Machine Learning Can Predict Total Death After Radiofrequency Ablation in Liver Cancer Patients

Jianhua Tong, Panmiao Liu, Muhuo Ji, Ying Wang, Qiong Xue, Jian-Jun Yang and Cheng-Mao Zhou

Department of Anesthesiology, Pain and Perioperative Medicine, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China.

ABSTRACT

OBJECTIVE: Over 1 million new cases of hepatocellular carcinoma (HCC) are diagnosed worldwide every year. Its prognosis remains poor, and the 5-year survival rate in all disease stages is estimated to be between 10% and 20%. Radiofrequency ablation (RFA) has become an important local treatment for liver cancer, and machine learning (ML) can provide many shortcuts for liver cancer medical research. Therefore, we explore the role of ML in predicting the total mortality of liver cancer patients undergoing RFA.

METHODS: This study is a secondary analysis of public database data from 578 liver cancer patients. We used Python for ML to establish the prognosis model.

RESULTS: The results showed that the 5 most important factors were platelet count (PLT), Alpha-fetoprotein (AFP), age, tumor size, and total bilirubin, respectively. Results of the total death model for liver cancer patients in test group: among the 5 algorithm models, the highest accuracy rate was that of gbm (0.681), followed by the Logistic algorithm (0.672); among the 5 algorithms, area under the curve (AUC) values, from high to low, were Logistic (0.738), DecisionTree (0.723), gbm (0.717), GradientBoosting (0.714), and Forest (0.693); Among the 5 algorithms, gbm had the highest precision rate (0.721), followed by the Logistic algorithm (0.714). Among the 5 algorithms, DecisionTree had the highest recall rate (0.642), followed by the GradientBoosting algorithm (0.571).

CONCLUSION: Machine learning can predict total death after RFA in liver cancer patients. Therefore, ML research has great potential for both personalized treatment and prognosis of liver cancer.

KEYWORDS: Treatment algorithms, radiofrequency ablation, hepatocellular carcinoma, individual outcomes, patient stratification

Introduction

Hepatocellular carcinoma (HCC) is the second-most deadly form of cancer in China, killing approximately 300 000 people every year. This accounts for 55% of liver cancer deaths worldwide. Surgical resection is one of the most effective treatments for liver cancer, but only 10% to 20% of patients are candidates for these treatments.[1] In recent years, with the development of minimally invasive and interventional techniques, radiofrequency ablation (RFA) has become an important local treatment for liver cancer because it is safe and involves minimal trauma.[2] It is also recognized as an effective surgical treatment for early HCC among patients with Barcelona Clinic Liver Cancer (BCLC) who do not meet the requirements. However, due to the high risk of liver cancer recurrence, the 5-year survival rate for HCC patients with tumors 5 cm in diameter before RFA is only 50%.[3] Moreover, targeted preventive measures for patients with liver cancer and patient stratification schemes for treatment algorithms customized to the specific conditions of patients with liver cancer can improve their prognosis,[4] such as application of liquid biopsy and multiparametric analysis in liver malignancy management. The economic and medical burden caused by liver cancer is also great, and patients with liver cancer need more effective treatment methods to prolong their survival times and improve their quality of life. Therefore, an effective prognostic evaluation tool for liver cancer patients after RFA is needed.

Machine learning (ML) provides another method for standard predictive modeling, which can solve the current limitations, and by making better use of "big data" for algorithm development, it has the potential to revolutionize medicine.[5-7] At present, ML is providing many shortcuts for medical research.[8] For example, ML methods can provide accurate hospital stay predictions for patients with heart disease. This can be used for clinical bed management and resource allocation[9]; studies have shown that ML can predict acute kidney injury after liver cancer resection[10]; ML can also predict persistent depressive symptoms in older adults.[11] Similarly, studies have shown that ML can provide individualized patient profile analysis, and build prediction models in the palliative treatment of malignant tumors in the liver.[12]

At present, there are few studies on the application of ML among liver cancer patients undergoing RFA. Therefore, this

study assesses 5 ML algorithms' ability to predict the prognosis of liver cancer patients undergoing RFA.

Methods

Patients

Our study is a secondary analysis of public database data from the BioStudies (https://www.ebi.ac.uk/biostudies/studies/S-EPMC6059486) (public) database. It includes 611 HCC patients who had been treated with RFA between January 2006 and December 2010. Finally, the study covers 578 enrolled patients whose ferritin serum levels were measured on admission before RFA. Radiofrequency ablation is typically performed in patients with 3 or fewer lesions less than 3 cm in diameter. The following variables and serum ferritin levels were used: sex, body mass index (BMI), aspartate aminotransferase (AST) level, age, hepatitis C antibody positive, platelet count (PLT), hepatitis B surface antigen positive, alcohol consumption, number of tumors, hemoglobin (HB), Child-Pugh classification, tumor size, and Alpha-fetoprotein (AFP) level.

Categorical variables were compared by Fisher's exact test, while continuous variables were compared by 1-way analysis of variance (parametric) or Kruskal–Wallis test (nonparametric). Survival follow-up ended on December 31, 2015. The difference was considered statistically significant when $P < .05$. We performed ML modeling with python, and Pearson's correlation analysis. The 5-ML model was performed with the following models: random forest, decision tree, Logistic regression, LightGBM, and gradient boosted decision trees (GBDT). Eighty percent of the data were divided into training groups for development, and 20% were verified by the test groups. See Appendix 1 Table 4 for the parameters used in ML in this study.

Results

1. Comparison of each basic index between the 2 groups: There was no significant difference in age between the nondeath group and the death group, nor between the training group and the test group ($P = .001$ and $P = .013$). Moreover, the platelet differences between the nondeath group and the death group in the training group and the test group were statistically significant ($P < .001$ and $P = .049$), and the rest are shown in Table 1.

2. The results of correlation analysis showed that age, tumor size, and AFP were proportional to the postoperative death of patients with liver cancer receiving RFA, and PLT was inversely proportional to the death outcome (Figure 1). These results showed that the 5 most important factors were PLT, AFP, age, tumor size, and total bilirubin, respectively (Figure 2).

3. Training group's effect on total postoperative death model for liver cancer patients: Among the 5 algorithm models, forest's accuracy was the highest (0.900), followed by the GradientBoosting algorithm (0.831); among the 5 algorithms, the AUC values were, from

high to low, Forest (0.971), Gradient Boosting (0.914), gbm (0.825), DecisionTree (0.748), and Logistic (0.739). Among the 5 algorithms, the highest precision rate was that of forest (0.904), followed by the GradientBoosting algorithm (0.795). Among the 5 algorithms, Forest had the highest recall rate (0.887), followed by Gradient-Boosting (0.874) (See Table 2 and Figure 3).

4. Results of the total death model for liver cancer patients in the test group: Among the 5 algorithm models, the highest accuracy rate was that of gbm (0.681), followed by the Logistic algorithm (0.672); among the 5 algorithms, the AUC values, from high to low, were Logistic (0.738), DecisionTree (0.723), gbm (0.717), Gradient-Boosting (0.714), and Forest (0.693); among the 5 algorithms, gbm had the highest precision rate (0.721), followed by the Logistic algorithm (0.714). Among the 5 algorithms, DecisionTree had the highest recall rate (0.642), followed by the GradientBoosting algorithm (0.571; see Table 3 and Figure 4).

Discussion

Radiofrequency ablation is considered the most effective first-line percutaneous ablation therapy.[13] Survival results for patients completely relieved by RFA have been shown to be equivalent to those of patients treated by hepatectomy,[14] with a 5-year recurrence rate that may be as high as 80%.[15] In clinical courses with frequent recurrence and retreatment, tumors have tended to be out of control, which has been the primary reason for the low long-term survival rates after ablation.[16] The results of our study showed that the 5 most important factors were PLT, AFP, age, tumor size, and total bilirubin. The results of the correlation analysis showed that age, tumor size, AFP, and death after RFA for liver cancer were proportional. In addition, PLT was inversely proportional to death outcome. After internal validation, all 5-ML algorithms could better predict the prognosis of liver cancer patients undergoing RFA.

Albumin-bilirubin grade has been introduced to assess liver function in liver cancer patients.[17] In HCC patients undergoing transarterial chemoembolizatio (TACE), the serum prealbumin-bilirubin score (PALBI) grades were predictive of postoperative overall survival.[18] Similarly, albumin-bilirubin grade could predict survival in HCC patients undergoing TACE.[19] Furthermore, PLT predicts functional recovery and complications after hepatectomy.[20] The results of this study also suggest that total bilirubin and platelets are important risk factors for the prognosis of liver cancer patients undergoing RFA.

Anemia is a risk factor for death among liver cancer patients.[21] Hemoglobin changes have also been shown to be associated with overall survival for a variety of malignancies, including lung, breast, colorectal, and liver cancers. This has also been reported in a cohort study of breast, colorectal, and liver cancer.[22] Also, the addition of AFP and ascites to the BCLC staging classification may improve prognostic

Table 1. Basic patient characteristics.

	TRAINING GROUP		P VALUE	TEST GROUP		P VALUE
	NONDEATH	DEATH		NONDEATH	DEATH	
Number of people	240	222		60	56	
Age (years)	67.7 ± 9.7	70.7 ± 8.7	0.001	66.0 ± 10.8	71.4 ± 9.5	0.013
Size (mm)	23.1 ± 8.7	24.2 ± 8.3	0.067	23.2 ± 8.3	26.0 ± 9.9	0.134
Height (cm)	159.3 ± 9.8	159.1 ± 9.1	0.942	161.2 ± 9.5	157.6 ± 10.1	0.034
Wight (kg)	59.6 ± 11.3	59.6 ± 11.7	0.813	60.6 ± 9.9	58.5 ± 10.3	0.126
Body mass index (kg/m²)	23.4 ± 3.3	23.5 ± 3.9	0.811	23.3 ± 3.1	23.6 ± 3.5	0.875
Serum albumin (g/dL)	3.8 ± 0.5	3.6 ± 0.4	<0.001	3.8 ± 0.5	3.5 ± 0.5	<0.001
Total bilirubin (mg/dL)	0.9 ± 0.4	1.1 ± 0.5	<0.001	0.9 ± 0.6	1.0 ± 0.6	0.052
AST (IU/L)	56.5 ± 36.7	62.1 ± 37.2	0.011	51.2 ± 30.1	63.0 ± 35.6	0.034
ALT (IU/L)	55.9 ± 50.8	53.3 ± 37.4	0.698	50.5 ± 31.2	53.2 ± 32.5	0.658
Platelet count (×109/L)	12.3 ± 5.3	10.5 ± 5.5	<0.001	13.3 ± 6.9	11.3 ± 6.6	0.049
Hemoglobin (g/dL)	12.9 ± 1.5	12.8 ± 1.6	0.394	13.2 ± 1.6	12.5 ± 1.8	0.025
Prothrombin activity (%)	88.3 ± 13.2	84.8 ± 13.0	0.002	90.0 ± 13.3	83.0 ± 15.3	0.005
PTINR	1.1 ± 0.1	1.1 ± 0.1	0.002	1.0 ± 0.2	1.1 ± 0.2	0.004
Creatinine	0.7 ± 0.2	0.8 ± 0.2	0.053	0.8 ± 0.2	0.8 ± 0.7	0.294
Serum ferritin level (ng/mL)	193.0 ± 249.8	197.8 ± 232.3	0.780	197.8 ± 194.7	212.7 ± 303.2	0.553
AFP	111.1 ± 459.8	129.5 ± 452.1	<0.001	131.8 ± 490.2	432.3 ± 2487.4	0.513
L3	5.1 ± 13.9	7.4 ± 15.7	<0.001	7.2 ± 16.0	5.9 ± 13.3	0.577
Male			0.299			0.289
No	90 (37.5%)	73 (32.9%)		21 (35.0%)	25 (44.6%)	
Yes	150 (62.5%)	149 (67.1%)		39 (65.0%)	31 (55.4%)	
HBsAg-positive only			<0.001			0.006
No	192 (80.0%)	209 (94.1%)		44 (73.3%)	52 (92.9%)	
Yes	48 (20.0%)	13 (5.9%)		16 (26.7%)	4 (7.1%)	
Anti HCVAb-positive only			0.122			0.092
No	75 (31.2%)	55 (24.8%)		25 (41.7%)	15 (26.8%)	
Yes	165 (68.8%)	167 (75.2%)		35 (58.3%)	41 (73.2%)	
Alcohol consumption			0.045			0.203
≤80 g/day	192 (80.0%)	159 (71.6%)		52 (86.7%)	41 (73.2%)	
>80 g/day	22 (9.2%)	21 (9.5%)		4 (6.7%)	6 (10.7%)	
None	26 (10.8%)	42 (18.9%)		4 (6.7%)	9 (16.1%)	
Child-Pugh score			<0.001			0.019
0	155 (64.6%)	86 (38.7%)		40 (66.7%)	20 (35.7%)	
1	57 (23.8%)	80 (36.0%)		12 (20.0%)	15 (26.8%)	
2	14 (5.8%)	38 (17.1%)		5 (8.3%)	10 (17.9%)	
3	11 (4.6%)	10 (4.5%)		2 (3.3%)	6 (10.7%)	
4	1 (0.4%)	7 (3.2%)		1 (1.7%)	3 (5.4%)	
5	2 (0.8%)	1 (0.5%)		0 (0.0%)	2 (3.6%)	

Abbreviations: AFP, alpha-fetoprotein; anti-HCVAb, anti-hepatitis C virus antibody; AST, aspartate aminotransferase; HBsAg, hepatitis B surface antigen; ALT, alanine aminotransferase; PTINR, prothrombin time-international normalized ratio.
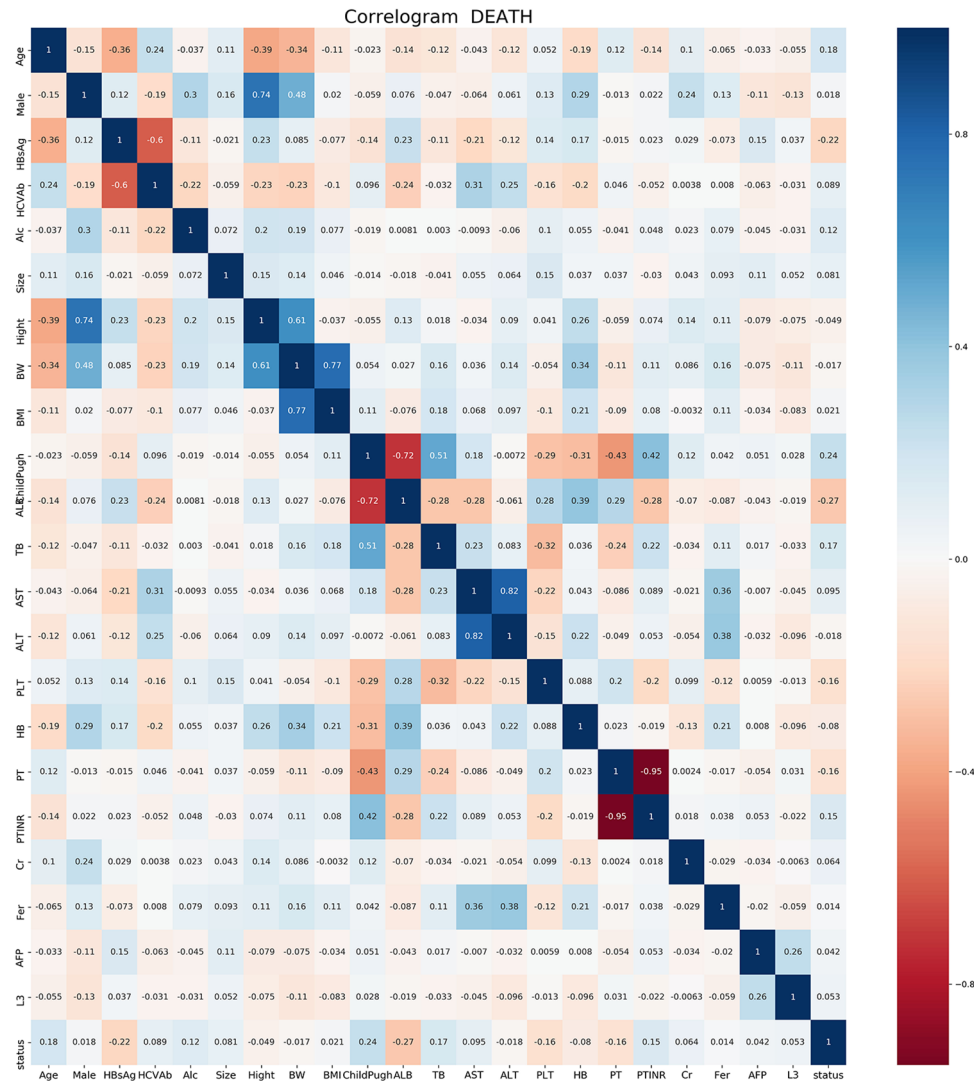
Figure 1. Factor correlations.

Table 2. Forecasted results for the training group.

	ACCURACY	PRECISION	RECALL	AUC
Logistic	0.686	0.679	0.658	0.739
DecisionTree	0.690	0.642	0.806	0.748
Forest	0.900	0.904	0.887	0.971
GradientBoosting	0.831	0.795	0.874	0.914
Gbm	0.755	0.742	0.752	0.825

Abbreviation: AUC, area under the curve.

predictors of early and mid-term liver cancer.[23] This finding is also supported by this study.

Patients with tumors >2 cm have consistently shown that the survival benefits of RFA exceed those of percutaneous ethanol injection (PEI).[16,24] Researchers have agreed that when a patient's tumor is ⩽2 cm, RFA is advantageous in liver cancer treatment.[25] The results of our study also suggest that tumor size is an important factor affecting the prognosis of liver cancer patients undergoing RFA.

Machine learning is instrumental for the multiomics and multiparametric analysis essential to improving the overall management of liver malignancies and individual outcomes. In the next step of ML and liver cancer research, ML schemes based on biomarkers of chronic inflammation may be a useful tool in prediction and prevention.[26] Moreover, to facilitate the interpretability of ML in liver cancer research, it would be of great significance and scientific value to incorporate multigroup learning.[27]
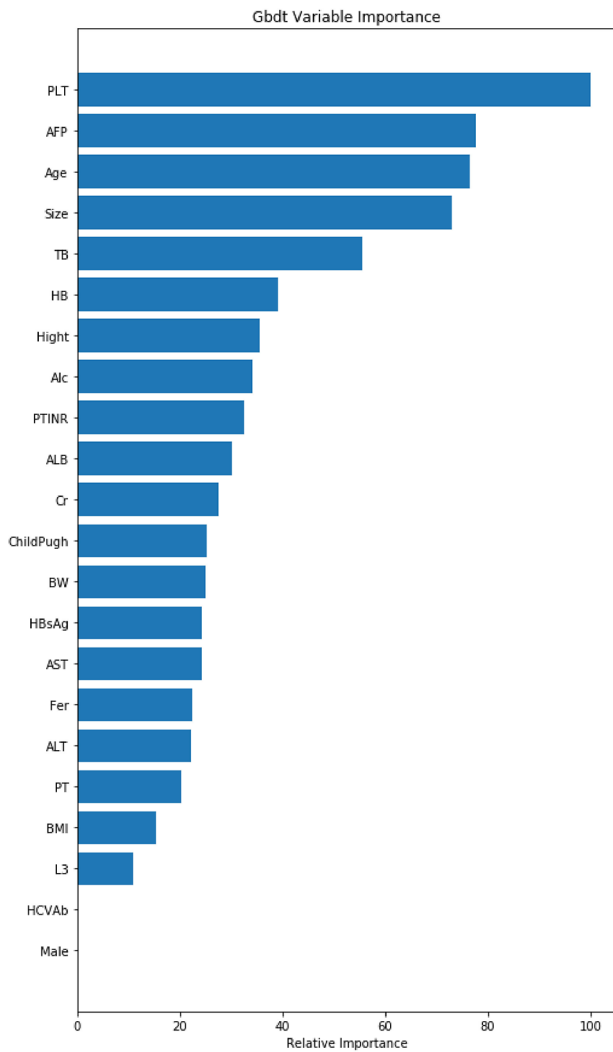
Figure 2. Variable importance of features included in the machine-learning algorithm for predicting postoperative death outcomes.
AFP indicates alpha-fetoprotein; AST, aspartate aminotransferase; BMI, body mass index; HB, hemoglobin; HBsAg, hepatitis B surface antigen; PLT, platelet count; TB, total bilirubin; PTINR, prothrombin time-international normalized ratio; ALT, aminoleucine transferase; ALB, albumin; PT, prothrombin time; BW, body weight.

Figure 3. Machine-learning algorithm predictions of postoperative death outcomes in the training group.
AUC indicates area under the curve, ROC, receiver operating characteristics.

Figure 4. Machine-learning algorithm predictions of postoperative death outcomes in the test group.
AUC indicates area under the curve, ROC, receiver operating characteristics.

Our study does have several limitations. First, since this was a retrospective study, it was impossible to collect all of the data from the HCC patients. This may have biased the model effect. Second, data on molecular factors was lacking. T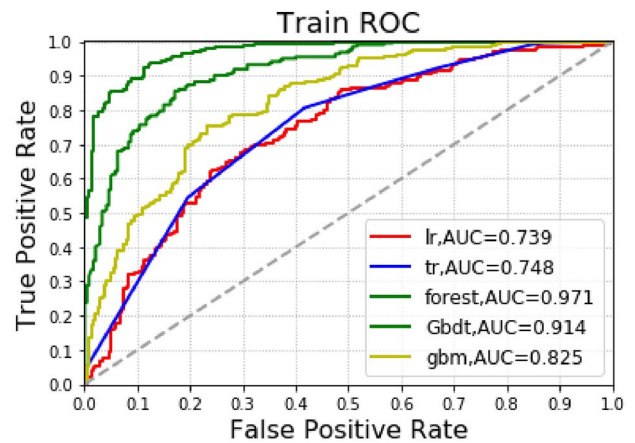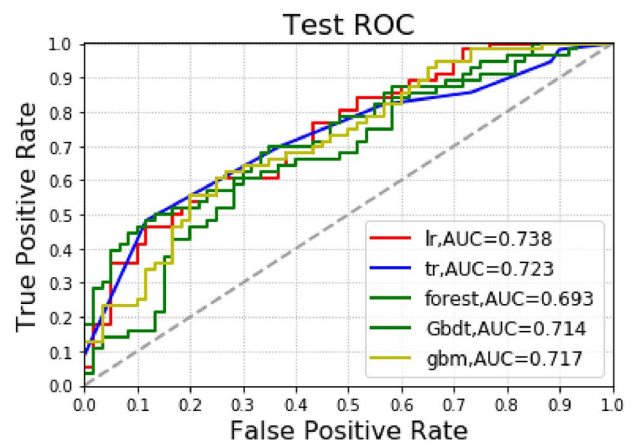his may have influenced patient prognosis. Finally, this study model has only been internally validated. Thus, further prospective and multicenter studies are needed in the future.

Conclusion

Our results show that each factor derived from the ML gbm algorithm accounts for the weight of death. This indicates that the 5 most important factors, ranked in order, were PLT, AFP, age, tumor size, and total bilirubin. Furthermore, ML algorithms

Table 3. Forecasted results for the testing group.

	ACCURACY	PRECISION	RECALL	AUC
Logistic	0.672	0.714	0.538	0.738
DecisionTree	0.664	0.690	0.642	0.723
forest	0.638	0.646	0.554	0.693
GradientBoosting	0.664	0.689	0.571	0.714
gbm	0.681	0.721	0.554	0.717

Abbreviation: AUC, area under the curve.

can improve prognosis prediction for liver cancer patients under-going RFA.

In sum, prevention, prediction, and individuation are of great significance to the prognosis of patients with liver cancer.[28] As such, ML research has great potential for person-alized treatment and prognosis of liver cancer.

Author Contributions
All authors contributed to this study's conceptualization. C.-M.Z. and J.-J.Y. contributed to project administration work. All authors contributed to writing, review, and editing the work.

Availability of Data and Materials
Data are available at the BioStudies Database, accession num-ber: S-EPMC6059486.

Ethics Approval and Consent to Participate
Our study did not require the approval of an ethics committee, as it was a secondary analysis of the BioStudies public database which is open to the public.

ORCID iD
Cheng-Mao Zhou iD https://orcid.org/0000-0001-5680-791X

REFERENCES
1. Hung H-H, Chiou Y-Y, Hsia C-Y, et al. Survival rates are comparable after radiofrequency ablation or surgery in patients with small hepatocellular carcino-mas. *Clin Gastroenterol Hepatol*. 2011;9:79-86.
2. Minami Y, Kudo M. Radiofrequency ablation of hepatocellular carcinoma: cur-rent status. *World J Radiol*. 2010;2:417-424.
3. Künzli BM, Abitabile P, Maurer CA. Radiofrequency ablation of liver tumors: actual limitations and potential solutions in the future. *World J Hepatol*. 2011;3:8-14.
4. Golubnitschaja O, Polivka J Jr, Yeghiazaryan K, Berliner L. Liquid biopsy and multiparametric analysis in management of liver malignancies: new concepts of the patient stratification and prognostic approach. *EPMA J*. 2018;9:271-285.
5. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375:1216-1219.
6. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inform*. 2002;35:352-359.
7. Berglund E, Lytsy P, Westerling R. Adherence to and beliefs in lipid-lowering medical treatments: a structural equation modeling approach including the necessity-concern framework. *Patient Educ Couns*. 2013;91:105-112.
8. Zhou C, Wang Y, Ji MH, Tong J, Yang JJ, Xia H. Predicting peritoneal metas-tasis of gastric cancer patients based on machine learning. *Cancer Control*. 2020;27:1073274820968900.
9. Daghistani TA, Elshawi R, Sakr S, Ahmed AM, Al-Thwayee A, Al-Mallah MH. Predictors of in-hospital length of stay among cardiac patients: a machine learning approach. *Int J Cardiol*. 2019;288:140-147.
10. Lei L, Wang Y, Xue Q, Tong J, Zhou CM, Yang JJ. A comparative study of machine learning algorithms for predicting acute kidney injury after liver cancer resection. *PeerJ*. 2020;8:e8583.
11. Hatton CM, Paton LW, McMillan D, Cussens J, Gilbody S, Tiffin PA. Predict-ing persistent depressive symptoms in older adults: a machine learning approach to personalised mental healthcare. *J Affect Disord*. 2019;246:857-860.
12. Goldstein E, Yeghiazaryan K, Ahmad A, Giordano FA, Fröhlich H, Golub-nitschaja O. Optimal multiparametric set-up modelled for best survival out-comes in palliative treatment of liver malignancies: unsupervised machine learning and 3 PM recommendations. *EPMA J*. 2020;11:505-515.
13. Poon TP, Fan ST, Tsang HF, Wong J. Locoregional therapies for hepatocellular carcinoma: a critical review from the surgeon's perspective. *Ann Surg*. 2002;235: 466-486.
14. Livraghi T, Bolondi L, Lazzaroni S, et al. Percutaneous ethanol injection in the treatment of hepatocellular carcinoma in cirrhosis. A study on 207 patients. *Can-cer*. 1992;69:925-929.
15. Llovet JM, Bruix J. Novel advancements in the management of hepatocellular carcinoma in 2008. *J Hepatol*. 2008;48:S20-S37.
16. Germani G, Pleguezuelo M, Gurusamy K, Meyer T, Isgrò G, Burroughs AK. Clini-cal outcomes of radiofrequency ablation, percutaneous alcohol and acetic acid injec-tion for hepatocelullar carcinoma: a meta-analysis. *J Hepatol*. 2010;52:380-388.
17. Johnson PJ, Berhane S, Kagebayashi C, et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence—based approach—the ALBI grade. *J Clin Oncol*. 2015;33:550-558.
18. Chan AWH, Leung HHW, Chong CCN, Chan SL. Validating the ALBI grade: its current and future use in HCC prognostication. *J Hepatol*. 2017;66: 661-663.
19. Liu PH, Hsu CY, Hsia CY, et al. ALBI and PALBI grade predict survival for HCC across treatment modalities and BCLC stages in the MELD Era. *J Gastro-enterol Hepatol*. 2017;32:879-886.
20. Takahashi K, Kurokawa T, Oshiro Y, Fukunaga K, Sakashita S, Ohkohchi N. Postoperative decrease in platelet counts is associated with delayed liver function recovery and complications after partial hepatectomy. *Tohoku J Exp Med*. 2016; 239:47-55.
21. Finkelmeier F, Bettinger D, Köberle V, et al. Single measurement of hemoglobin predicts outcome of HCC patients. *Med Oncol*. 2014;31:806.
22. Wan S, Lai Y, Myers RE, et al. Post—diagnosis hemoglobin change associates with overall survival of multiple malignancies—results from a 14-year hospital-based cohort of lung, breast, colorectal, and liver cancers. *BMC Cancer*. 2013; 13:340.
23. Gomaa AI. Ascites and alpha-fetoprotein improve prognostic performance of Bar-celona Clinic Liver Cancer staging. *World J Gastroenterol*. 2015;21:5654-5662.
24. Lin SM. Randomised controlled trial comparing percutaneous radiofrequency ther-mal ablation, percutaneous ethanol injection, and percutaneous acetic acid injection to treat hepatocellular carcinoma of 3 cm or less. *Gut*. 2005;54:1151-1156.
25. Lencioni RA, Allgaier HP, Cioni D, et al. Small hepatocellular carcinoma in cirrhosis: randomized comparison of radio-frequency thermal ablation versus percutaneous ethanol injection. *Radiology*. 2003;228:235-240.
26. Qian S, Golubnitschaja O, Zhan X. Chronic inflammation: key player and bio-marker-set to predict and prevent cancer development and progression based on individualized patient profiles. *EPMA J*. 2019;10:365-381.
27. Lu M, Zhan X. The crucial role of multiomic approach in cancer research and clinically relevant outcomes. *EPMA J*. 2018;9:77-102.
28. Janssens JP, Schuster K, Voss A. Preventive, predictive, and personalized medi-cine for effective and affordable cancer care. *EPMA J*. 2018;9:113-123.

Appendix 1

Table 4. Functions, packages, and tuning parameters used in Anaconda for each machine learning algorithm.

ALGORITHM	CLASSIFIER	PACKAGE	TUNING PARAMETERS
Logistic regression	LogisticRegression	from sklearn.linear_model import LogisticRegression	penalty='l2', tol=0.000001, C=0.1, fit_intercept=True,intercept_scaling=1, class_weight=None,max_iter=100, multi_class='ovr',verbose=0, warm_start=False,n_jobs=1
DecisionTree	DecisionTreeClassifier	from sklearn.tree import DecisionTreeClassifier	splitter='best', max_depth=3, min_samples_split=30, min_samples_leaf=2, min_weight_fraction_leaf=0.01
forest	RandomForestClassifier	from sklearn.ensemble import RandomForestClassifier	n_estimators=50, n_jobs=-1, min_samples_split=20, min_samples_leaf=2, random_state=41
GradientBoosting	GradientBoostinglassifier	from sklearn.ensemble import GradientBoostinglassifier	learning_rate=0.2, n_estimators=20, max_depth=3, min_samples_split=20, min_samples_leaf=5
gbm	lgb.LGBMClassifier	lightgbm 2.2.0	boosting_type='gbdt', objective='binary',metrics='auc',learning_rate=0.1, n_estimators=100, max_depth=2, bagging_fraction=0.5, feature_fraction=0.5