Supplemental Online Content

Famularo S, Donadon M, Cipriani F, et al; the HE.RC.O.LE.S. Group. Machine learning predictive model to guide treatment allocation for recurrent hepatocellular carcinoma after surgery. *JAMA Surg*. Published online December 28, 2022. doi:10.1001/jamasurg.2022.6697

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods 1: variable definitions.

Presence of cirrhosis was evaluated prior to treatment by hepatologists and confirmed by pathology. The severity of cirrhosis was measured by Child-Pugh Score. Number, size and localization of recurrent nodules were estimated by CT or MRI scan at the time of recurrence diagnosis. Extra-hepatic relapse was considered in any case a lesion was evident outside the liver (e.g. local lymphnodes, right kidney etc).

eMethods 2: Statistical analysis in detail.

Formal sample size calculation was not performed because data were taken from a nationwide registry and all available information of patients meeting inclusion and exclusion criteria was retrieved for the analysis.

The problem of missing values was tackled using Multiple Imputation by Chained Equations (MICE) method (van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations inR. *Journal of Statistical Software.*;45 . Epub ahead of print 2011. DOI: 10.18637/jss.v045.i03). We imputed 20 datasets using predictive mean matching for continuous variables, logistic regression for binary and multinomial regression for categorical variables with more than two levels. Candidate features to be included as independent variables in the prediction model were included in the imputation model with the further addition of the completely observed treatment variable and bivariate survival outcome (death indicator and log-transformed follow-up time) (White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med.* 2009;28:1982–1998).

To ease the selection of predictive features and the development of the predictive model we averaged the imputed datasets (mean values across imputations for numerical values and mode for categorical ones) to create a single complete dataset.

Selection of features that play a role as prognostic factors for SAR but also as treatment modifiers was performed by fitting a Cox model with LASSO (Least Absolute Shrinkage and Selection Operator) penalization (Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. 1997;16:385–395). Ten-fold cross-validation was used to choose the optimal value of the shrinkage parameter "lambda" (Simon N, Friedman J, Hastie T, et al. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *J Stat Softw*. 2011;39:1–13.). Continuous variables were modelled flexibly using fractional polynomials (Sauerbrei W, Meier-Hirmer C, Benner A, et al. Multivariable regression model building by using fractional polynomials: Description of SAS, STATA and R programs. *Comput Stat Data Anal*. 2006;50:3464–3485).

All features that were retained by this procedure (coefficient different from zero at the optimal lambda) were used in the subsequent development of the prediction model. These features were: age at recurrence (years), presence of cirrhosis, number of recurrent nodules, size of the biggest recurrent nodule (mm), single or bilobar recurrence, intra or extra hepatic recurrence and time from first surgery to recurrence (months). Univariate Cox models were fitted to show the prognostic effect of these 7 selected variables (number of recurrent nodules and size of the biggest recurrent nodule were dichotomized using appropriate cut-offs). Moreover, in order to illustrate the possible role of these variables as treatment effect-modifiers we evaluated the association between treatment and SAR by univariate Cox model fitted on the subgroups defined by the levels of the selected features

(again, number of recurrent nodules and size of the biggest recurrent nodule were dichotomized using appropriate cutoffs).

The prediction model was built as a standard (not penalized) Cox model with all second order interactions of treatment with the features selected by LASSO. Again, continuous variables were modelled flexibly using fractional polynomials. The performance of the model in terms of discrimination was evaluated drawing Receiver Operating Characteristics (ROC) curves at fixed time points (3 and 5 years) and computing the Area under the Curve (AUC) index, also using bootstrap to correct for over-optimism (Gerds TA, Schumacher M. Efron-type measures of prediction error for survival analysis. *Biometrics*. 2007;63:1283–1287). Calibration plots and the Brier score index were used to check the calibration of the model (Gerds TA, Schumacher M. Consistent estimation of the expected Brier score in general survival models with right-censored event times. *Biom J.* 2006;48:1029–1040).

An external validation of the model was also performed using data of patients with HCC recurrence after surgery taken from another national cohort (ITALICA) and from a Japanese cohort (Tokyo University Hospital). In each external cohort, the following procedure was applied. First, the prognostic index of the model was calculated using data of the validation set and the variable coefficients estimated on the derivation set. A recalibration was then performed by fitting a Cox model on the validation set including the prognostic index as the only covariate (Royston and Altman: External validation of a Cox prognostic model: principles and methods. *BMC Medical Research Methodology* 2013;13:33). Form this model, time-dependent ROC curves at 3 and 5 years were created

and the AUC index was calculated to obtain the discrimination level of the model in each external validation set.

For each patient, the model estimates were used to predict the potential SAR at 3 and 5 years under each treatment. Subsequently, the potentially optimal treatment within patient was determined as the one leading to the highest predicted SAR.

To describe the characteristics of the patients who most benefit from each treatment, the distribution of the features among groups of optimal treatment was compared.

Considering the possible combinations of the 7 features included, the model was able to deal with potentially infinite different patients risk profiles, and consequently it was built up in an R Shiny web application to let users calculate the potential SAR in time under each treatment for every profile of interest.

Moreover, to create a simple algorithm guiding the choice of the best potential treatment we also re-fitted the same prediction model described before but using the dichotomized features number of recurrent nodules=1 or >1 and size of the biggest recurrent nodule<50 mm, instead of their continuous version. Then, in order to show the best potential treatment according to each risk profile, we considered only the 5 features showing the largest impact as treatment effect-modifiers (age< 75 or ≥75, cirrhosis yes/no, number of recurrent nodules=1 or >1, single/bilobar recurrence and intra/extra hepatic recurrence) and we created 2^5=32 risk profiles from all the possible combinations of the levels of these features (the remaining two features were fixed at: size of the biggest recurrent nodule<50 mm, time from first surgery to recurrence=15 months). We then predicted the potential SAR at 3 years under each treatment, according to the model, for all the risk profiles, and we defined the optimal treatment as the one leading to the highest SAR. When the

difference in SAR between two treatments is less than 5%, both treatments were labelled as optimal. A tree-diagram based on these results was drawn to express a possible algorithm of optimal treatment choice. The same process (predicted SAR for all risk profiles, comparison of SAR under each treatment and tree-diagram drawing) was repeated for the potential SAR at 5 years.

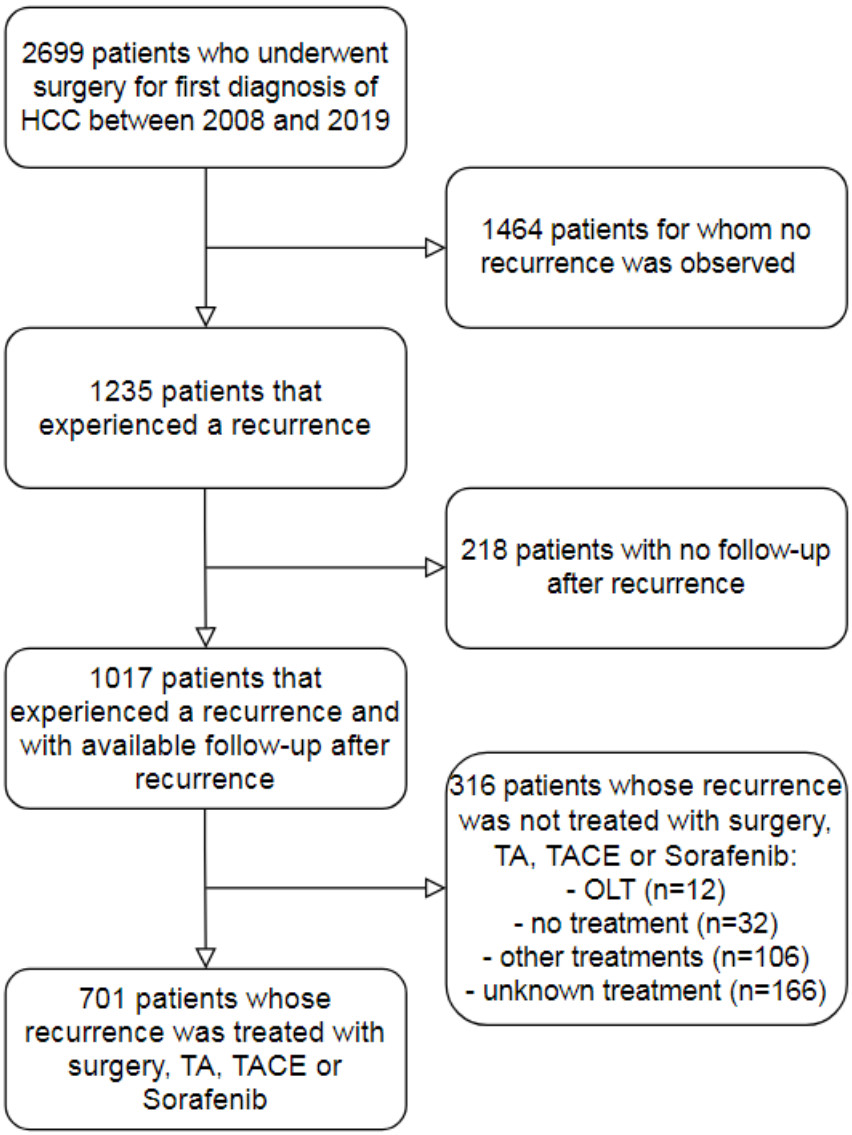
Finally, considering again the best potential treatment for each patient based on the SAR predicted by the model with the dichotomized version of number of recurrent nodules and size of the biggest recurrent nodule, an alluvial plot was drawn to illustrate the typical profile of the patients who most benefit from each treatment.

All the analyses were carried out using R (version 4.0.3).

eMethods 3. Patients excluded.

Two hundred and eighteen patients were excluded because no information on the follow-up after recurrence was available. The characteristics of these patients were similar to those included in the final analysis (median age 73 years, 23% females, 69% with cirrhosis, median number of nodules 2, median size 2 cm, 35% treated with CUR, 41% with SOR and 24% with TACE). Three hundred and sixteen patients were not treated with CUR, SOR or TACE after recurrence.

eFigure 1. Flow-chart depicting the enrolment process from the original dataset.

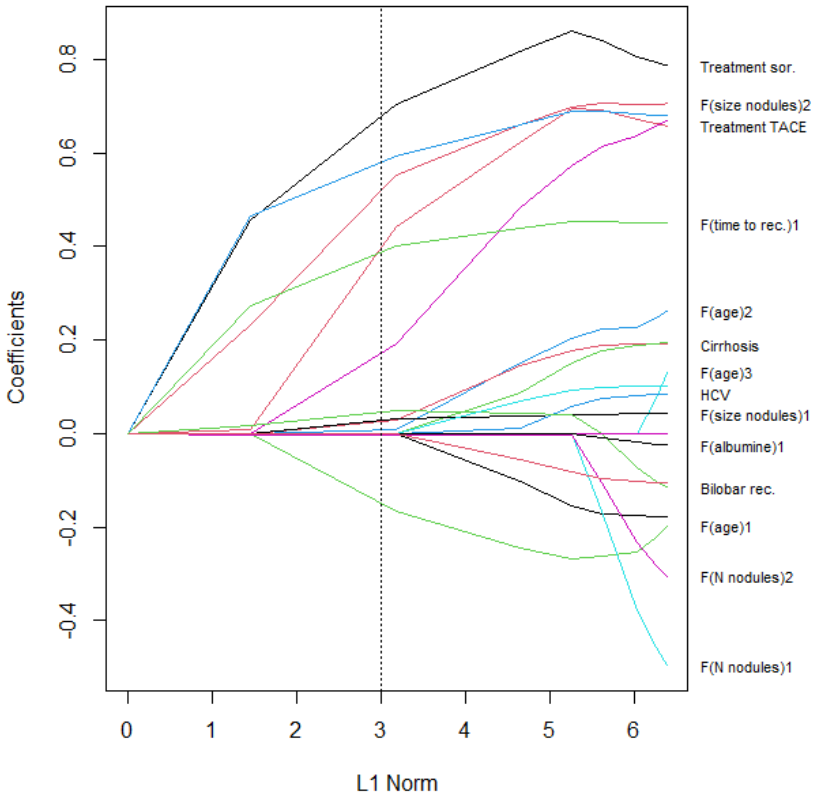


eFigure 2A-B. Lasso model for variable selection. Variables initially included in the model are: Treatment, Age, gender, Cirrhosis, HBV, HCV, n nodules, size nodules, bilobar recurrence, extra-hepatic recurrence, microvascular invasion, platelets, albumin, bilirubin, time to recurrence**.** Variables selected are: Treatment, Age, Cirrhosis, n nodules, size nodules, bilobar recurrence, extra-hepatic recurrence, time to recurrence.

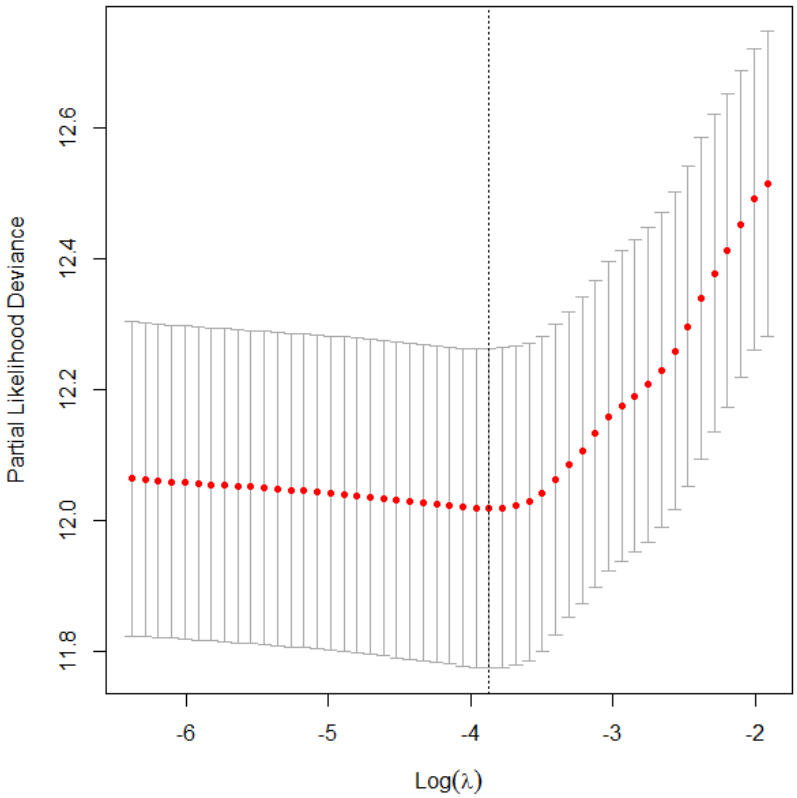
A) paths of the coefficients as a function of the L1 norm

B) cross-validated deviance curve as a function of the shrinkage parameter lambda

A:



B:



eTable 1. Univariate Cox model to evaluate the association of features with SAR

Features	**HR (95%CI)**	**p-value**
Treatment sorafenib vs curative	3.678 (2.664;5.078)	<0.001
Treatment TACE vs curative	1.995 (1.431;2.782)	<0.001
Age > 75 vs ≤75	1.231 (0.938;1.617)	0.134
Cirrhosis yes vs no	1.109 (0.838;1.467)	0.469
N. rec. nodules > 1 vs 1	2.151 (1.638;2.826)	<0.001
Size rec. nodules ≥5 vs <5	2.892 (2.145;3.900)	<0.001
Bilobar rec. yes vs no	1.713 (1.318;2.227)	0.001
Extra-hepatic rec. yes vs no	2.946 (2.187;3.968)	<0.001
F(time to recurrence)	1.804 (1.450;2.244)	<0.001
Portal hypertension yes vs no	1.186 (0.800;1.757)	0.396
Charlson Comorbidity Index, per point	1.001 (0.919;1.091)	0.978
ASA score 3-4 vs 1-2	1.041 (0.782;1.385)	0.784
MELD score, per point	0.991 (0.919;1.070)	0.826
First surgery major vs minor	1.267 (0.929;1.729)	0.135
Log(total bilirubin), per mg/dL point	0.794 (0.645;0.977)	0.029
Albumin, per g/dL	0.862 (0.667;1.116)	0.260
Platelets, per 10^3/ml	1.001 (0.999;1.002)	0.440
Edmonson grading of the first tumor 3-4 vs 1-2	1.616 (1.232;2.119)	0.001
Child Pugh A vs no cirrhosis	1.096 (0.833;1.441)	0.513
Child Pugh B vs no cirrhosis	0.715 (0.330;1.549)	0.394
Micro vascular invasion vs no	1.209 (0.915;1.597)	0.182

F: function

eTable 2. Variables considered for the selection based on the Cox model with LASSO penalization. Those with coefficient different from 0 were retained to be included in the prediction model (continuous variables modeled with spline functions were retained if at least one spline base had a coefficient other than 0).

Features	Coefficient
Treatment sorafenib vs curative	0.689
Treatment TACE vs curative	0.416
F(age)1	-0.154
F(age)2	0
F(age)3	0
F(age)4	0.131
Gender	0
Cirrhosis yes vs no	0.015
HBV yes vs no	0
HCV yes vs no	0
F(N nodules)1	0
F(N nodules)2	-0.004
F(size nodules)1	0.03
F(size nodules)2	0.535
Bilobar rec. yes vs no	0.052
Extra-hepatic rec. yes vs no	0.584
Microvascular invasion vs no	0
F(platelets)1	0
F(albumine)1	0
F(bilirubin)1	0
F(time to rec.)1	0.392

F: function

Features	Coefficient
Treatment sorafenib vs curative	0.577289
Treatment TACE vs curative	0.259759
F(age)1	-0.15613
F(age)2	0
F(age)3	0
F(age)4	0.043637
Gender	0
Cirrhosis yes vs no	0.012399
HBV yes vs no	0
HCV yes vs no	0
F(N nodules)1	0
F(N nodules)2	-0.80138
F(size nodules)1	0.027748
F(size nodules)2	0.453399

Bilobar rec. yes vs no	0.069202
Extra-hepatic rec. yes vs no	0.570956
Microvascular invasion vs no	0
F(platelets)	0
F(albumine)	0
F(bilirubin)	0
F(time to rec.)	0.34842
Major vs minor surgery	0
F(MELD score)	0
Portal hypertension yes vs no	0
ASA score 3-4 vs 1-2	0
F(CCI)	0

eTable 3. Description of patients of the ITALICA validation set.

Variables	Overall N=295	Treatment (%)			p
		curative 121 (41.0)	sorafenib 54 (18.3)	TACE 120 (40.7)	
Age (median [IQR])	68.75 [59.39, 73.88]	69.43 [61.55, 73.90]	66.62 [56.53, 73.74]	68.53 [58.83, 73.74]	0.558
Sex M (%)	241 (81.7)	94 (77.7)	48 (88.9)	99 (82.5)	0.2
HBV (%)	47 (17.6)	18 (16.4)	10 (22.2)	19 (17.0)	0.667
HCV (%)	135 (49.3)	51 (45.1)	24 (50.0)	60 (53.1)	0.485
Cirrhosis (%)	238 (81.5)	96 (79.3)	36 (69.2)	106 (89.1)	0.006
Child Grade B (%)	23 (10.0)	7 (6.6)	3 (8.1)	13 (14.8)	0.154
BCLC first diagnosis (%)					0.003
0	22 (8.2)	10 (9.1)	0 (0.0)	12 (11.0)	
A	152 (56.7)	65 (59.1)	21 (42.9)	66 (60.6)	
B	63 (23.5)	23 (20.9)	15 (30.6)	25 (22.9)	
C	30 (11.2)	11 (10.0)	13 (26.5)	6 (5.5)	
D	1 (0.4)	1 (0.9)	0 (0.0)	0 (0.0)	
Major hepatectomy (%)	35 (13.6)	9 (8.3)	9 (22.5)	17 (15.6)	0.061
Number of recurrent nodules (median [IQR])	1.00 [1.00, 3.00]	1.00 [1.00, 2.00]	2.00 [1.00, 4.00]	2.00 [1.00, 3.00]	<0.001
Size of recurrent nodules cm (median [IQR])	1.80 [1.50, 2.70]	1.80 [1.50, 2.30]	2.80 [1.70, 4.00]	1.80 [1.30, 2.62]	0.003
Bilobar recurrence (%)	48 (19.3)	16 (14.7)	11 (33.3)	21 (19.6)	0.058
Localization of Recurrence (%)					<0.001
intrahepatic	231 (89.9)	98 (91.6)	27 (64.3)	106 (98.1)	
extrahepatic	12 (4.7)	5 (4.7)	6 (14.3)	1 (0.9)	
both	14 (5.4)	4 (3.7)	9 (21.4)	1 (0.9)	
Time from surgery to recurrence months (median [IQR])	492 [213, 932]	624 [365, 1104]	289 [123, 668.50]	471 [184, 856.75]	<0.001
Death (%)	143 (48.5)	49 (40.5)	37 (68.5)	57 (47.5)	0.003
Time from recurrence to death/end fup months (median [IQR])	25.64 [10.95, 46.28]	33.25 [15.02, 56.79]	12.23 [4.27, 25.13]	28.89 [16.73, 50.15]	<0.001

eTable 4. Description of patients of the Tokyo University Hospital validation set.

Variables	Overall N=422	Treatment (%)			p
		curative 254 (60.2)	sorafenib 5 (1.2)	TACE 163 (38.6)	
Age (median [IQR])	67 [58, 73]	66 [57, 72]	61 [58, 63]	69 [60.5, 74]	0.003
Sex M (%)	322 (76.3)	197 (77.6)	5 (100.0)	120 (73.6)	0.298
HBV (%)	90 (21.3)	54 (21.3)	2 (40.0)	34 (20.9)	0.588
HCV (%)	249 (59.0)	146 (57.5)	3 (60.0)	100 (61.3)	0.735
Cirrhosis (%)	272 (64.5)	167 (65.7)	5 (100.0)	100 (61.3)	0.163
Child Grade B (%)	50 (11.8)	24 (9.4)	0 (0.0)	26 (16.0)	0.095
BCLC first diagnosis (%)					<0.001
0	52 (12.3)	39 (15.4)	2 (40.0)	11 (6.7)	
A	255 (60.4)	165 (65.0)	0 (0.0)	90 (55.2)	
B	81 (19.2)	39 (15.4)	1 (20.0)	41 (25.2)	
C	34 (8.1)	11 (4.3)	2 (40.0)	21 (12.9)	
Major hepatectomy (%)	74 (17.5)	40 (15.7)	2 (40.0)	32 (19.6)	0.246
Non anatomic resection (%)	254 (60.2)	171 (67.3)	3 (60.0)	80 (49.1)	0.001
Microvascular invasion (%)	105 (24.9)	60 (23.6)	0 (0.0)	45 (27.6)	0.284
Number of recurrent nodules (median [IQR])	2.00 [1.00, 3.00]	1.00 [1.00, 2.00]	3.00 [2.50, 3.50]	3.00 [1.00, 5.00]	<0.001
Size of recurrent nodules cm (median [IQR])	1.5 [1.0, 2.0]	1.5 [1.0, 2.0]	1.0 [1.0, 1.0]	1.5 [1.0, 2.0]	0.312
Bilobar recurrence (%)					<0.001
No	153 (36.3)	97 (38.2)	1 (20.0)	55 (33.7)	
Yes	252 (59.7)	146 (57.5)	1 (20.0)	105 (64.4)	
Not applicable	17 (4.0)	11 (4.3)	3 (60.0)	3 (1.8)	
Localization of Recurrence (%)					<0.001
intrahepatic	394 (93.4)	239 (94.1)	0 (0.0)	155 (95.1)	
extrahepatic	19 (4.5)	12 (4.7)	4 (80.0)	3 (1.8)	
both	9 (2.1)	3 (1.2)	1 (20.0)	5 (3.1)	
Time from surgery to recurrence months (median [IQR])	15.84 [6.69, 35.71]	20.90 [9.85, 39.93]	6.02 [2.73, 17.17]	10.82 [3.90, 24.18]	<0.001
Death (%)	200 (47.4)	89 (35.0)	1 (20.0)	110 (67.5)	<0.001
Time from recurrence to death/end fup months (median [IQR])	30.61 [14.42, 54.56]	39.01 [22.19, 60.93]	9.80 [3.85, 10.33]	22.89 [7.19, 37.45]	<0.001

eTable 5. Comparison of the characteristics of patients among treatments in the ITALICA (left part) and Tokyo University Hospital (right part) validation sets. Mean difference (for continuous variables) or RR (for categorical variables) with 95% confidence intervals are reported.

Variables	ITALICA validation set			Tokyo University Hospital validation set		
	Sorafenib vs Curative	TACE vs Sorafenib	TACE vs Curative	Sorafenib vs Curative	TACE vs Sorafenib	TACE vs Curative
Age (median [IQR])	-2.54 (-5.93;0.84)	1.76 (-1.78;5.30)	-0.79 (-3.45;1.88)	-6.50 (-15.86;2.85)	9.99 (1.39;18.59)	3.49 (1.41;5.57)
Sex M (%)	1.14 (1.00;1.31)	0.93 (0.82;1.05)	1.06 (0.94;1.20)	1.29 (1.21;1.38)	0.74 (0.67;0.81)	0.95 (0.85;1.06)
HBV (%)	1.36 (0.68;2.71)	0.76 (0.39;1.51)	1.04 (0.58;1.87)	1.88 (0.63;5.65)	0.52 (0.17;1.59)	0.98 (0.67;1.44)
HCV (%)	1.11 (0.78;1.57)	1.06 (0.76;1.48)	1.18 (0.90;1.54)	1.04 (0.51;2.15)	1.02 (0.49;2.11)	1.07 (0.91;1.25)
Cirrhosis (%)	0.87 (0.71;1.07)	1.29 (1.06;1.56)	1.12 (1.01;1.25)	1.52 (1.39;1.66)	0.61 (0.54;0.69)	0.93 (0.80;1.08)
Child Grade B (%)	1.23 (0.33;4.50)	1.82 (0.55;6.02)	2.24 (0.93;5.36)	-	-	1.69 (1.00;0.06)
BCLC first diagnosis B (%)	1.46 (0.84;2.55)	0.75 (0.43;1.29)	1.10 (0.66;1.81)	1.30 (0.22;7.70)	1.26 (0.21;7.41)	1.64 (1.11;2.42)
BCLC first diagnosis C (%)	2.65 (1.28;5.50)	0.21 (0.08;0.51)	0.55 (0.21;1.44)	9.24 (2.73;31.26)	0.32 (0.10;1.01)	2.97 (1.47;6.01)
Major hepatectomy (%)	2.70 (1.15;6.32)	0.69 (0.34;1.43)	1.87 (0.87;4.01)	2.54 (0.84;7.71)	0.49 (0.16;1.50)	1.25 (0.82;1.90)
Number of recurrent nodules (median [IQR])	1.65 (0.83;2.46)	-0.83 (-1.79;0.14)	0.82 (0.31;1.33)	1.07 (-4.61;6.74)	2.13 (-6.08;10.34)	3.20 (2.39;4.00)
Size of recurrent nodules cm (median [IQR])	1.11 (0.61;1.62)	-1.14 (-1.67;-0.60)	-0.02 (-0.38;0.33)	-0.79 (-2.41;0.83)	0.82 (-0.90;2.53)	0.03 (-0.20;0.26)
Bilobar recurrence (%)	2.27 (1.17;4.40)	0.59 (0.32;1.09)	1.34 (0.74;2.42)	0.83 (0.21;3.34)	1.31 (0.33;5.27)	1.09 (0.94;1.27)
Localization of Recurrence extrahepatic or both (%)	4.25 (2.02;8.95)	0.05 (0.01;0.22)	0.22 (0.05;1.00)	16.93 (10.36;27.67)	0.05 (0.02;0.10)	0.83 (0.36;1.92)

Time from surgery to recurrence months (median [IQR])	-17.08 (-25.41;-8.79)	9.44 (2.46;16.43)	-7.64 (-14.20;-1.11)	-12.81 (-33.70;8.07)	4.68 (-16.30;25.66)	-8.13 (-12.77;-3.49)

eTable 6. Comparison of derivation set with ITALICA and Tokyo University Hospital validation sets.

Variables	Derivation set N=701	Validation set ITALICA N=295	ITALICA vs derivation set Mean difference or RR (95%CI)	Validation set Tokyo University Hospital N=422	Tokyo University Hospital vs derivation set Mean difference or RR (95%CI)
Treatment					
Curative	293 (41.8)	121 (41.0)	0.98 (0.83;1.15)	254 (60.2)	1.44 (1.28;1.62)
Sorafenib	188 (26.8)	54 (18.3)	0.68 (0.52;0.89)	5 (1.2)	0.03 (0.14;0.08)
TACE	220 (31.4)	120 (40.7)	1.30 (1.09;1.55)	163 (38.6)	1.23 (1.04;1.44)
Age (median [IQR])	72.55 [65.72, 76.89]	68.75 [59.39, 73.88]	-4.12 (-5.51;-2.73)	67 [58, 73]	-5.59 (-6.82;-4.35)
Sex M (%)	550 (78.5)	241 (81.7)	1.04 (0.97;1.11)	322 (76.3)	0.97 (0.91;1.04)
HBV (%)	133 (19.3)	47 (17.6)	0.91 (0.67;1.23)	90 (21.3)	1.10 (0.87;1.40)
HCV (%)	321 (46.5)	135 (49.3)	1.06 (0.92;1.22)	249 (59.0)	1.27 (1.13;1.42)
Cirrhosis (%)	457 (65.8)	238 (81.5)	1.23 (1.15;1.34)	272 (64.5)	0.98 (0.90;1.07)
Child Grade B (%)	28 (6.3)	23 (10.0)	1.57 (0.93;2.67)	50 (11.8)	1.87 (1.20;2.91)
BCLC first diagnosis B (%)	98 (21.5)	63 (23.5)	1.09 (0.83;1.44)	81 (19.2)	0.98 (0.69;1.16)
BCLC first diagnosis C (%)	72 (15.8)	31 (11.6)	0.71 (0.48;1.06)	34 (8.1)	0.51 (0.35;0.75)
Major hepatectomy (%)	138 (21.5)	35 (13.6)	0.64 (0.45;0.89)	74 (17.5)	0.82 (0.63;1.06)
Number of recurrent nodules (median [IQR])	1.00 [1.00, 3.00]	1.00 [1.00, 3.00]	-0.30 (-0.62;0.02)	2.00 [1.00, 3.00]	0.72 (0.25;1.18)
Size of recurrent nodules cm (median [IQR])	2.00 [1.50, 3.00]	1.80 [1.50, 2.70]	-1.26 (-1.87;-0.65)	1.5 [1.0, 2.0]	-1.72 (-2.32;-1.13)

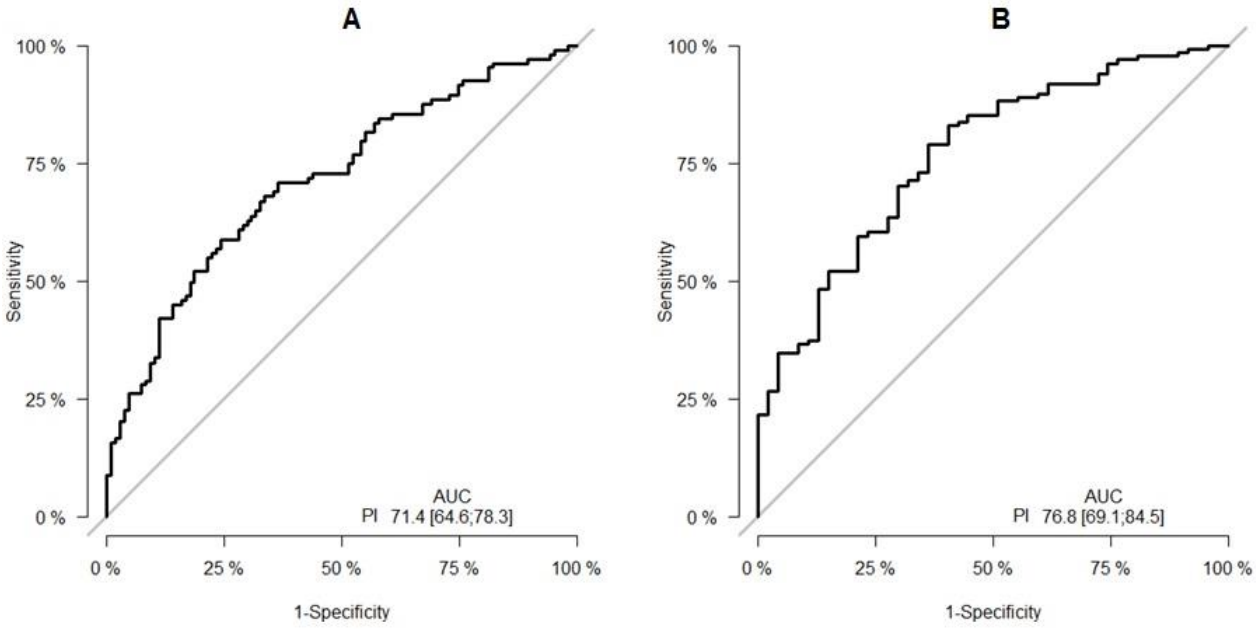
Bilobar recurrence (%)	148 (30.0)	48 (19.3)	0.64 (0.48;0.86)	252 (59.7)	2.07 (1.78;2.42)
Localization of Recurrence extrahepatic or both (%)	100 (14.6)	26 (10.1)	0.69 (0.46;1.04)	28 (6.6)	0.45 (0.30;0.68)
Time from surgery to recurrence months (median [IQR])	15.02 [6.53, 29.49]	16.13 [6.98, 30.56]	3.06 (-0.31;6.43)	15.84 [6.69, 35.71]	3.55 (0.84;6.26)
Death (%)	234 (33.4)	143 (48.5)	2.02 (1.78;2.28)	200 (47.4)	1.42 (1.23;1.64)
Time from recurrence to death/end fup months (median [IQR])	16.36 [7.02, 33.90]	25.64 [10.95, 46.28]	9.54 (5.93;13.15)	30.61 [14.42, 54.56]	14.72 (11.32;18.13)

eTable 7. Comparison of features among groups of treatment actually received (on the right), and among groups of Best Potential Treatments (i.e. the treatment leading to the highest SAR for each patient; on the left). Mean difference (for continuous variables) or RR (for categorical variables) with 95% confidence intervals are reported.

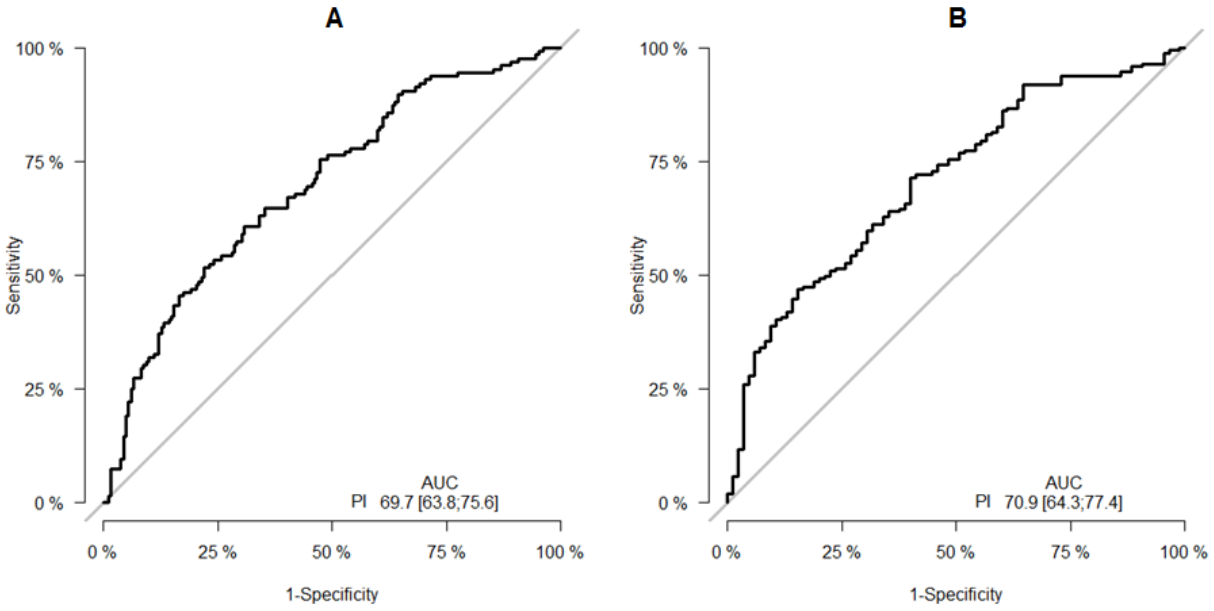
	Treatment Received			**Best Potential Treatment**		
Features	**Sorafenib vs Curative**	**TACE vs Sorafenib**	**TACE vs Curative**	**Sorafenib vs Curative**	**TACE vs Sorafenib**	**TACE vs Curative**
Age at rec, years	-2.20 (-3.92;0.48)	2.32 (0.41;4.22)	0.12 (1.52;1.76)	7.16 (4.70;9.62)	0.76 (-1.63;3.15)	7.92 (4.78;11.05)
Gender F	1.08 (0.77;1.51)	0.78 (0.53;1.13)	0.84 (0.59;1.19)	1.43 (0.93;2.20)	0.69 (0.32;1.49)	0.99 (0.50;1.95)
Cirrhosis	0.98 (0.86;1.11)	0.97 0.84;1.12)	0.94 (0.83;1.07)	1.39 (1.26;1.54)	0.32 (0.19;0.55)	0.45 (0.27;0.76)
Child B	0.96 (0.38;2.41)	1.28 (0.50;3.26)	1.23 (0.53;2.81)	-	-	-
N Rec Nodules	2.34 (1.89;2.79)	-0.65 (-1.30;0.005)	1.69 (1.26;2.13)	-0.65 (-1.32;0.02)	-0.27 (-0.73;0.19)	-0.91 (-1.77;-0.06)
N Rec Nodules >1	3.03 (2.42;3.78)	0.91 (0.79;1.04)	2.74 (2.19;3.45)	1.00 (0.78;1.29)	1.08 (0.75;1.56)	1.08 (0.81;1.45)
Bilobar rec (%)	3.74 (2.55;5.48)	0.85 (0.65;1.12)	3.19 (2.19;4.64)	2.89 (2.32;3.60)	0.96 (0.72;1.29)	2.79 (2.13;3.66)
Size ≥5cm	2.95 (1.79;4.86)	0.64 (0.41;1.01)	1.90 (1.10;3.26)	0.75 (0.34;1.63)	0.28 (0.04;2.22)	0.21 (0.03;1.45)
Size, cm	2.45 (1.02;3.88)	-2.27 (-4.21;-0.32)	0.18 (-1.19;1.56)	-0.79 (-2.61;1.03)	-1.18 (-1.84;-0.52)	-1.97 (-4.29;0.35)
MVI	1.20 (0.96;1.50)	0.88 (0.69;1.11)	1.05 (0.83;1.33)	0.76 (0.51;1.12)	1.12 (0.62;2.02)	0.85 (0.53;1.35)
Extra-hepatic rec	3.53 (2.35;5.31)	0.15 (0.08;0.28)	0.54 (0.28;1.06)	2.84 (1.93;4.16)	-	-
TTR	-4.40 (-8.01;-0.80)	-1.80 (-5.35;1.74)	-6.21 (-9.65;-2.76)	-1.92 (-7.23;3.39)	12.98 (3.25;22.72)	11.06 (4.31;17.82)

F: female; N: number; Rec: recurrence; MVI: microvascular invasion; TTR: time-to-recurrence (from first surgery)

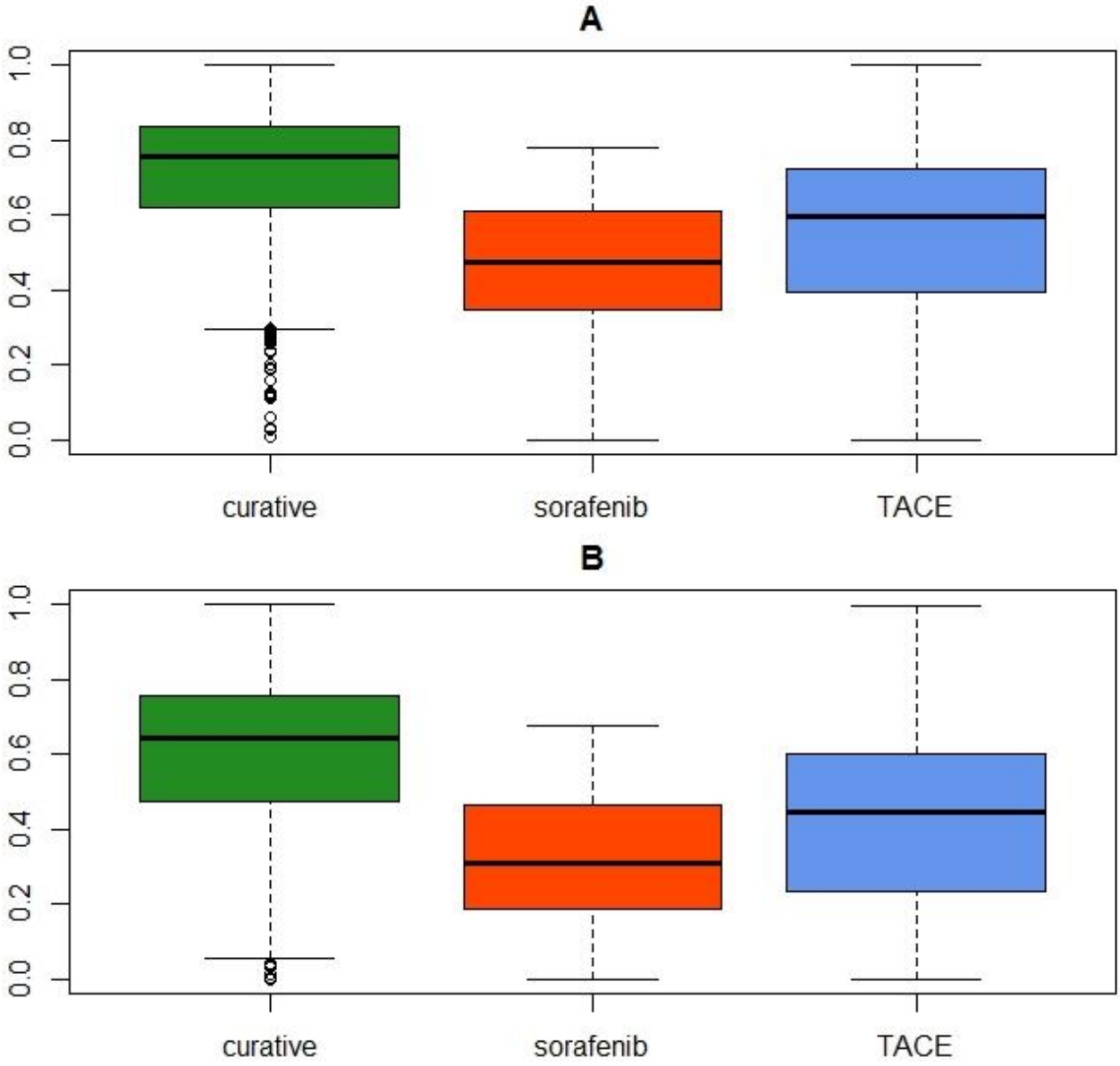
eFigure 3. Time-dependent ROC curves to validate the predictive model for SAR at (A) 3 and (B) 5 years in the external ITALICA cohort. AUC index with 95% confidence interval is reported.



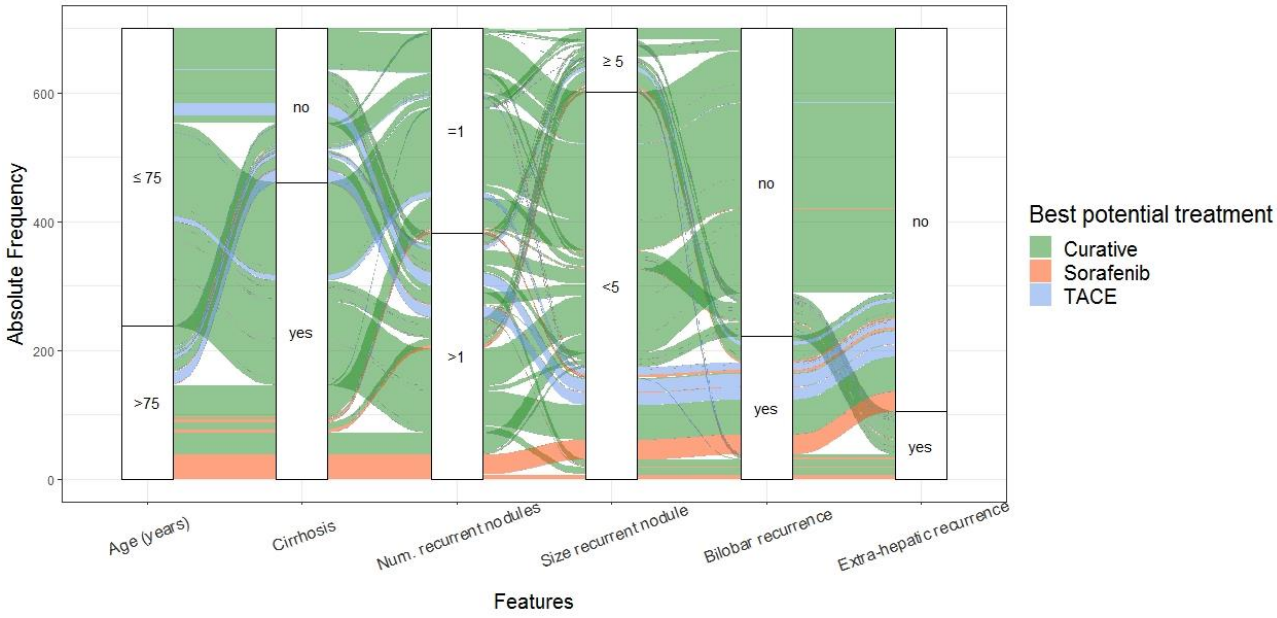
eFigure 4. Time-dependent ROC curves to validate the predictive model for SAR at (A) 3 and (B) 5 years in the external Tokyo University Hospital cohort. AUC index with 95% confidence interval is reported.



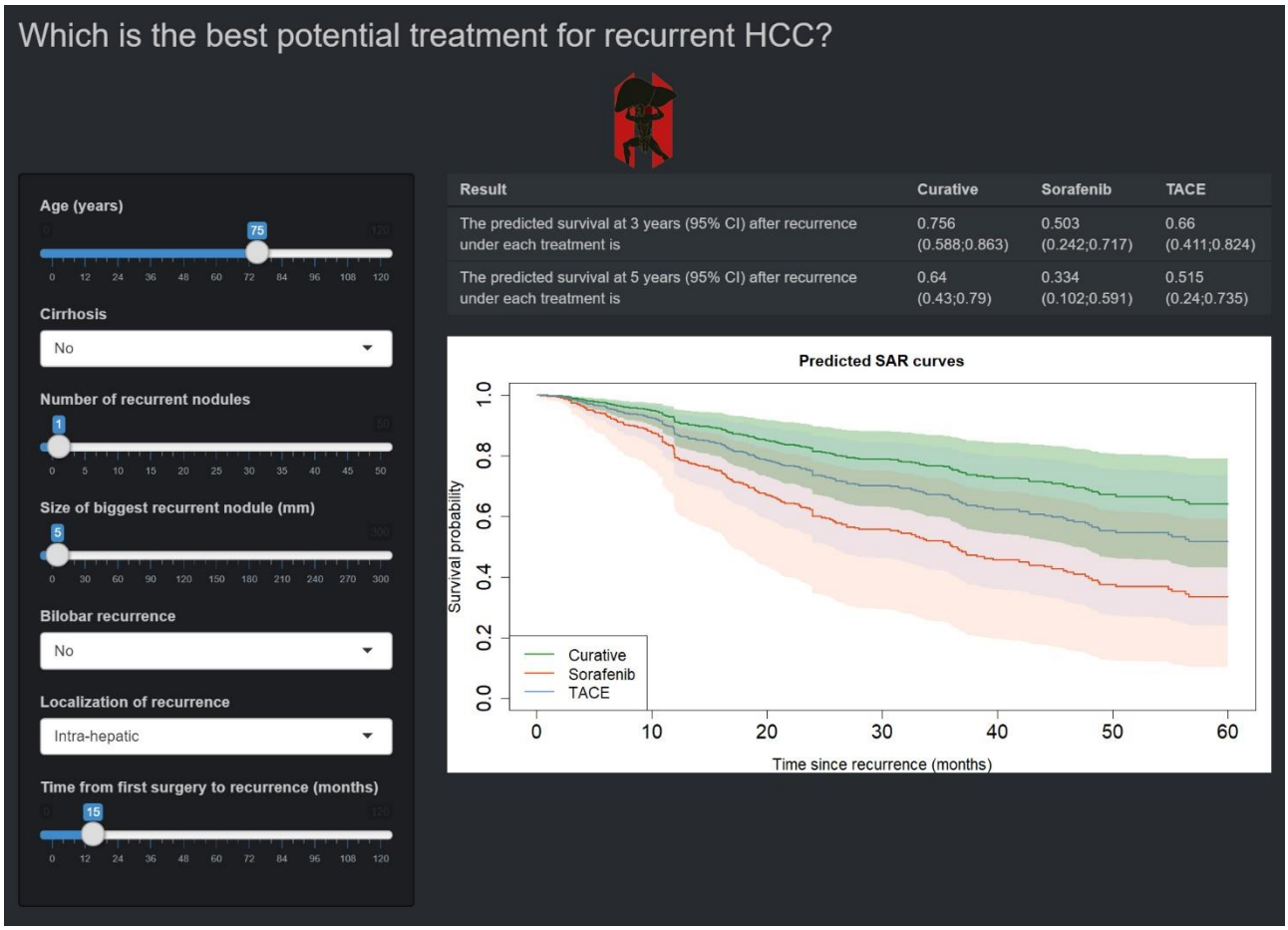
eFigure 5. Box-plot comparison of the distribution of potential SAR under the three considered treatments after application of the algorithm at A) 36 months and B) 60 months.



eFigure 6. Alluvial plot showing the features composition of each BPT group. Each vertical bar is a feature and the size of the bar indicates the absolute frequency of patients with that feature level. The streamlines describe the frequency of patients with a particular combination of features for each BPT group.

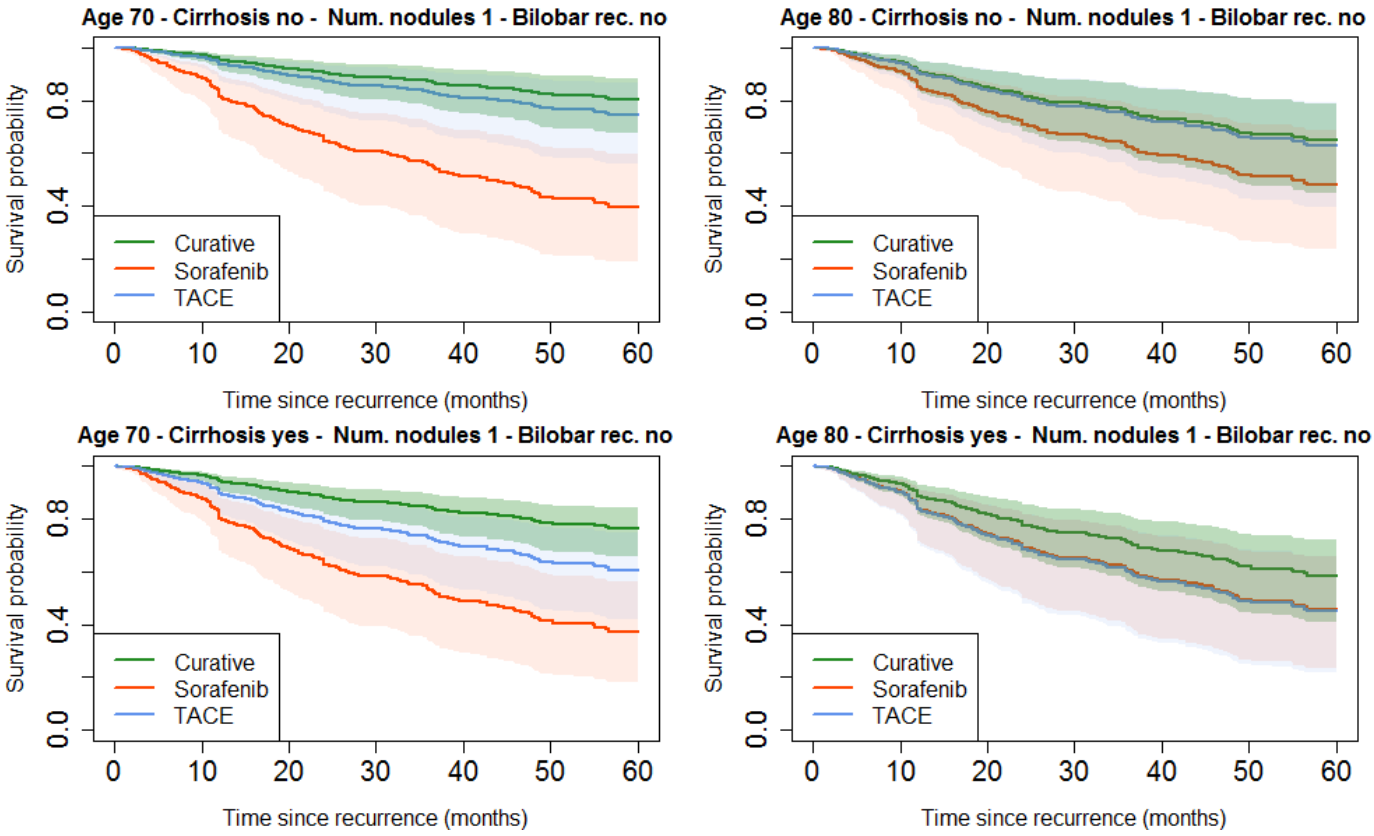


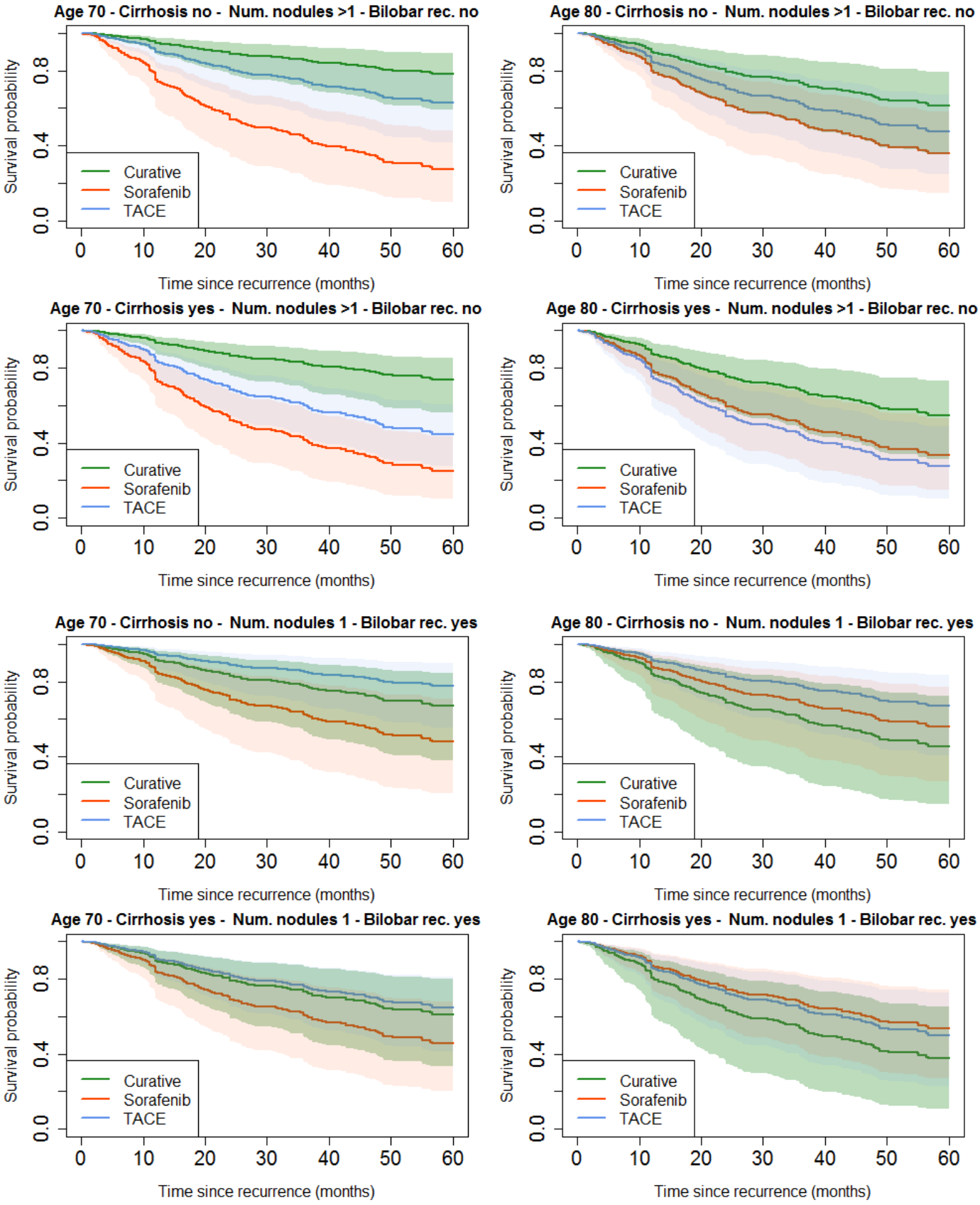
eFigure 7. A snapshot from the web-app available at https://recurrence.hercolesgroup.eu is provided.

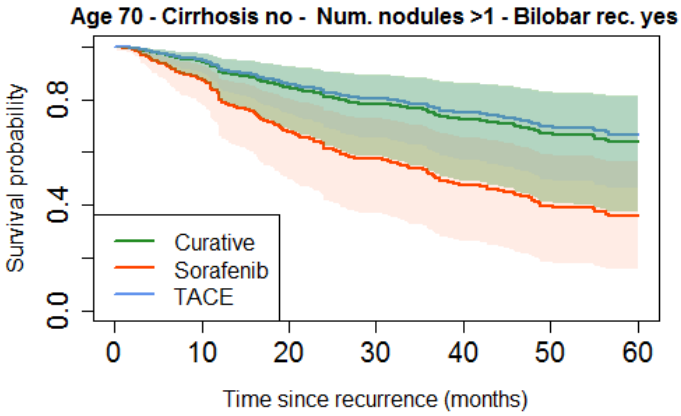


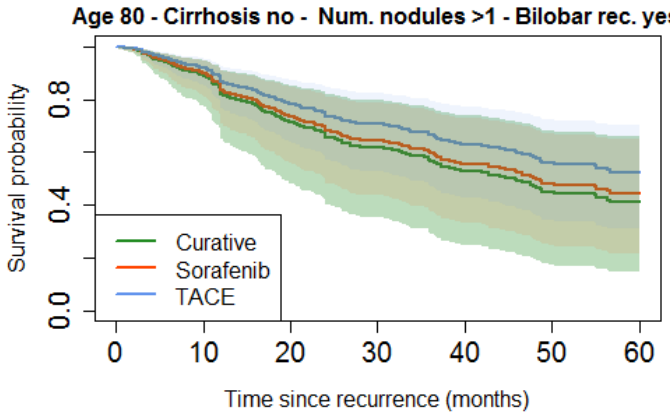
eFigure 8. Predicted SAR according to all profiles and under each treatment.

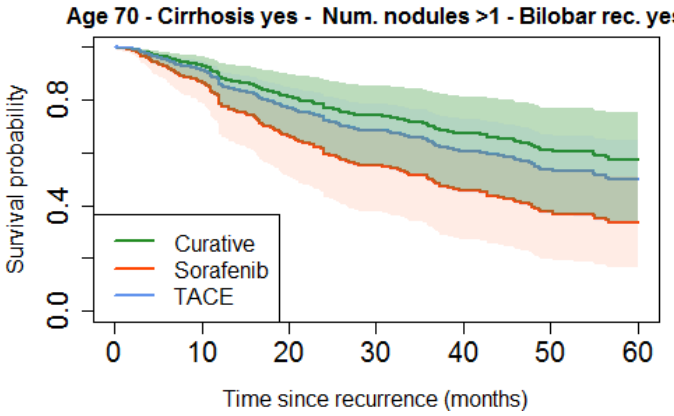
Predicted SAR curves for 16 risk profiles, according to the combination of 4 varying features (Age, Cirrhosis, Number of recurrent nodules and Bilobar recurrence) and 3 fixed features (Size of recurrent nodule <5, Hepatic recurrence and Time from first surgery to recurrence 15 months).

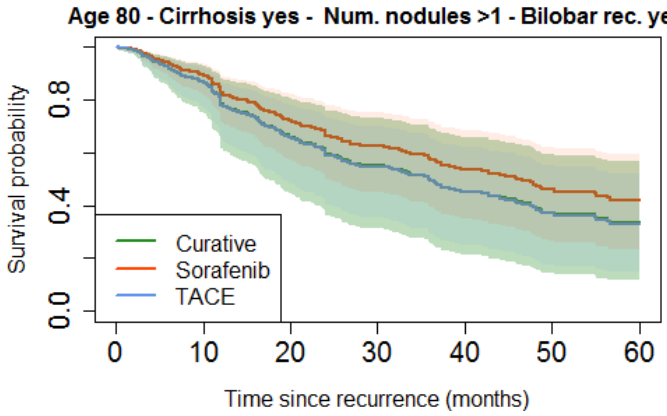




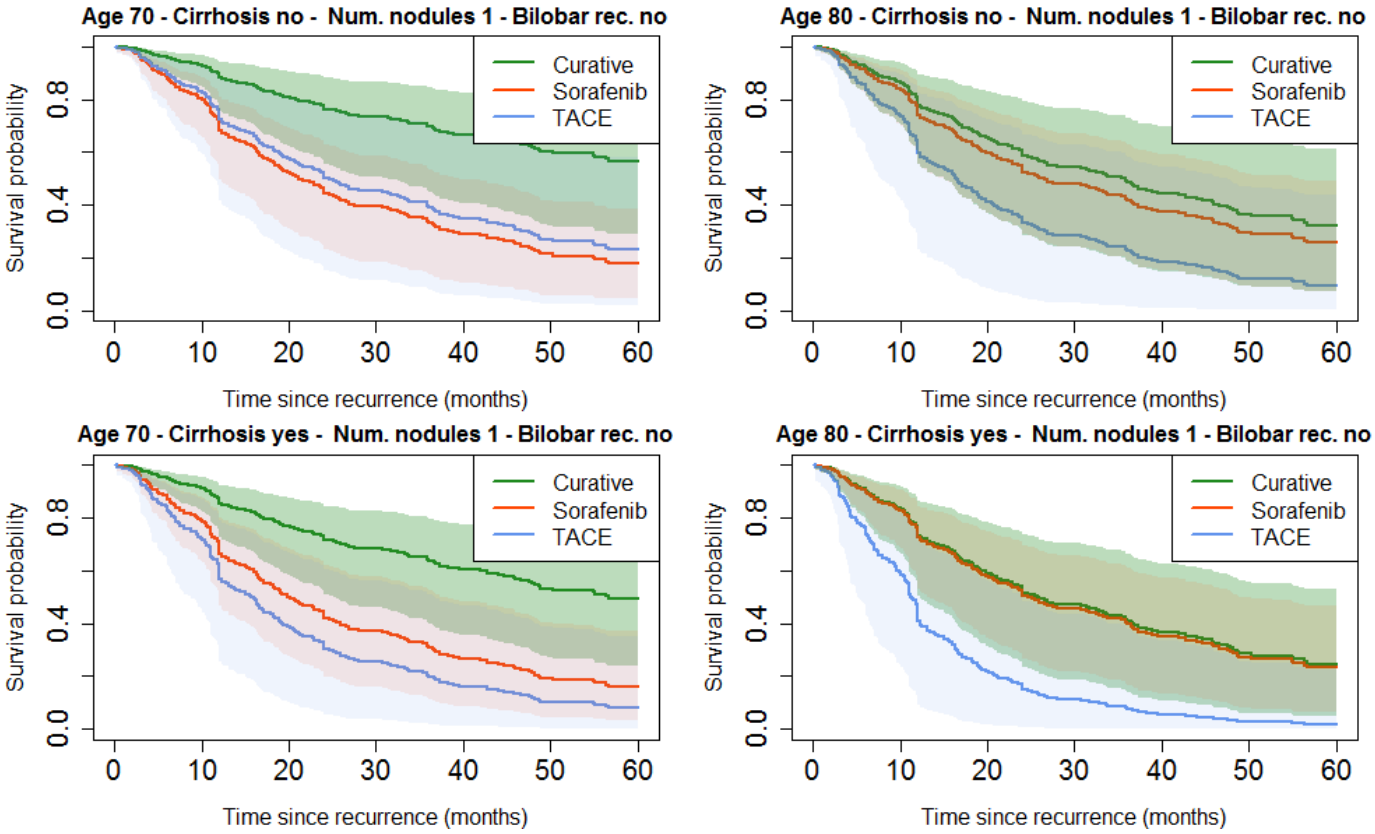


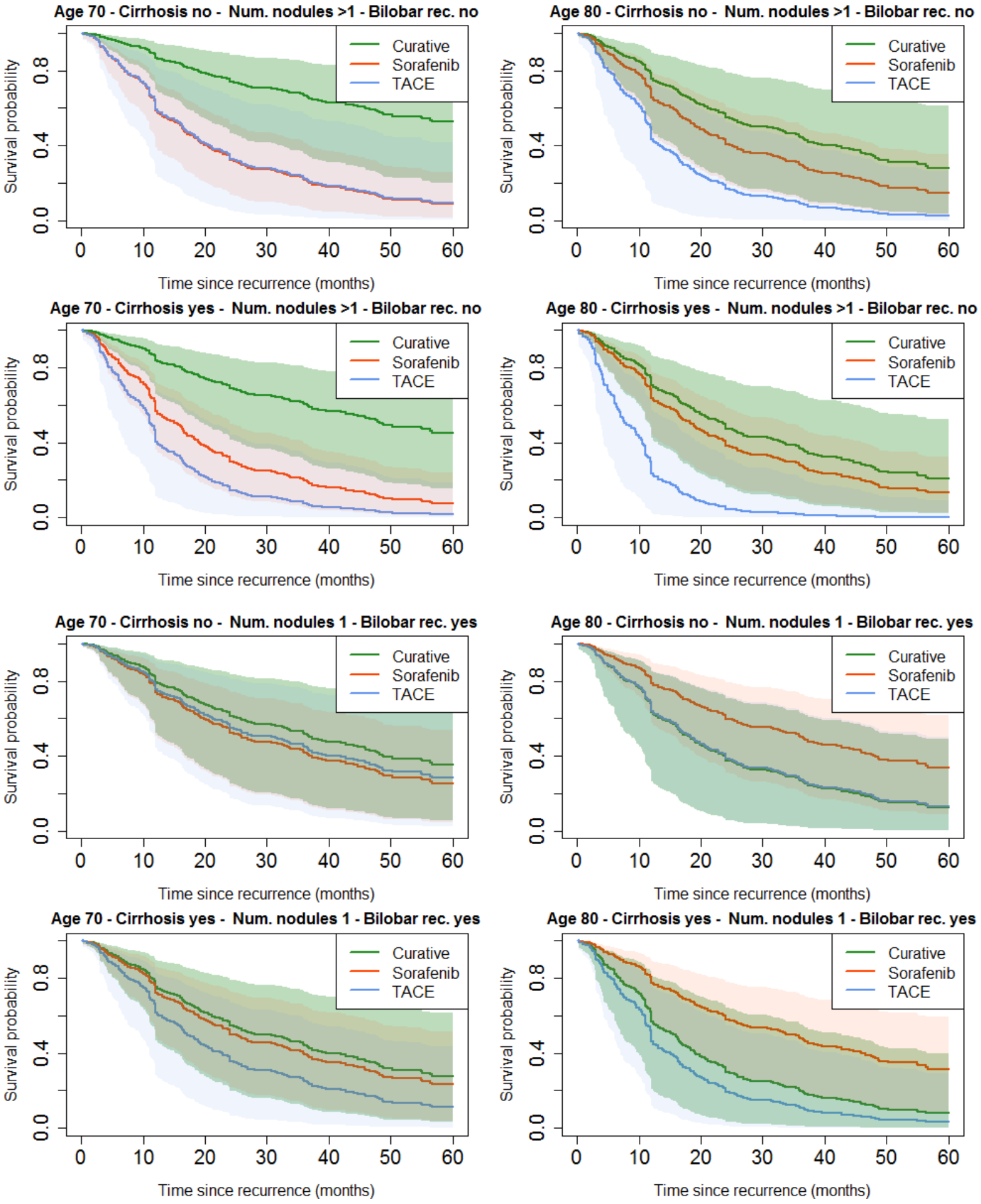


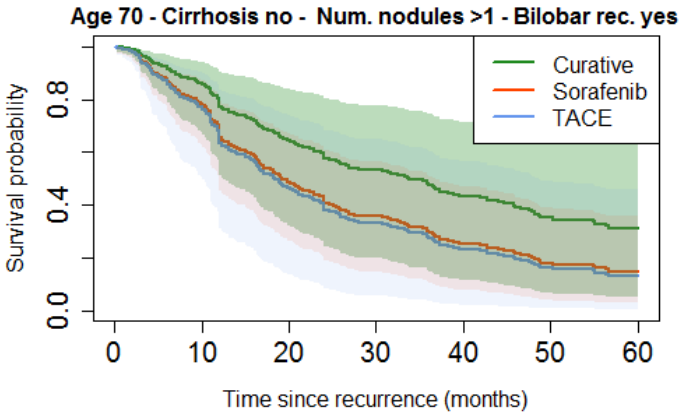


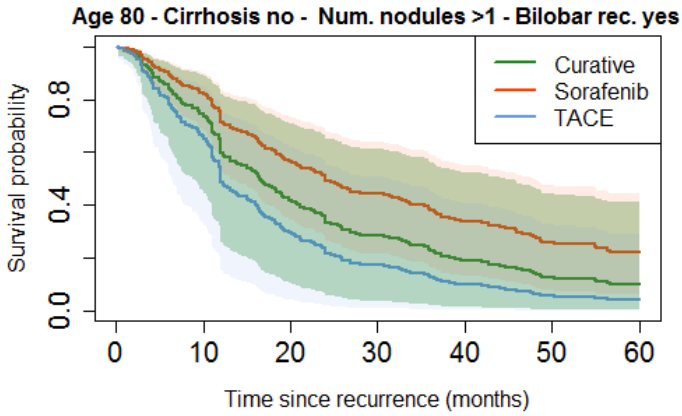


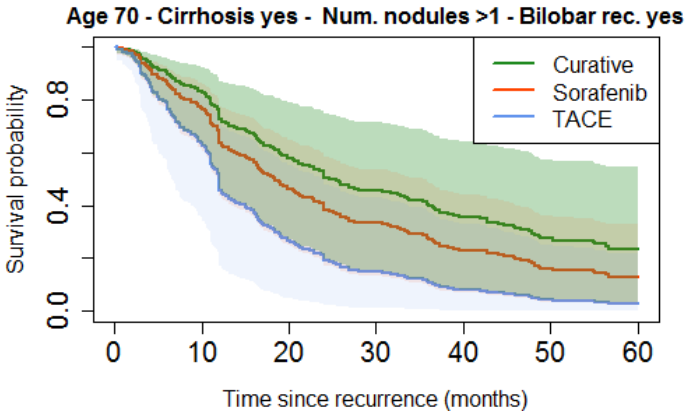
eFigure 9. Predicted SAR according to all profiles and under each treatment. Predicted SAR curves for 16 risk profiles, according to the combination of 4 varying features (Age, Cirrhosis, Number of recurrent nodules and Bilobar recurrence) and 3 fixed features (Size of recurrent nodule <5, Extra-hepatic recurrence and Time from first surgery to recurrence 15 months).

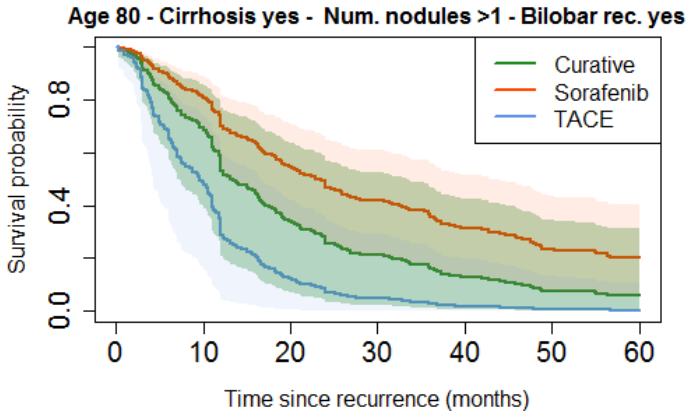




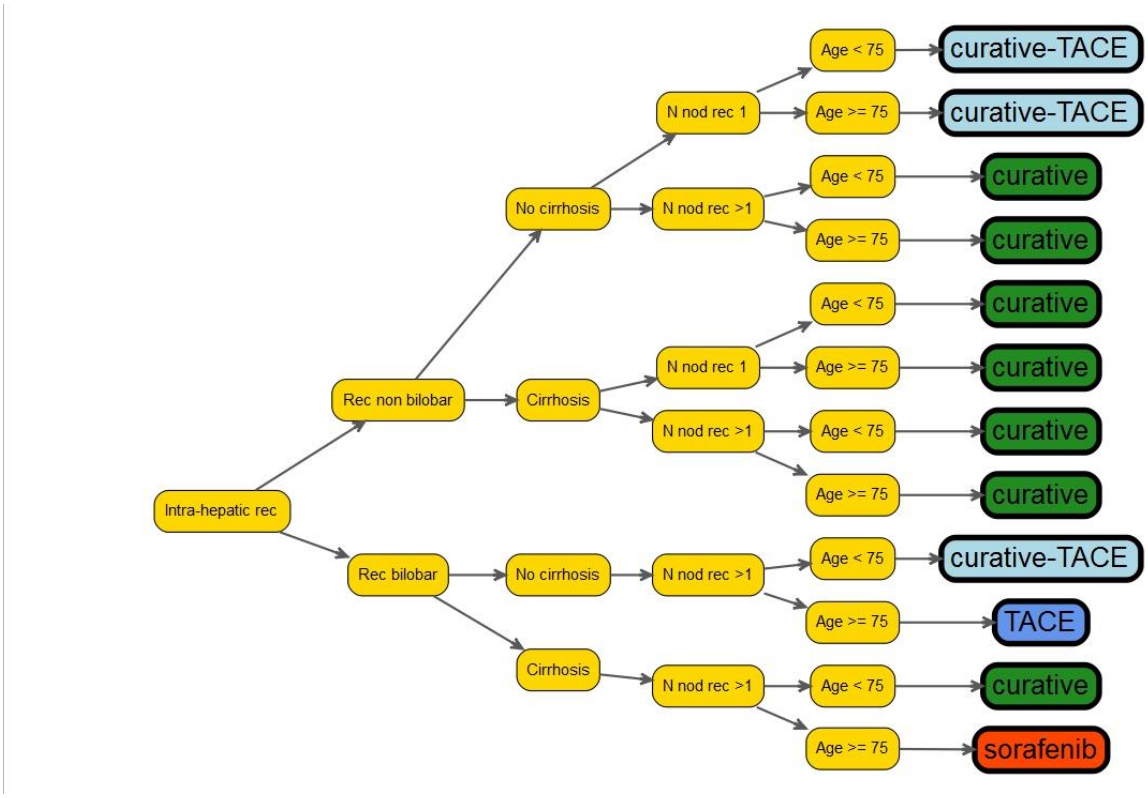








eFigure 10. Algorithm based on SAR at 36 months for intrahepatic recurrence.



eFigure 11. Algorithm based on SAR at 36 months for extrahepatic recurrence.