

Supplementary: INSIGHT: Spatially resolved survival modelling from routine histology crosslinked with molecular profiling reveals prognostic epithelial-immune axes in stage II/III colorectal cancer

Variable	Group	TCGA+SR386 (n=678)	PETACC3 (n=1155)	Janssen (n=80)	IMMUcan (n=15)	Orion (n=27)
Age,years		68 (59-76)	60 (53-67)	68 (62-78)	72 (51-83)	60 (48-72)
Sex						
	Male	355 (52%)	643 (56%)	19 (24%)	7 (47%)	13 (48%)
	Female	332 (48%)	512 (44%)	16 (20%)	8 (53 %)	14 (52%)
	Missing	0 (0%)	0 (0%)	45 (56%)	0 (0 %)	0 (0)
MSI Status						
	MSS	593 (87%)	939 (81%)	65 (81%)	5 (33%)	20 (74%)
	MSI High	79 (12%)	140 (12%)	14 (18%)	10 (67%)	3 (11%)
	Missing	6 (1%)	76 (7%)	1 (1%)	0 (0%)	4 (15%)
Stage						
	Stage 1	0 (0%)	0 (0%)	4 (5%)	1 (7%)	0 (0%)
	Stage 2	332 (49%)	345 (30%)	19 (24%)	8 (53%)	9 (33%)
	Stage 3	346 (51%)	810 (70%)	10 (13%)	6 (40%)	18 (67%)
	Stage 4	0 (0)	0 (0%)	2 (2%)	0 (0%)	0 (0%)
	Missing	0 (0%)	0 (0%)	45 (56%)	0 (0%)	0 (0%)

Table 1: Characteristics of the datasets used in this paper. For age we report its median and Inter Quartile Range (IQR). For remaining variable we report total patient counts of each subgroup as well as their overall percentage.

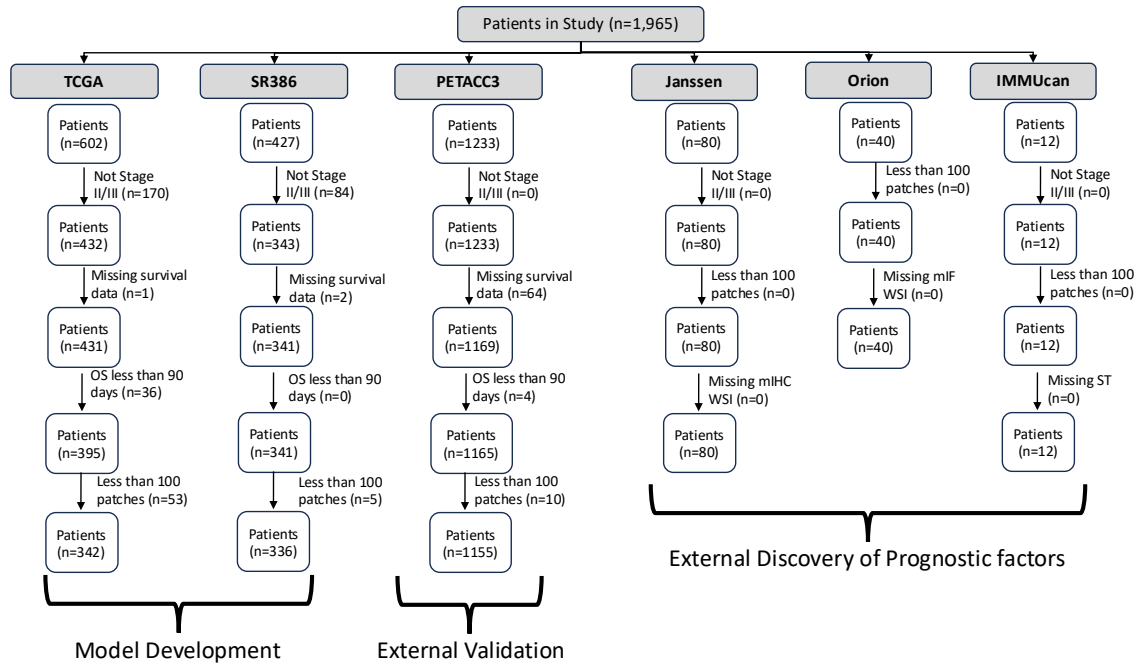


Figure 1: CONSORT diagram of all datasets used in this study as well as the filtering criteria used to exclude patients.

Variable	Group	TCGA+SR386		PETACC3	
		HR (95% CI)	p-value	HR (95% CI)	p-value
Stage	Stage II	ref.		ref.	
	Stage III	2.31 (1.55-3.47)	p<0.005	2.58 (1.6-3.98)	p<0.005
MSI-status	MSS	ref.		ref.	
	MSI High	0.67 (0.34-1.33)	p=0.25	0.68 (0.32-1.46)	p=0.33
Sex	Female	ref.		ref.	
	Male	1.04 (0.7-1.51)	p=0.84	1.27 (0.93-1.74)	p=0.14
Age	>60	ref.		ref.	
	≤60	0.59 (0.37-0.94)	p=0.03	0.68 (0.50-0.93)	p=0.01

Table 2: Hazard ratio between different subpopulations within training set (TCGA+SR386) and external validation set (PETACC3). Here we report the hazard ratio (HR) with its 95% confidence interval (CI) as well as the log-rank p-value.