Supplementary

Table S1 Demographic and clinical characteristics of the patients enrolled in derivation and validation of clinical models

Characteristic	Derivation (n=1,114)	Validation (n=184)
Age, years	53 (45 to 59)	51 (45 to 59)
Sex, female	182 (16.3)	24 (13.0)
Diabetes mellitus	90 (8.1)	20 (10.9)
ALT, U/mL	35.0 (25.0 to 52.0)	39.0 (27.0 to 62.5)
AST, U/mL	34.0 (25.0 to 50.0)	41.0 (29.8 to 65.0)
Total bilirubin, μmol/L	14.0 (11.0 to 17.0)	14.9 (12.0 to 19.4)
Prothrombin time, sec	12 (11 to 13)	12 (12 to 13)
Albumin, g/L	42.0 (39.0 to 45.0)	41.3 (38.3 to 43.8)
AFP, ng/mL	84.5 (5.5 to 1210)	172.8 (10.8 to 1000)
CEA, ng/mL	2.0 (2.0 to 4.0)	2 (1.2 to 3.1)
HBV infection	982 (88.2)	166 (90.2)
HCV infection	10 (0.9)	4 (2.2)
Liver cirrhosis	773 (69.4)	159 (86.4)
Tumor size, cm	4.0 (3.0 to 6.0)	5.2 (3.9 to 8.0)
Tumor number		
Single	1010 (90.7)	153 (83.2)
Double or multiple	104 (9.3)	31 (16.8)
Tumor location		
Segment 1 (middle)	18 (1.6)	6 (3.3)
Segment 2–4 (left)	318 (28.5)	59 (32.1)
Segment 5–8 (right)	778 (69.8)	119 (64.7)
Vascular invasion	90 (8.1)	53 (28.8)
Lymph node metastasis	28 (2.5)	2 (1.1)
TNM stage		
I–II	966 (86.7)	128 (69.6)
III–IV	148 (13.3)	56 (30.4)

Data are median (IQR) or n (%). ALT, alanine aminotransferase; AST, aspartate aminotransferase; AFP, alpha fetoprotein; CEA, carcinoembryonic antigen; HBV, hepatitis B virus; HCV, hepatitis C virus.

http://dx.doi.org/10.21037/atm-20-4828

Table S2 Demographic and clinical characteristics of the patients enrolled in derivation and validation of clinical models

Characteristic	TCGA-LIHC (n=374)	ICGC-LIRI-JP (n=193)	GSE116174 (n=64)
Age, years	61 (52 to 69)	69 (62 to 74)	54 (49 to 62)
Sex, female	121 (32.4)	49 (25.4)	6 (9.4)
HBV infection	159 (42.5)	54 (28.0)	47 (73.4)
Alcohol consumption	72 (19.3)	115 (59.6)	13 (20.3)
Smoking	15 (4.0)	114 (59.1)	31 (48.4)
Vascular invasion	16 (4.3)	61 (31.6)	29 (45.3)
TNM stage			
I–II	260 (74.3)	123 (63.7)	53 (82.8)
III–IV	90 (25.7)	70 (36.3)	11 (17.2)

Data are median (IQR) or n (%). HBV, hepatitis B virus.

Table S3 Univariable and multivariable analyses of significant prognostic factors for the overall survival in the derivation set (EHBH, n=1,114)

Variable	Input type	Univariable, P value	Multivariable	
			HR (95% CI)	P value
ALT, U/mL	Continuous	0.007	NS	NS
AST, U/mL	Continuous	<0.001	NS	NS
Total bilirubin, μmol/L	Continuous	<0.001	1.01 (1.00–1.01)	0.001
Prothrombin time, sec	Continuous	<0.001	1.14 (1.04–1.24)	0.003
Albumin, g/L	Continuous	0.001	NS	NS
AFP, ng/mL	Continuous	0.023	NS	NS
Tumor size, cm	Continuous	<0.001	1.16 (1.12–1.19)	<0.001
Tumor number	Single *vs.* others	<0.001	1.48 (1.18–1.86)	0.001
Tumor location	Left *vs.* others	0.003	NS	NS
Vascular invasion	Present *vs.* absent	<0.001	2.22 (1.65–2.99)	<0.001
Lymph node metastasis	Present *vs.* absent	<0.001	1.99 (1.22–3.24)	0.006
TNM stage	Continuous (I–IV)	<0.001	1.48 (1.31–1.66)	<0.001

HR, hazard ratio; CI, confidence interval; ALT, alanine aminotransferase; NS, not significant (P>0;05); AST, aspartate aminotransferase; AFP, alpha fetoprotein.

http://dx.doi.org/10.21037/atm-20-4828

Table S4 Univariable and multivariable analyses of clinical factors for the overall survival in the derivation set (TCGA-LIHC, n=374)

Variable	Input type	Univariable, P value	Multivariable	
			HR (95% CI)	P value
Age, years	Continuous	0.091	NA	NA
Gender	Male *vs.* female	0.201	NA	NA
Alcohol consumption	Present *vs.* absent	0.033	0.97 (0.52–1.79)	0.921
Smoking	Present *vs.* absent	0.756	NA	NA
HBV infection	Present *vs.* absent	0.465	NA	NA
Vascular invasion	Present *vs.* absent	0.040	1.67 (0.79–3.53)	0.183
TNM stage	Continuous (I–IV)	<0.001	1.52 (1.19–1.94)	0.001

HR, hazard ratio; CI, confidence interval; NA, not applicable; HBV, hepatitis B virus.

http://dx.doi.org/10.21037/atm-20-4828