Supplementary

Table S1 Data sources

ID	Platform	Sample Type	Sample size (case:control)	Data type
TCGA-STAD		Tumor and adjacent tissues	407 (375: 32)	RNAseq
GSE84437	GPL6947	Tumor	433	mRNA array
GSE66229	GPL570	Tumor	300	mRNA array
GSE28541	GPL13376	Tumor	40	mRNA array

Table S2 Gene Ontology functional enrichment analysis of differentially expressed liquid–liquid phase separation-related genes

	ID	Description	GeneNumber	type	type_order		
1	GO:0034063	stress granule assembly	10	biological process	stress granule assembly		
2	GO:1903311	regulation of mR metabolic process	21	biological process	regulation of mR metabolic process		
3	GO:0050684	regulation of mR processing	12	biological process	regulation of mR processing		
4	GO:0008380	R splicing	20	biological process	R splicing		
5	GO:0000377	R splicing	via transesterification reactions	17	biological process	R splicing	via transesterification reactions
6	GO:0000398	mR splicing	via spliceosome	17	biological process	mR splicing	via spliceosome
7	GO:0000375	R splicing	via transesterification reactions	17	biological process	R splicing	via transesterification reactions
8	GO:0000380	alternative mR splicing	via spliceosome	9	biological process	alternative mR splicing	via spliceosome
9	GO:0006403	R localization	13	biological process	R localization		
10	GO:0051236	establishment of R localization	12	biological process	establishment of R localization		
11	GO:0048024	regulation of mR splicing	via	10	biological process	regulation of mR splicing	via
12	GO:0006270	D replication initiation	7	biological process	D replication initiation		
13	GO:0030522	intracellular receptor signaling pathway	16	biological process	intracellular receptor signaling pathway		
14	GO:0043484	regulation of R splicing	12	biological process	regulation of R splicing		
15	GO:0140694	non-membrane-bounded organelle assembly	17	biological process	non-membrane-bounded organelle assembly		
16	GO:0010494	cytoplasmic stress granule	13	cellular component	cytoplasmic stress granule		
17	GO:0035770	ribonucleoprotein granule	18	cellular component	ribonucleoprotein granule		
18	GO:0036464	cytoplasmic ribonucleoprotein granule	17	cellular component	cytoplasmic ribonucleoprotein granule		
19	GO:0000792	heterochromatin	11	cellular component	heterochromatin		
20	GO:0098687	chromosomal region	20	cellular component	chromosomal region		
21	GO:0031519	PcG protein complex	8	cellular component	PcG protein complex		
22	GO:0000781	chromosome	telomeric region	11	cellular component	chromosome	telomeric region
23	GO:0035102	PRC1 complex	5	cellular component	PRC1 complex		
24	GO:0000228	nuclear chromosome	12	cellular component	nuclear chromosome		
25	GO:0016607	nuclear speck	15	cellular component	nuclear speck		
26	GO:0140693	molecular condensate scaffold activity	8	molecular function	molecular condensate scaffold activity		
27	GO:0003688	D replication origin binding	6	molecular function	D replication origin binding		
28	GO:0140030	modification-dependent protein binding	12	molecular function	modification-dependent protein binding		
29	GO:0030674	protein-macromolecule adaptor activity	17	molecular function	protein-macromolecule adaptor activity		
30	GO:1990841	promoter-specific chromatin binding	8	molecular function	promoter-specific chromatin binding		
31	GO:0043021	ribonucleoprotein complex binding	11	molecular function	ribonucleoprotein complex binding		

Table S3 Predictive accuracy (concordance index with 95% confidence interval) of the prognostic model in The Cancer Genome Atlas and external validation cohorts

	Cohort	C_index	Lower_95	Upper_95	CI_String
1	TCGA_Train	0.653	0.594	0.711	0.653 (0.594-0.711)
2	TCGA_Test	0.562	0.465	0.659	0.562 (0.465-0.659)
3	GSE84437	0.573	0.536	0.61	0.573 (0.536-0.61)
4	GSE66229	0.528	0.481	0.576	0.528 (0.481-0.576)
5	GSE28541	0.462	0.353	0.571	0.462 (0.353-0.571)

https://dx.doi.org/10.21037/tcr-2025-1-2846

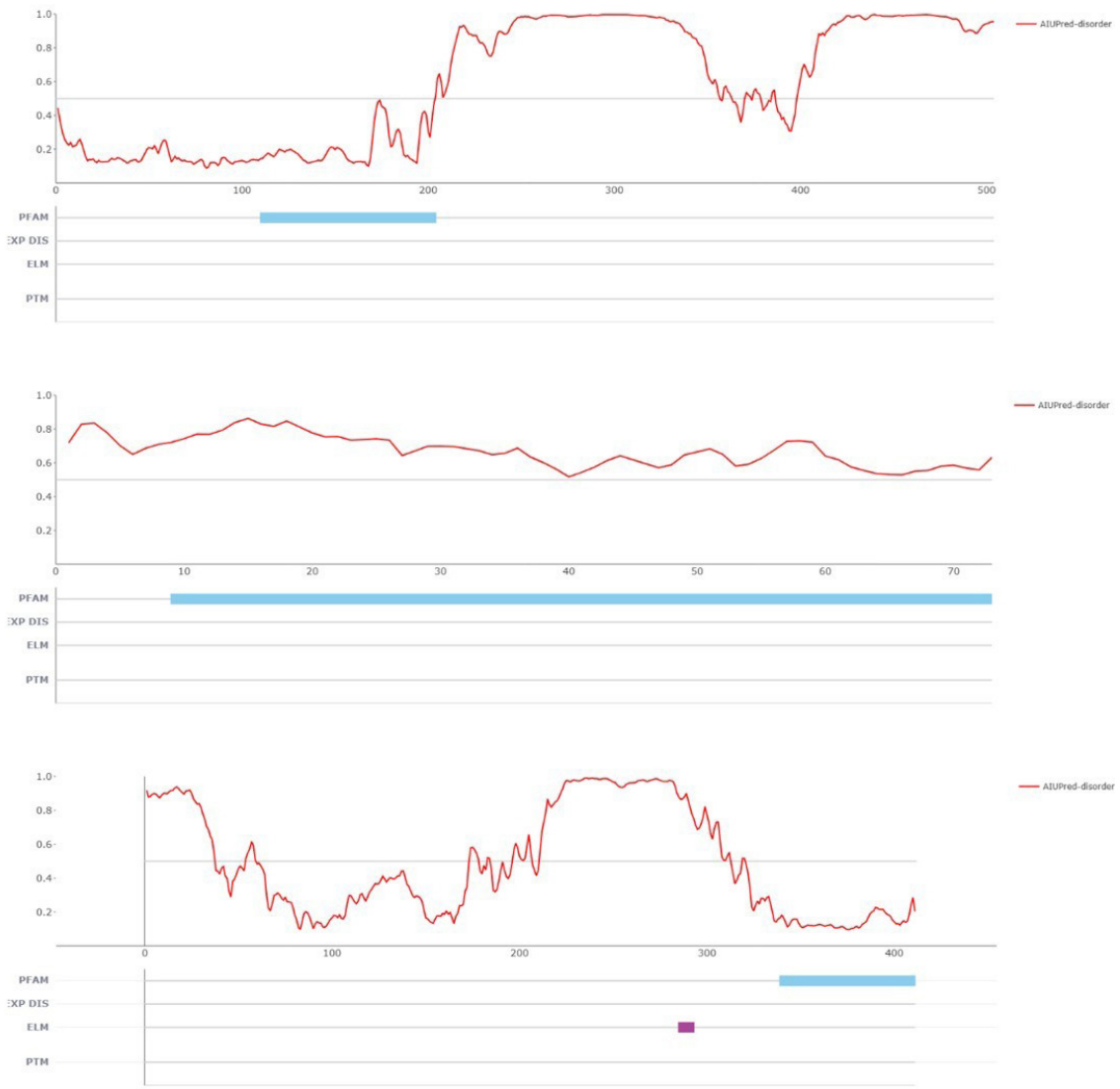

Figure S1 Sequence-based prediction of intrinsic disorder in model proteins. (A-C) Intrinsic disorder profiles of *DOK7* (A), *GNG11* (B), and *ZFYVE27* (C) predicted by the Intrinsically Unstructured Protein Predictor. The x-axis represents amino acid position; the y-axis indicates disorder probability. The red curve indicates the predicted disorder score, with values >0.5 considered intrinsically disordered. Annotated features are displayed below each plot: light blue bars denote Pfam domains, purple bars indicate Eukaryotic Linear Motifs (ELMs), blue segments represent predicted disordered regions (XP DIS), and green markers indicate post-translational modification (PTM) sites.

https://dx.doi.org/10.21037/tcr-2025-1-2846

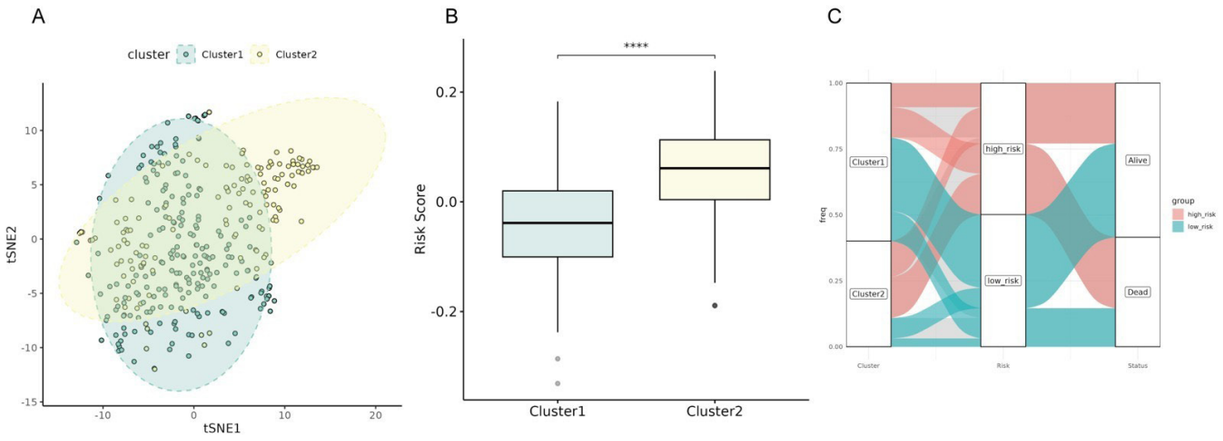

Figure S2 Relationship between clusters, risk scores, and survival status in the The Cancer Genome Atlas stomach adenocarcinoma cohort. (A) t-distributed Stochastic Neighbor Embedding plot illustrating the separation between the two clusters. (B) Boxplot comparing the risk scores between different clusters. Cluster 2 shows significantly higher risk scores than Cluster 1 (****, P<0.0001). (C) Sankey diagram showing the distribution of patients from each cluster into high-risk and low-risk groups and their corresponding survival status. High-risk patients are depicted in red and low-risk patients in blue, with survival outcomes (alive or dead) represented on the right.

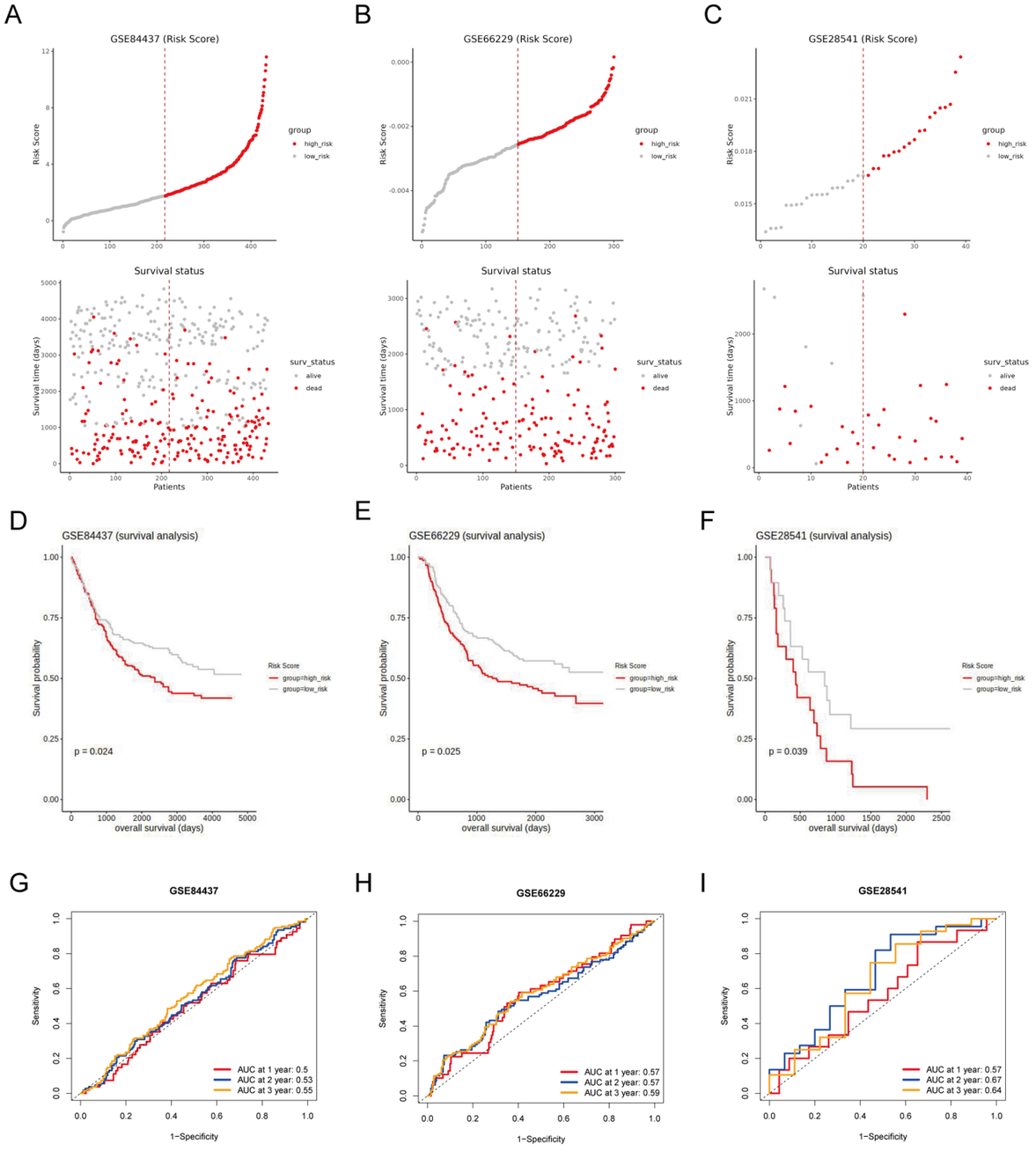

Figure S3 Validation of the prognostic value of the risk model in external gastric cancer cohorts. (A-C) Risk factor distribution and survival status plots for GSE84437 (A), GSE66229 (B), and GSE28541 (C). (D-F) Kaplan-Meier survival curves for GSE84437 (D), GSE66229 (E), and GSE28541 (F). (G-I) Time-dependent ROC curves for GSE84437 (G), GSE66229 (H), and GSE28541 (I) at 1 year, 2 years, and 3 years. area under the curve values are indicated.

https://dx.doi.org/10.21037/tcr-2025-1-2846

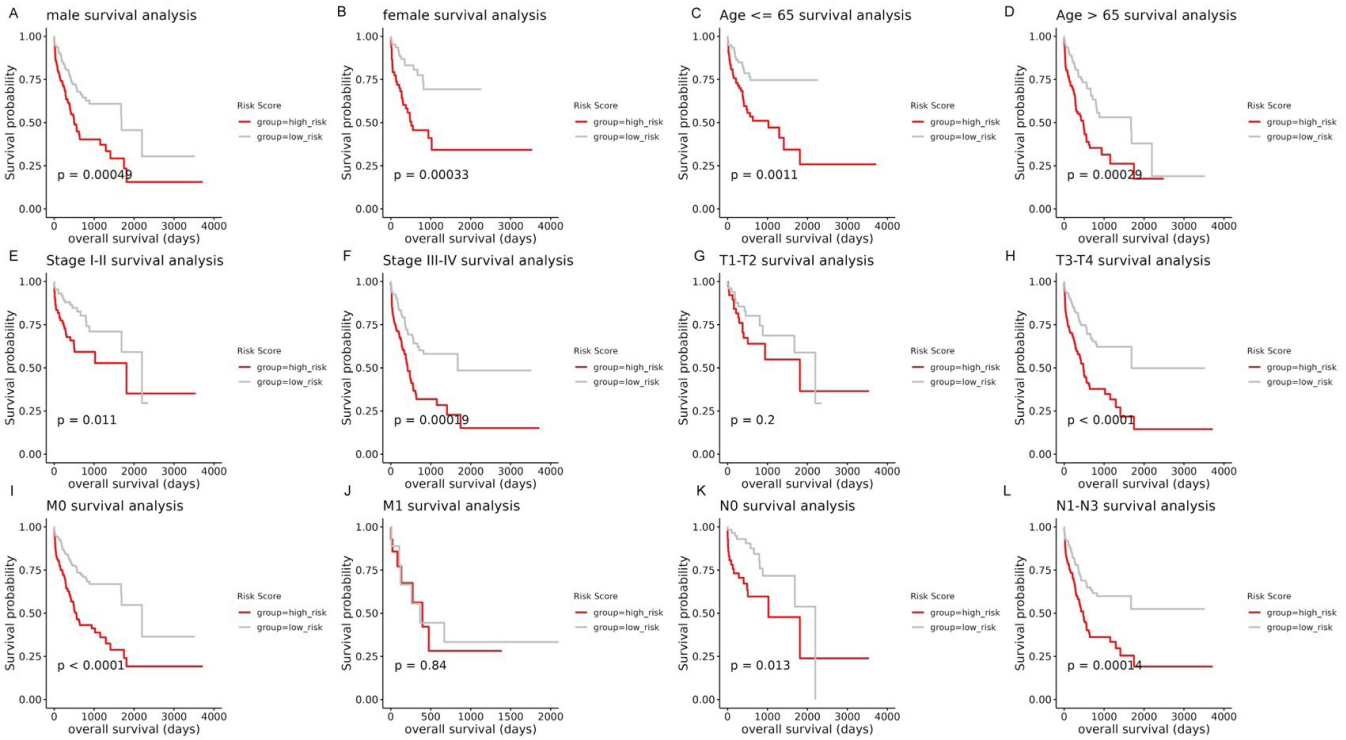

Figure S4 Survival analysis of high-risk vs. low-risk groups across various clinical and demographic subgroups. High-risk patients consistently showed significantly worse overall survival compared to low-risk patients across multiple subgroups, including males (A), females (B), different age groups (C,D), early and advanced stages (E,F), larger tumors (H), no metastasis (I), and both no (K) and with lymph node involvement (L). Exceptions were observed in subgroups with smaller tumors (G) and those with metastasis (J).

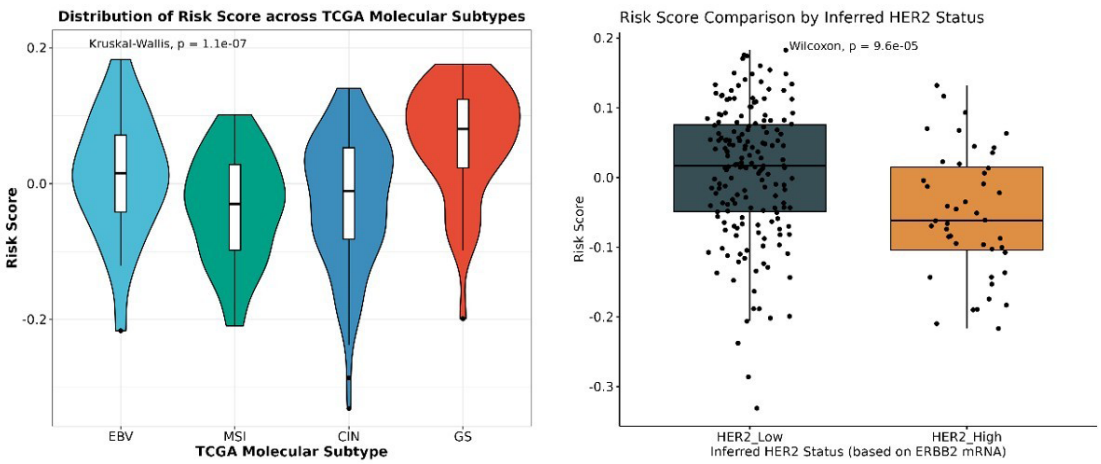

Figure S5 RiskScore distribution among The Cancer Genome Atlas stomach adenocarcinoma (TCGA-STAD) molecular subtypes and HER2-defined groups. (A) Distribution of RiskScore among TCGA-STAD molecular subtypes, including Epstein-Barr virus (EBV), microsatellite instability (MSI), chromosomal instability (CIN), and genomically stable (GS). Statistical significance was assessed using the Kruskal-Wallis test. (B) Comparison of RiskScore between human epidermal growth factor receptor 2 (HER2)-defined groups. Patients were stratified into HER2 high- and low-expression groups based on mRNA levels. Statistical significance was evaluated using the Wilcoxon rank-sum test.

https://dx.doi.org/10.21037/tcr-2025-1-2846

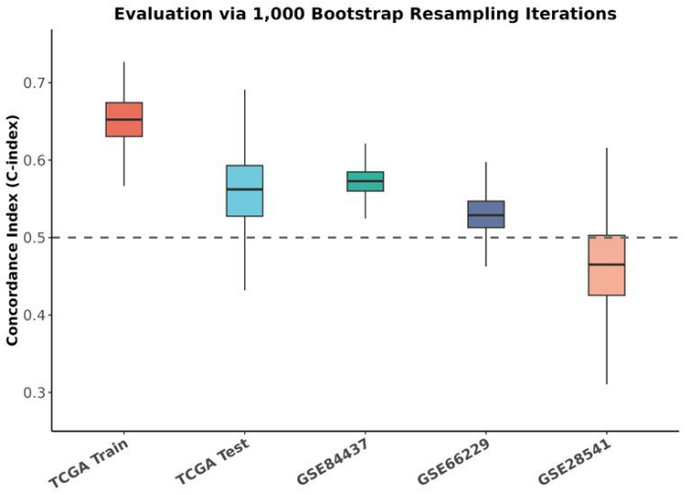

Figure S6 Bootstrap validation of model discrimination across cohorts. Distribution of concordance index (C-index) values obtained from 1,000 bootstrap resampling iterations in the The Cancer Genome Atlas (TCGA) training set, TCGA testing set, and external validation cohorts (GSE84437, GSE66229, and GSE28541). The dashed line indicates the reference value of 0.5.

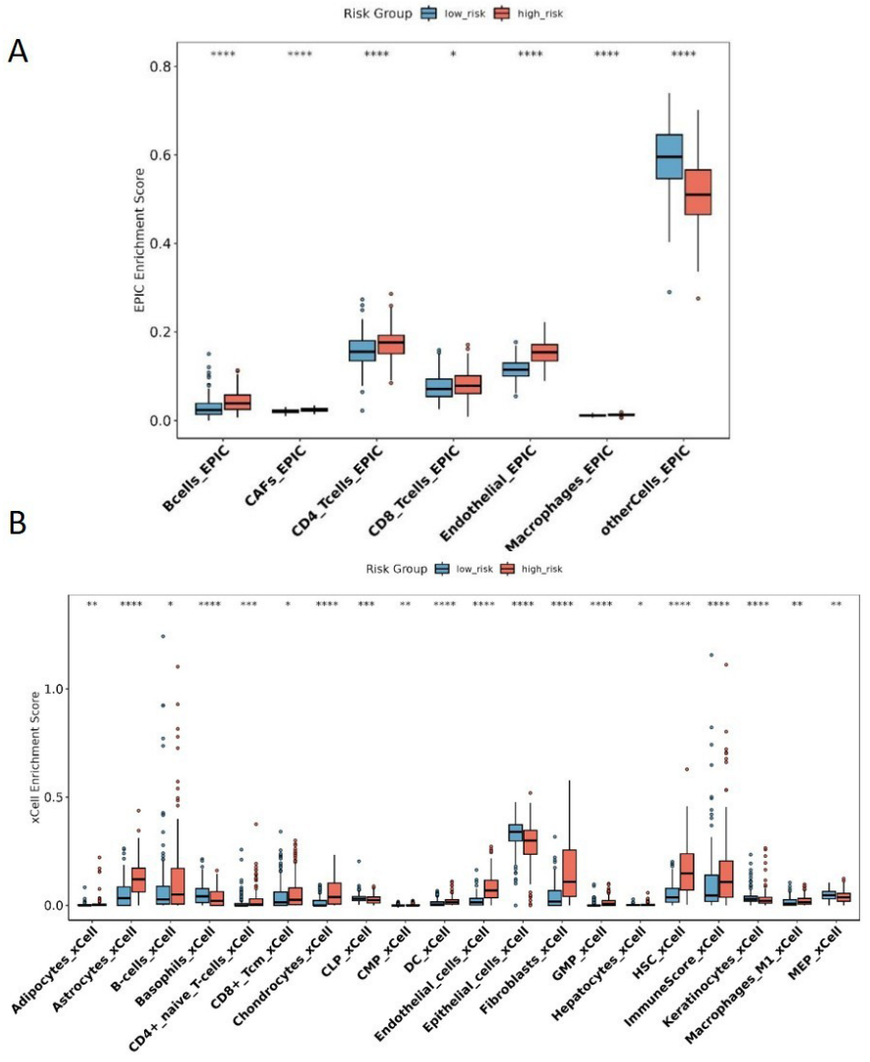

Figure S7 Cross-validation of immune infiltration using xCell and Estimating the Proportions of Immune and Cancer cells (EPIC) algorithms. (A) EPIC analysis comparing the relative abundance of cancer-associated fibroblasts (CAFs), endothelial cells, macrophages, and B cells between high- and low-risk groups. (B) xCell analysis comparing enrichment scores of fibroblasts, endothelial cells, macrophages, CD4+ T cells, and overall ImmuneScore between high- and low-risk groups. Data are presented as boxplots with group comparisons performed using the Wilcoxon rank-sum test. *, P<0.05; **, P<0.01; ***, P<0.001; ****, P<0.0001.

https://dx.doi.org/10.21037/tcr-2025-1-2846