

## URINARY BIOMARKERS FOR EARLY DETECTION OF PANCREATIC CANCER

Pancreatic cancer is a type of cancer that begins in the tissues of pancreas and it is one of the deadliest cancers, with less than 10% patients surviving for 5 years. The most common type of pancreatic cancer is the Pancreatic Ductal Adenocarcinoma (PDAC). It is often asymptomatic until it has spread to other organs, so its early detection is very rare as there are no useful biomarkers in the initial stage. Debernardi et al. [1] have validated certain urinary biomarkers for early detection of PDAC and it helps in predicting the disease before it is diagnosed.

I plan to use Mechanistic Data Science to solve this problem with the six steps given below:

### **1. Multimodal Data Collection:**

The data will be collected from the abovementioned paper of Debernardi et al. Here we have 590 data points.

sample_id	patient_cohort	sample_origin	age	sex	diagnosis	stage	benign_sample_diagnosis	plasma_CA19_9	creatinine	LYVE1	REG1B	TFF1	REG1A
S1	Cohort1	BPTB	33	F	1			11.7	1.83222	0.893219	52.94884	654.2822	1262
S10	Cohort1	BPTB	81	F	1				0.97266	2.037585	94.46703	209.4883	228.407
S100	Cohort2	BPTB	51	M	1			7	0.78039	0.145589	102.366	461.141	
S101	Cohort2	BPTB	61	M	1			8	0.70122	0.002805	60.579	142.95	
S102	Cohort2	BPTB	62	M	1			9	0.21489	0.00086	65.54	41.088	
S103	Cohort2	BPTB	53	M	1				0.84825	0.003393	62.126	59.793	

### **2. Extracting Mechanistic Features:**

According to the paper, the key features are four urinary biomarkers: creatinine, LYVE1, REG1B, and TFF1.

- Creatinine is a protein that is often used as an indicator of kidney function.
- YVLE1 is lymphatic vessel endothelial hyaluronan receptor 1, a protein that may play a role in tumor metastasis
- REG1B is a protein that may be associated with pancreas regeneration
- TFF1 is trefoil factor 1, which may be related to regeneration and repair of the urinary tract

### **3,4. Dimension reduction & Reduced order model:**

The extracted four features and least square optimization for each data set will be used for reducing dimension and thus a reduced order model will be created.

### **5. Regression and Classification:**

I plan to use SVM to build a predictive model, to differentiate between different stages 3(pancreatic cancer), 2(non-cancerous pancreas condition) and 1(healthy).

### **6. System and Design:**

If we provide the urinary biomarkers as input, the model will detect if the patient is healthy or if the patient can have cancer (before developing critical symptoms).

[1]Debernardi, S., O'Brien, H., Algahmdi, A. S., Malats, N., Stewart, G. D., Plješa-Ercegovac, M., Costello, E., Greenhalf, W., Saad, A., Roberts, R., Ney, A., Pereira, S. P., Kocher, H. M., Duffy, S., Blyuss, O., & Crnogorac-Jurcevic, T. (2020). A combination of

urinary biomarker panel and PANCRISK score for earlier detection of pancreatic cancer: A case–control study. *PLOS Medicine*, 17(12). <https://doi.org/10.1371/journal.pmed.1003489>