Table S1 The definition of tumors' stage by SEER

Stage	Invasion area
Localized	Adjacent ipsilateral lobe
	Confined to hilus
	Main stem bronchus (without involvement of the carina), including extension from other part of lung
	Minimally invasive adenocarcinoma
	Superficial tumor with invasive component limited to bronchial wall, with or without proximal extension to main stem bronchus
Regional	Atelectasis/obstructive pneumonitis
	Blood vessel(s): aorta, azygos vein, pulmonary artery or vein superior vena cava
	Brachial plexus
	Carina from lung
	Chest wall (thoracic wall)
	Compression of esophagus or trachea not specified as direct extension
	Diaphragm
	Esophagus
	Mediastinum
	Nerve(s): cervical sympathetic (Horner's syndrome), recurrent laryngeal (vocal cord paralysis), vagus, phrenic nerve
	Pancoast tumor (superior sulcus syndrome)
	Visceral pleura invasion (PL1, PL2 or PL3)
	Pericardium
	Pulmonary ligament
	Separate tumor nodule(s) in the same lobe as the primary
	Trachea
	Lymph node(s): N1 or N2
Distant	Adjacent rib
	Contralateral lung
	Contralateral main stem bronchus
	Heart
	Inferior vena cava
	Neural foramina
	Pericardial nodules or pleural effusion (malignant) (ipsilateral, contralateral, bilateral)
	Pleural tumor foci or nodules on ipsilateral lung (separate from direct extension) or contralateral lung
	Skeletal muscle, skin of chest, sternum, vertebra(e) (vertebral body) or another distant metastasis
	Lymph node: N3

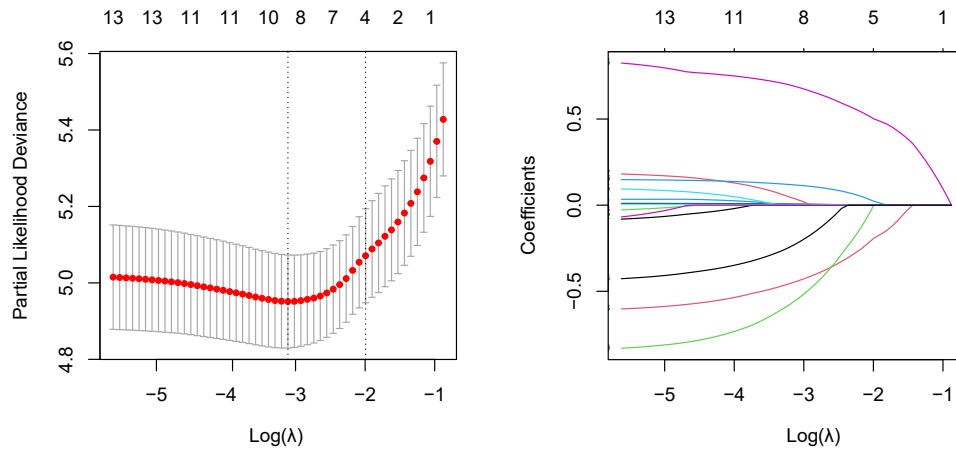SEER, The Surveillance, Epidemiology, and End Results database.

Figure S1 The LASSO regression result used to decide λ value. LASSO, Least Absolute Shrinkage and Selection Operator.

Table S2 The OS time of the other four subtypes of PSC patients

Subtypes	Median [IQR] (months)	Range (months)
Pleomorphic carcinoma (N=757)	7 [2, 27]	1–203
Giant cell carcinoma (N=636)	4 [1, 13]	1–227
Pulmonary blastoma (N=68)	28 [8, 106]	1–226
Carcinosarcoma (N=461)	7 [2, 24]	1–219

OS, overall survival; PSC, pulmonary sarcomatoid carcinoma; IQR, interquartile range.

Table S3 Life table of pulmonary spindle cell carcinoma (n=833)

Time (months)	N.Risk	N.Event	Survival probability	Lower 95% CI	Upper 95% CI
1	833	195	76.6%	73.8%	79.5%
2	598	68	67.9%	64.7%	71.2%
3	512	52	61.0%	57.7%	64.5%
4	447	41	55.4%	52.0%	59.0%
5	401	29	51.4%	48.0%	55.0%
6	364	38	46.0%	42.6%	49.7%
7	322	16	43.7%	40.3%	47.4%
8	301	15	41.6%	38.2%	45.2%
9	283	12	39.8%	36.4%	43.5%
10	269	12	38.0%	34.7%	41.7%
11	253	9	36.7%	33.3%	40.3%
12	237	9	35.3%	32.0%	38.9%
13	226	7	34.2%	30.9%	37.8%
14	212	10	32.6%	29.3%	36.2%
15	200	7	31.4%	28.2%	35.0%
16	193	2	31.1%	27.9%	34.7%
17	190	5	30.3%	27.1%	33.9%
18	182	4	29.6%	26.4%	33.2%
19	176	6	28.6%	25.4%	32.2%
20	170	3	28.1%	24.9%	31.7%
21	166	5	27.3%	24.1%	30.8%
22	157	7	26.0%	22.9%	29.6%
23	146	3	25.5%	22.4%	29.0%
24	143	2	25.2%	22.1%	28.7%
25	141	1	25.0%	21.9%	28.5%
26	138	2	24.6%	21.6%	28.1%
27	134	1	24.4%	21.4%	27.9%
29	130	2	24.1%	21.0%	27.5%
31	128	1	23.9%	20.8%	27.3%
32	127	1	23.7%	20.6%	27.2%
33	123	2	23.3%	20.3%	26.8%
38	118	1	23.1%	20.1%	26.6%
39	116	1	22.9%	19.9%	26.4%
40	114	1	22.7%	19.7%	26.2%
41	112	1	22.5%	19.5%	26.0%
43	109	1	22.3%	19.3%	25.7%
44	108	2	21.9%	18.9%	25.3%
46	106	2	21.5%	18.5%	24.9%
47	102	1	21.2%	18.3%	24.7%
50	99	1	21.0%	18.1%	24.5%
53	98	1	20.8%	17.9%	24.3%
55	93	1	20.6%	17.7%	24.0%
60	87	2	20.1%	17.2%	23.6%
67	80	1	19.9%	16.9%	23.3%
71	77	1	19.6%	16.7%	23.1%
75	75	1	19.4%	16.4%	22.8%
82	69	1	19.1%	16.2%	22.5%
94	60	1	18.8%	15.8%	22.2%
95	59	1	18.4%	15.5%	21.9%
96	58	1	18.1%	15.2%	21.6%
111	47	1	17.7%	14.8%	21.2%
144	31	1	17.2%	14.2%	20.8%
145	30	1	16.6%	13.5%	20.3%
146	29	1	16.0%	12.9%	19.8%
154	24	1	15.3%	12.2%	19.3%

CI, confidence interval; N.Risk, the number of people at risk; N.Event, the number of observed events.