THE UNIVERSITY OF TEXAS
MD Anderson ~~Cancer~~ Center
Making Cancer History®

Predicting Time to Ovarian Carcinoma Recurrence Using Protein Markers

TCGA 2nd Annual Scientific Symposium

Ji-Yeon Yang, jyang7@mdanderson.org
Bioinformatics and Computational Biology
UT MD Anderson Cancer Center, Houston, TX

Mentor: Roel Verhaak

Background

- Standard of care for ovarian cancer: resection, platinum based chemotherapy.

- For about 25% of patients, the tumor does not respond to platinum therapy.

TCGA	RESISTANT	SENSITIVE	TOTAL
	90 (31.1%)	199 (68.9%)	289

- Identification of platinum resistant patients at time of diagnosis may result in a change in therapeutic approach.

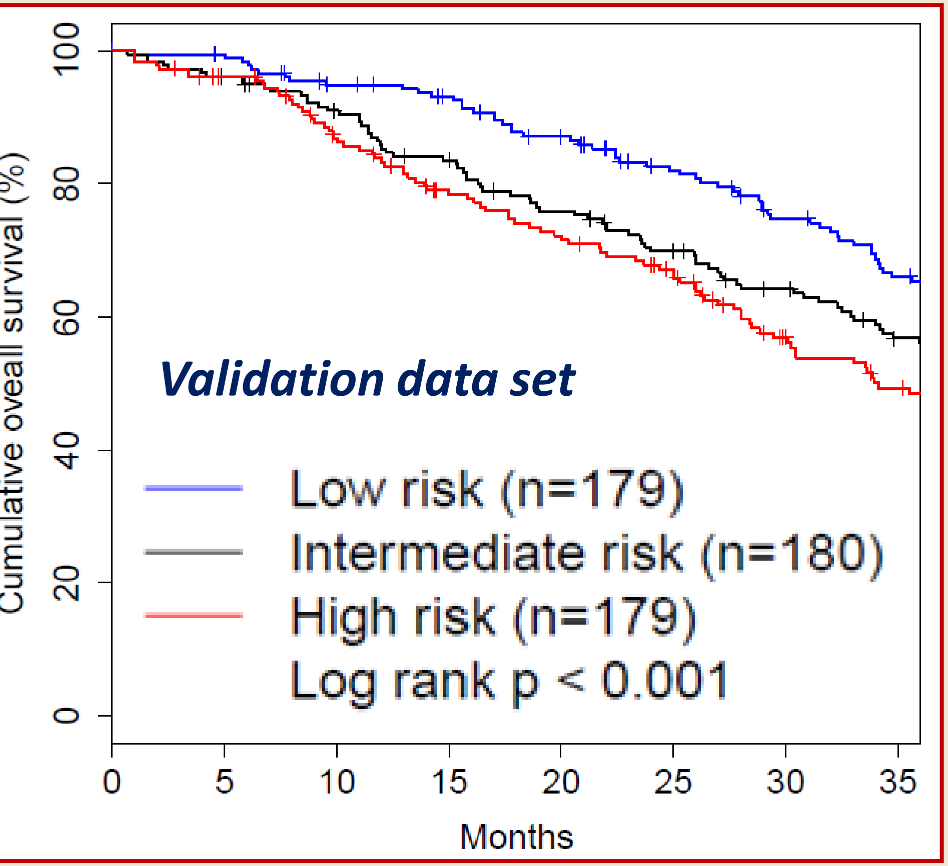
CLOVAR model predicts outcome in ovarian cancer

- ‘CLassification of OVARian cancer’ (CLOVAR) predicts survival using gene expression signatures.
- Using TCGA, subtypes and survival gene expression signatures were identified, which would provide a prognostic model.

Verhaak RG et al., *JCI 2012 (in press).*

CLOVAR model predicts outcome in ovarian cancer

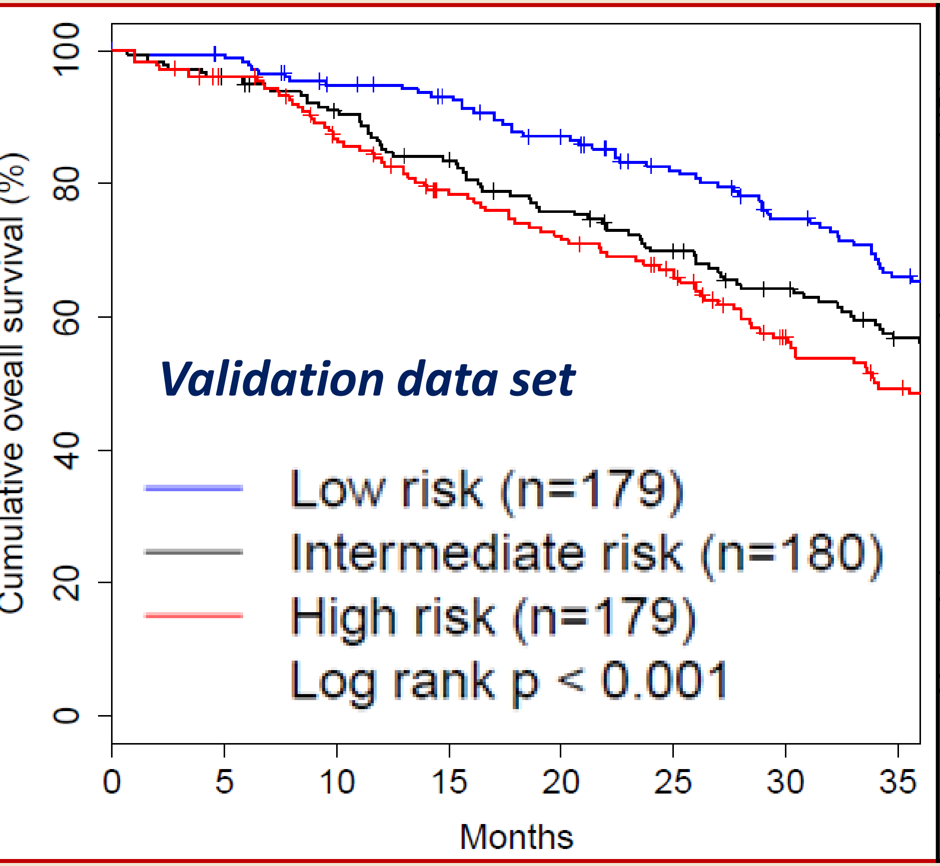
- ‘CLassification of OVARian cancer’ (CLOVAR) predicts survival using gene expression signatures.

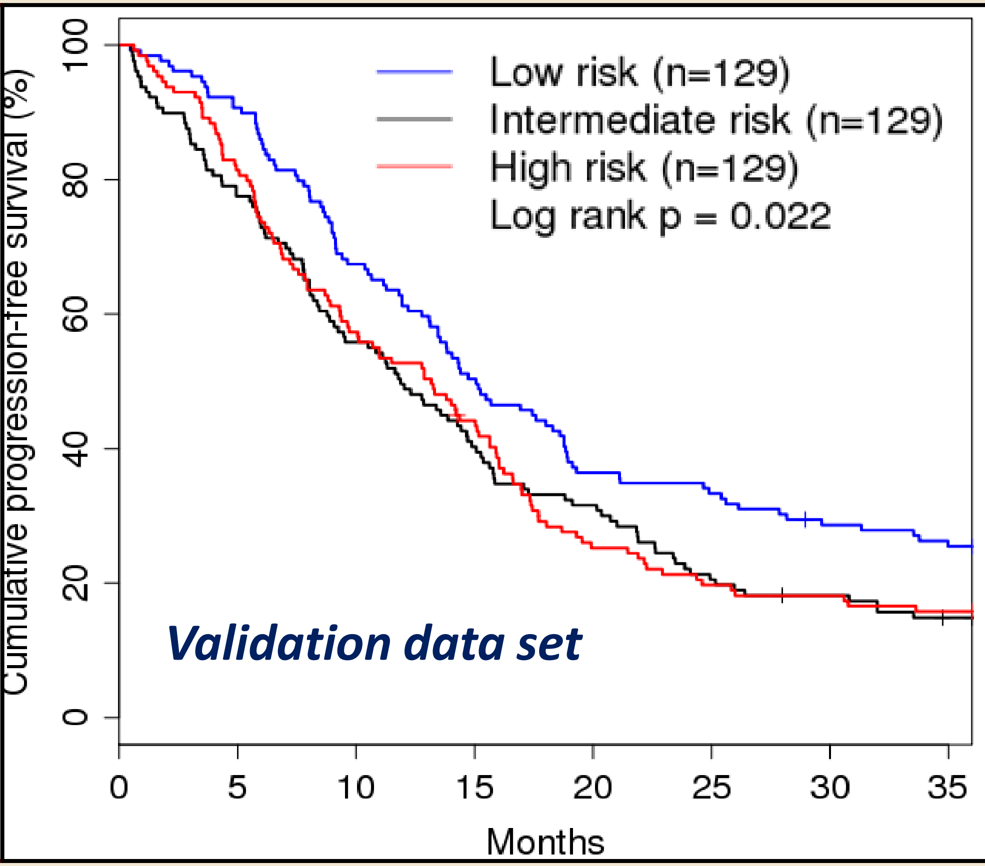


Verhaak RG et al., *JCI 2012 (in press).*

CLOVAR is less effective to predict progression free survival

- ‘CLassification of OVARian cancer’ (CLOVAR) predicts survival using gene expression signatures.





Verhaak RG et al., *JCI 2012 (in press).*

Can we use protein markers to predict PFS?

- Aim to develop a predictor of platinum resistance that is based on protein markers.

- RPPA (reverse phase protein arrays):

  172 proteins and phosphoproteins in 412 TCGA samples with serous ovarian cancer.

  222 cases were included in the model construction (non-missing values for PFS, advanced stage)

PROVAR

PRotein-driven index of OVARian cancer

TCGA

Training set (n=222 samples)

LASSO

Selected 9 proteins most associated with PFS & estimated β

PROVAR =

$$\hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \ldots + \hat{\beta}_9 X_9$$

Feature Selection & Estimation

LASSO (Least Absolute Shrinkage and Selection Operator)

L1-constrained (lasso) Cox regression (Tibshirani,1997), sparse interpretable models by shrinking some variables to exactly zero.

Feature Selection & Estimation

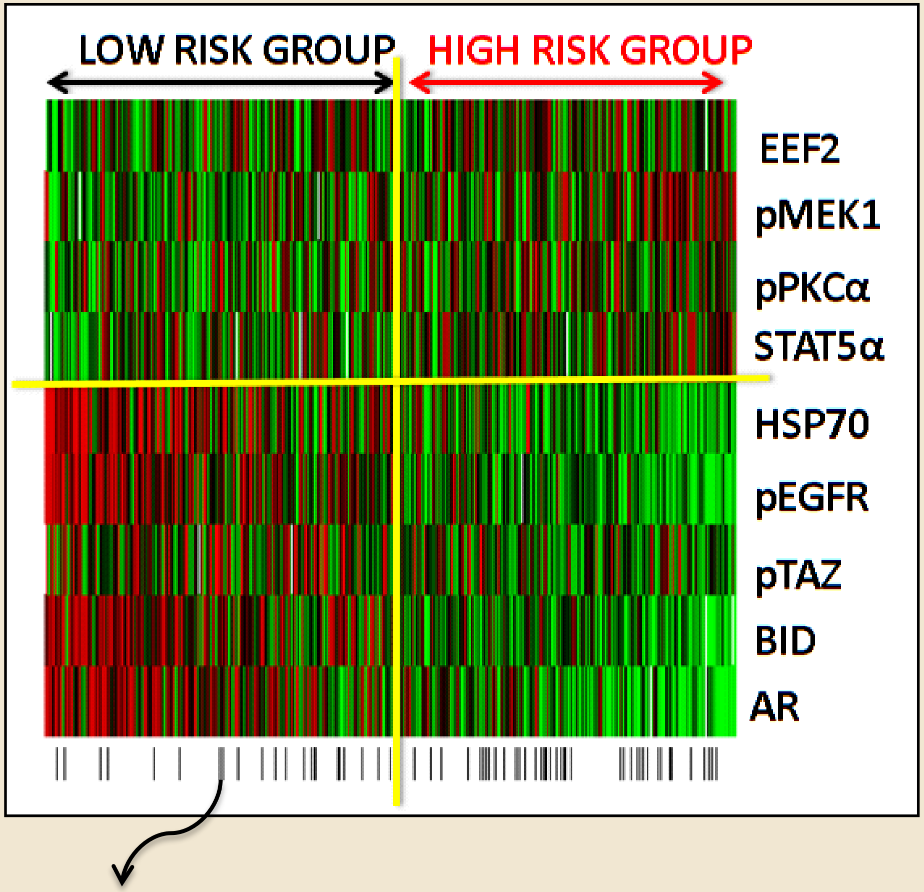
LASSO (Least Absolute Shrinkage and Selection Operator)

L1-constrained (lasso) Cox regression (Tibshirani,1997), sparse interpretable models by shrinking some variables to exactly zero.

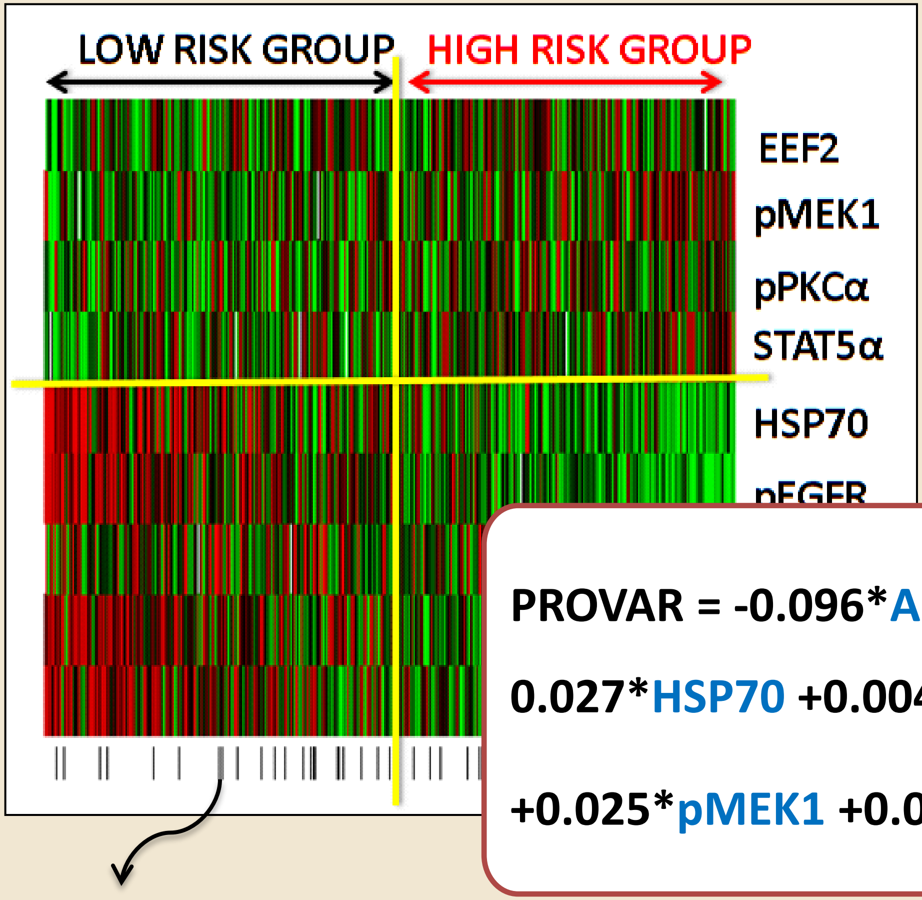
AR	BID	pTAZ	pEGFR	HSP70	STAT5α	pPKCα	pMEK1	EEF2
-0.096	-0.075	-0.058	-0.052	-0.027	0.004	0.022	0.025	0.029

9 proteins most associated with progression-free survival.

Numbers, lasso coefficients.



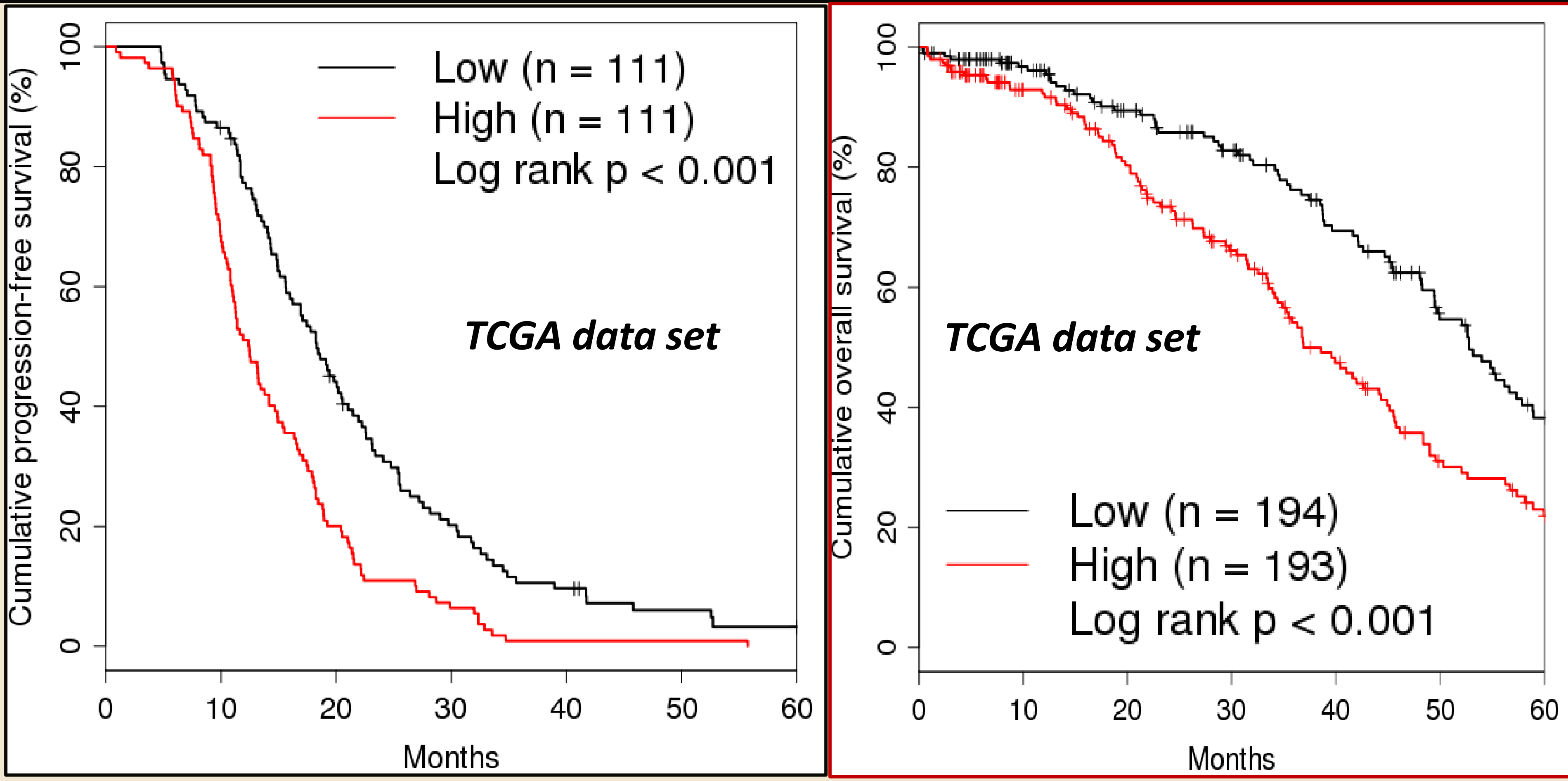
platinum resistant



platinum resistant

PROVAR = -0.096*AR -0.075*BID -0.058*pTAZ -0.052*pEGFR -0.027*HSP70 +0.004*STAT5.ALPHA +0.022*pPKC.ALPHA +0.025*pMEK1 +0.029*EEF2

PROVAR is predictive of both OS & PFS in the TCGA data set



Progression-free survival

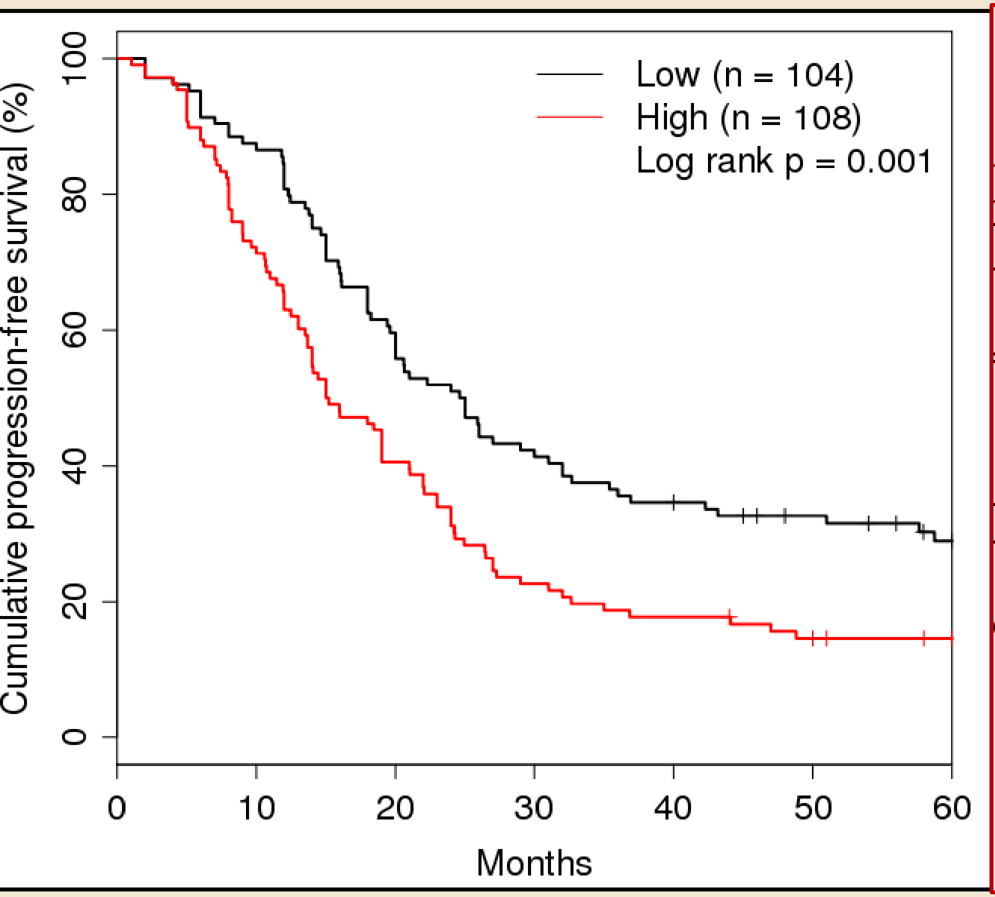
Overall survival

Validation of PROVAR in an independent data set

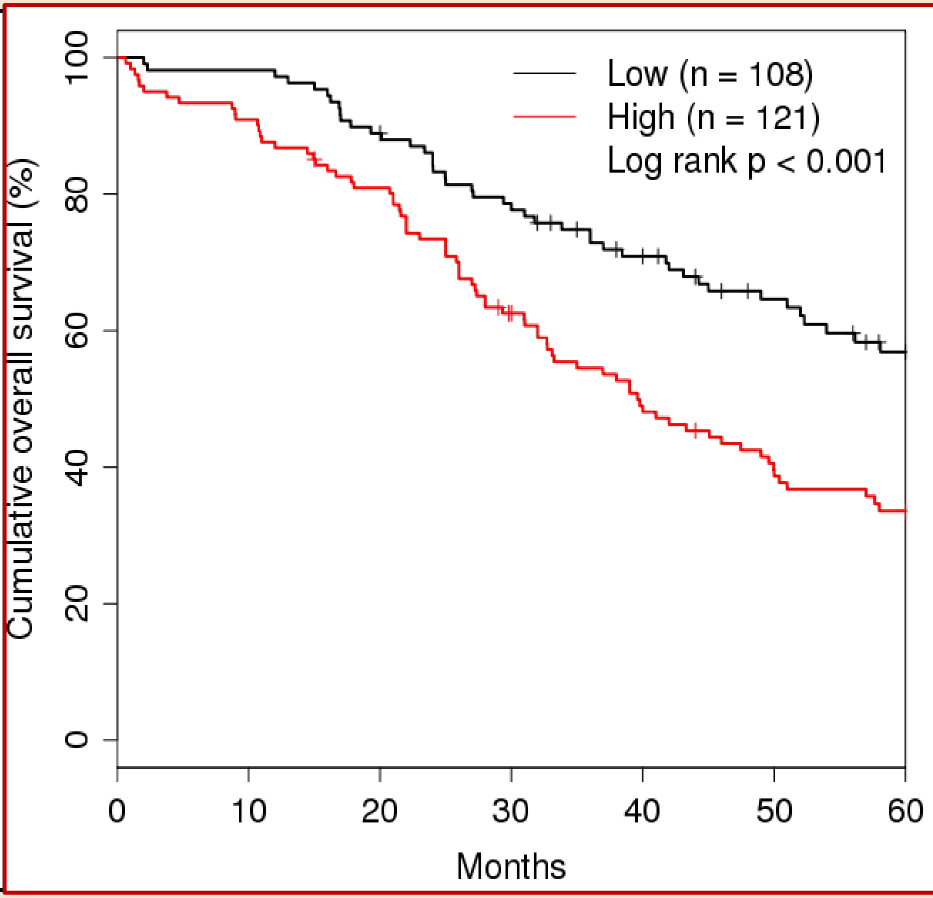
- 229 high-grade serous samples from Japan and Philadelphia.

- Expression levels of 144 proteins and phosphoproteins were measured by RPPA.

PROVAR is predictive of time to tumor recurrence in an independent dataset.



Progression-free survival



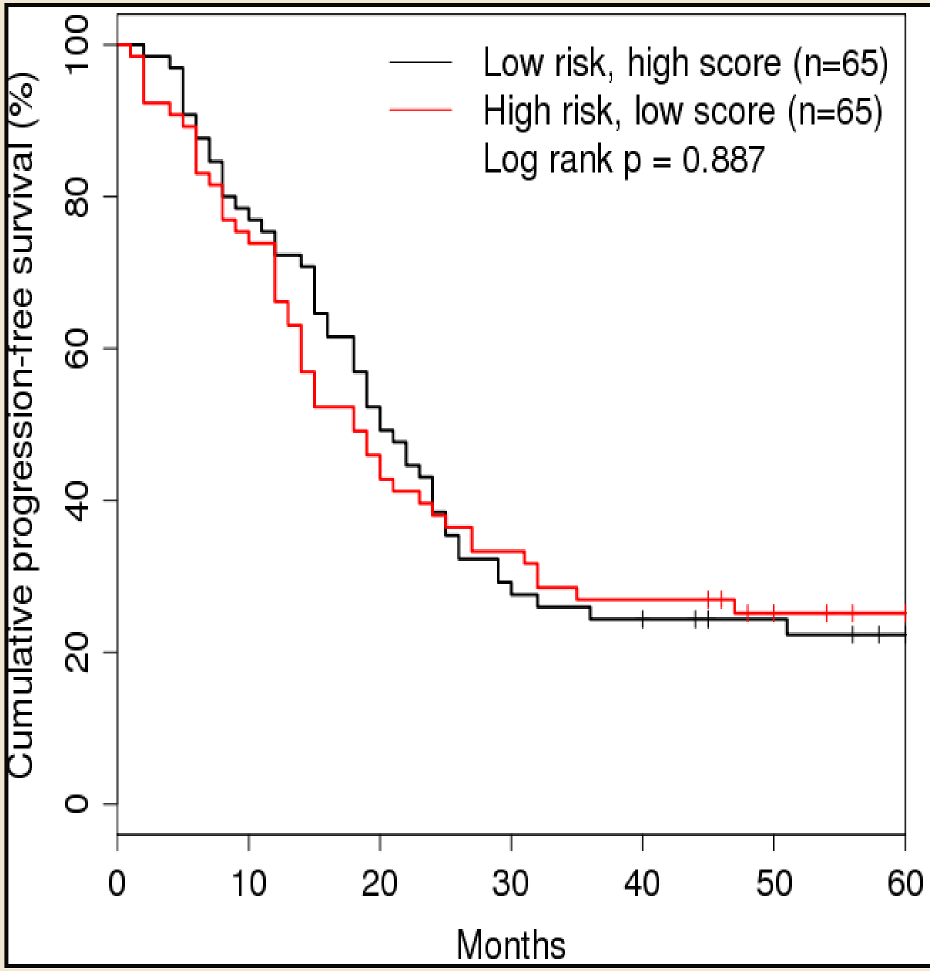
Overall survival

PROVAR (protein-driven) vs. CLOVAR (gene-driven)

n=130 samples with gene expression data available

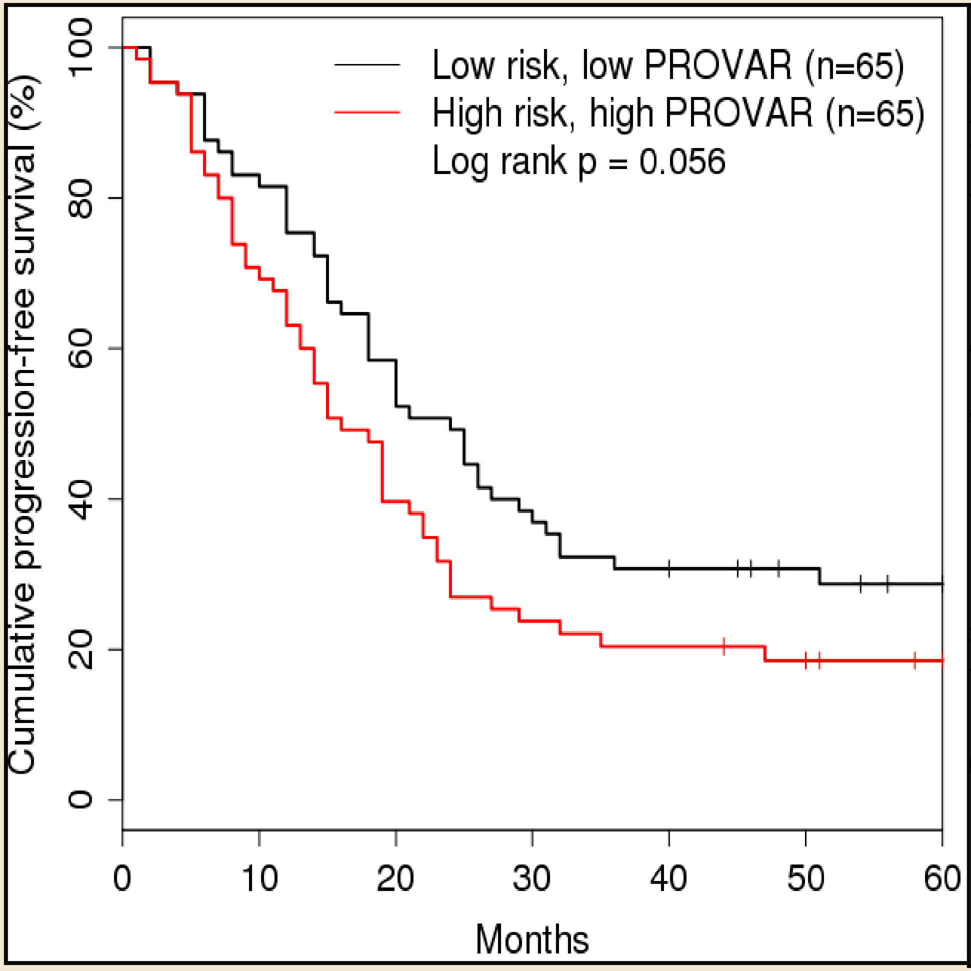
PROVAR (protein-driven) vs. CLOVAR (gene-driven): PFS, n=130 samples with gene expression data available

CLOVAR

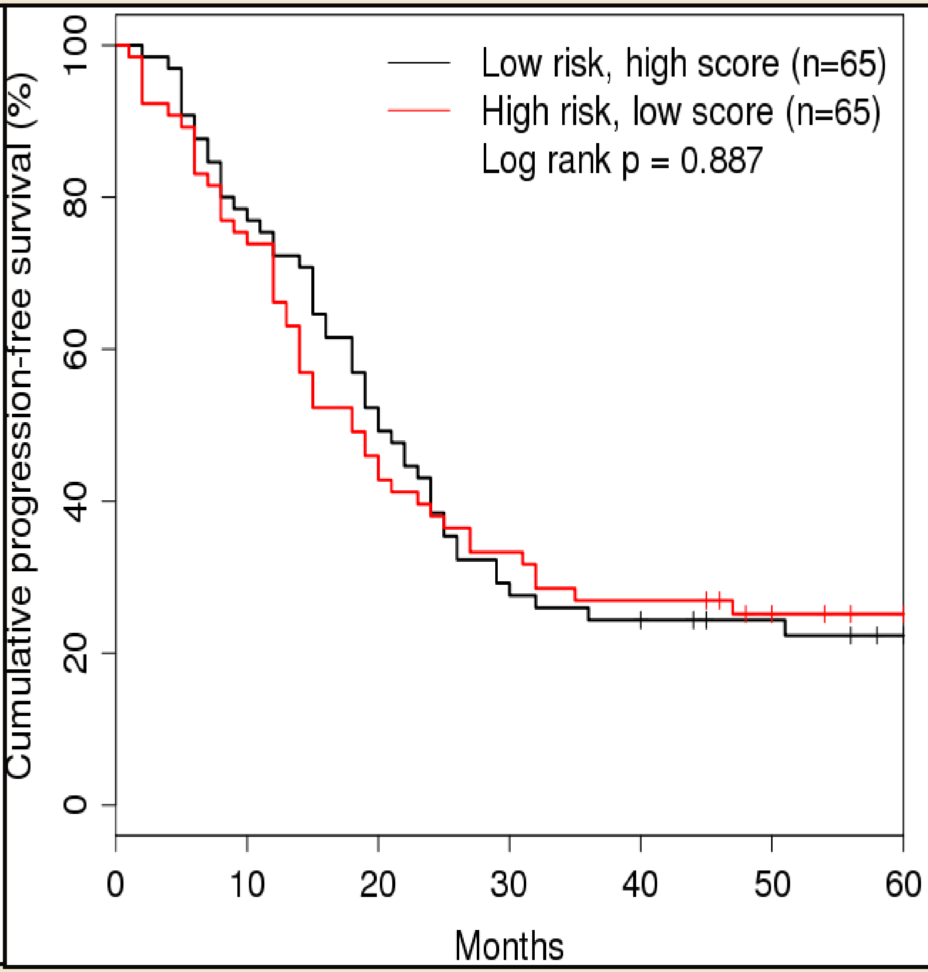


PROVAR (protein-driven) vs. CLOVAR (gene-driven): PFS, n=130 samples with gene expression data available

PROVAR



CLOVAR



Robustness of the nine proteins markers

Robustness of proteins markers

Using the validation samples,

COX2	VASP	CYCLINB1	pNFKB	AR
-0.006	0.057	0.050	-0.028	-0.014

Numbers are lasso coefficients.

Robustness of proteins markers

Using the validation samples,

COX2	VASP	CYCLINB1	pNFKB	AR
-0.006	0.057	0.050	-0.028	-0.014

Using the TCGA samples,

BID	pEGFR	HSP70	EEF2	pPKCα	STAT5α	pTAZ	pMEK1	AR
-0.075	-0.052	-0.027	0.029	0.022	0.004	-0.058	0.025	-0.096

Numbers are lasso coefficients.

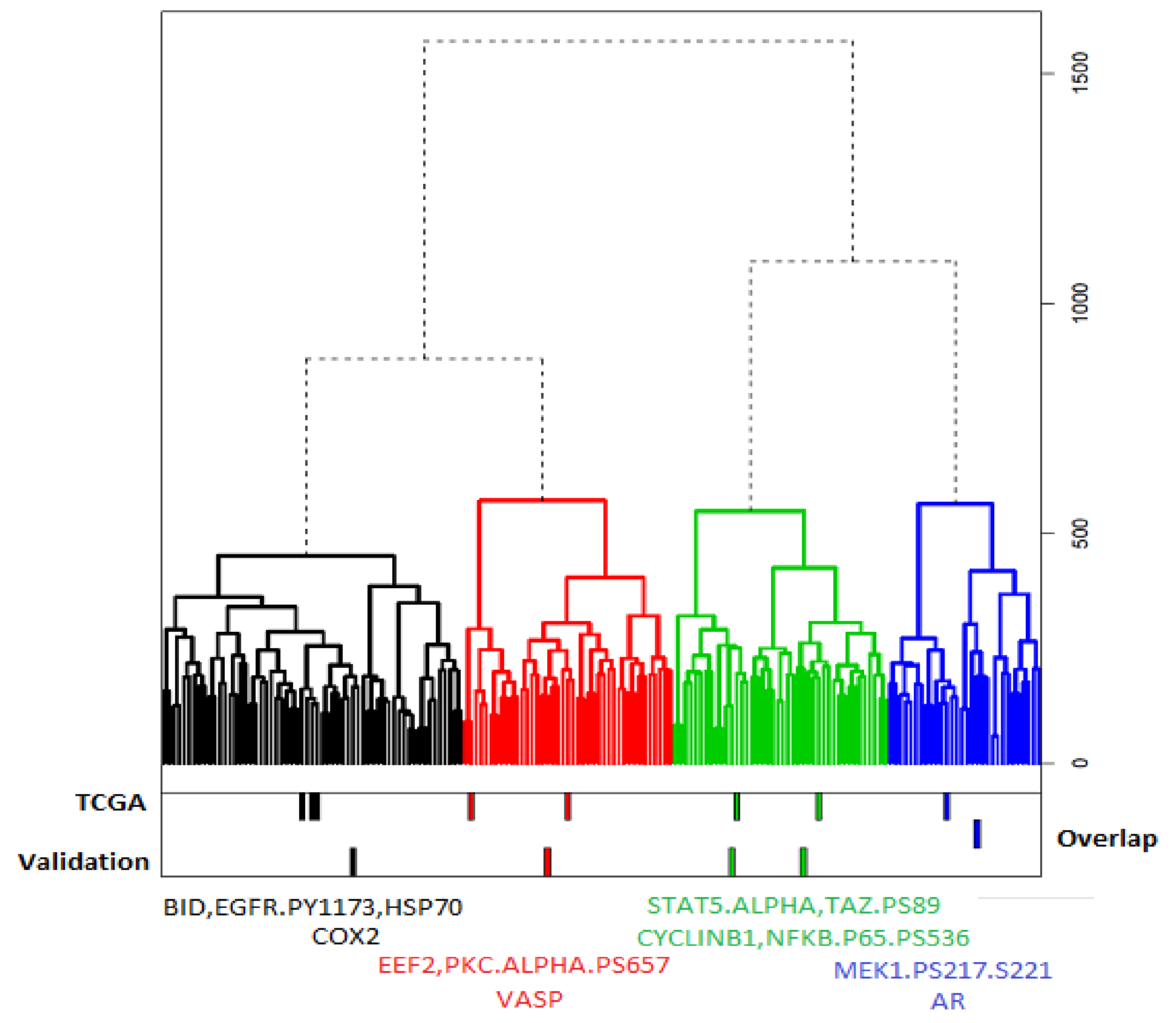
Hierarchical clustering of 172 proteins from TCGA

[Cluster1]
BID, pEGFR, HSP70
COX2

[Cluster2]
EEF2, pPKCα
VASP

[Cluster3]
STAT5α, pTAZ
CYCLINB1, pNFKB

[Cluster4]
pMEK1, AR
(AR: overlap)



Conclusions

- We developed a 'PRotein-driven index of OVARian cancer', PROVAR using progression-free survival,
- and successfully validated its discriminative ability to predict both progression-free and overall survival in high-grade serous ovarian cancers
- Unlike genetic signatures in previous studies that often contained a large number of genes, PROVAR is simple but still predictive of progression and survival, making it useful in clinical practice.

Acknowledgements

- Verhaak Lab (virtuon.nl)

Kosuke Yoshihara, Siyuan Zheng, Hoon Kim, Wandaliz Torres-Garcia, Rahul Vegesna, Emmanuel Martinez, **Roel Verhaak.**

- Collaborators

Gordon Mills (MDACC), Yiling Lu (MDACC), Kenichi Tanaka (Niigata Medical Center Hospital), Masayuki Hatae (Kagoshima City Hospital), Hideaki Masuzaki (Nagasaki Univ), Hiroaki Itamochi (Tottori Univ), Masashi Takano (National Defense Medical College), Kimio Ushijima (Kurume University School of Medicine), Roshan Agarwal (Imperial College London), Janos Tanyi (Hospital of the University of Pennsylvania).

- The Cancer Genome Atlas Research Network
- Genome Data Analysis Center (GDAC) at MD Anderson Cancer Center