

SURVIVAL ANALYSIS

Poster: **Integration Strategies for Multi-Omics Survival Analysis**

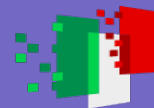
Author: **Francesca Calanca**



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



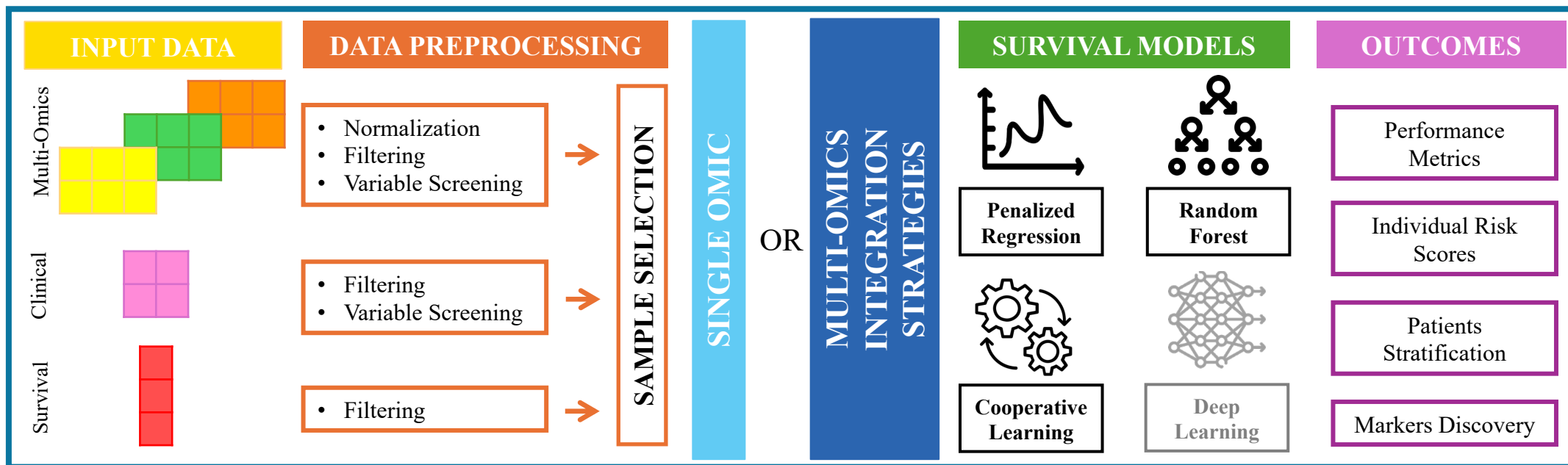
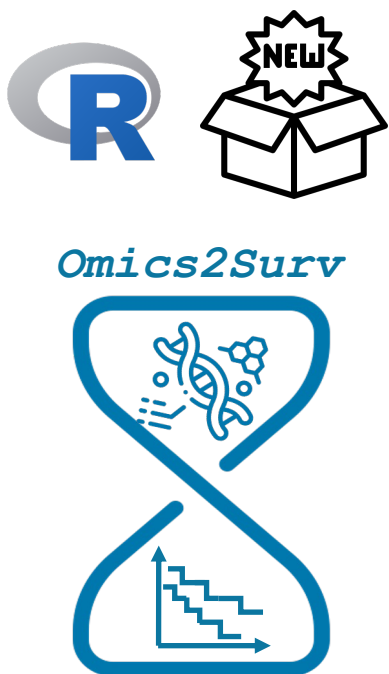
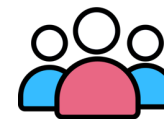
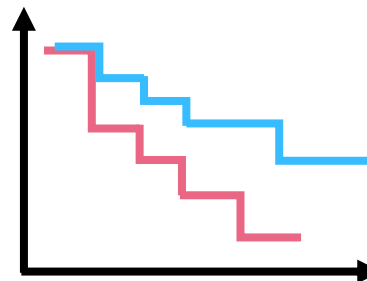
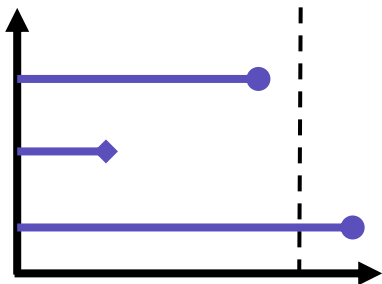
Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



CNR-IAC
Istituto
per le Applicazioni
del Calcolo

This work was supported by PRIN 2022 PNRR P2022BLN38 project, "Computational approaches for the integration of multi-omics data" funded by European Union - Next Generation EU, CUP B53D23027810001.

with Multi-Omics Data



**PENALIZED
REGRESSION**

**COOPERATIVE
LEARNING**

**RANDOM
FOREST**

PENALIZED REGRESSION

Cox Proportional Hazard Model

Hazard function

$$h(t \mid X_i) = h_0(t) \exp(X_i \beta)$$

Partial log-likelihood

$$\ell(\beta) = \sum_{i: \delta_i=1} \left(X_i^\top \beta - \log \sum_{j \in R_i} \exp(X_j^\top \beta) \right)$$

Penalized objective

$$\hat{\beta} = \arg \max_{\beta} \{ \ell(\beta) - \lambda P(\beta) \}$$

Penalizations

$$\left\{ \begin{array}{l} P(\beta) = \|\beta\|_1 \quad (\text{LASSO}) \\ P(\beta) = \alpha \|\beta\|_1 + (1 - \alpha) \|\beta\|_2^2 \quad (\text{Elastic Net}) \\ P(\beta) = \sum_j w_j |\beta_j| \quad (\text{Adaptive LASSO}) \\ P(\beta) = \alpha \|\beta\|_1 + (1 - \alpha) \beta^\top L \beta \quad (\text{Network-based}) \end{array} \right.$$

glmnet¹

NETSURVPROX



- Handles high-dimensional omics data ($p \gg n$)
- Produces interpretable linear risk scores
- Automatic variable selection

+

Note | This is also applicable to AFT models

COOPERATIVE LEARNING

RANDOM FOREST

COOPERATIVE LEARNING

Cooperative Cox Proportional Hazard Model¹

Hazard function

$$h(t \mid U_i, Z_i) = h_0(t) \exp(U_i \beta_U + Z_i \beta_Z)$$

Penalized
objective

$$\max_{\beta_U, \beta_Z} \ell_{U,Z}(\beta_U, \beta_Z) - \rho \|U \beta_U - Z \beta_Z\|_2^2 - \lambda \|\beta_U\|_1 / \hat{\beta}_U + \|\beta_Z\|_1 / \hat{\beta}_Z$$



Encourages different omics to “agree” on predictive risk scores

Agreement parameter



Note | Support only 2 omics at the moment



Note | This is also applicable to AFT models²

¹ Hahn, G., Prokopenko, D., Hecker, J., Lutz, S.M. et al. Prediction of disease-free survival for precision medicine using cooperative learning on multi-omic data. Briefings in Bioinformatics 25(4). (2024)

² Angelini C., De Canditiis D., De Feis L., Iuliano A. Cooperative AFT models for multi-omics data integration

RANDOM FOREST



The group of covariates corresponding to one specific data type is denoted as a 'block'

Variables selected



$$mtry = (\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_M})$$

Splitting rule

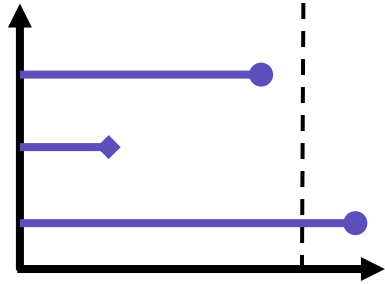
$$\text{Log-rank}(s) = \frac{(O_L - E_L)^2}{\text{Var}(O_L)}$$

- Nonlinear effects, interactions captured
- Robust to high-dimensional and noisy omics data
- No proportional hazards assumption

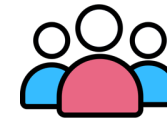
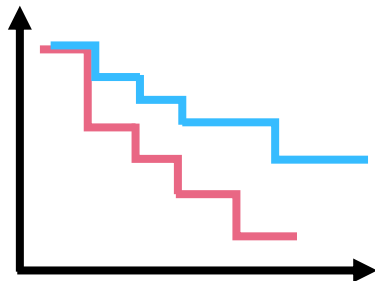
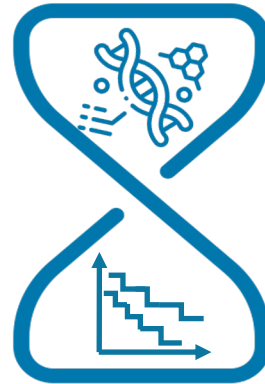
THE PACKAGE

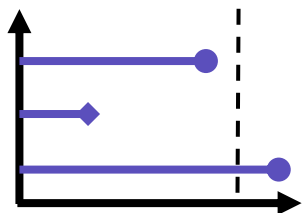
		Survival Model	R Package	Integration Strategy	Omics supported
 Omics2Surv	COX based	Penalized Regression	glmnet	Early + Late	2+
		Random Forest	blockForest	Joint	2+
		Cooperative Learning	coop	Joint	2
 WORK IN PROGRESS	COX based	Penalized Regression	NETSURVPROX	Early + Late	2+
		Deep Learning	AUTOSURV	Joint	
	AFT based	Penalized Regression	NETSURVPROX	Early + Late	2
		Cooperative Learning	AFTCoop	Joint	

COMPARATIVE ANALYSIS



Omics2Surv





DATASETS



LinkedOmics

<https://www.linkedomics.org>



Expression (Gene level)



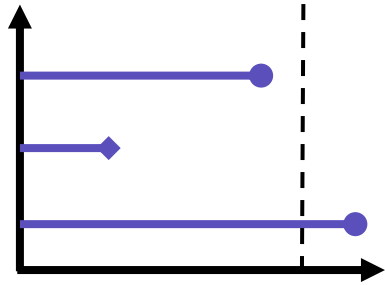
Copy number change (Gene level)

Dataset	Number of patients (n)	RNA features Raw(p) Screened(m)		CNV features (p) (m)		Early Integration features
Lower Grade Glioma (LGG)	478	20086	2562	24776	3716	6278
Pan-Kidney Cohort (KIPAN)	851	20190	2499	24776	3716	6215
Head and Neck Squamous Cell Carcinoma(HNSC)	505	20164	2513	24776	3716	6229

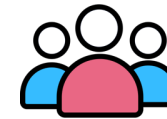
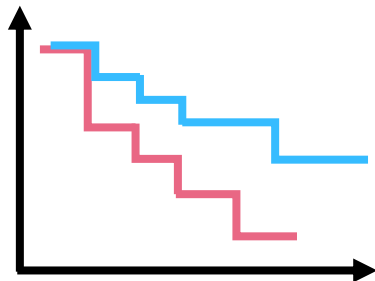
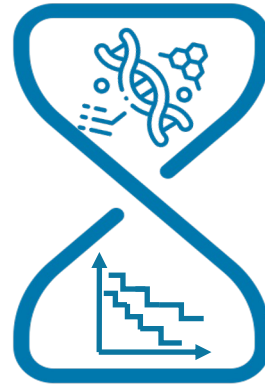
SUMMARY

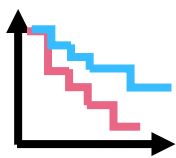
Analysis Modality				Early Integration					Late Integration	
Dataset	Train Test		Method	Alpha	Lambda min	(d)	C-index Train	C-index Test	C-index Train	C-index Test
LGG	382	96	LASSO	1	0,0716	20	0,8859	0,8530	0,5600	0,5954
			EN	0.2	0,4022	52	0,8768	0,8661	0,5450	0,5862
			ADAPTIVE	1	0,2290	50	0,9221	0,8705	0,5289	0,5744
KIPAN	681	170	LASSO	1	0,0403	53	0,8015	0,7886	0,5164	0,3634
			EN	0.8	0,1215	13	0,7589	0,7645	0,5201	0,4001
			ADAPTIVE	1	0,2545	224	0,9580	0,8197	0,5225	0,4006
HNSC	404	101	LASSO	1	0,0616	54	0,8048	0,6174	0,4902	0,4811
			EN	0,9	0,0610	70	0,8205	0,6315	0,4893	0,4670
			ADAPTIVE	1	0,1316	222	0,9889	0,6444	0,4951	0,4623
				Joint Integration			Joint Integration			
			Method	C-index Train	C-index Test	Method	C-index Train	C-index Test		
			Random Forest	0,8172	0,8206	Cooperative	0,4055	0,4928		
			Random Forest	0,7262	0,7677	Cooperative	0,4742	0,5		
			Random Forest	0,5663	0,5901	Cooperative	0,4978	0,5		

RESULTS



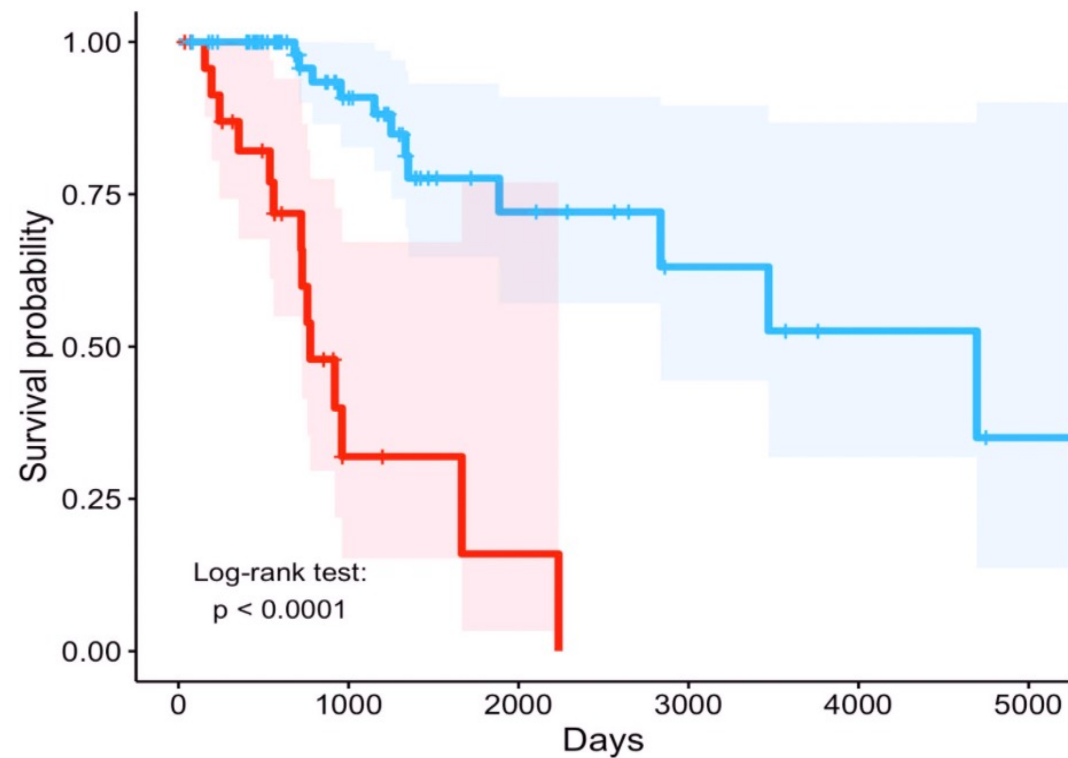
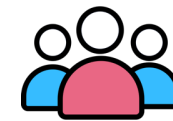
Omics2Surv





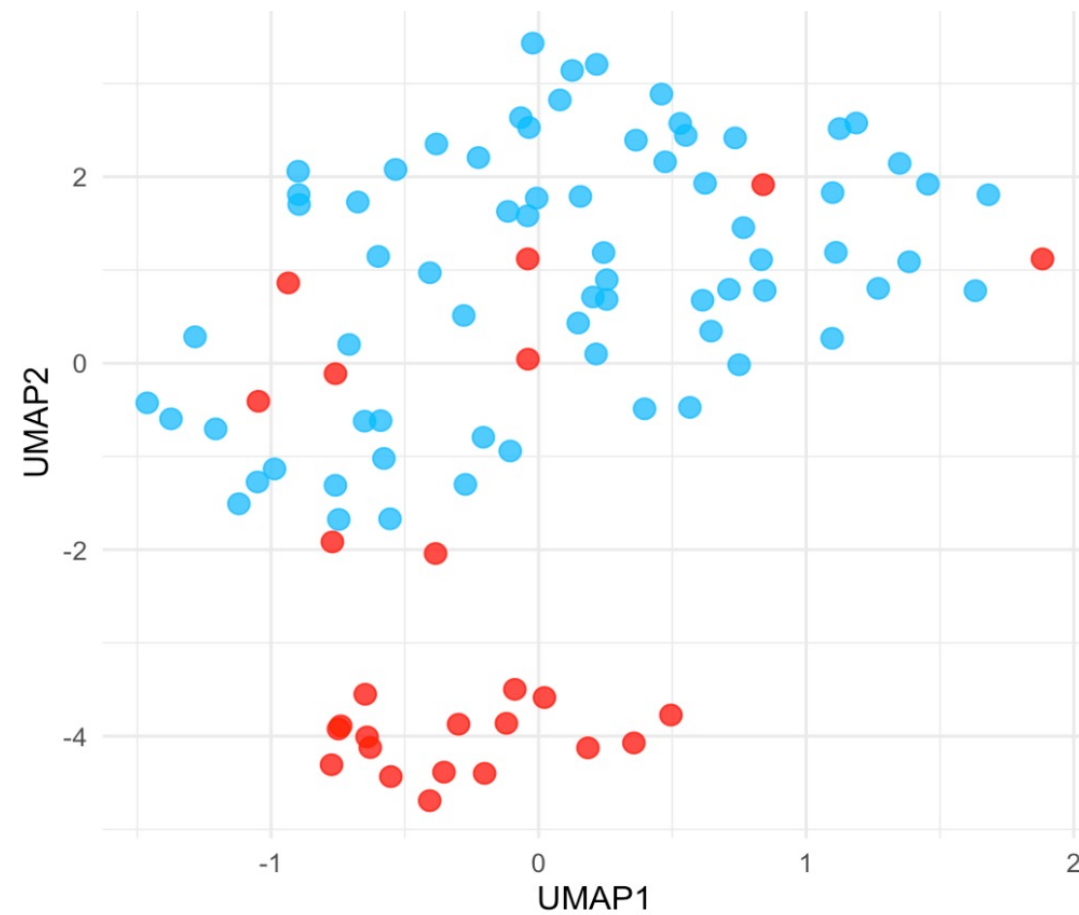
LGG

Patient stratification based on early integration with adaptive lasso-cox results



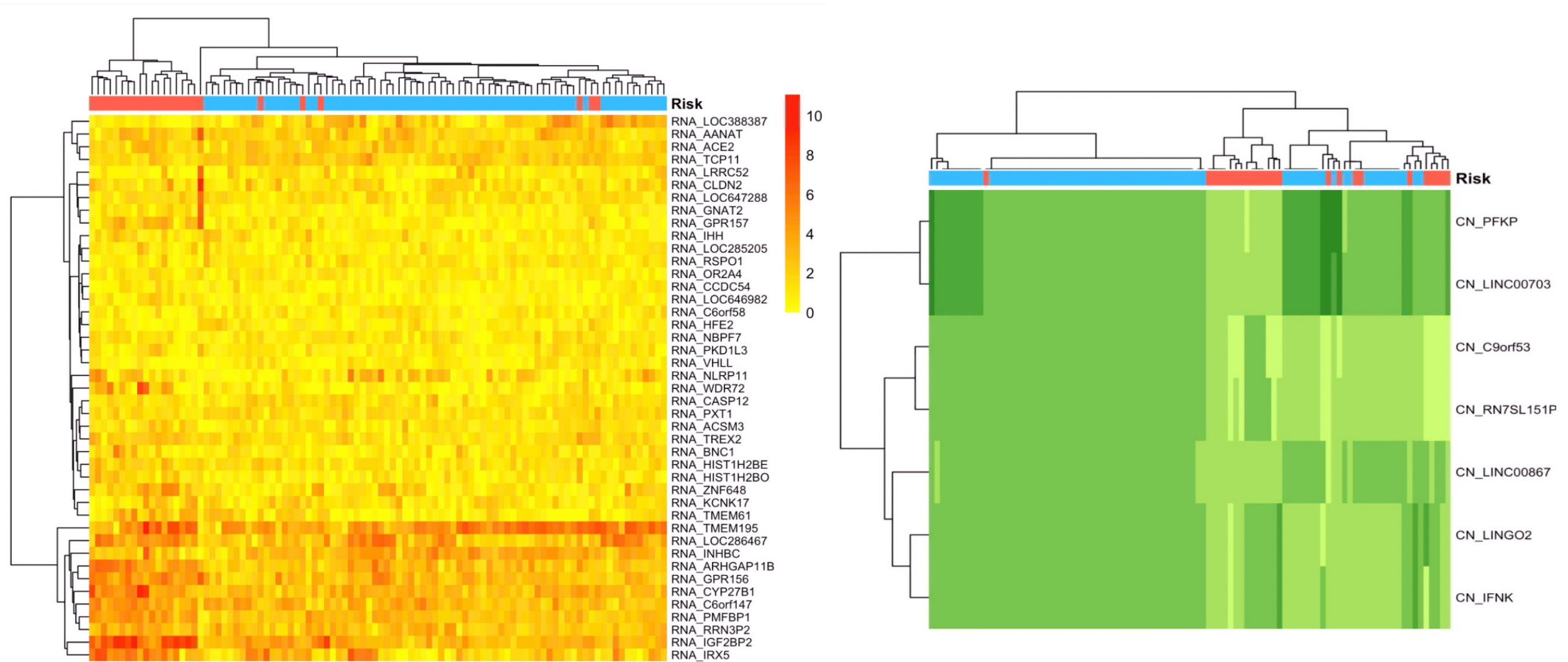
Number at risk		Days					
Risk groups	High risk	25	3	1	0	0	0
	Low risk	71	35	13	6	3	1
		0	1000	2000	3000	4000	5000

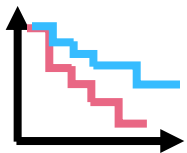
Risk groups
— High risk
— Low risk



LGG

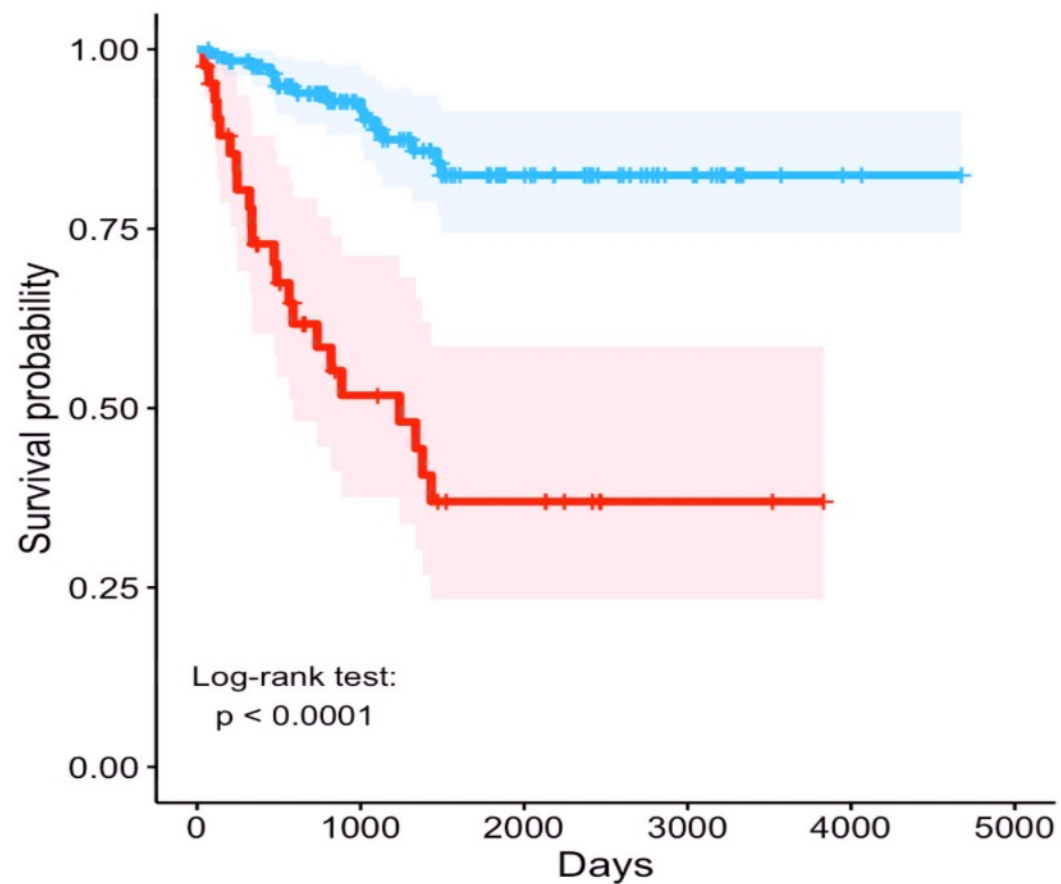
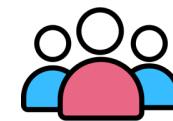
Heatmaps based on model's selected features





KIPAN

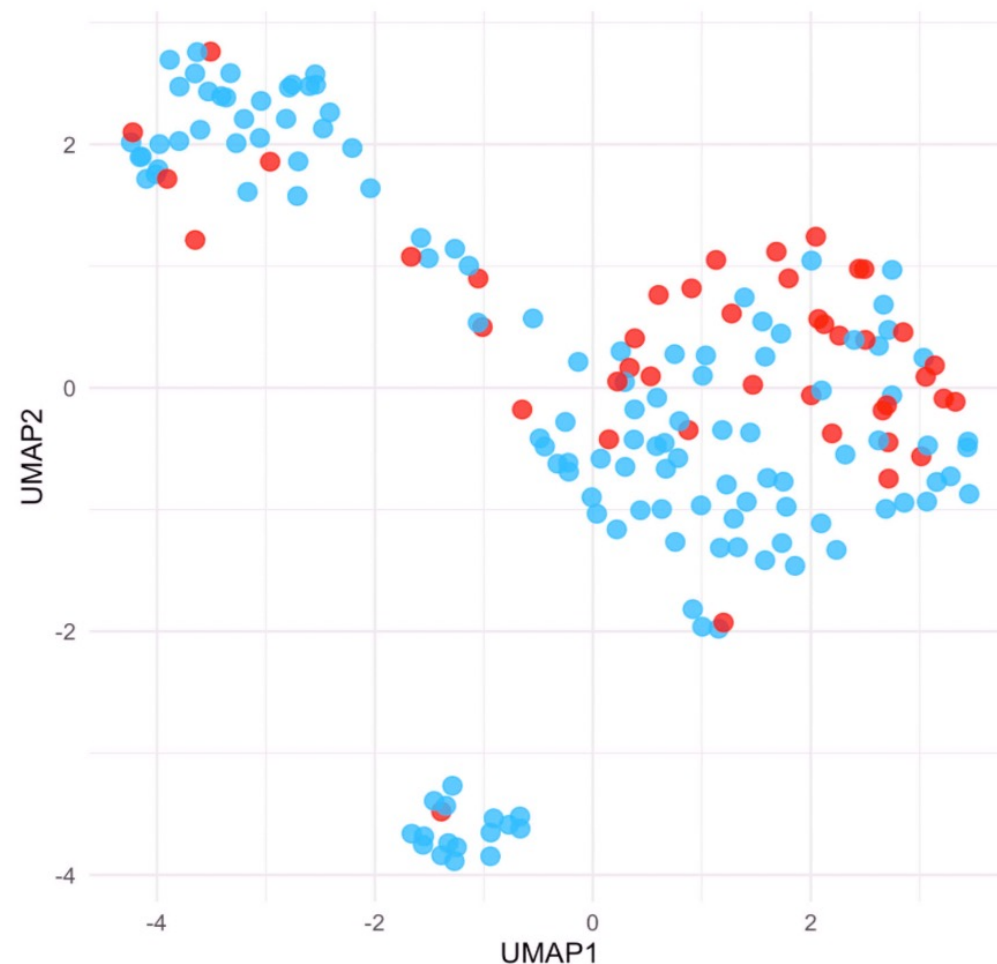
Patient stratification based on early integration with adaptive lasso-cox results



Risk groups

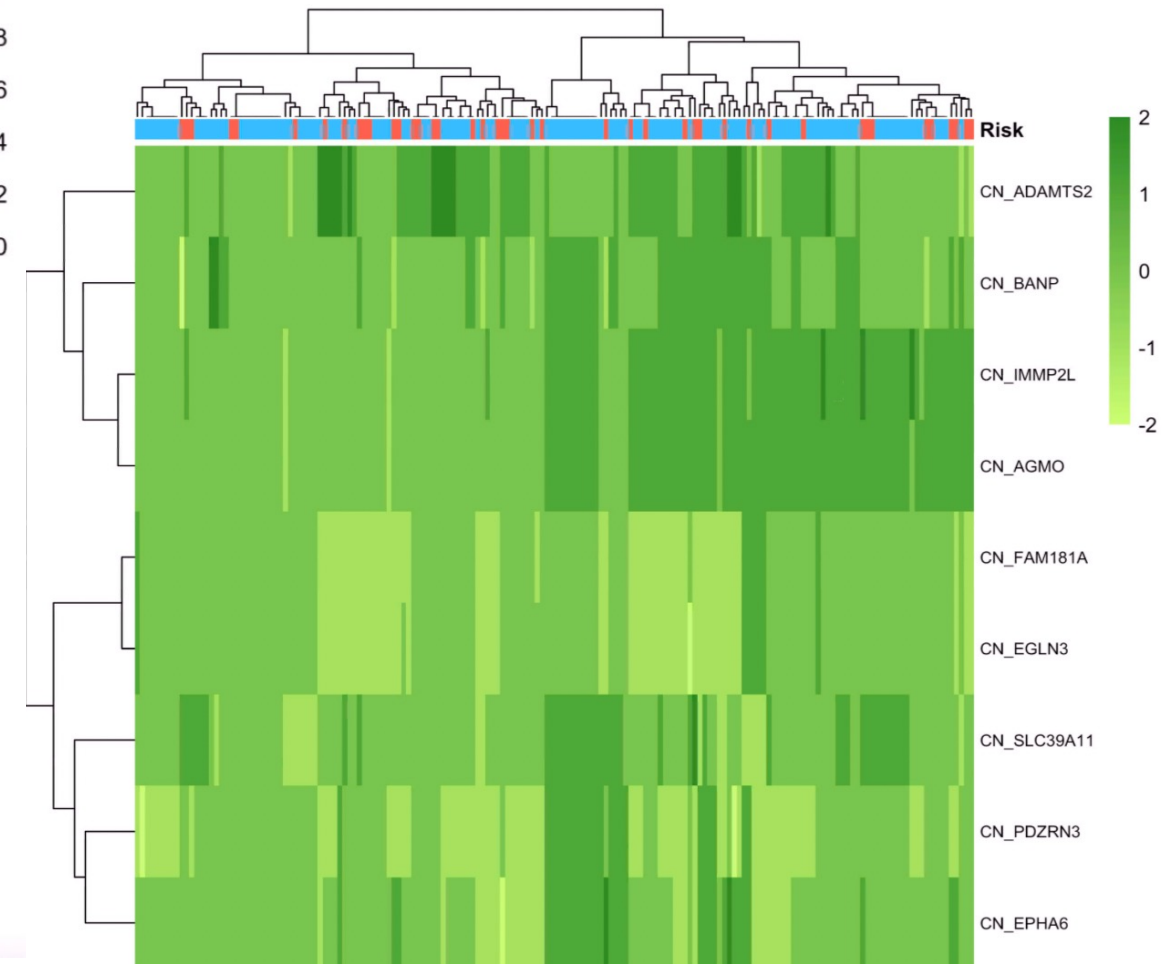
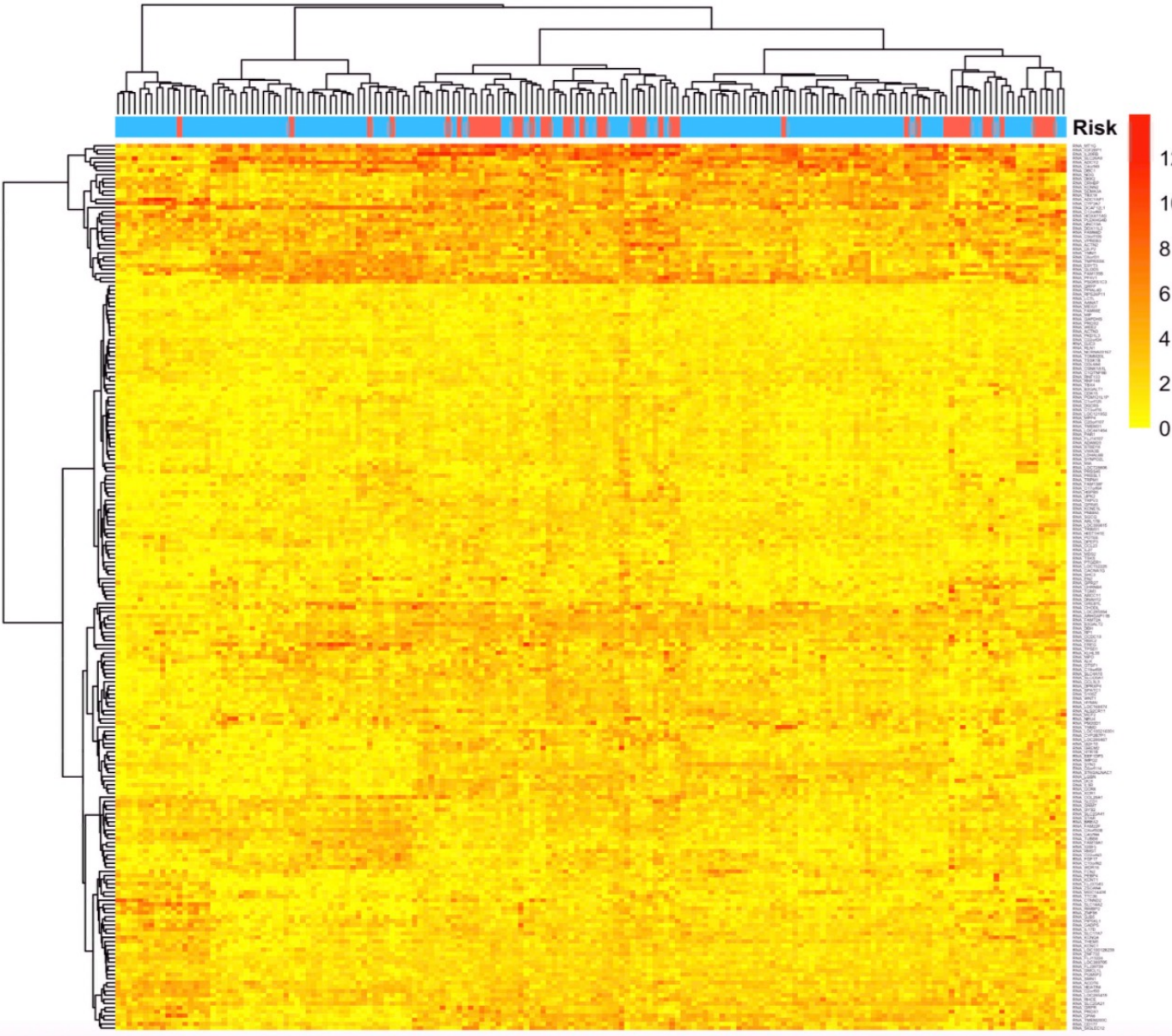
Number at risk					
High risk	43	15	8	2	0
Low risk	127	72	33	14	2
	0	1000	2000	3000	4000

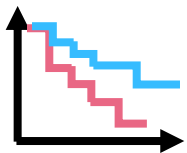
Days



KIPAN

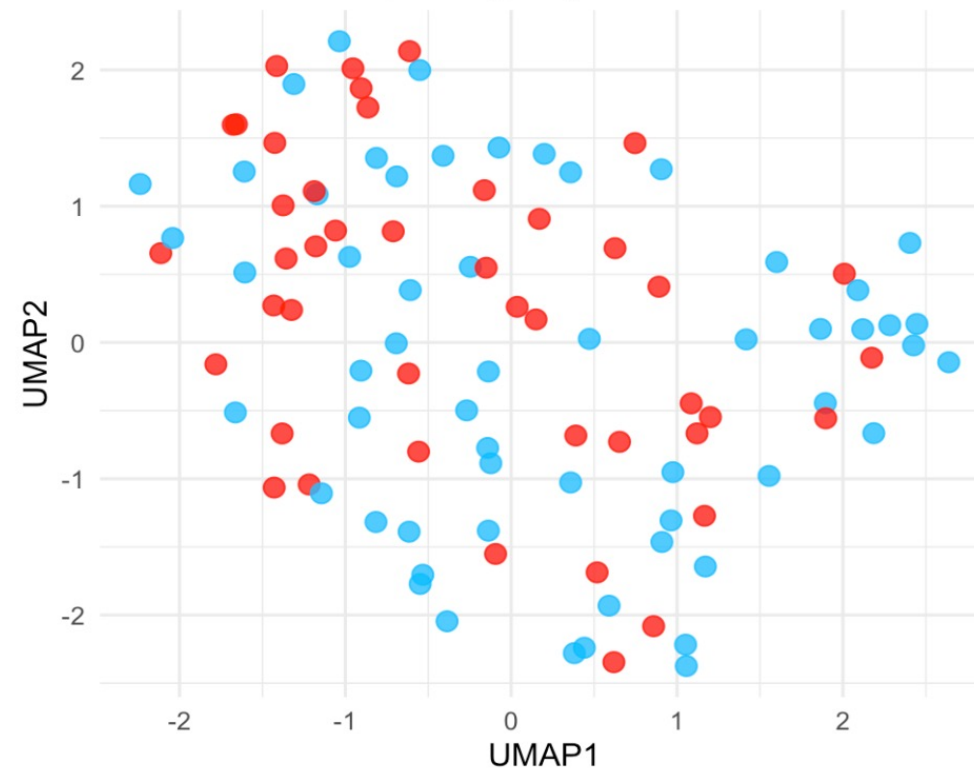
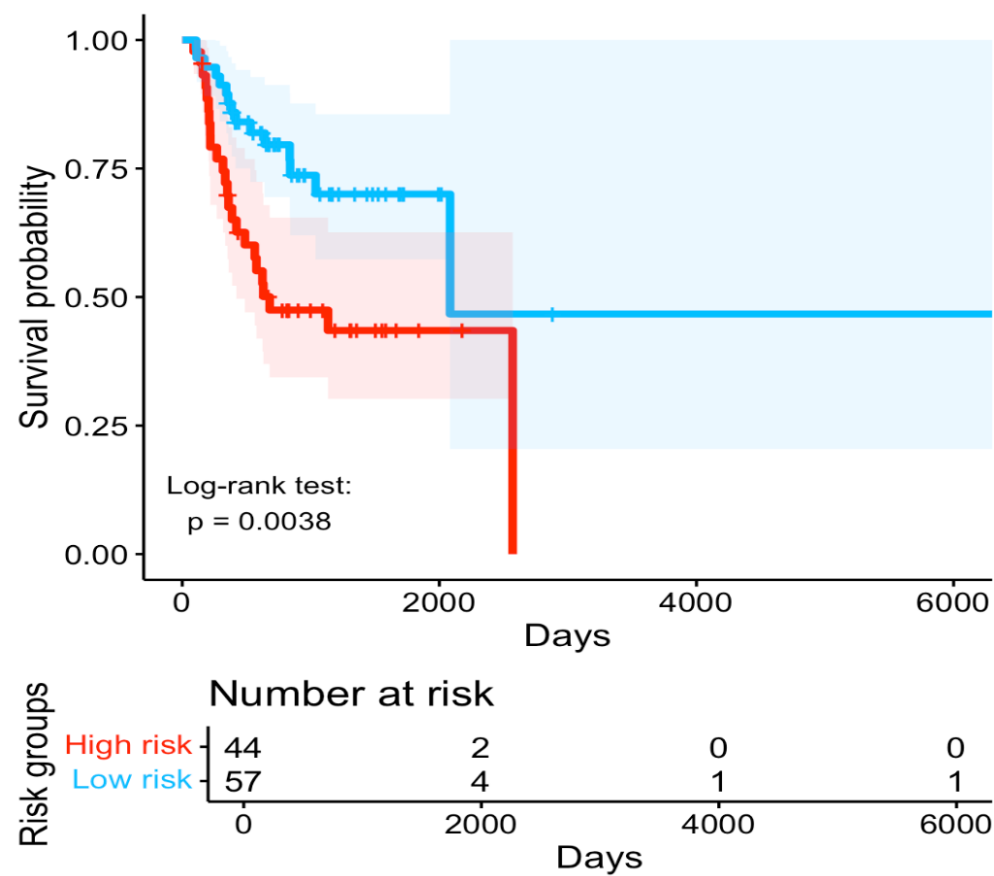
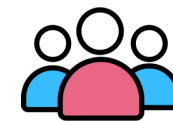
Heatmaps based on model's selected features





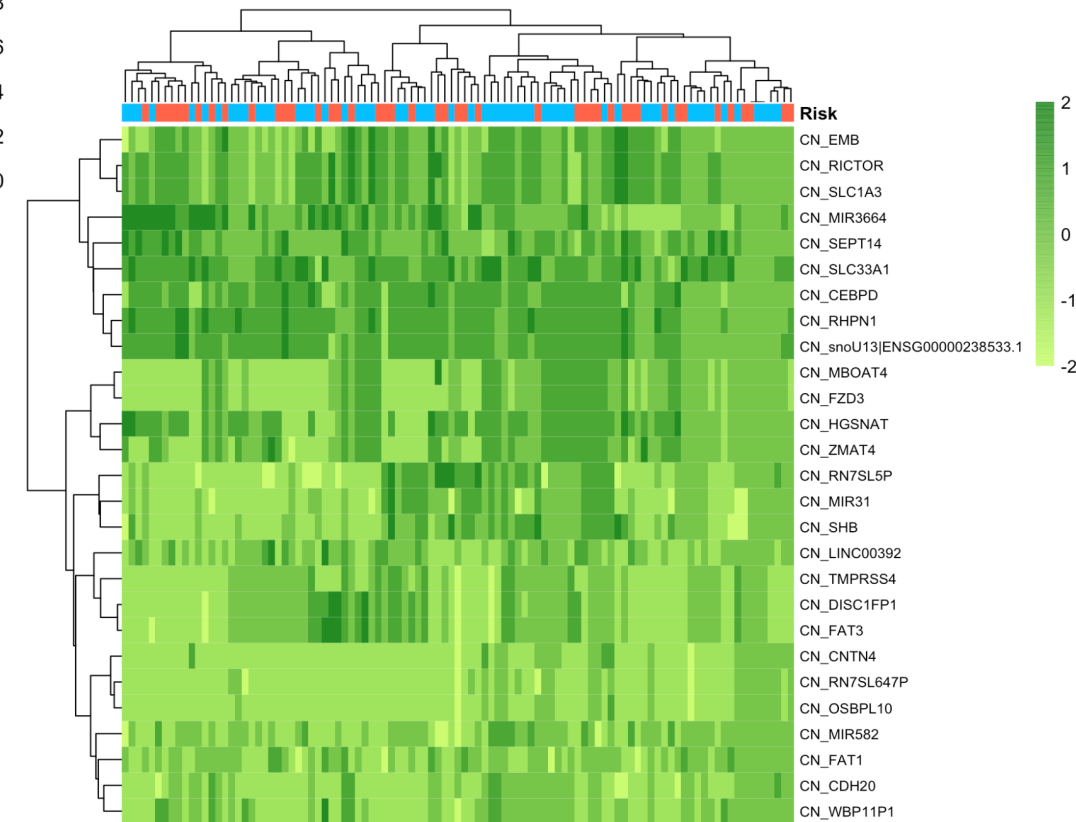
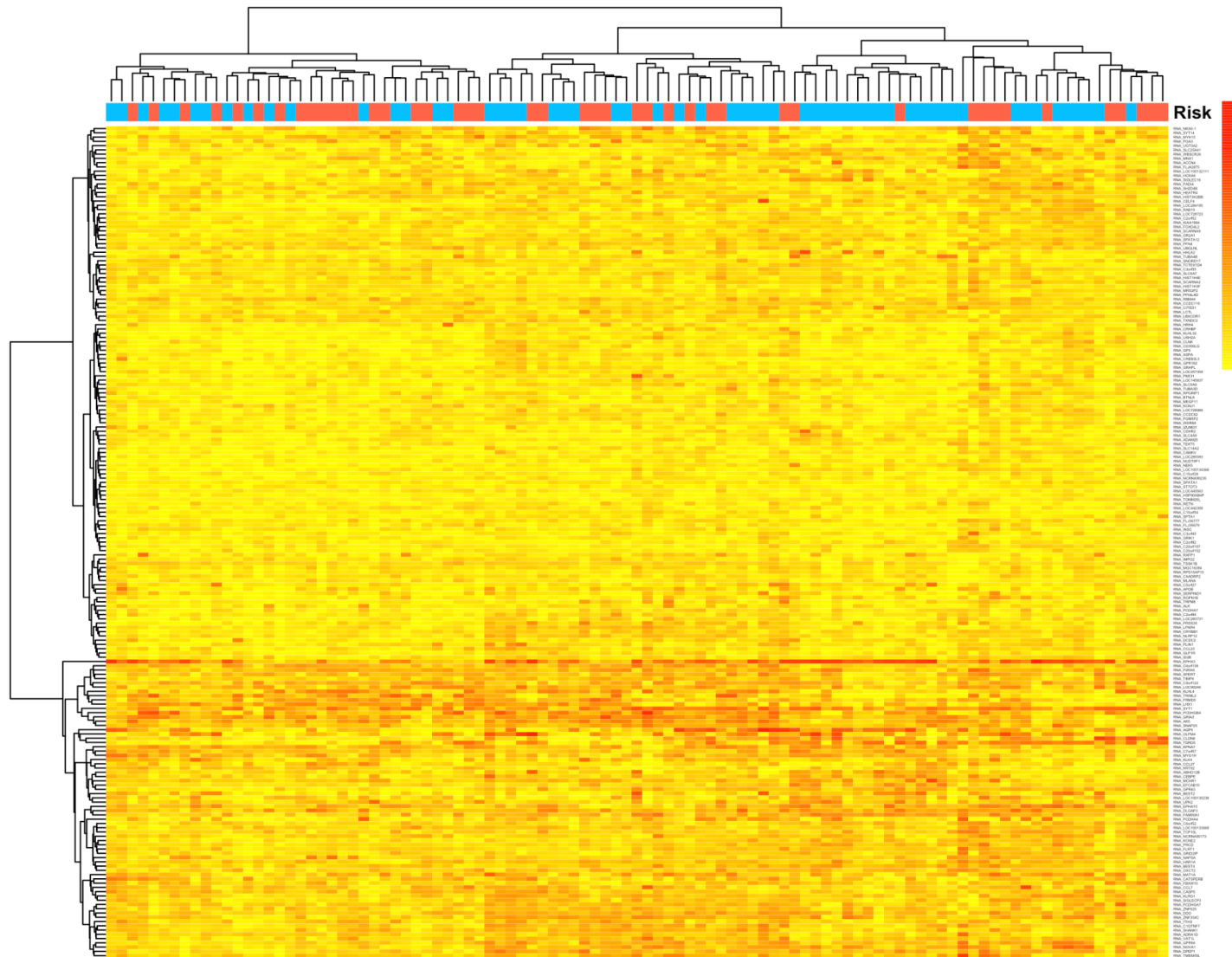
HNSC

Patient stratification based on early integration with adaptive lasso-cox results



HNSC

Heatmaps based on model's selected features



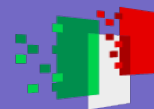
THANKS FOR YOUR ATTENTION



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



CNR-IAC
Istituto
per le Applicazioni
del Calcolo

This work was supported by PRIN 2022 PNRR P2022BLN38 project, "Computational approaches for the integration of multi-omics data" funded by European Union - Next Generation EU, CUP B53D23027810001.