

Table S1. Clinical information of training set and validation sets

Clinical characteristics	GSE32062 (N=260)	TCGA (N=419)	GSE49997 (N=204)	GSE8842 (N=83)	GSE26712 (N=185)	GSE31245 (N=58)
Age at diagnosis (mean±SD)	-	59.49±11.46	57.66±11.82	51.28±13.35	-	-
Clinical stage (I/II/III/IV/-)	0/0/204/56	0/23/331/62/3	0/9/154/31/10	-	-	-
Neoplasm histologic grade (G1/G2/ G3/G4/-)	0/131/129/0	1/52/357/1/8	0/50/143/0/11	-	-	-
Tumor recurrence (Yes/No/-)	193/67	222/24/173	124/70/10	21/62	-	-
Recurrence free survival time	26.87±22.99	18.21±13.03	19.06±11.22	76.69±41.69	-	-
Death (Dead/Alive/-)	121/139	233/185/1	57/137/10	15/66/2	129/24/32	24/34
Overall survival days (mean±SD)	44.81±25.01	34.07±27.61	25.72±11.71	76.69±41.69	42.54±32.59	42.52±16.15

SD: Standard deviation

Table S2. Univariable and multivariate cox regression analysis for clinical factors

Clinical factors	Univariable cox regression		Multivariable cox regression	
	<i>P</i> values	HR(95%CI)	<i>P</i> value	HR (95%CI)
Clinical stage (III/IV)	0.0132	1.515 (1.088-2.109)	0.0194	1.49 (1.067-2.082)
Neoplasm histologic grade (G2/G3)	0.329	1.151 (0.868-1.527)	0.5121	1.10 (0.827-1.465)

HR: Hazard ratio; CI: Confidence interval.

Table S3. The predict results of the pathway-based model in training sets and validation sets

Datasets	Validation in PFS				Validation in OS			
	LogRank <i>P</i> value	AUC	High risk (mean±SD)	Low risk (mean±SD)	LogRank <i>P</i> value	AUC	High risk (mean±SD)	Low risk (mean±SD)
GSE32062	<0.001	0.992	13.74±11.71	21.61±20.16	0.010	0.918	38.82±26.58	45.74±23.39
TCGA	0.030	0.960	16.22±12.97	20.19±12.83	0.007	0.932	34.29±26.38	39.70±27.74
GSE8842	0.046	0.909	68.29±40.07	85.11±42.07	0.227	0.912	68.29±40.07	85.11±42.07
GSE49997	0.002	0.946	16.08±8.98	22.03±12.43	0.006	0.943	23.52±10.82	27.93±12.20
GSE26712	-	-	-	-	0.012	0.930	36.40±28.38	48.76±35.48
GSE31245	-	-	-	-	0.024	0.848	38.89±16.05	46.14±15.70

SD: Standard deviation; PFS: Progression-free survival; OS: overall survival; AUC: Area under the curve.

Table S4. The predict results of the gene-based model in training sets and validation sets

Datasets	Validation in PFS				Validation in OS			
	LogRank <i>P</i> value	AUC	High risk (mean ± SD)	Low risk (mean ± SD)	LogRank <i>P</i> value	AUC	High risk (mean ± SD)	Low risk (mean ± SD)
GSE32062	<0.001	0.954	19.00±18.09	34.74±24.68	<0.001	0.943	41.71±25.29	47.92±24.41
TCGA	0.039	0.894	16.77±13.47	19.64±12.46	0.634	0.856	36.61±27.96	37.39±26.43
GSE8842	0.565	0.675	74.9±43.81	78.49±39.93	0.297	0.700	74.9±43.81	78.49±39.93
GSE49997	0.002	0.888	16.58±10.21	21.53±11.69	0.066	0.882	23.65±12.12	27.79±10.96
GSE26712	-	-	-	-	0.353	0.882	39.12±32.07	45.92±32.96
GSE31245	-	-	-	-	0.262	0.704	40.69±15.43	44.34±16.91

SD: Standard deviation; PFS: Progression-free survival; OS: overall survival; AUC: Area under the curve.

Table S5. The predict results of comprehensive (pathways and clinical factor)-based model in training sets and validation sets

Datasets	Validation in PFS				Validation in OS			
	LogRank <i>P</i> value	AUC	High risk (mean±SD)	Low risk (mean±SD)	LogRank <i>P</i> value	AUC	High risk (mean±SD)	Low risk (mean±SD)
GSE32062	<0.001	0.918	20.09±19.26	33.68±24.44	0.001	0.928	42.42±26.34	47.2±23.44
TCGA	0.002	0.902	16.21±12.86	20.20±12.94	0.030	0.868	34.82±27.18	39.73±26.96
GSE4997	<0.001	0.898	15.76±8.49	22.35±12.61	0.005	0.898	23.23±10.77	28.02±12.18

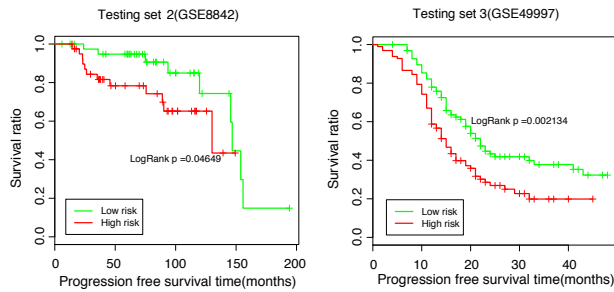
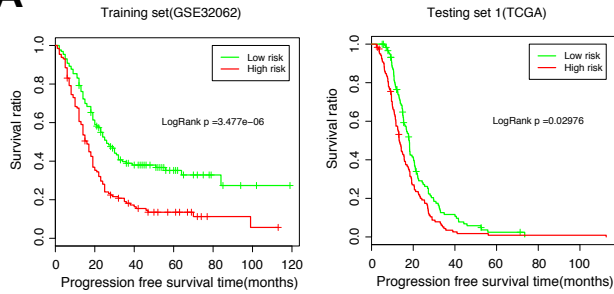
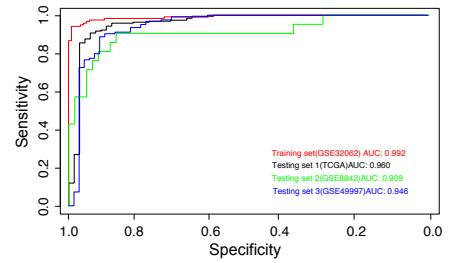
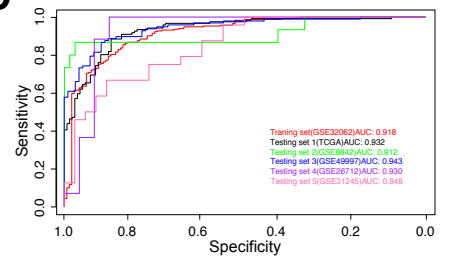
SD: Standard deviation; PFS: Progression-free survival; OS: overall survival; AUC: Area under the curve.

Supplemental Figure S1 Kaplan-Meier survival analysis and Receiver Operating Characteristic (ROC) curves for the pathway model.

(A) Kaplan-Meier survival curves and (B) ROC curves of progression-free survival (PFS). (C) Kaplan-Meier survival curves and (D) ROC curves of overall survival (OS). The abscissa represents sensitivity and the ordinate represents specificity.

Supplemental Figure S2 Kaplan-Meier survival analysis and Receiver Operating Characteristic (ROC) curves for the gene model.

(A) Kaplan-Meier survival curves and (B) ROC curves of progression-free survival (PFS). (C) Kaplan-Meier survival curves and (D) ROC curves of overall survival (OS). The abscissa represents sensitivity and the ordinate represents specificity.

A**B****D****C**