

**Table S1** The selecting code from the SEER database

Inclusion criteria and exclusion criteria	Program selecting codes
Included cases of liver cancer from SEER data base	Included {Site recode ICD-O-3/WHO 2008} = 'C22.0'
Included patients with hepatocellular carcinoma	Included {Histologic Type ICD-O-3} = '8170', '8171', '8172', '8173', '8174', '8175'
Included liver cancer as the first primary malignant tumor	Included {Sequence number} = 'one primary only' or '1st of 2 or more primaries'
Included patients with hepatocellular carcinoma diagnosed by pathology	Included {Diagnostic confirmation} = 'Positive histology'
Included cases with definite information about intrahepatic vascular invasion	Included {CS-extension} = '100', '150', '250', '270', '350', '370', '380', '390', '400', '630', '635', '638'
Included cases with active follow-up time	Included {Type of follow-up time} = 'Active follow-up'
Excluded patients with unknown follow-up time	Excluded {Survival months flag} = 'Incomplete dates are available' or 'Not calculated' or 'Complete dates with 0 days of survival'
Excluded patients with unknown cause of death	Excluded {SEER cause-specific death classification} = 'Dead (missing/unknown COD)'
Excluded patients with unknown race	Excluded {Race recode (White, Black, Other)} = 'Unknown'
Excluded patients under 20 years	Exclude {Age recode with <1 years old} = '0-19 years'
Excluded patients with useless T stage	Excluded {Derived AJCC T, 7th ed (2010+)} = 'NA', 'T0', 'Tx'
Excluded patients with useless N stage	Excluded {Derived AJCC N, 7th ed (2010+)} = 'NA', 'Nx'
Excluded patients with unknown M stage	Excluded {Derived AJCC M, 7th ed (2010+)} = 'NA'
Excluded patients with unknown surgical approach	Excluded {RX Summ--Surg Prim Site (1998+)} = '90', '99'
Excluded patients with unknown radiation information	Excluded {Radiation recode} = 'Recommended, unknown if administered'
Excluded patients with unknown tumor size	Excluded {CS-tumor size} = '990', '991', '992', '993', '994', '995', '996', '999'