test.

The primary limitation of this study is that the same patient samples were used as training and test sets. That said, given that the LPCP test measures biomarkers with single-cell resolution, each training/test data point can be viewed, and thought of, as a single cell which enables the randomized grouping of 70% of cells from the patient population to train MLDSAs and subsequent blinded utilization of the remaining distinct 30% of cells as the test set. Importantly, a major limitation of the current study is that test data sets (single cell data) are taken from the same samples that were included in our test set and work is being undertaken to train this model on a wider set of samples, expanding the range of samples for which a more robust prediction can be generated based on completely different training and test sets. Further, this study only collected patient's samples from 10 different sites within the United States and only analyzed 237 patient samples to generate sensitivity and specificity measurements in a short-term, randomized prospective study using blinded P-SAP results after analyzing RP surgical samples. Additionally, another limitation of the presented study is that clinical patient data, such as presurgical stage and GS, age, and race, were not incorporated into the MLDSAs. Finally, given our observation that tumor-trained MLDSAs were unable to predict specific P-SAP features when analyzing field/nontumor samples, the LPCP test requires tumor rich sample to provide accurate predictions. Results from a study to train new MLDSAs on field, or nontumor, are forthcoming.

New and emerging molecular and genomic tests aimed at risk stratifying patients have used the prediction of adverse pathology as the end point of their respective proof-of-concept studies followed by subsequent long-term outcome studies. This proof-of concept study demonstrated that the LPCP test and its suite of phenotypic biomarkers[15] improved upon prediction of P-SAP features and existing PCa RS measures. Analysis of 237 patients yielded strong performance with the ability to predict if a GS 3 + 3/PCGG 1 or GS 3 + 4/PCGG 2 patient will experience P-SAP (with AUCs>0.80) without the use of other patient data such as GS, patient age, or PSA level. Addition of these demographic and clinical data may in fact further increase the performance of LPCP and is the subject of continued investigation. The LPCP test's strong performance may be attributed to not only the measurement of biomarkers with single-cell resolution, but the ability to measure dynamic biomarkers and cell behavior objectively with machine vision (Fig. 2) as well as unbiased biomarker ranking by machine learning algorithms (Fig. 3) that predict P-SAP features and risk stratify GS 3 + 3/PCGG 1 and GS 3 + 4/PCGG2 patients (Figs. 2C and 4).[15] Indeed, LPCP is the first in a growing class of tests to measure live-cell, dynamic, and biomarkers.[17] Interestingly, comparison of predictive performance when using biopsy tissue from noncancerous tissue demonstrated that LAPP, and MAPP based predictions are specific to cancerous tissue (Supplemental information, Supplemental Figures 7 and 8).

Scores derived from phenotypic biomarkers assessing tumor aggressiveness (ie, LAPP), metastatic invasion potential (ie, MAPP), or combined as General Adverse Pathology Potential (GAPP) (Supplemental Figure 3) may then provide objective and actionable insight for treatment selection for GS 3 + 3/PCGG 1 and GS 3 + 4/PCGG 2 patients whether it be active surveillance or definitive therapy with surgical intervention or radiotherapy.

This novel live-cell phenotypic diagnostic platform provides unique, and previously inaccessible, cellular and molecular biomarker information to assess temporal and spatial cellular and subcellular features that predict pathobiological behavior and adverse pathology. The phenotypic biomarkers measured using the LPCP test resulted in statistically significant predictions of P-SAP features in RP specimens. The LPCP test, as a member of a new class of diagnostic/prognostic tests, generated clinically predictive scores: LAPP, and MAPP, for the RS of PCa, and has the potential to improve personalized prognosis and treatment guidance in men with newly diagnosed prostate cancer.

The strong performance demonstrated by this novel live-cell phenotypic biomarker test in predicting adverse pathologies at the time of RP supports a similar follow-up study from patient biopsy samples prior to RP. A multi-center prostate needle biopsy is planned to confirm the findings of the presented proof of concept study that utilized prostate tissue from RP specimens. This will further support the LPCP test's ability to better understand the dynamic and heterogeneous biology of PCa toward precision RS and aid in personalized treatment decision-making.

Acknowledgment. In addition to Lahey Hospital and Medical Center, American Medical Professionals of New York, and Urology Place of San Antonio, tissue samples were provided by the NCI Cooperative Human Tissue Network (CHTN). Authors respectfully acknowledge: Dr. Hui-May Chu for thoughtful review of the manuscript and Dr. Pravin Chaturvedi, Mr. Robert Gottileb, and Ms. Mary Beth Lisman for discussion as well as Dr. Mani Foroohar, and Dr. Travis B Sullivan. MSM acknowledges Jenn A. Manak, Chloe A. Manak, Gabriel H. Manak, Paul L. Manak, Linda L. Manak, and Paul W. Manak for support. JSV acknowledges Sarah H. Varsanik for support. ACC acknowledges Coimbatore S. Chandersekaran, Achamma C. Chandersekaran, Angela B. Cravens Chander, and Indira E.A. Chander for support.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.1016/j.urology.2018.06.068.

References

1. He Y, Gu J, Strom S, Logothetis CJ, Kim J, Wu X. 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.
2. Gonzalgo 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.
3. Miller DC, Gruber SB, Hollenbeck BK, Montie JE, Wei JT. Incidence of initial local therapy among men with lower-risk prostate cancer in the United States. *J Natl Cancer Inst*. 2006;98:1134–1141.
4. Egevad L, Delahunt B, Srigley JR, Samaratunga H. International Society of Urological Pathology (ISUP) grading of prostate cancer - An ISUP consensus on contemporary grading. *APMIS*. 2016;124:433–435.
5. Epstein JI, Zelefsky MJ, Sjoberg DD, et al. A contemporary prostate cancer grading system: a validated alternative to the gleason score. *Eur Urol*. 2016;69:428–435.
6. Cuzick J, Swanson GP, Fisher G, et al. Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. *Lancet Oncol*. 2011;12:245–255.
7. Klein EA, Cooperberg MR, Magi-Galluzzi C, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. *Eur Urol*. 2014;66:550–560.
8. Little J, Wilson B, Carter R, et al. Multigene panels in prostate cancer risk assessment: a systematic review. *Genet Med*. 2016;18:535–544.
9. Mendhiratta N, Meng X, Taneja SS. Using multiparametric MRI to 'personalize' biopsy for men. *Curr Opin Urol*. 2015;25:498–503.
10. Ross AE, Feng FY, Ghadessi M, et al. A genomic classifier predicting metastatic disease progression in men with biochemical recurrence after prostatectomy. *Prostate Cancer Prostatic Dis*. 2014;17:64–69.
11. Scarpato KR, Barocas DA. Use of mpMRI in active surveillance for localized prostate cancer. *Urol Oncol*. 2016;34:320–325.
12. Sternberg IA, Vela I, Scardino PT. Molecular profiles of prostate cancer: to treat or not to treat. *Annu Rev Med*. 2016;67:119–135.
13. Zhuang L, Johnson MT. How precisely can prostate cancer be managed? *Int Neurourol J*. 2016;20:S120–S130.
14. Epstein JI, Allsbrook Jr. WC, Amin MB, Egevad LL, Committee IG. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol*. 2005;29:1228–1242.
15. Manak MS, Varsanik JS, Hogan BJ, et al. Live-cell phenotypic-biomarker microfluidic assay for the risk stratification of cancer patients via machine learning. *Nature Biomedical Engineering*. 2018.
16. 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.
17. Sant GR, Knopf KB, Albala DM. Live-single-cell phenotypic cancer biomarkers-future role in precision oncology? *npj Precision Oncology*. 2017;1:7.