Pacific Northwest
NATIONAL LABORATORY

Application of advanced proteomic capabilities to identify a tissue-based classifier predicting distant metastasis or biochemical recurrence

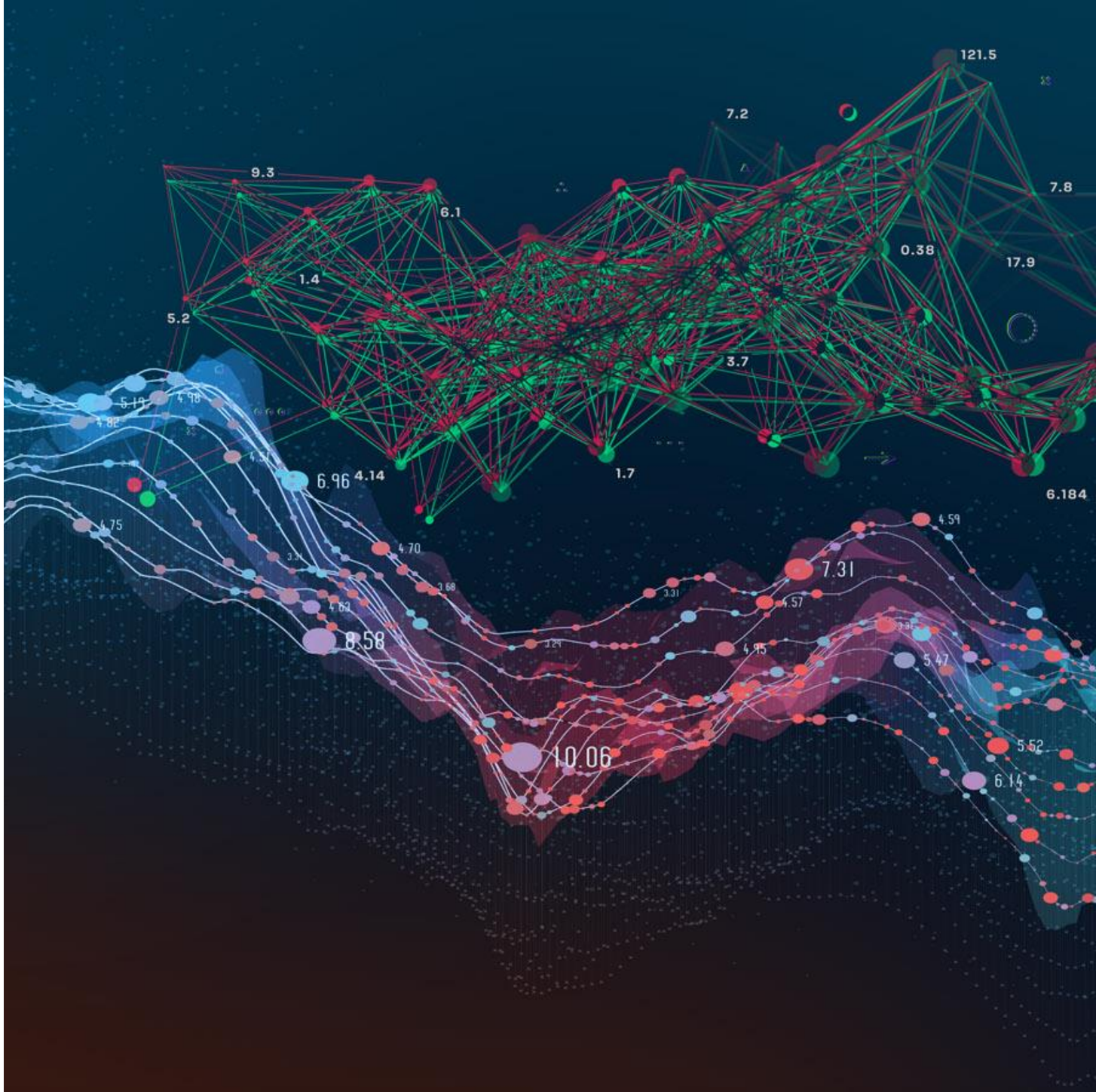
Tao Liu

Senior Staff Scientist

EDRN GU Collaborative Group Meeting

June 30, 2020

U.S. DEPARTMENT OF ENERGY



The early detection problem in prostate cancer: discriminating indolent from aggressive disease

- Although ~40% of screen-detected prostate cancers are indolent, advanced-stage PCa is a lethal disease with 5-year survival rates around 29%
- Identification of biomarkers for early detection of aggressive disease is a key challenge
- *What is/are the intended use and/or clinical utility of your markers or assays?*
  - **Improved risk stratification** based on the initial diagnostic biopsy to assign patients to active surveillance or radical prostatectomy
  - **Improved accuracy** for predicting the likelihood of aggressive disease following radical prostatectomy

Identifying a proteomic classifier from genomic candidates using PRISM-SRM and RP tissues

52-protein panel for verification

PCa prognosis associated genes	Other PCa associated genes	Other cancer related genes
AKT1	AMACR	BRAF
ANXA2	CRISP3	CAMKK2
AR	pan-ERG	EGFR
AURKA	ERG8	HIF1A
CCND1	ETV1	HPN (TMPRSS1)
CDKN1A	FOLH1 (PSMA)	HSPB1
EZH2	HOXC6	MMP2
FGFR1	KLK2	MMP9
MUC1	KLK3 (PSA)	PDGFRB
MUC6	KLK11	PIK3CA
MYC	MYO6	PLA2G7
MYCN	NPY	ODC1
NCOA2	PSGR	RAF1
PMP22	SPARC	SERPINI1
SMAD4	TWIST1	STAT3
SPINK1		TGFB1
SPP1 (OPN)		TP53
TFF3		TPM2
		VEGFA

Oncotarget **2019**, 10, 6466–6483

- Radical prostatectomy FFPE samples from 338 cases selected and provided by CPDR (blinded analysis)
  - **distant metastasis** (DM): ≥ 1 year post-RP
  - **biochemical recurrence** (BCR): ≥ 1 year post-RP
  - **no progression**: no progression after ≥ 10 years post-RP
- Highly sensitive, multiplexed PRISM-SRM assays

Descriptive statistics for study cohort and clinical base models

To evaluate the improvement of performance over two standard of care (SOC) base models:

- **Biopsy** SOC base model (at diagnosis)

= Age + Race + NCCN risk strata

- Best for predicting outcomes that involve time points *prior to surgery*

- **Pathology** SOC base model (post-RP)

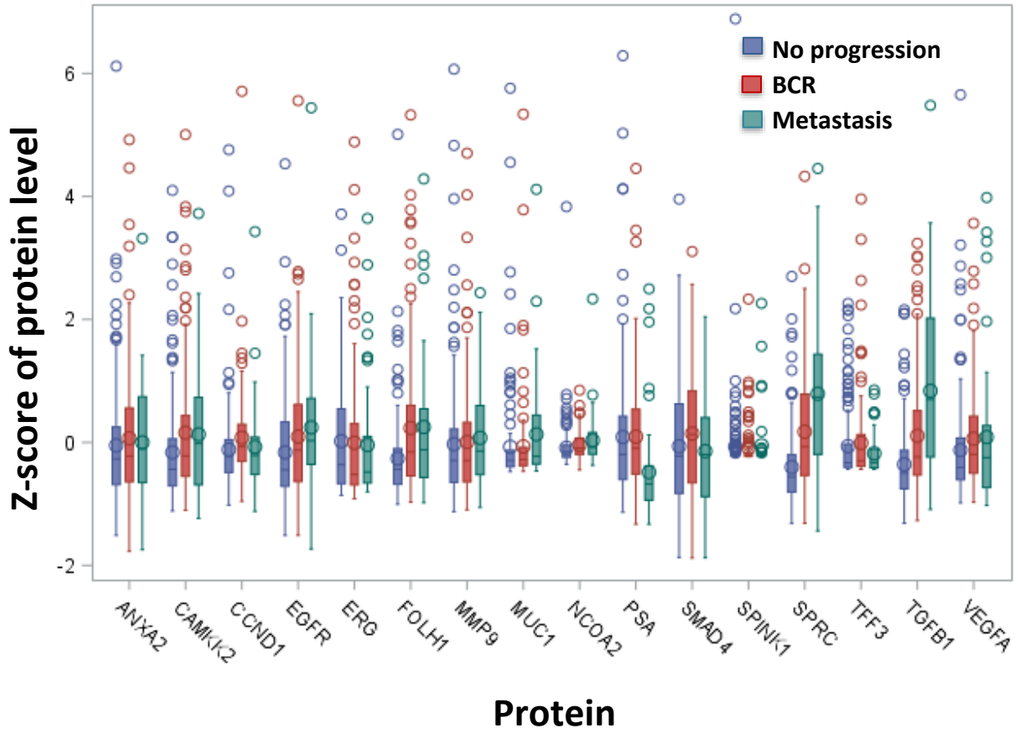
= Pathologic T stage + Grade Group + Surgical margin

- Best for predicting outcomes that involve time points *post-RP*

Variable	Total	Non-event	BCR	Metastasis	*P* value
N	338	**161**	**124**	**53**	
Age at diagnosis (years)					
Mean (SD)	59.5 (7.7)	59.0 (8.1)	59.2 (7.7)	61.7 (5.9)	0.0897
Time from diagnosis to RP (months)					
Median (range)	2.3 (0.2-21)	2.2 (0.2-21)	2.5 (0.2-9)	2.0 (0.7-10)	0.4689
Race					
AA	120 (35.6)	55 (34.2)	48 (39.0)	17 (32.1)	
CA & Other	217 (64.4)	106 (65.8)	75 (61.0)	36 (67.9)	0.5882
PSA at diagnosis (ng/mL)					
<10	262 (78.0)	133 (83.6)	90 (72.6)	39 (73.6)	
10-20	59 (17.6)	25 (15.7)	25 (20.2)	9 (17.0)	
>20	15 (4.5)	1 (0.6)	9 (7.3)	5 (9.4)	**0.0062**
Clinical T stage					
T1-T2a	274 (82.0)	134 (85.4)	107 (86.3)	33 (62.3)	
T2b-T2c	52 (15.6)	22 (14.0)	15 (12.1)	15 (28.3)	
T3a-T4	8 (2.4)	1 (0.6)	2 (1.6)	5 (9.4)	**0.0005**
Biopsy grade					
6 or less	182 (58.3)	100 (70.9)	68 (57.1)	14 (26.9)	
=7	95 (30.4)	35 (24.8)	41 (34.4)	19 (36.5)	
8-10	35 (11.2)	6 (4.3)	10 (8.4)	19 (36.5)	**<.0001**
NCCN risk					
Low	125 (40.6)	69 (50.7)	46 (38.3)	10 (19.2)	
Intermediate	134 (43.5)	59 (43.4)	55 (45.8)	20 (38.5)	
High	49 (15.9)	8 (5.9)	19 (15.8)	22 (42.3)	**<.0001**
Pathological T stage					
pT2	174 (52.6)	119 (74.4)	46 (37.4)	9 (18.8)	
pT3-4	157 (47.4)	41 (25.6)	77 (62.6)	39 (81.2)	**<.0001**
GG					
GG1	31 (9.3)	18 (11.2)	13 (10.6)	0	
GG2	105 (31.6)	77 (48.1)	27 (22.0)	1 (2.0)	
GG3	6 (1.8)	2 (1.2)	4 (3.2)	0	
GG4	124 (37.4)	54 (33.8)	49 (39.8)	21 (42.9)	
GG5	66 (19.9)	9 (5.6)	30 (24.4)	27 (55.1)	**<.0001**
Surgical margin					
Negative	209 (63.7)	126 (79.2)	62 (51.2)	21 (43.8)	
Positive	119 (36.3)	33 (20.8)	59 (48.8)	27 (56.2)	**<.0001**
Post-RP Follow-up (months)					
Median (range)	150 (18-253)	156 (121-252)	129 (18-229)	124 (24-253)	**<.0001**

PRISM-SRM analysis of candidate proteins in RP tissues for correlation with clinical outcome

338 FFPE samples

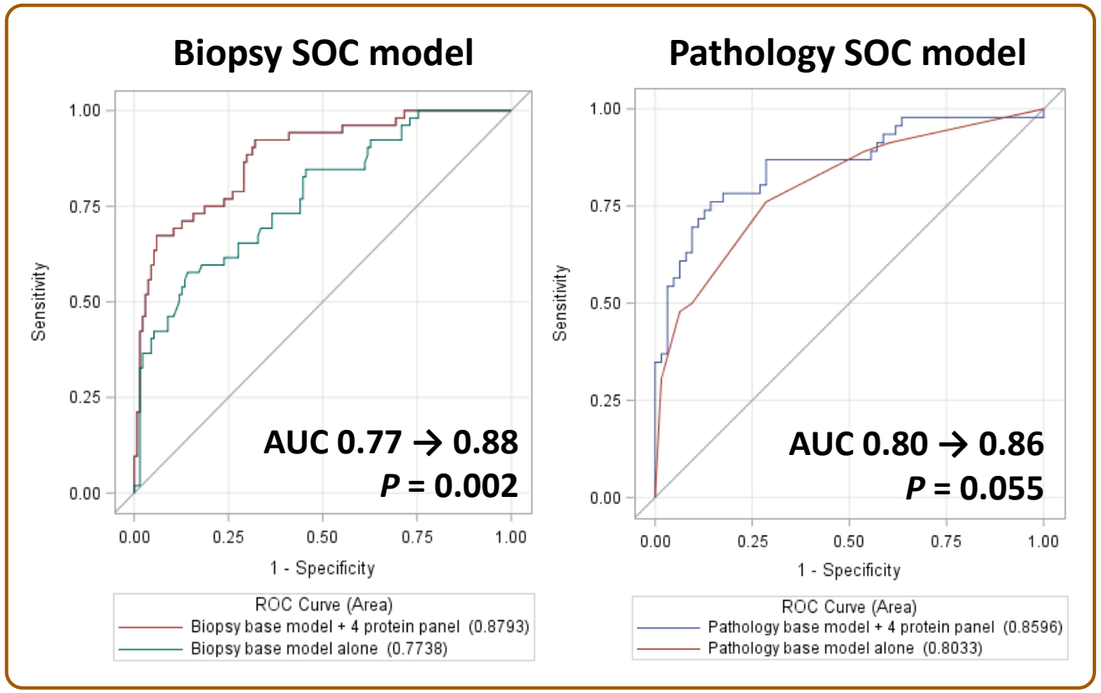


Univariable logistic regression and ROC curve analysis

Protein	DM vs. non-event		BCR vs. non-event		GG (3-5 vs. 1-2)	
	AUC	*P* value	AUC	*P* value	AUC	*P* value
ANXA2	0.535	0.741	0.538	0.341	0.499	0.692
CAMKK2	0.591	0.051	0.604	0.009	0.667	**<.001**
CCND1	0.532	0.166	0.624	0.037	0.592	0.034
EGFR	0.628	0.012	0.578	0.035	0.653	**<.001**
ERG	0.543	0.668	0.546	0.830	0.482	0.708
FOLH1	0.653	**0.001**	0.627	**<.001**	0.657	**<.001**
MMP9	0.562	0.518	0.511	0.770	0.554	0.643
MUC1	0.570	0.461	0.474	0.603	0.506	0.200
NCOA2	0.637	0.095	0.613	0.225	0.670	**0.001**
PSA	0.730	**0.001**	0.529	0.955	0.608	0.005
SMAD4	0.511	0.622	0.526	0.092	0.521	0.383
SPINK1	0.486	0.207	0.548	0.535	0.547	0.470
SPARC	0.800	**<.001**	0.695	**<.001**	0.715	**<.001**
TFF3	0.541	0.174	0.472	0.578	0.492	0.751
TGFB1	0.788	**<.001**	0.649	**<.001**	0.705	**<.001**
VEGFA	0.528	0.168	0.601	0.040	0.573	0.009

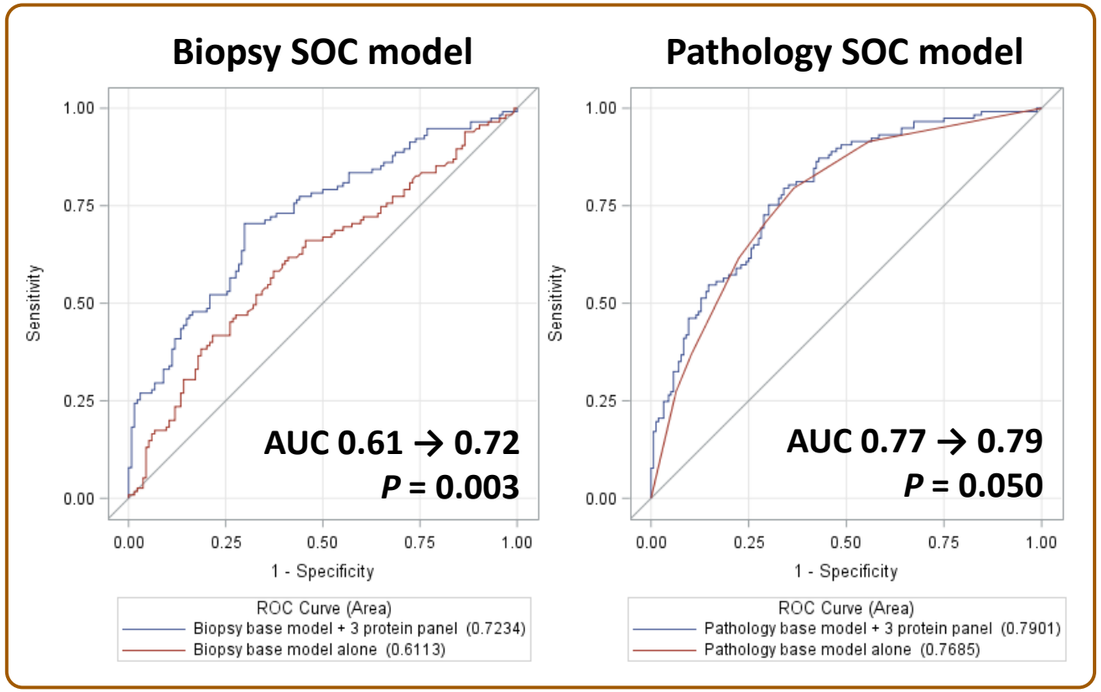
Initial performance evaluation: protein panels for predicting DM and BCR

DM



Protein panel = TGFB1, SPARC, PSA, FOLH1

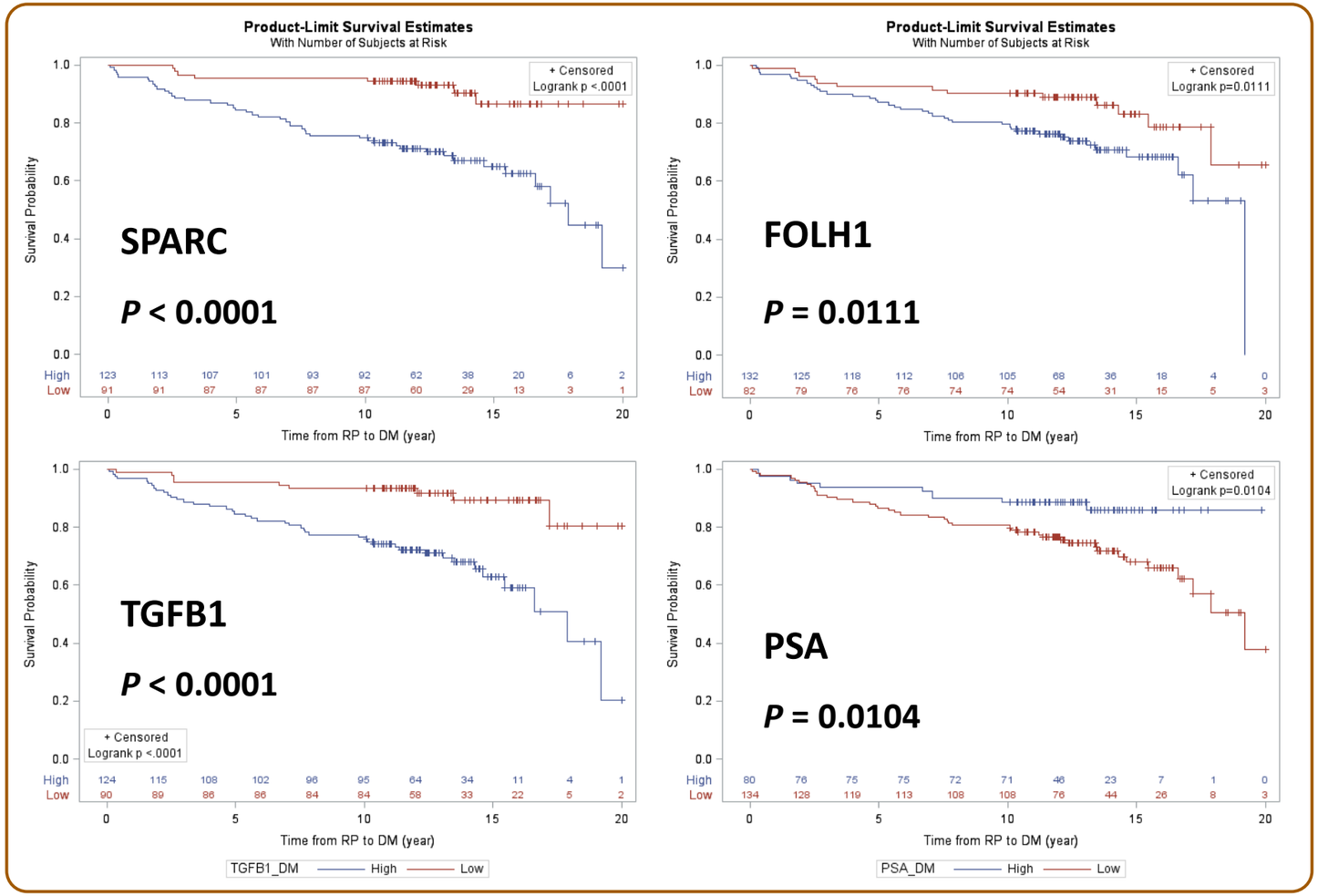
BCR



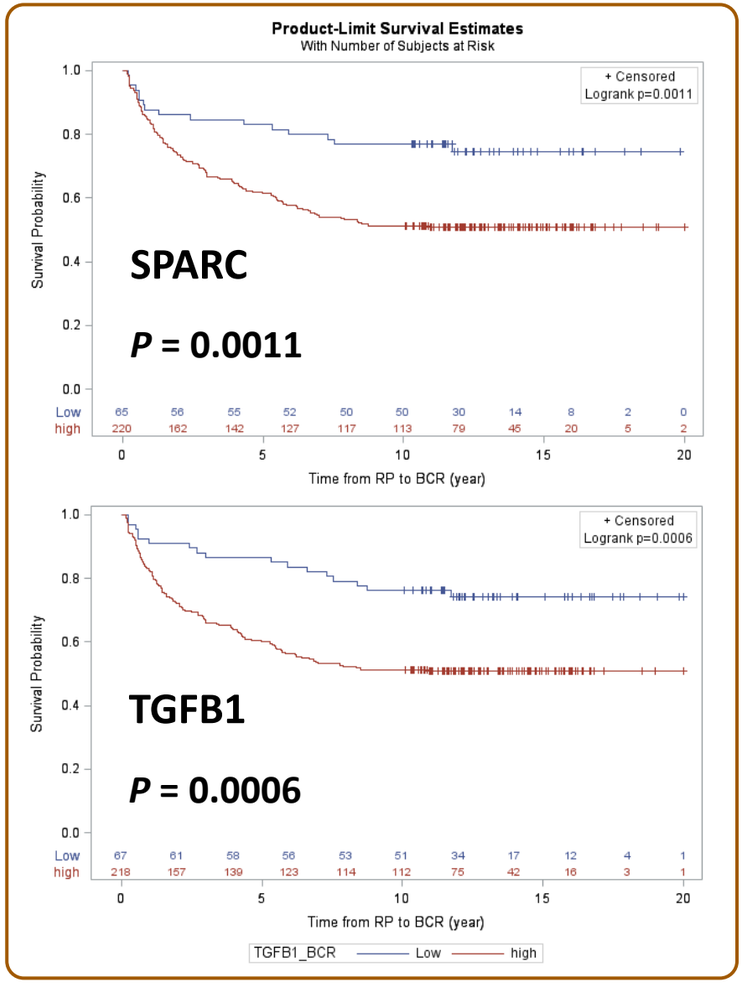
Protein panel = TGFB1, SPARC, FOLH1

Kaplan-Meier DM- and BCR-free survival curves across high versus low groups

DM



BCR



All 4 proteins have >0.40 specificity and >0.80 NPV

2 proteins have >0.30 specificity and >0.70 NPV

Training and testing analysis of proteomic classifier for predicting DM

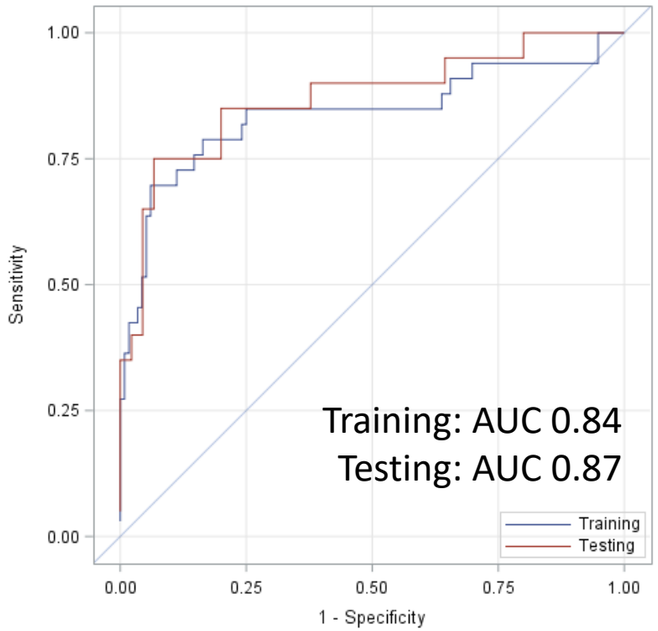
- The training and testing cohorts
  - 214 patients (53 DM, and 161 non-events) were used to develop a biomarker panel classifier to predict DM
  - **70/30** randomly split into training and testing cohorts

- Biomarker selection, classifier construction and optimal cutoff development in training cohort
  - **Biomarker selection**: Univariable logistic regression (*P*<0.05; AUC>0.65): **TGFB1, SPARC, PSA**, **FOLH1**, and **CAMKK2**
  - **Protein classifier construction**: Fitted multivariable logistic regression model to get parameter estimates for 5-protein panel classifier construction in predicting DM
  - **Classifier threshold**: Bootstrapped multivariable logistic regression (with 1000 replicates) to search for optimal cutoff (maximizes sensitivity, with at least **90% NPV** and at least **35% SPC**)

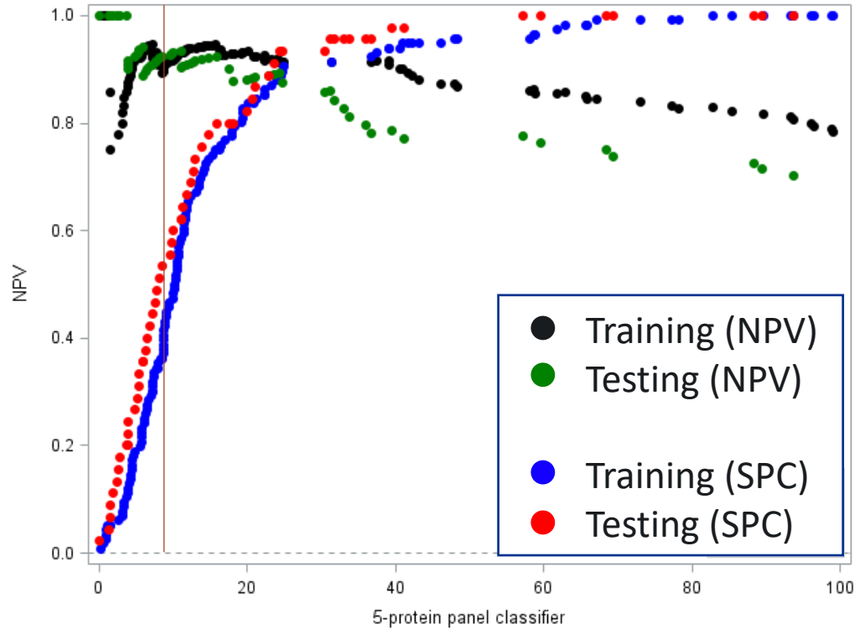
- Testing protein classifier and its threshold in testing cohort
  - **Protein classifier testing**: Applied parameter estimates generated in training cohort to construct protein classifier, with and without adding it to the SOC models
  - **Classifier threshold testing**: Tested threshold in predicting DM-free survival using univariable Kaplan-Meier curve and log-rank test, as well as adding it to the SOC models, using multivariable Cox proportional hazards analysis

Performance of a 5-protein classifier in predicting DM: training and testing analysis

ROC curves



Training (n=149); Testing (n=65)



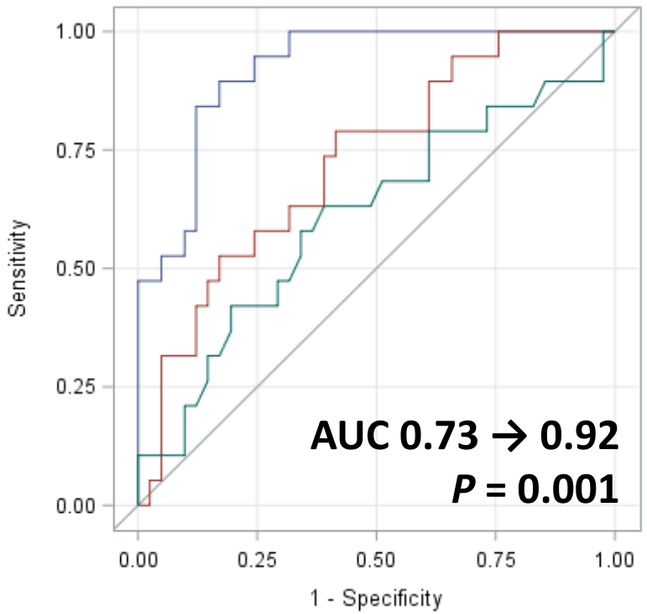
Performance of the optimal cutoff of DM risk score (8.3)

5-protein Classifier	Training				Testing			
Threshold	NPV (95% CI)	Sens (95% CI)	SPC (95% CI)	PPV (95% CI)	NPV	Sens	SPC	PPV
8.3	0.913 (0.911-0.915)	0.879 (0.875-0.883)	0.362 (0.350-0.374)	0.282 (0.274-0.289)	**0.923**	0.900	**0.533**	0.462

Protein classifier = TGFB1, SPARC, PSA, FOLH1, CAMKK2

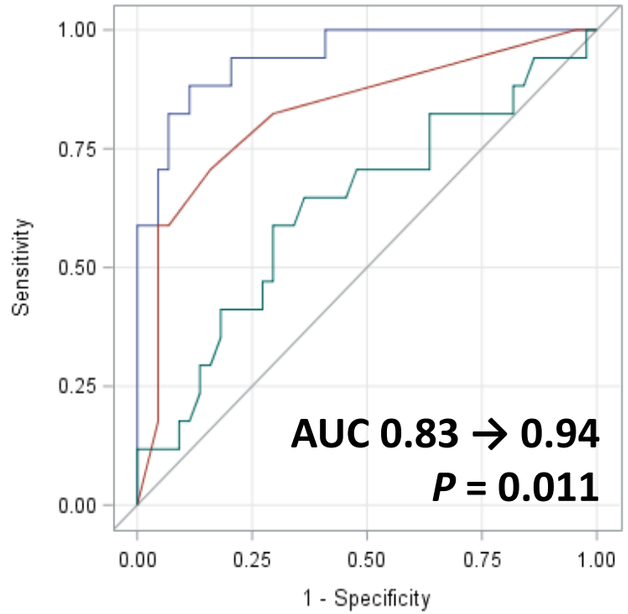
Adding 5-protein classifier to SOC models to predict DM in testing cohort

Biopsy SOC model



	AUC	95% CI		p-value
Biopsy SOC alone	0.7253	0.59	0.86	Ref
Biopsy SOC + protein classifier	0.9217	0.86	0.99	0.001
Diagnosis PSA alone	0.6085	0.44	0.77	0.226

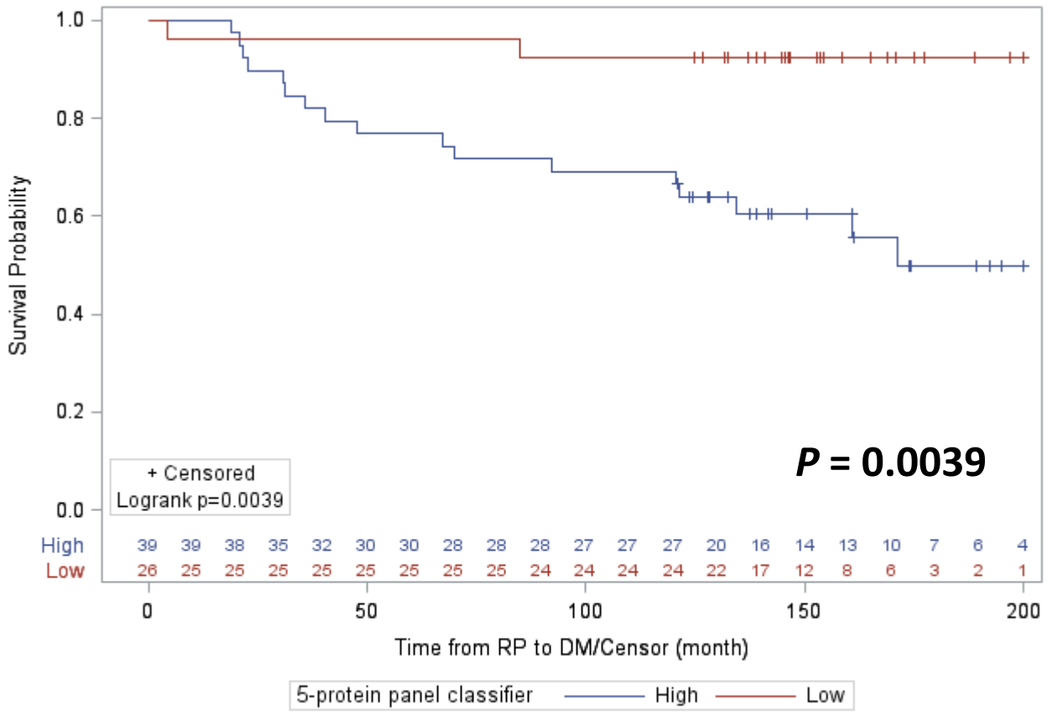
Pathology SOC model



	AUC	95% CI		p-value
Pathology SOC alone	0.8269	0.71	0.95	Ref
Pathology SOC + protein classifier	0.9439	0.89	1	0.011
Diagnosis PSA alone	0.629	0.46	0.8	0.023

Protein classifier = TGFB1, SPARC, PSA, FOLH1, CAMKK2

Kaplan-Meier DM-free survival curves across high versus low groups in testing cohort



High: DM score >8.3; Low: DM score <8.3

N = 214: 53 DM, and 161 non-events

Multivariable Cox proportional hazards model predicting DM by adding 5-protein panel classifier to SOC models in the testing cohort

Biopsy SOC model

Variable	Model 1*			Model 2**		
	HR	95% CI	P	HR	95% CI	P
Age at diagnosis	1.00	0.93-1.07	0.898	1.03	0.96-1.11	0.407
Race (AA vs. CA)	0.94	0.33-2.74	0.916	1.59	0.54-4.64	0.396
Risk (intermediate vs. low)	2.31	0.69-7.76	0.176	1.49	0.41-5.47	0.545
Risk (high vs. low)	4.68	1.14-19.22	**0.032**	2.29	0.52-10.16	0.274
5-protein panel classifier	5.09	1.11-23.38	**0.036**	1.03	1.02-1.05	**<.001**

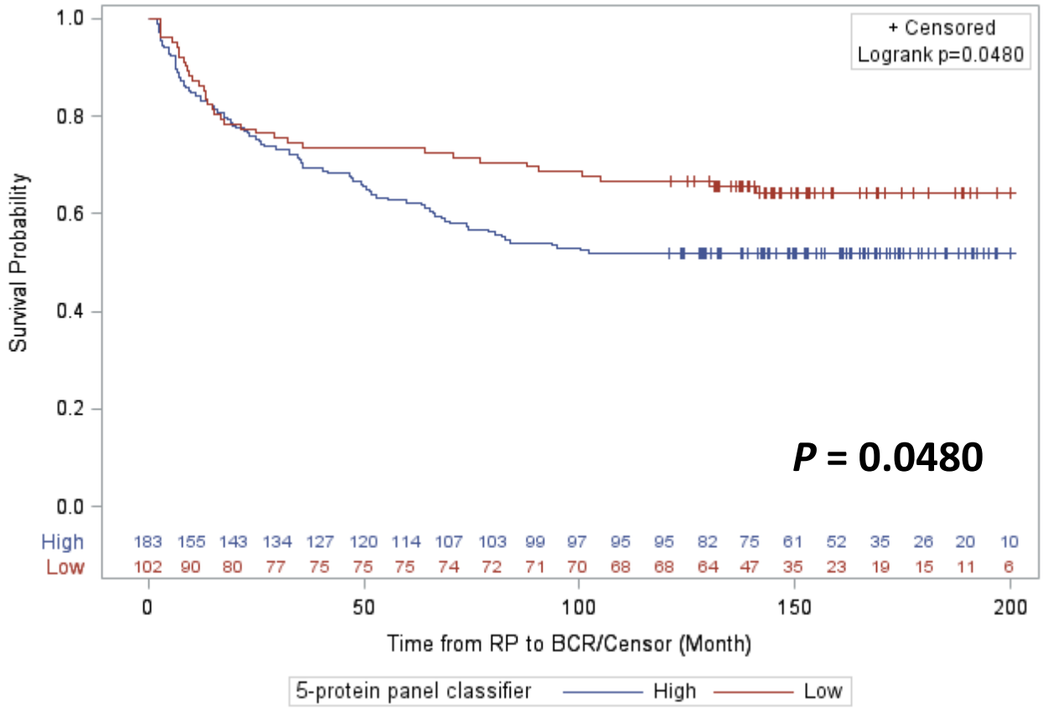
Pathology SOC model

Variable	Model 1*			Model 2**		
	HR	95% CI	P	HR	95% CI	P
Pathology T (pT3 vs. pT2)	2.54	0.78-8.27	0.122	1.94	0.52-7.15	0.321
GG (GG5 vs. GG1-4)	3.42	1.17-10.03	**0.025**	2.04	0.52-8.04	0.309
Surgical margin (pos vs. neg)	1.31	0.47-3.68	0.603	1.23	0.42-3.57	0.705
5-protein panel classifier	3.71	0.82-16.88	**0.089**	1.02	1.01-1.05	**0.018**

*model 1: Classifier was dichotomized at threshold of 8.3
**model 2: Classifier was continuous

Kaplan-Meier BCR-free survival curves across high versus low groups in testing cohort

BCR



High: BCR score >8.3; Low: BCR score <8.3

N = 285: 124 BCR and 161 non-events

Multivariable Cox proportional hazards model predicting BCR by adding 5-protein panel classifier to SOC models in the testing cohort

Biopsy SOC model

Variable	Model 1*			Model 2**		
	HR	95% CI	P	HR	95% CI	P
Age at diagnosis	1.00	0.98-1.02	0.998	1.00	0.97-1.02	0.723
Race (AA vs. CA)	1.18	0.81-1.71	0.399	1.20	0.82-1.74	0.346
Risk (intermediate vs. low)	1.25	0.84-1.86	0.271	1.09	0.73-1.64	0.667
Risk (high vs. low)	2.35	1.36-4.07	**0.002**	1.78	1.01-3.15	**0.045**
5-protein panel classifier	1.25	0.83-1.86	0.284	1.02	1.01-1.03	**<.001**

Pathology SOC model

Variable	Model 1*			Model 2**		
	HR	95% CI	P	HR	95% CI	P
Pathology T (pT3 vs. pT2)	2.48	1.64-3.75	**<.001**	2.25	1.49-3.42	**<.001**
GG (GG5 vs. GG1-4)	2.21	1.43-3.42	**<.001**	1.93	1.23-3.04	**0.004**
Surgical margin (pos vs. neg)	1.77	1.22-2.59	**<.001**	1.67	1.14-2.45	**0.008**
5-protein panel classifier	0.96	0.62-1.46	0.833	1.01	1.00-1.02	0.072

*model 1: Classifier was dichotomized at threshold of 8.3
**model 2: Classifier was continuous

Pacific Northwest NATIONAL LABORATORY

Comparison of performance of commercial prostate cancer tests and the 5-protein classifier

Test	Biomarker Source	Assay Targets	Clinical Endpoint	Results	References
Oncotype DX Genomic Health	Biopsy Very low and low NCCN risk	12+5 RNA targets (four pathways)	Predicts adverse pathology, MET	AUC = 0.73 combined with CAPRA-S or NCCN	Klein et al, Eur Urol, 2014; Cullen et al, Eur Urol, 2015
Prolaris Myriad Genetics	Biopsy Prostatectomy Very low and low NCCN risk	31+15 RNA targets (cell cycle pathway)	Predicts prostate specific death, MET	Biopsy: HR = 1.65 Prostatectomy: HR = 1.77	Cuzick et al, Lancet Oncol, 2011; Cooperberg et al, J Clin Oncol, 2013
Decipher GenomeDX	Prostatectomy	22 RNA targets (multiple pathways)	Predicts adverse pathology, MET	AUC = 0.75-0.79 combined with CAPRA-S or NCCN	Karnes et al, J Urol, 2013 (AUC = 0.79); Spratt et al, J Clin Oncol, 2017 (C-Index = 0.81); Spratt et al, J Clin Onc, 2018 (C-Index = 0.84)
5-protein classifier	Prostatectomy	5 proteins	Predicts adverse pathology, MET	**HR = 5.09**; **AUC = 0.92** combined with age, race and NCCN	Current study

Summary

- PRISM-SRM analysis identified a 5-protein classifier (**TGFB1, SPARC, PSA, FOLH1, CAMKK2**) discriminating between “metastatic progression” and “no progression” tumors, and between “BCR” and “no progression” tumors
- Adding the proteomic classifier to SOC models significantly enhanced the prediction value on PCa progression, especially for the **biopsy SOC model** in predicting **distant metastasis** (AUC increased from 0.73 to **0.92**, *P* = 0.001; **NPV = 0.92**)
- An initial comparison with the other existing commercial tests provides evidence of comparable or superior performance using the 5-protein classifier
- Potential clinical application: early prediction of distant metastasis and BCR using **biopsy tissue** specimens to stratify patients for prostatectomy vs. surveillance

The state of your marker/s. In which phase is your marker? Did you move your marker/s form development to verification, etc.?

- Phase II/III (moving from “assay validation for detecting established disease” to “detecting disease early using retrospective (longitudinal) samples”)
- Needs further validation in independent cohort (and earlier timepoint samples, e.g., biopsy, serum/plasma)

Candidate Biomarker	Discovery			Pre-validation	Validation
	Discovery	Predictive Analysis	Assay Refinement	Blinded Limited Cross-Sectional	Large Cross-Sectional
Panel of proteins (TGFB1, SPARC, FOLH1, KLK3, CAMKK2) (Tissue)					

Other updates and input

- *What was accomplished in this cycle of EDRN?*
  - Developed and initially verified a tissue-based 5-protein classifier for predicting DM and BCR (*Cancers. 2020, 12(5):E1268*)
  - Developed and initially applied nanoscale to single-cell proteomics for studying heterogeneity
- *What is the progress and the current state of your trans-network collaborative projects (collaborative projects funded through set-aside funds)?*
  - Collaborative projects on ovarian cancer (Dr. Skates) and breast cancer (Dr. Paulovich) have been completed
  - Collaborative project on lung cancer (Dr. Massion) is near completion
- *How does the Network structure benefit your biomarkers development and validation study (how is it better than R01 mechanism)?*
  - The highly collaborative nature of the EDRN network provided much needed access to different cohorts and types of samples for biomarker development and further validation
    - The structure of EDRN enabled the collaboration with CPDR and access to a unique cohort of PCa patient tumor samples, which was necessary to discover and characterize the tissue-based 5-protein classifier
    - Access to biostatistical expertise at the DMCC improved the rigor of the statistics used to define the classifier
- *Each investigator should submit 1-2 bullet points as potential areas that should be included in the EDRN GU renewal discussions.*
  - Early detection of clonal malignant lesions within biopsies by nanoproteomics/single-cell analysis
  - Exploration of tumor heterogeneity on prognosis by integration of biomarker data with MRI-guided biopsy and nanoproteomics/single-cell analysis

Acknowledgements

BRRC at PNNL

Yuqian Gao, Yi-Ting Wang, Tujin Shi, Athena Schepmoes, Thomas Fillmore, Richard Smith, Karin Rodland, Tao Liu

CPDR

Jennifer Cullen, Gyorgy Petrovics, Yongmei Chen, Claire Kuo, Denise Young, Yingjie Song, Isabell Sesterhenn, Inger Rosner, Albert Dobi, Shiv Srivastava

EDRN GU Collaborative Group

NCI

Jacob Kagan, Richard Mazurchuk, Sudhir Srivastava