Liver Cancer

Research Article

Liver Cancer
DOI: 10.1159/000530078

Received: December 14, 2022
Accepted: March 6, 2023
Published online: May 25, 2023

Attempt to Establish Prognostic Predictive System for Hepatocellular Carcinoma Using Artificial Intelligence for Assistance with Selection of Treatment Modality

Atsushi Hiraoka[a] Takashi Kumada[b, c] Toshifumi Tada[c] Hidenori Toyoda[c] Kazuya Kariyama[d] Takeshi Hatanaka[e] Satoru Kakizaki[f] Atsushi Naganuma[g] Ei Itobayashi[h] Kunihiko Tsuji[i] Toru Ishikawa[j] Hideko Ohama[a] Fujimasa Tada[a] Kazuhiro Nouso[d]
on behalf of the Real-life Practice Experts for HCC (RELPEC) Study Group

[a]Gastroenterology Center, Ehime Prefectural Central Hospital, Matsuyama, Japan; [b]Department of Nursing, Gifu Kyoritsu University, Ogaki, Japan; [c]Department of Gastroenterology and Hepatology, Ogaki Municipal Hospital, Ogaki, Japan; [d]Department of Hepatology, Okayama City Hospital, Okayama, Japan; [e]Department of Gastroenterology, Gunma Saiseikai Maebashi Hospital, Maebashi, Japan; [f]Department of Clinical Research, National Hospital Organization Takasaki General Medical Center, Takasaki, Japan; [g]Department of Gastroenterology, National Hospital Organization Takasaki General Medical Center, Takasaki, Japan; [h]Department of Gastroenterology, Asahi General Hospital, Asahi, Japan; [i]Center of Gastroenterology, Teine Keijinkai Hospital, Sapporo, Japan; [j]Department of Gastroenterology, Saiseikai Niigata Hospital, Niigata, Japan

Abstract

Introduction: Because of recent developments in treatments for hepatocellular carcinoma (HCC), methods for determining suitable therapy for initial or recurrent HCC have become important. This study used artificial intelligence (AI) findings to establish a system for predicting prognosis of HCC patients at time of reoccurrence based on clinical data as a reference for selection of treatment modalities. ***Methods:*** As a training cohort, 5,701 observations obtained at the initial and each subsequent treatment for recurrence from 1,985 HCC patients at a single center from 2000 to 2021 were used. The validation cohort included 5,692 observations from patients at multiple centers obtained at the time of the initial treatment. An AI calculating system (PRAID) was constructed based on 25 clinical factors noted at each treatment from the training cohort, and then predictive prognostic values for 1- and 3-year survival in both cohorts were evaluated. ***Results:*** After exclusion of patients lacking clinical data regarding albumin-bilirubin (ALBI) grade or tumor-node-metastasis stage of the Liver Cancer Study Group of Japan, 6th edition (TNM-LCSGJ 6th), ALBI-TNM-LCSGJ 6th (ALBI-T) and modified ALBI-T scores confirmed that prognosis for patients in both cohorts was similar. The area under the curve for prediction of both 1- and 3-year survival in the validation cohort was 0.841 (sensitivity 0.933 [95% CI: 0.925–0.940], specificity 0.517 [95% CI: 0.484–0.549]) and 0.796 (sensitivity 0.806 [95% CI: 0.790–0.821], specificity 0.646 [95% CI: 0.624–0.668]), respectively. ***Conclusion:*** The present PRAID system

karger@karger.com
www.karger.com/lic

Correspondence to:
Atsushi Hiraoka, hirage@m.ehime-u.ac.jp

might provide useful prognostic information related to short and medium survival for decision-making regarding the best therapeutic modality for both initial and recurrent HCC cases.

Introduction

Several reports have noted development of therapies for hepatocellular carcinoma (HCC), including surgical resection (SR) [1], ablation including radio frequency ablation (RFA) [2], transcatheter arterial chemoembolization (TACE) [3], systemic therapy including immunotherapy [4], and multi-targeting agents [5, 6]. Nevertheless, HCC was recently reported to be the sixth most commonly diagnosed cancer worldwide and the third leading cause of cancer death [7]. As for reasons for the poor prognosis of HCC patients, high recurrence rates and coexistence of basal liver diseases (e.g., chronic viral hepatitis infection, alcohol abuse) are considered to be important factors.

Guidelines and therapeutic treatment algorithms for HCC have been reported [8–14], though all propose the same strategies for treating initial and recurrent HCC. Different from other malignant tumors, the prognosis of an HCC patient depends not only on the therapeutic modality employed but also on clinical factors, such as hepatic reserve function [15, 16], tumor burden, malignant potential [17], and remaining hepatic reserve volume, as well as others. Because of recent developments in systemic therapy for unresectable HCC including intermediate-stage cases [18], consideration of suitable treatments for cases with unresectable status, not only at the time of the initial diagnosis but also at the time of recurrence, has become important for prolonging survival. There is no doubt that the level of invasiveness in association with maintenance of liver reserve and antitumor effects related to treatment are important factors when considering strategies for treatment planning.

The present study was conducted to establish a prediction system using artificial intelligence (AI) for short-term (1 year) and medium-term (3 years) prognosis of HCC patients at the time of recurrence based on routine clinical information, including treatment modality, for use as a reference when selecting treatment for obtaining a better prognosis.

Materials and Methods

A total of 1,985 HCC patients with a variety of clinical courses diagnosed and treated at Ehime Prefectural Central Hospital from 2000 to 2021 were enrolled as the training cohort. Data related to each treatment for either initial or recurrent HCC given to a single patient were treated as one dataset. Following exclusion of patients who remained alive with a survival period of less than 1 year or within 3 years from the last relevant treatment, 4,944 cases including those of recurrence were analyzed as models of survival after 1 year, while 4,264 cases were analyzed as models of survival after 3 years (Fig. 1). Cases that fit those models but have an observation period shorter than 1 or 3 years, respectively, were excluded from survival analysis because clinical factors were used for evaluation of predictive factors of survival at 1 and 3 years. To confirm predictive values for analysis as a validation cohort, data including clinical factors from 5,692 HCC patients at the time of the initial treatment were obtained from multiple collaborating institutions.

Underlying Liver Disease

Positive anti-hepatitis C virus (anti-HCV) findings were considered to indicate that HCC was due to hepatitis C virus (HCV), whereas HCC due to hepatitis B virus (HBV) was determined when the HBV surface antigen was positive. For patients with a history of alcohol abuse (≥60 g/day) [19, 20], underlying liver disease was judged as related to alcohol. In the present study, we treated chronic hepatitis viral-infected patients with alcohol abuse as those with viral etiologies.

Liver Function Assessment

For assessment of hepatic reserve function, albumin-bilirubin (ALBI) score [21, 22] was used along with modified ALBI grade (mALBI), for which ALBI grade 2 is divided into two subgrades (2a and 2b) using an ALBI score of −2.27 as the cutoff value [23].

HCC Diagnosis and Treatment

HCC diagnosis was based on an increasing course of alpha-fetoprotein (AFP) as well as dynamic CT [24], magnetic resonance imaging [25, 26], and/or pathological findings obtained during the clinical course. BCLC stage [27] and tumor-node-metastasis (TNM) staging, determined based on TNM staging for HCC conducted by the Liver Cancer Study Group of Japan, 6th edition (LCSGJ 6th) [28] (TNM-LCSGJ 6th), were utilized for evaluations of tumor progression.

Conventional Integrated Scoring System for Predicting Prognosis

Japan integrated staging (JIS) score, a combination of Child-Pugh class and TNM-LCSGJ 6th [15], as well as ALBI-TNM-LCSGJ 6th (ALBI-T) score, a combination of ALBI grade and TNM-LCSGJ 6th [22], and mALBI-T, a combination of mALBI grade and TNM-LCSGJ 6th [23], were used for comparing prognoses between the training and validation cohorts.

AI System and Analyses

The XGBoost software library was used as a Python package (https://xgboost.readthedocs.io/en/stable/python/python_intro.html) for development of an AI system to establish a prognosis prediction system, which was termed the prediction of prognosis of HCC by AI for treatment decision-making (PRAID) system (Fig. 2) (https://livercancer.onrender.com/). Development of the system was done by Amane Saito of Amatechno Co., Ltd.

Available clinical information for the 25 factors at the time of administration and/or diagnosis was analyzed for use in the present analysis, as shown as follows:

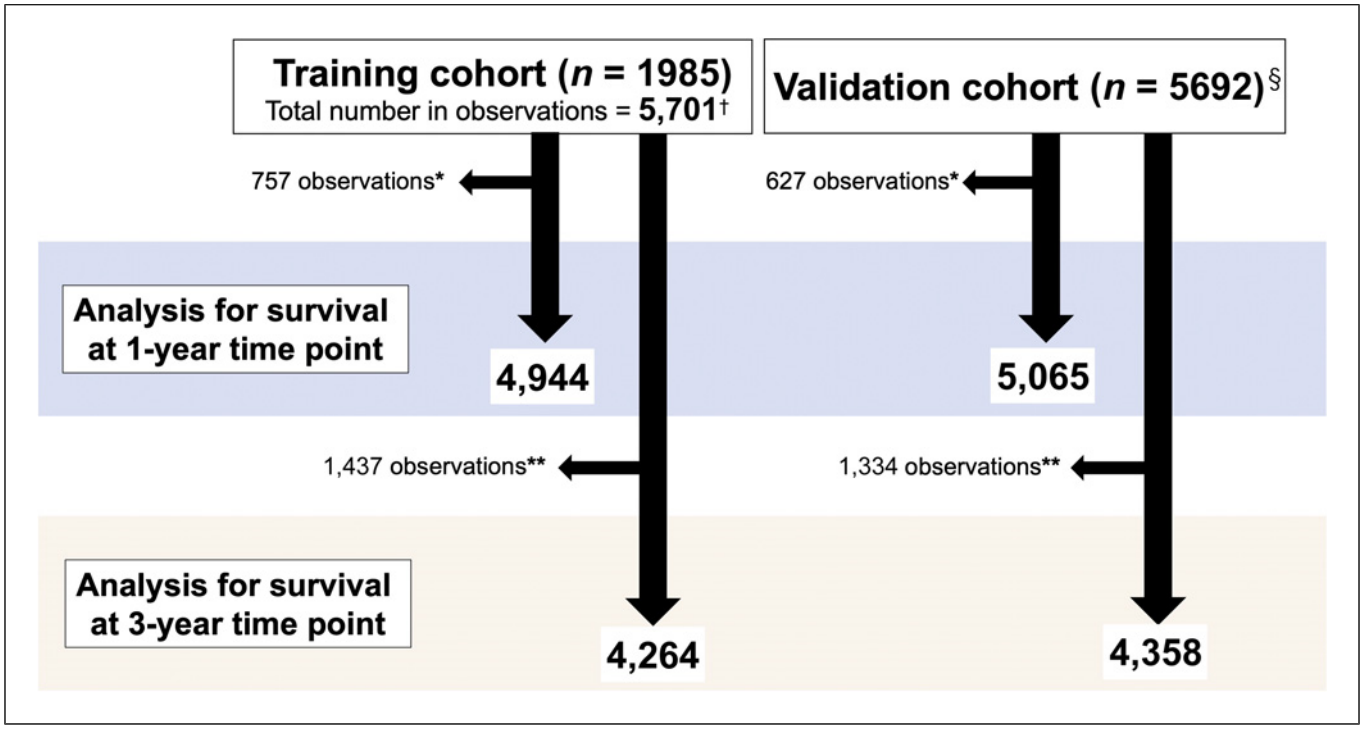

Fig. 1. Flow of patient selection in present study for building PRAID system. †Total dataset from all treatments for initial and recurrent HCC cases in training cohort of 1,985 patients. §Dataset from initial treatments performed for validation cohort. *Dataset of relevant surviving patients with survival period <1 year from last treatment. **Dataset of relevant surviving patients with survival period <3 years from last treatment.

- Basic clinical factors included patient age, gender, body mass index, Eastern Cooperative Oncology Group performance status (ECOG PS), positive for HBV antigen, positive for anti-HCV, alcohol-related liver disease if not positive for HBV antigen or anti-HCV, and hepatic reserve function (ALBI score used as absolute value).
- Clinical factors related to tumors included maximum intrahepatic tumor size (cm) and number (0, 1, 2, 3, 4, 5, 6; >5 or diffuse type considered as 6), cumulative tumor burden (intrahepatic tumor number x maximum intrahepatic tumor size (cm) including relevant treatment), with cumulative number of tumors treated in the past including at the time of relevant treatment also used as an explanatory variable, positive for extrahepatic metastasis, positive for portal vein tumor thrombosis shown by imaging (none, Vp1, Vp2, Vp3, Vp4), positive for AFP (cutoff value 100 ng/mL), fucosylated AFP (AFP-L3) (cutoff value 10%), and positive for PIVKA-II (cutoff value 100 mAU/mL).
- Clinical factors related to therapy included the main therapeutic modality selected (SR, ablation, TACE, systemic treatment, radiotherapy) at the time of the related treatment, cumulative number of past treatments, and period since most recent treatment. For expedience, an intrahepatic tumor number >5 or diffuse type was treated as 6 tumors. When TACE and RFA were given in combination, RFA was the main therapeutic modality.

Because of the existence of nonlinearity in medical data and the corresponding lack of data when utilizing results obtained in clinical practice, as well as the significance of "variable importance" when interpreting analysis results, a machine learning method termed gradient boosting decision tree (GBDT) was utilized. GBDT was performed using the XGBoost software library to decide which branch to assign to missing values during training, a strong point because complements for data that are lacking are not necessary [29–31]. We have not handled autocorrelation in the present analysis because the GBDT model is a nonlinear model, unlike generalized linear model, and it can avoid multicollinearity problems [32]. In this analysis, we used Bayesian optimization and tuned the hyperparameters (online suppl. material; for all online suppl. material, see https://doi.org/10.1159/000530078).

Based on the explanatory variables described above, a model was created to predict the probability of survival after 1 and 3 years, and then values obtained with the Shapley Additive Explanations (SHAP) method were used to analyze the effects of each explanatory variable on the predicted value. SHAP values were evaluated by importing the SHAP module (https://shap.readthedocs.io/en/latest/index.html) into Python, and then the model was trained to minimize binary logarithmic loss (LogLoss). GBDT hyperparameters were determined by tuning the model so as to minimize average LogLoss from a five-part cross-validation. Because the GBDT model in this case minimizes LogLoss, the probabilities are well calibrated from a mathematical point of view [33]. It was confirmed that there is no significant deviation between the prognostic value and the actual establishment in reliability diagram (online suppl.

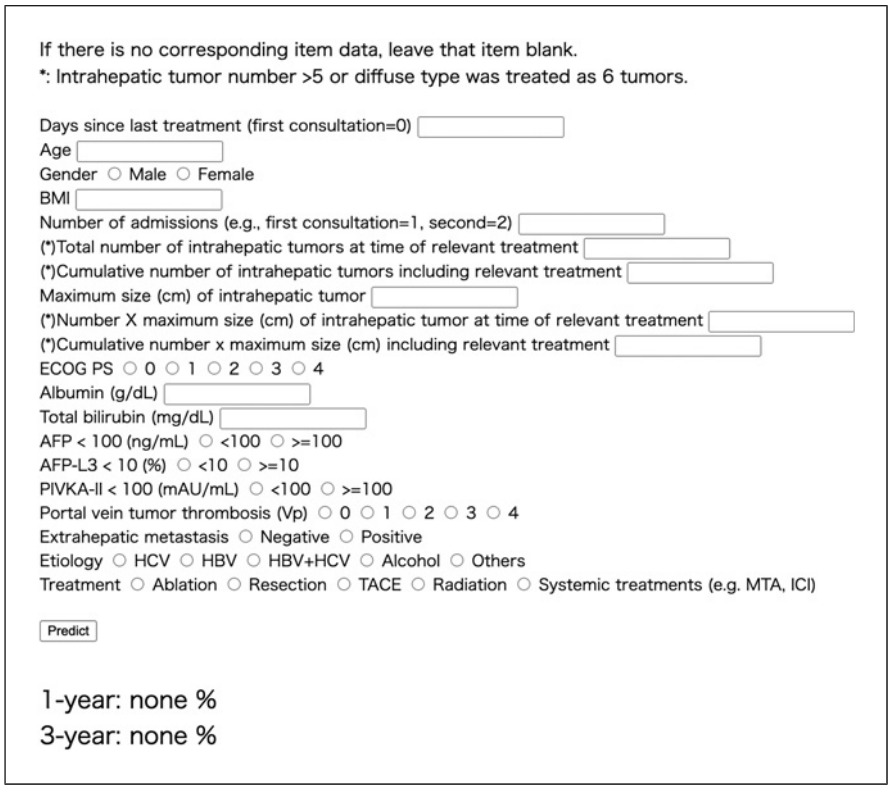

Fig. 2. Interface for PRAID system (https://livercancer.onrender.com/). Instead of albumin-bilirubin (ALBI) score, serum albumin and total bilirubin are used with this system to calculate ALBI score for analysis. BMI, body mass index; ECOG PS, Eastern Cooperative Oncology Group performance status; AFP, alpha-fetoprotein; AFP-L3, fucosylated AFP; PIVKA-II, protein-induced vitamin K absence or antagonist-II; Vp, portal vein tumor thrombosis; HCV, hepatitis C virus; HBV, hepatitis B virus; TACE, transcatheter arterial chemoembolization; MTA, multi-targeting agent; ICI, immune checkpoint inhibitor; BSC, best supportive care.

Fig. 1), and calibration was not performed in the present analysis. The analysis code is available to the public on the web (https://github.com/amapen/LiverCancerAnalysis). Calculations were performed to determine the receiver operating characteristic curve and area under the curve (AUC). When values are shown as median, interquartile range is also presented.

Results

Clinical features of the enrolled 1,985 patients in the training cohort are shown in Table 1. Median age was 71 years, and 1,423 (71.7%) were male. Hepatitis viral infection-related HCC was noted in 1,392 (70.1%). Of 1,985, 379 did not have chronic viral infection or past history of alcohol abuse. Of the 379, nonalcoholic fatty liver disease or nonalcoholic steatohepatitis might be main basal disease in 352, while 27 were due to autoimmune hepatitis (n = 8), primary biliary cholangitis (n = 14), glycogen storage disease (n = 3), Wilson's disease (n = 1), and post-Fontan procedure (n = 1). As for mALBI grade, which was 1 in 751, 2a in 422, 2b in 639, and 3 in 173, curative treatments (liver transplantation, SR, ablation) were performed in 1,255 (63.2%). The median observation period was 29.0 months (interquartile range 6.9–63.8) (Table 1). Clinical features of the 5,692 patients without history of HCC treated at multiple institutions used for validation analysis are shown in Table 2.

The training cohort (n = 1,985), which had 5,701 datasets including from the time of initial diagnosis of HCC and recurrence, were used to make the "PRAID system." For the 1-year survival analysis, 757 of the total 5,701 datasets were excluded because they were alive but had not yet reached 1 year, while 1,437 sets were excluded from 3-year survival because they were alive but had not yet reached 3 years. The validation cohort (n = 5,692) treated at multiple institutions was used for validation testing. Of those, 627 and 1,334 sets were excluded from 1- and 3-year survival, respectively, because they were alive but had not yet reached those time periods (Fig. 1). A GBDT model was adopted for the present study because that was more accurate as compared to a neural network model (data not shown). SHAP values for 1- and 3-year survival were evaluated for highlighting the importance of clinical factors, with the amount of positive or negative impact on the predicted value

Table 1. Clinical features of training cohort at initial treatment including cases lacking some data (n = 1,985)

Age, years*	71 (64–77)
Male (%)	1,423 (71.7)
ECOG PS (0:1:2:3:4:ND)	1,289:229:92:72:35:268
Etiology (HCV:HBV:HBV+HCV:alcohol:others)	1,193:180:19:214:379
ALBI score*	−2.33 (−2.78 to −1.94)
mALBI grade (1:2a:2b:3)	751:422:639:173
AFP, ng/mL*	17.2 (5.6–153.8) (ND 17)
AFP-L3, %*	3.7 (0.5–17.6) (ND 223)
PIVKA-II, mAU/mL*	103 (29–1,331) (ND 39)
Intrahepatic tumor size, maximum, cm*	2.8 (1.8–5.0)
Number of intrahepatic tumors. 1:2:3:4:5:≥5 or diffuse type	1241:334:105:57:37:211
Vp (0:1:2:3:4)	1,790:21:53:62:59
EHM positive, n (%)	106 (5.3)
TNM-LCSGJ 6th (I:II:III:IV)	427:855:430:273
Initial treatment (LT:SR:ablation:TACE:chemotherapies[†]:others including BSC)	1:597:658:365:71:293 ([†]6:7:7:41:1:9)
Treatment for recurrence (SR:ablation:TACE:chemotherapy[††]:others including BSC)	147:933:2,146:425:65 ([††]31:52:136:9:13:6:160:12:6)
Number of treatments per patient*	2 (1–5)
Observation period, months*	29.0 (6.9–63.8)

ECOG PS, Eastern Cooperative Oncology Group performance status; HCV, hepatitis C virus; HBV, hepatitis B virus; ND, no data patients' number; ALBI score, albumin-bilirubin score; mALBI grade, modified ALBI grade; AFP, alpha-fetoprotein; AFP-L3, fucosylated AFP; PIVKA-II, protein-induced vitamin K absence or antagonist-II; Vp, portal vein tumor thrombosis; EHM, extrahepatic metastasis; TNM-LCSGJ 6th, tumor-node-metastasis stage of Liver Cancer Study Group of Japan, 6th edition; LT, liver transplantation; SR, surgical resection; TACE, transcatheter arterial chemoembolization; chemotherapy, chemotherapy including systemic treatments; HAIC, hepatic arterial infusion chemotherapy; ICI, immune checkpoint inhibitor; BSC, best supportive care. *Median (interquartile range).
[†]Atezolizumab plus bebacizumab:lenvatinib:sorafenib:HAIC:clinical trial:others.
[††]Atezolizumab plus bebacizumab:lenvatinib:sorafenib:regorafenib:ramucirumab:cabozantinib:HAIC:clinical trial with ICI:others.

shown in Figure 2. In Figure 3, colors from red to blue are linked to greater or lesser value of each explanatory variable.

Following exclusion of patients lacking clinical data related to Child-Pugh score, ALBI score, and/or TNM-LCSGJ 6th stage, prognoses of those in the training and validation cohorts according to JIS, ALBI-T, and mALBI-T scores are shown in Figure 4, with median survival time results presented in Table 3. Both cohorts showed similar overall survival for each score with the integrated scoring system.

After dividing the training cohort into 2 groups (training group [n = 3,955] and test group [n = 989]) randomly to develop the present system with AI, PRAID system was developed. The AUC value for prediction of 1-year survival was 0.875 (confusion matrix [CM]: true negative [TN] 100, false positive [FP] 115, false negative [FN] 29, true positive [TP] 745) (sensitivity 0.963 [95% CI: 0.947–0.974], specificity 0.465 [95% CI: 0.400–0.532]) (Fig. 5a) and that for prediction of 3-year survival was also 0.885 (CM: TN 365, FP 83, FN 80, TP 325) (sensitivity 0.802 [95% CI: 0.761–0.838], specificity 0.815 [95% CI: 0.776–0.848]) (Fig. 5b), while those values for the validation cohort were 0.841 (CM: TN 467, FP 437, FN 280, TP 3,881) (sensitivity 0.933 [95% CI: 0.925–0.940], specificity 0.517 [95% CI: 0.484–0.549]) (Fig. 5c), 0.796 (CM: TN 1,174, FP 642, FN 494, TP 2,048) (sensitivity 0.806 [95% CI: 0.790–0.821], specificity 0.646 [95% CI: 0.624–0.668]) (Fig. 5d), respectively. When the subanalysis of prediction of prognosis for patients at initial treatments of the training cohort was performed, the AUC value for prediction of 1-year survival was 0.928 [CM: TN 175, FP 121, FN 24, TP 1,303] (sensitivity 0.982 [95% CI: 0.973–0.988], specificity 0.591 [95% CI: 0.534–0.646]) and that for prediction of 3-year survival was also 0.913 (CM: TN 265, FP 31, FN 314, TP 1,013) (sensitivity 0.763 [95% CI: 0.740–0.785], specificity 0.895 [95% CI: 0.855–0.925]).

Discussion

In addition to hepatic reserve function and tumor burden, frequent recurrence of HCC over a short period is often noted by clinicians even after performing curative

Table 2. Clinical features of HCC patients from multiple institutions at time of initial treatment used for validation analysis including cases lacking complete data (n = 5,692)

Age, years*	70 (63–77)
Male (%)	4,244 (74.6)
ECOG PS (0:1:2:3:4:ND)	3,535:385:150:100:17:1,505
Etiology (HCV:HBV:HBV+HCV:alcohol:others:ND)	3,193:775:44:464:1,205:11
ALBI score*	−2.36 (−2.74 to −1.93)
mALBI grade (1:2a:2b:3)	2,006:1,162:2,069:455
AFP, ng/mL*	15.0 (5.3–121.0) (ND 518)
AFP-L3, %*	2.2 (0.5–16.9) (ND 1,850)
PIVKA-II, mAU/mL*	59.3 (21.0–774.5) (ND 990)
Intrahepatic tumor size, maximum size, cm*	2.7 (1.8–5.0)
Number of intrahepatic tumors (1:2:3:4:5:≥6 + diffuse type)	3,106:1,126:538:214:214:494
Vp (0:1:2:3:4)	5,166:105:89:186:146
EHM positive, n (%)	302 (6.5)
TNM-LCSGJ 6th (I:II:III:IV)	1,318:2,132:1,470:772
Treatment (SR:ablation:TACE:chemotherapy†:other including BSC:ND)	1,320:1,999:1,522:256:587:8 (†1:24:31:197:3:0)
Observation period, months*	31.2 (10.8–64.8)

ECOG PS, Eastern Cooperative Oncology Group performance status; HCV, hepatitis C virus; HBV, hepatitis B virus; ND, no data patients' number; ALBI score, albumin-bilirubin score; mALBI grade, modified ALBI grade; AFP, alpha-fetoprotein; AFP-L3, fucosylated AFP; PIVKA-II, protein-induced vitamin K absence or antagonist-II; Vp, portal vein tumor thrombosis; EHM, extrahepatic metastasis; TNM-LCSGJ 6th, tumor-node-metastasis stage of Liver Cancer Study Group of Japan, 6th edition; LT, liver transplantation; SR, surgical resection; TACE, transcatheter arterial chemoembolization; chemotherapy, chemotherapy including systemic treatments; HAIC, hepatic arterial infusion chemotherapy; ICI, immune checkpoint inhibitor; BSC, best supportive care.
*Median (interquartile range). †Atezolizumab plus bebacizumab:lenvatinib:sorafenib:HAIC:clinical trial with ICI:others.

treatments. The present PRAID system provided good prediction of prognosis in the present training cohort for cases with initial as well as recurrent tumors. Although the validation group only included clinical data obtained at the time of the initial diagnosis, validation analysis of that large cohort treated at multiple institutions with the present system showed similar predictive prognostic values. As a result, comparable prognostic probability values for 1- and 3-year survival were found for both cohorts (AUC $\geqq$0.796).

Progression in development of a surveillance system for HCC in high-risk patients as well as therapeutic modalities (e.g., SR, ablation, TACE, systemic chemotherapy) has contributed to improve the prognosis of affected patients in Japan, with median survival period shown to be the best worldwide [34]. In addition, clinical practice therapeutic guidelines and algorithms have been proposed [8–14], though those are based on clinical evidence following initial treatments of HCC patients. Existing guidelines list several treatment options for a single category; the response and outcome to each treatment option should be different because each patient has different hepatic reserve function conditions, tumor factors, clinical recurrence course. Although a few recent publications have noted that SR and RFA can serve as curative therapeutic options with similar efficacy for recurrent HCC [35, 36], treatment selection remains very difficult for clinicians at any time of the clinical course, as aggressive treatment is often a double-edged sword, as the patient might have potential for rapid progression of HCC or possibility of a decline of hepatic function. Therefore, an ability to refer to the availability of short- and medium-term prognostic indicators would be very useful for determining the best treatment for individual cases.

Machine learning provides a technological foundation for AI [37], especially deep learning using multilayer neural networks. Deep learning is a very powerful learning algorithm that has produced revolutionary results in a variety of fields [38], including the medical imaging field, where a convolutional neural network architecture is being used for research and development [39]. Prognostic and course prediction models for different diseases can also be created based on a variety of factors [40–44]. On the other hand, in the field of medicine, it is difficult to collect data from a large group of patients (e.g., tens of thousands) due to need to obtain consent and ethical considerations.

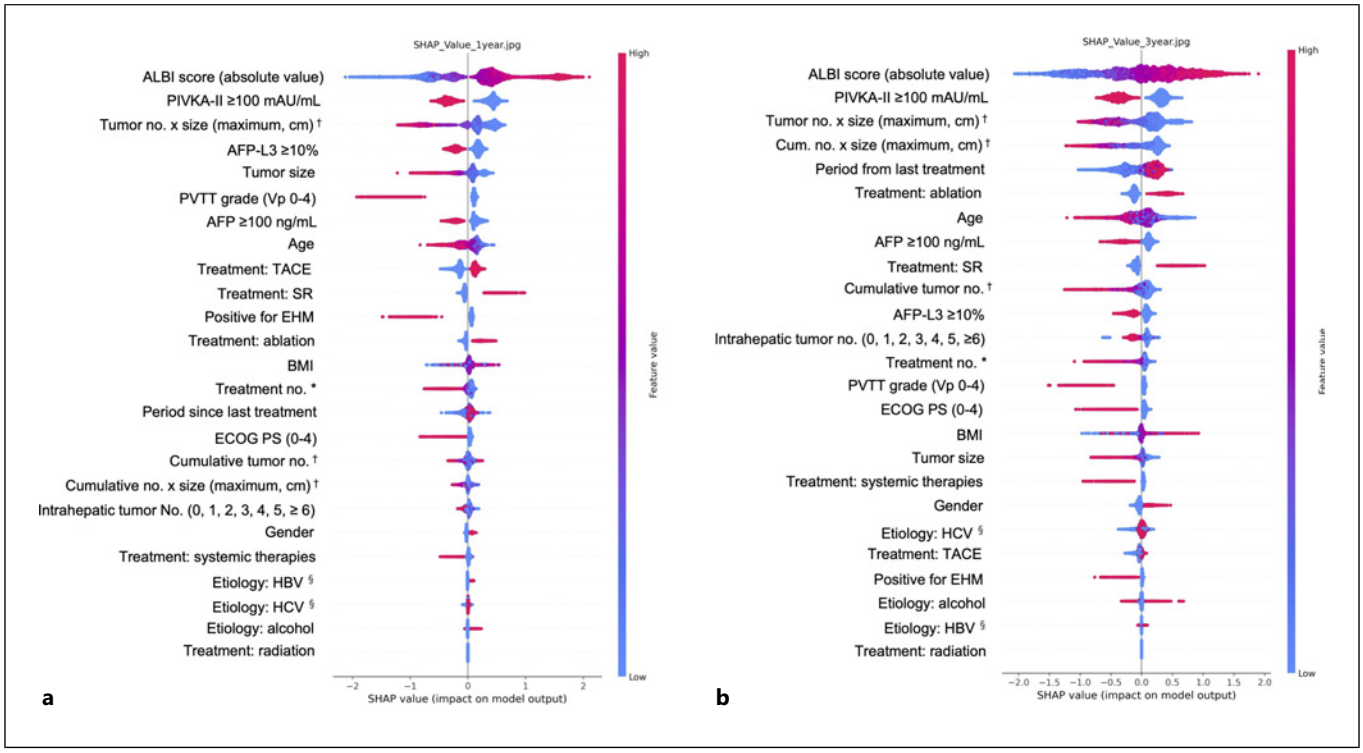

Fig. 3. a, **b** SHAP value at time of treatment for 1- and 3-year survival. † Number of tumors (0, 1, 2, 3, 4, 5, 6) and/or size of intrahepatic tumor including relevant treatment were used. Intrahepatic tumor number >5 or diffuse type was treated as 6 tumors. Cum: cumulative. * Total treatment number including relevant treatment. § When patients were double-positive for HBsAg and anti-HCV, both factors were inputted as positive into each box. HBsAg, hepatitis B virus antigen.

Therefore, we used a GBDT method [45–47] in the present study, as the sample size was limited. For reference, in this analysis, GBDT, neural networks, and generalized linear model were used to create models and compare the AUC of the training cohort, with GBDT showing the best results (data not shown).

At the time of diagnosis of HCC recurrence, many patients have already undergone several prior treatments for a recurrent condition during their clinical course. Several clinicians have noted that the period from the most recent recurrence becomes shorter over time. This phenomenon may be related to the characteristics of HCC, as affected patients often have intrahepatic metastasis, and a cumulative tumor burden might have a significant effect on prognosis. Indeed, the present findings indicated that previous cumulative tumor burden and period from the most recent treatment can have an impact on prognosis in decision-making for therapeutic modality. Pantel et al. [48] reported that early recurrence of HCC is often affected by circulating tumor cells (CTCs). More recently, cancer stem cells (CSCs) among CTCs have been hypothesized to have an important role in recurrence and metastasis [49], while another study reported that preoperative measurement of blood CSC level can help with prediction of recurrence after SR [50]. However, there are no established clinical methods for detection of CTCs or CSCs. Cumulative tumor burden and period from the most recent treatment at the time of recurrence might be alternative factors related to the existence of those cell types. The present PRAID system developed with AI included important clinical factors such as cumulative tumor burden (intrahepatic tumor number x intrahepatic tumor size) and is considered to be a useful decision-making tool regarding therapeutic modalities for HCC recurrence.

This study has some limitations. First, since 2005, treatment has mainly followed the guidelines of the Japanese Society of Hepatology (JSH) [8–10], but the training cohort whose entire clinical course was analyzed belonged to a single institution. Second, the number of enrolled patients was thought not to be

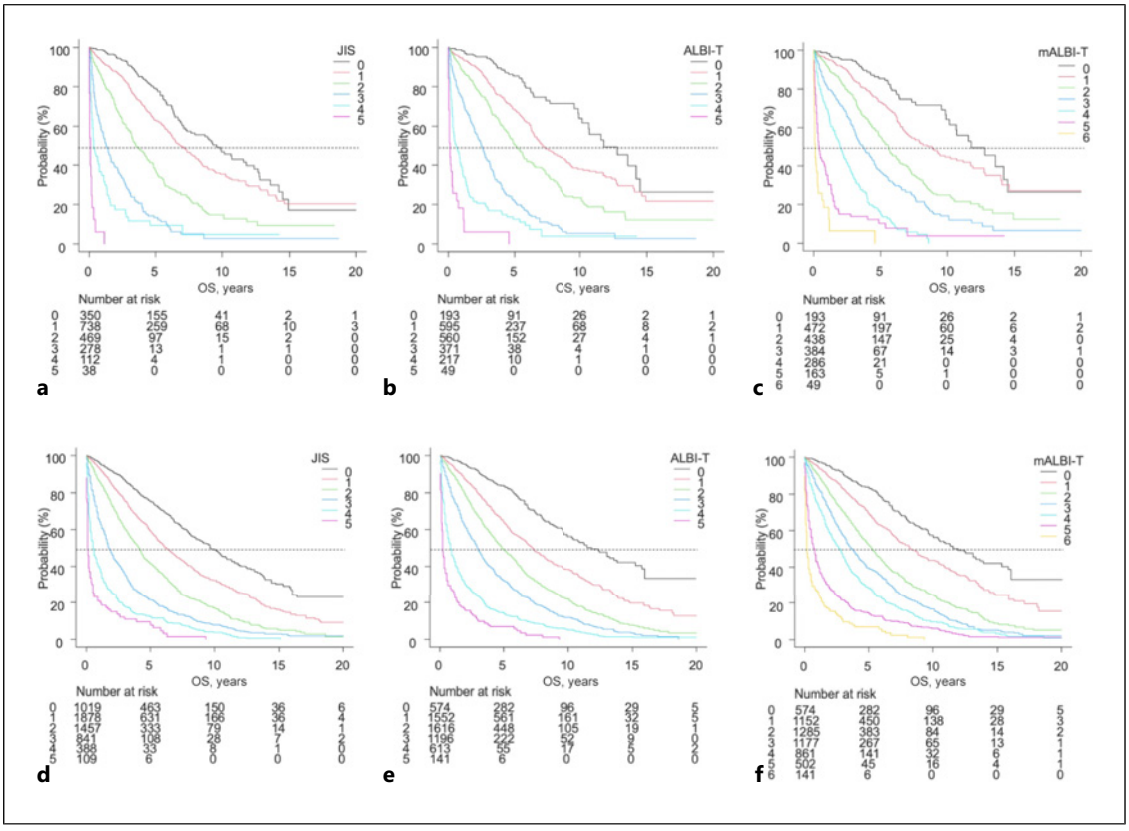

Fig. 4. a–f Survival period (years) of training and validation cohorts according to JIS, ALBI-T, and mALBI-T scores. JIS, Japan integrated scoring; ALBI-T, albumin-bilirubin (ALBI) grade and TNM stage of Liver Cancer Study Group of Japan 6th (TNM-LCSGJ 6th); mALBI-T, modified ALBI and TNM-LCSGJ 6th. **a**–**c** Training cohort. **d**–**f** Validation cohort.

Table 3. Prognoses according to JIS, ALBI-T, and mALBI-T in training and validation cohorts

JIS score	Score 0	Score 1	Score 2	Score 3	Score 4	Score 5	
Training	9.6 (7.3–11.8)	6.9 (6.2–7.7)	3.5 (3.2–4.3)	1.3 (1.0–1.7)	0.4 (0.3–0.7)	0.09 (0.04–0.18)	
Validation	9.7 (8.8–10.6)	6.2 (5.8–6.7)	4.1 (3.7–4.3)	1.7 (1.6–2.0)	0.5 (0.4–0.7)	0.1 (0.1–0.2)	
ALBI-T score	Score 0	Score 1	Score 2	Score 3	Score 4	Score 5	
Training	11.8 (10.2–14.5)	7.1 (6.5–8.4)	5.1 (4.6–5.6)	2.5 (2.1–2.8)	0.6 (0.4–0.9)	0.13 (0.08–0.29)	
Validation	11.9 (10.0–13.9)	7.2 (6.7–7.8)	4.8 (4.5–5.2)	2.9 (2.7–3.2)	0.8 (0.7–1.0)	0.2 (0.17–0.3)	
mALBI-T score	Score 0	Score 1	Score 2	Score 3	Score 4	Score 5	Score 6
Training	11.8 (10.2–14.5)	8.5 (7.1–10.7)	5.6 (4.9–6.2)	3.8 (3.2–4.4)	2.0 (1.6–2.4)	0.4 (0.3–0.8)	0.13 (0.08–0.29)
Validation	11.9 (10.0–13.9)	8.2 (7.4–9.1)	5.2 (4.9–5.6)	3.7 (3.3–4.1)	2.7 (2.4–3.0)	0.2 (0.17–0.3)	0.14 (0.08–0.19)

Values indicate median survival, years (95% CI). JIS score, Japan integrated scoring; ALBI-T score, albumin-bilirubin (ALBI) grade and tumor-node-metastasis stage of Liver Cancer Study Group of Japan, 6th edition (TNM-LCSGJ 6th); mALBI-T score, modified ALBI and TNM-LCSGJ 6th.

yet sufficient. Third, while the validation cohort was treated at multiple centers, their clinical data were only from the time of the initial treatment. Fourth, few of the enrolled patients underwent liver transplantation, as that is rare in Japan because of a lack of donors. Fifth, although the present analysis was based on the

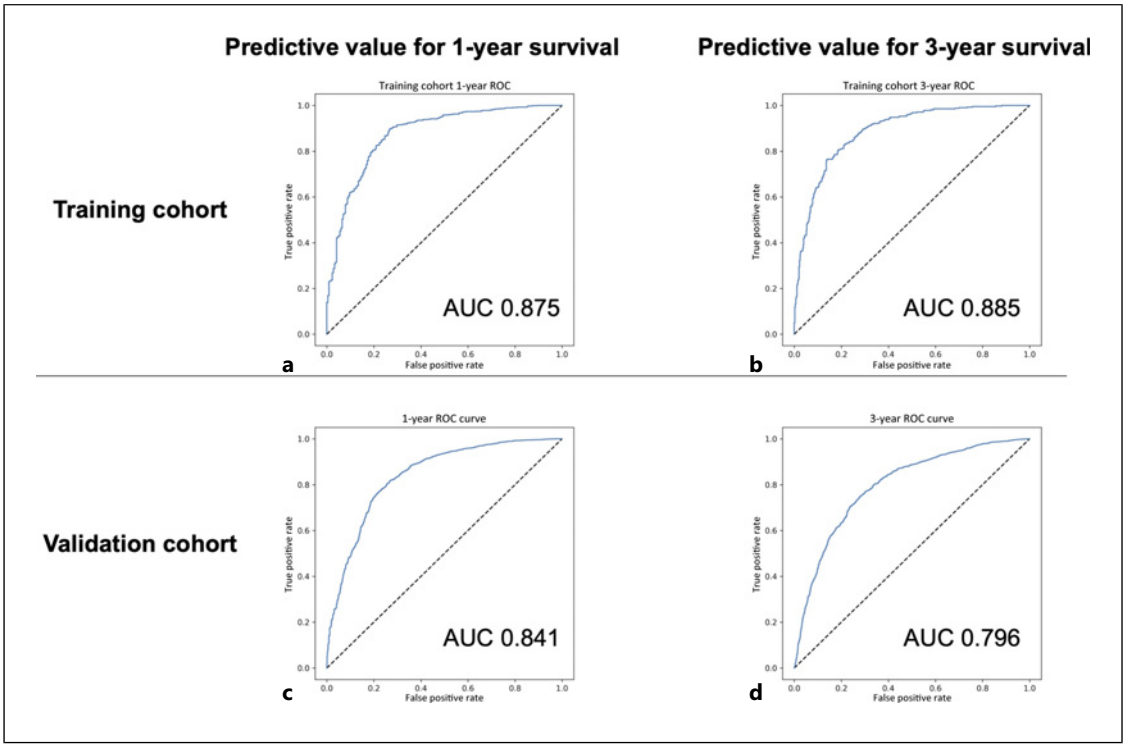

Fig. 5. Receiver operating characteristic curves for 1- and 3-year survival. In the training cohort, the area under the curve (AUC) for prediction of 1-year survival was 0.875 (sensitivity 0.963 [95% CI: 0.947–0.974], specificity 0.465 [95% CI: 0.400–0.532]) (**a**), while that for 3-year survival was also 0.885 (sensitivity 0.802 [95% CI: 0.761–0.838], specificity 0.815 [95% CI: 0.776–0.848]) (**b**). Using the validation cohort, the AUC for prediction of 1-year survival was 0.841 (sensitivity 0.933 [95% CI: 0.925–0.940], specificity 0.517 [95% CI: 0.484–0.549]) (**c**) and for 3-year survival was 0.796 (sensitivity 0.806 [95% CI: 0.790–0.821], specificity 0.646 [95% CI: 0.624–0.668]) (**d**).

same concept as the Kaplan-Meier method and eliminated patients who were lost to follow-up at 1 and 3 years, there may be a bias due to exclusion of the lost to follow-up patients at these times. Sixth, recent developments in laparoscopic SR methods [51] and future development of systemic chemotherapy will likely have a large impact on prognosis in recurrent HCC cases. Finally, it should also be noted that interpretation of answers obtained with AI by physicians as well as patients is an issue that should be carefully discussed, as that is directly related to concerns regarding responsibility and ethics. It is recommended that the present system be reanalyzed and restructured every 3 to 5 years according to progression of available treatments for HCC so as to obtain greater accuracy for prediction. It may also be important to determine whether a less invasive treatment should be selected when different treatments in the PRAID system predict a similar prognosis.

In conclusion, the present proposed PRAID system developed with AI is considered to provide useful prognostic information for decision-making regarding therapeutic modalities at the time of initial diagnosis as well as recurrence. Of course, this is a prototype system for prognostication at this time, but we think it shows the potential for AI to create optimized therapeutic algorithms in the future. A validation study with other cohorts including data from examinations at initial and recurrent stages will be necessary.

Acknowledgments

Analyses using the AI system were performed by Amane Saito (Amatechno Co., Ltd., Fukuoka, Japan), who also serves as a data analyst for the Softbank Hawks professional baseball team (Fukuoka, Japan) (e-mail: mynameisamane@gmail.com).

Statement of Ethics

The entire study protocol was approved by the Institutional Ethics Committee of Ehime Prefectural Central Hospital (No. 04-06). After receiving official approval, this study was conducted as a retrospective analysis of database records based on the Guidelines for Clinical Research issued by the Ministry of Health and Welfare of Japan. All procedures were done in accordance with the Declaration of Helsinki. The data were made anonymous before

analysis to protect patient privacy. Written informed consent was obtained from all patients before treatment, and this study received ethical approval for use of an opt-out methodology based on low risk to the participants.

Conflict of Interest Statement

Atsushi Hiraoka, MD, PhD, received lecture fees from Chugai, Bayer, and Eli Lilly. None of the other authors have potential conflicts of interest to declare.

Funding Sources

No funding was received for this study.

Author Contributions

A.H. and T.K. conceived the study, drafted the text, and participated in its design and coordination. A.H., T.T., H.T., K.K., T.H., S.K., A.N., E.I., K.T., T.I., H.O., F.T., and K.N. performed data curation. All authors have read and approved the final version of the manuscript.

Data Availability Statement

Due to the nature of this research, the participants could not be contacted regarding whether the findings could be shared publicly; thus, supporting data, including datasets generated and/or analyzed for the current study, are not publicly available. Further inquiries can be directed to the corresponding author.

References

1 Miyagawa S, Makuuchi M, Kawasaki S, Kakazu T. Criteria for safe hepatic resection. Am J Surg. 1995 Jun;169(6):589–94.
2 Shiina S, Tateishi R, Arano T, Uchino K, Enooku K, Nakagawa H, et al. Radiofrequency ablation for hepatocellular carcinoma: 10-year outcome and prognostic factors. Am J Gastroenterol. 2012 Apr;107(4):569–77. quiz 78.
3 Takayasu K, Arii S, Ikai I, Omata M, Okita K, Ichida T, et al. Prospective cohort study of transarterial chemoembolization for unresectable hepatocellular carcinoma in 8510 patients. Gastroenterology. 2006 Aug;131(2):461–9.
4 Galle PR, Finn RS, Qin S, Ikeda M, Zhu AX, Kim TY, et al. Patient-reported outcomes with atezolizumab plus bevacizumab versus sorafenib in patients with unresectable hepatocellular carcinoma (IMbrave150): an open-label, randomised, phase 3 trial. Lancet Oncol. 2021 Jul;22(7):991–1001.
5 Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med Overseas Ed. 2008 Jul 24;359(4):378–90.
6 Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet. 2018 Feb 9;391(10126):1163–73.
7 Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021 May;71(3):209–49.
8 Makuuchi M, Kokudo N, Arii S, Futagawa S, Kaneko S, Kawasaki S, et al. Development of evidence-based clinical guidelines for the diagnosis and treatment of hepatocellular carcinoma in Japan. Hepatol Res. 2008 Jan;38(1):37–51.
9 Kokudo N, Hasegawa K, Akahane M, Igaki H, Izumi N, Ichida T, et al. Evidence-based clinical practice guidelines for hepatocellular carcinoma: the Japan society of Hepatology 2013 update (3rd JSH-HCC guidelines). Hepatol Res. 2015 Jan;45(2).
10 Kudo M, Kawamura Y, Hasegawa K, Tateishi R, Kariyama K, Shiina S, et al. Management of hepatocellular carcinoma in Japan: JSH consensus statements and recommendations 2021 update. Liver cancer. 2021 Jun;10(3):181–223.
11 Yau T, Tang VY, Yao TJ, Fan ST, Lo CM, Poon RT. Development of Hong Kong Liver Cancer staging system with treatment stratification for patients with hepatocellular carcinoma. Gastroenterology. 2014 Jun;146(7):1691–700. e3.
12 Omata M, Cheng AL, Kokudo N, Kudo M, Lee JM, Jia J, et al. Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. Hepatol Int. 2017 Jul;11(4):317–70.
13 Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, et al. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. J hepatology. 2022 Mar;76(3):681–693.
14 Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, et al. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. J Hepatol. 2022 Mar;76(3):681–93.
15 Kudo M, Chung H, Haji S, Osaki Y, Oka H, Seki T, et al. Validation of a new prognostic staging system for hepatocellular carcinoma: the JIS score compared with the CLIP score. Hepatology. 2004 Dec;40(6):1396–405.
16 Hiraoka A, Kumada T, Kudo M, Hirooka M, Tsuji K, Itobayashi E, et al. Albumin-Bilirubin (ALBI) grade as part of the evidence-based clinical practice guideline for HCC of the Japan society of Hepatology: a comparison with the liver damage and child-pugh classifications. Liver cancer. 2017 Jun;6(3):204–15.
17 Hiraoka A, Michitaka K, Kumada T, Izumi N, Kadoya M, Kokudo N, et al. Prediction of prognosis of intermediate-stage HCC patients: validation of the tumor marker score in a nationwide database in Japan. Liver cancer. 2019 Oct;8(5):403–11.
18 Kudo M. A new era of systemic therapy for intermediate and advanced stage hepatocellular carcinoma. Hepatobiliary Surg Nutr. 2020 Aug;9(4):530–3.
19 European Association for the Study of Liver. EASL clinical practical guidelines: management of alcoholic liver disease. J Hepatol. 2012 Aug;57(2):399–420.
20 Mittal S, El-Serag HB. Epidemiology of hepatocellular carcinoma: consider the population. J Clin Gastroenterol. 2013 Jul;47 Suppl(0):S2–6.
21 Johnson PJ, Berhane S, Kagebayashi C, Satomura S, Teng M, Reeves HL, et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach-the ALBI grade. J Clin Oncol. 2015 Feb 20;33(6):550–8.
22 Hiraoka A, Kumada T, Michitaka K, Toyoda H, Tada T, Ueki H, et al. Usefulness of albumin-bilirubin grade for evaluation of prognosis of 2584 Japanese patients with hepatocellular carcinoma. J Gastroenterol Hepatol. 2016 May;31(5):1031–6.
23 Hiraoka A, Michitaka K, Kumada T, Izumi N, Kadoya M, Kokudo N, et al. Validation and potential of albumin-bilirubin grade and prognostication in a nationwide survey of 46,681 hepatocellular carcinoma patients in Japan: the need for a more detailed evaluation of hepatic function. Liver cancer. 2017 Nov;6(4):325–36.

24 Bruix J, Sherman M. Management of hepatocellular carcinoma. Hepatology. 2005 Nov; 42(5):1208–36.
25 Di Martino M, Marin D, Guerrisi A, Baski M, Galati F, Rossi M, et al. Intraindividual comparison of gadoxetate disodium-enhanced MR imaging and 64-section multidetector CT in the Detection of hepatocellular carcinoma in patients with cirrhosis. Radiology. 2010 Sep;256(3):806–16.
26 Sano K, Ichikawa T, Motosugi U, Sou H, Muhi AM, Matsuda M, et al. Imaging study of early hepatocellular carcinoma: usefulness of gadoxetic acid-enhanced MR imaging. Radiology. 2011 Dec;261(3):834–44.
27 Llovet JM, Villanueva A, Marrero JA, Schwartz M, Meyer T, Galle PR, et al. Trial design and endpoints in hepatocellular carcinoma: AASLD consensus conference. Hepatology. 2021 Jan;73(Suppl 1):158–91.
28 The Liver Cancer Study Group of Japan. The general rules for the clinical and pathological study of primary liver cancer. 6th ed. Tokyo: Kanehara; 2015. p. 26.
29 Rusdah DA, Murfi H. XGBoost in handling missing values for life insurance risk prediction. SN Appl Sci. 2020 2020/07/06;2(8):1336.
30 Latief MA, Bustamam A, Siswantining T. Performance evaluation XGBoost in handling missing value on classification of hepatocellular carcinoma gene expression data. 20204th international conference on informatics and computational sciences (ICICoS)2020. p. 1–6.
31 Aydin ZE, Ozturk ZK. Performance analysis of XGBoost classifier with missing data. Manch J Artif Intell Appl Sci (MJAIAS). 2021;2(02):2021.
32 Yang L, Liang Y, Zhu Q, Chu X. Machine learning for inference: using gradient boosting decision tree to assess non-linear effects of bus rapid transit on house prices. Ann GIS. 2021;27(3):273–84.
33 Cearns M, Hahn T, Clark S, Baune BT. Machine learning probability calibration for high-risk clinical decision-making. Aust N Z J Psychiatry. 2020 Feb;54(2):123–6.
34 Kudo M. Surveillance, diagnosis, treatment, and outcome of liver cancer in Japan. Liver cancer. 2015;4(1):39–50.
35 Song KD, Lim HK, Rhim H, Lee MW, Kim YS, Lee WJ, et al. Repeated hepatic resection versus radiofrequency ablation for recurrent hepatocellular carcinoma after hepatic resection: a propensity score matching study. Radiology. 2015 May;275(2):599–608.
36 Ohama H, Hiraoka A, Tada F, Kato K, Fukunishi Y, Yanagihara E, et al. Comparison of surgical resection and percutaneous ultrasonographic guided radiofrequency ablation for initial recurrence of hepatocellular carcinoma in early stage following curative treatment. Cancers. 2022 Nov 10;14(22):5524. in press.
37 Wang S, Summers RM. Machine learning and radiology. Med Image Anal. 2012 Jul;16(5): 933–51.
38 LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015 May 28;521(7553):436–44.
39 Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017 Feb 2;542(7639):115–8.
40 Kurasawa H, Hayashi K, Fujino A, Takasugi K, Haga T, Waki K, et al. Machine-learning-based prediction of a missed scheduled clinical appointment by patients with diabetes. J Diabetes Sci Technol. 2016 May;10(3):730–6.
41 Sato M, Morimoto K, Kajihara S, Tateishi R, Shiina S, Koike K, et al. Machine-learning approach for the development of a novel predictive model for the diagnosis of hepatocellular carcinoma. Sci Rep. 2019 May 30; 9(1):7704.
42 Hsu TMH, Schawkat K, Berkowitz SJ, Wei JL, Makoyeva A, Legare K, et al. Artificial intelligence to assess body composition on routine abdominal CT scans and predict mortality in pancreatic cancer- A recipe for your local application. Eur J Radiol. 2021 Sep;142:109834.
43 Ren Q, Zhu P, Li C, Yan M, Liu S, Zheng C, et al. Pretreatment computed tomography-based machine learning models to predict outcomes in hepatocellular carcinoma patients who received combined treatment of trans-arterial chemoembolization and tyrosine kinase inhibitor. Front Bioeng Biotechnol. 2022;10:872044.
44 Liu Z, Li H, Dang Q, Weng S, Duo M, Lv J, et al. Deep learning for prediction of hepatocellular carcinoma recurrence after resection or liver transplantation: a discovery and validation study. Cell Mol Life Sci. 2022 Jun; 79(11):577–89.
45 Kruse C, Eiken P, Vestergaard P. Machine learning principles can improve hip fracture prediction. Calcif Tissue Int. 2017 Apr; 100(4):348–60.
46 Taylor RA, Moore CL, Cheung KH, Brandt C. Predicting urinary tract infections in the emergency department with machine learning. PLoS One. 2018;13(3):e0194085.
47 Babajide Mustapha I, Saeed F. Bioactive molecule prediction using extreme gradient boosting. Molecules. 2016 Jul 28;21(8):983.
48 Pantel K, Alix-Panabières C, Riethdorf S. Cancer micrometastases. Nat Rev Clin Oncol. 2009 Jun;6(6):339–51.
49 Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. Nature. 2001 Nov 1;414(6859):105–11.
50 Fan ST, Yang ZF, Ho DW, Ng MN, Yu WC, Wong J. Prediction of posthepatectomy recurrence of hepatocellular carcinoma by circulating cancer stem cells: a prospective study. Ann Surg. 2011 Oct;254(4):569–76.
51 Kaibori M, Hiraoka A, Matsui K, Matsushima H, Kosaka H, Yamamoto H, et al. Predicting complications following surgical resection of hepatocellular carcinoma using newly developed neo-glasgow prognostic score with ALBI grade: comparison of open and laparoscopic surgery cases. Cancers. 2022 Mar 9; 14(6):1402.