

Supplementary

Table S1 R packages

R package	Source	Function
table1	CRAN Repository	table1
randomForestSRC	CRAN Repository	rfsrc.fast, predict
survminer	CRAN Repository	survfit, survival
survival	CRAN Repository	coxph
timeROC	CRAN Repository	timeROC, plot
Hmisc	CRAN Repository	rccorrcens
CPE	CRAN Repository	phcpe
stdca	MSKCC website	stdca
pec	CRAN Repository	pec, ibs, R2
rms	CRAN Repository	cph, calibrate
nricens	CRAN Repository	nricens
survIDINRI	CRAN Repository	IDI.INF, IDI.INF.OUT
shiny	CRAN Repository	EventReactive, shinyServer

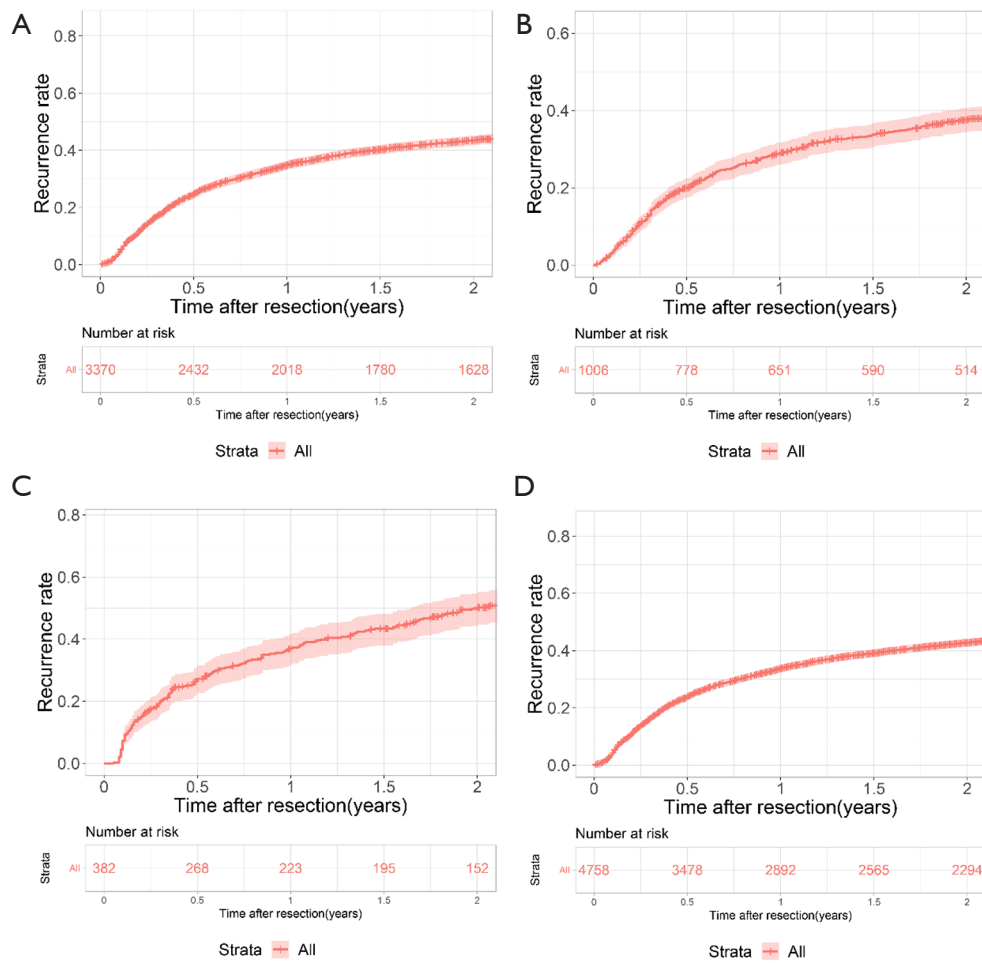


Figure S1 Kaplan-Meier curves for early recurrence. (A) Training cohort, (B) internal validation cohort, (C) external validation cohort, (D) the entire cohort.

Table S2 Risk reclassification table for all patients in predicting early recurrence probabilities by the RSF and ERASL model

ERASL model	RSF model			Total
	<50% risk	50–85% risk	≥85% risk	
Patients with early recurrence (n=1,935)				
<50% risk	592	103	23	718
50–85% risk	207	423	320	950
≥85% risk	0	53	214	267
Total	799	579	557	1,935
Patients without early recurrence (n=2,287)*				
<50% risk	1,756	73	2	1,831
50–85% risk	180	214	33	427
≥85% risk	0	17	12	29
Total	1,936	304	47	2,287

*, lost to follow-up or dead within 2 years of surgery were excluded. RSF, random survival forests; ERASL, early recurrence after surgery for liver tumor.

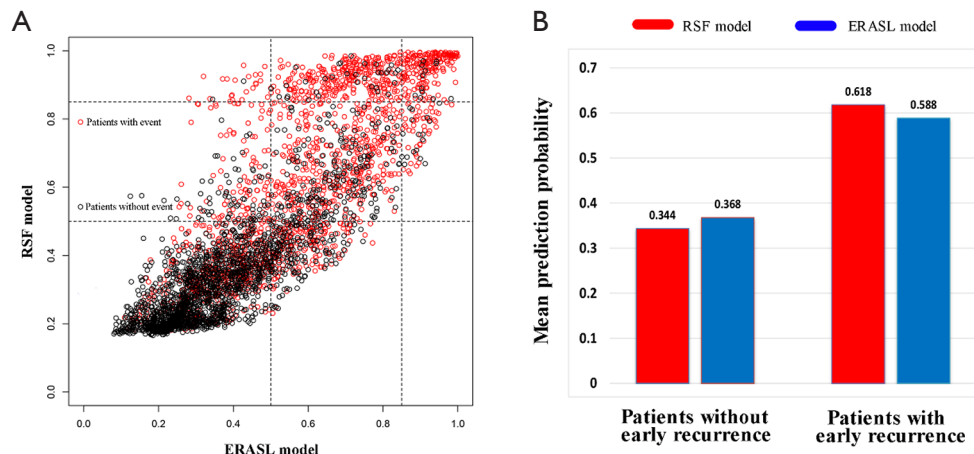


Figure S2 NRI and IDI. (A) Scatter plot for the entire cohort in predicting early recurrence probabilities by the RSF and ERASL models. (Red circle, early recurrence; black circle, without early recurrence). (B) Histogram for the entire cohort with early recurrence and without early recurrence in mean prediction probability by the RSF model (red) and the ERASL model (blue). NRI, net reclassification improvement; IDI, integrated discrimination improvement; RSF, random survival forests; ERASL, early recurrence after surgery for liver tumor.

Table S3 2-year recurrence rates, hazard ratio and P-value according to the three risk groups as defined by the RSF model

Cohort	Risk groups	n	2-year recurrence rate, % (95% CI)	Hazard ratio (95% CI)	P value
Training	Low	1,685	22.5 (20.5, 24.5)	1	
	Intermediate	1,179	58.7 (55.6, 61.5)	2.527 (2.423, 2.632)	<0.0001
	High	506	86.5 (82.1, 89.8)	6.699 (6.566, 6.831)	<0.0001
Internal	Low	484	17.2 (13.7, 20.7)	1	
	Intermediate	402	48.3 (42.9, 53.1)	3.084 (2.851, 3.318)	<0.0001
	High	120	85.2 (76.6, 90.6)	11.655 (11.371, 11.938)	<0.0001
External	Low	158	24.0 (16.7, 30.6)	1	
	Intermediate	157	61.4 (52.6, 68.6)	3.366 (3.012, 3.720)	<0.0001
	High	67	86.7 (74.0, 93.2)	7.712 (7.308, 8.115)	<0.0001
Entire	Low	2,327	21.6 (19.8, 23.2)	1	
	Intermediate	1,738	56.5 (54.0, 58.9)	2.638 (2.547, 2.730)	<0.0001
	High	693	86.2 (82.7, 89.0)	7.423 (7.309, 7.537)	<0.0001

RSF, random survival forests; CI, confidence interval.

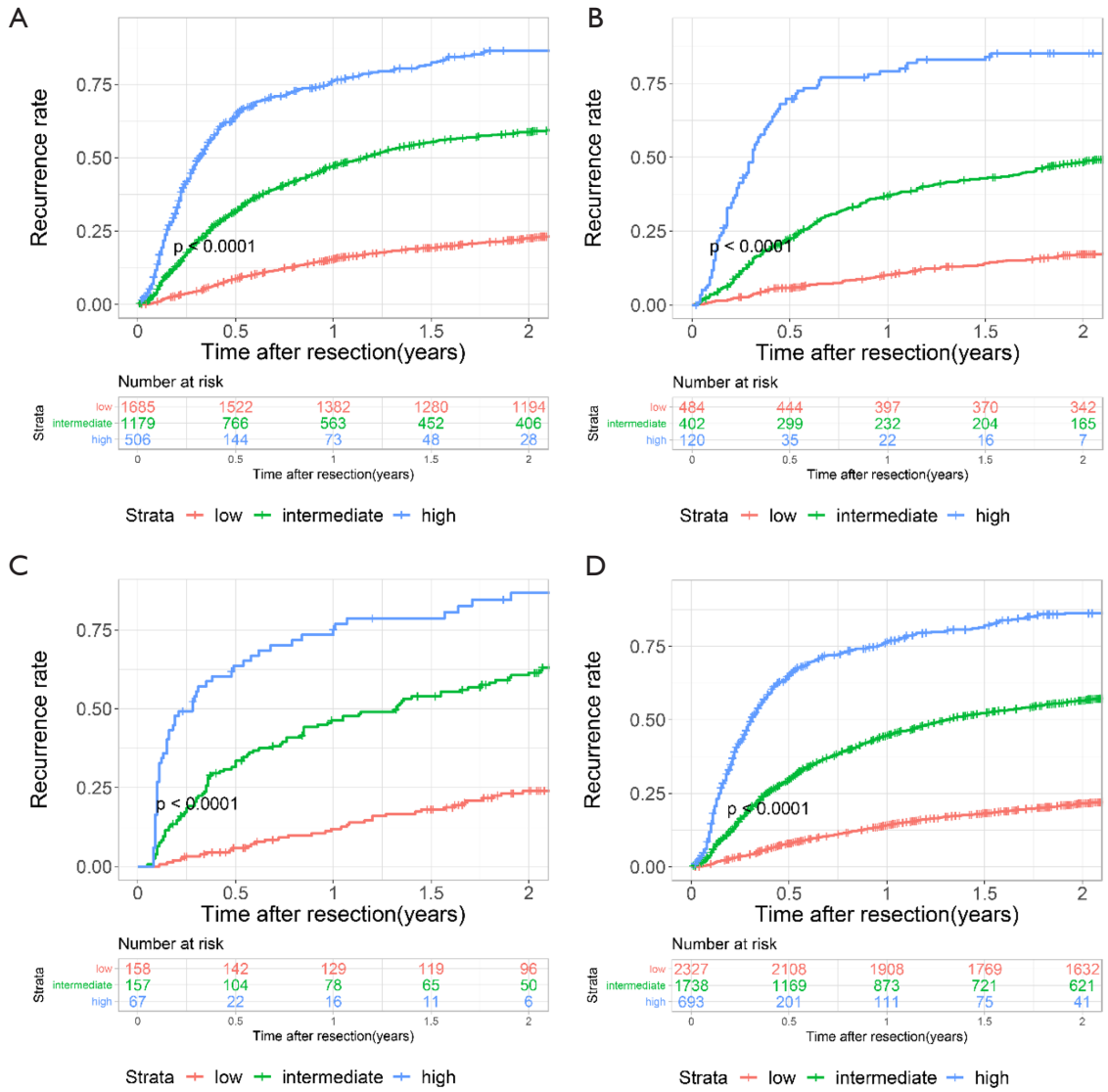


Figure S3 Kaplan-Meier curves for early recurrence in the low-risk, intermediate-risk, and high-risk groups defined by the RSF model. (A) Training cohort, (B) internal validation cohort, (C) external validation cohort, (D) the entire cohort.

RSF model: Early Recurrence Prediction for patients undergoing R0 liver resection

Age(years) :

Gender:

Etiology:

PLT(10E9/L) :

ALB(g/L) :

TBIL(μmol/L) :

AFP(ng/mL) :

Tumor size(cm) :

Tumor number :

Microvascular invasion :

Macrovascular invasion :

Edmondson grade :

Tumor capsular :

Satellite nodules :

Liver cirrhosis :

<<< RSF-model score: 71.39 >>>

<<< Risk-group: high-risk >>>

Recurrence-free probability within 2 years

Month 3	Month 6	Month 9	Month 12	Month 18	Month 24
0.66	0.44	0.33	0.26	0.18	0.14

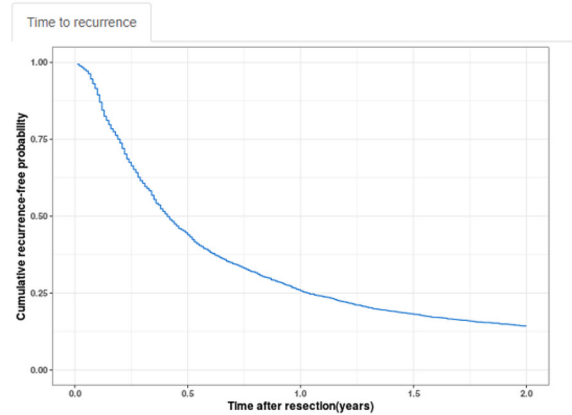


Figure S4 Screenshot of the online prediction tool based on the RSF model.