








Artificial intelligence method to predict overall survival of hepatocellular carcinoma

 Cem Simsek¹,  Deniz Can Guven²,  Taha Koray Sahin³,  Ibrahim Emir Tekin³,  Ozlem Sahan³,
 Hatice Yasemin Balaban¹,  Suayib Yalcin²

¹Department of Gastroenterology, Hacettepe University School of Medicine, Ankara, Turkey; ²Department of Medical Oncology, Hacettepe University School of Medicine, Ankara, Turkey; ³Department of Internal Medicine, Hacettepe University School of Medicine, Ankara, Turkey

Abstract

Background and Aim: Hepatocellular carcinoma (HCC) is a complex disease with heterogeneous outcomes influenced by disease- and patient-related factors. The prediction of outcomes requires a comprehensive approach, and artificial intelligence could provide a feasible means of estimating HCC outcomes. This study was designed to assess the viability of a machine learning model to predict survival in HCC patients.

Materials and Methods: HCC patient data with at least 5 years of follow-up were retrospectively reviewed. Patients with accessible data on the primary liver disease, tumor and laboratory values at the time of diagnosis, and length of survival were included. A gradient boosting machine learning algorithm was constructed to predict patient survival at 6 time points.

Results: A total of 100 HCC patients (80% male) with a median overall survival of 43 months (range: 0.7–256 months) were included. The survival rate for 6, 12, 24, 36, 60, and 120 months was 88%, 81%, 67%, 60%, 40%, and 11%, respectively. The mean area under the curve of the model prediction was 0.92 (0.061) for >6 months, 0.81 (0.107) for >1 year, 0.78 (0.11) for >2 years, 0.81 (0.083) for >3 years, 0.82 (0.079) for >5 years, 0.81 (0.96) for >8 years, and 0.66 (0.14) for >10 years.

Conclusion: The machine learning model successfully predicted short- and long-term survival of patients with HCC.

Keywords: Artificial intelligence; hepatocellular carcinoma; survival.

Introduction

Hepatocellular cancer (HCC) is the sixth most common cancer and the third leading cause of cancer death globally.^[1] In addition, the mortality and incidence rates are expected to increase in the next 10 years.^[2] ^{3]} Therefore, new management strategies are needed using approaches

guided by the genomic, molecular, and other features of the tumor, as well as patient characteristics and hepatic functional reserve.

The term artificial intelligence refers to computer programs that try to reproduce human cognitive functions, such as learning and problem solving. Machine learning was originally developed as a branch of artificial intelligence and can be defined as algorithms that create predictive models and detect behavior patterns. Machine learning methods are used extensively in virtually every field of medicine and are now also used to address areas with knowledge gaps in HCC, such as carcinogenesis, screening, non-invasive diagnosis, post-treatment recurrence, and treatment selection. A greater ability to predict outcomes and survival in HCC patients is an essential step to consolidate our knowledge and guide case management. It has thus far been relatively understudied in HCC and deserves our attention.

The objective of this study was to test the ability of a machine learning model to predict survival probability in HCC patients for different time points. The research hypothesis was that the model would provide acceptable predictions of survival at specified time points.

Materials and Methods

This research was approved by the Hacettepe University Institutional Review Board on March 20201 (GO 21/351).

Patient Selection

HCC cases from the institutional database of a tertiary referral university hospital in Turkey were retrospectively reviewed. Patients with histologically or clinically proven HCC and initial radiological, clinical, laboratory data, and survival documentation were included, and patients with an indefinite final diagnoses of cirrhosis or HCC were excluded. In order to assess 5-year overall survival, patients who were diagnosed before January 2015 were included. All of the cases that met the criteria were included in order to test the model's maximum performance.

Data Collection and Variables

Cases from the hospital electronic medical database with relevant International Classification of Diseases (ICD) codes (ICD-10 C22.0: Liver cell carcinoma, C22.8: Malignant neoplasm of liver, primary, C22.9: Malignant neoplasm of liver, not specified as primary or secondary) were selected. The cases were reviewed by 3 authors (CS, TKS, OS) to confirm histological or clinical and radiological proof of HCC. Clinical

How to cite this article: Simsek C, Guven DC, Sahin TK, Tekin IE, Sahan O, Balaban HY, Yalcin S. Artificial intelligence method to predict overall survival of hepatocellular carcinoma. *Hepatology Forum* 2021; 2(2):64–68.

Received: April 26, 2021; **Accepted:** May 06, 2021; **Available online:** May 21, 2021

Corresponding author: Cem Simsek; Hacettepe Universitesi Tip Fakultesi, Gastroenteroloji Anabilim Dalı, Ankara, Turkey
Phone: +1 443 467 55 45; **e-mail:** cemgsimsek@gmail.com



OPEN ACCESS
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

© Copyright 2021 by Hepatology Forum - Available online at www.hepatologyforum.org



variables, such as performance status and the presence of ascites or hepatic encephalopathy were retrospectively retrieved from patient files and electronic records, as well as the laboratory values (albumin, bilirubin, creatinine, international normalized ratio, alanine transaminase, aspartate transaminase, alkaline phosphatase, gamma-glutamyl transferase [GGT], hemoglobin, platelet count) recorded closest to the time of diagnosis. Child-Pugh-Turcotte (CPT) classification information was also collected. Computed tomography (CT) or magnetic resonance imaging (MRI) scan reports were used to record the number of lesions, the diameter of the largest lesion and the presence of vascular invasion at the time of diagnosis. All of the patients were classified according to the Barcelona Clinic Liver Cancer (BCLC) and Okuda scoring systems. Treatment data (chemotherapy, transarterial chemo- or radio-embolization, radiofrequency ablation, hepatic resection, transplantation) were also retrospectively collected.

Outcomes and Statistical Analysis

The primary outcomes were the machine learning model predictions of >6-, 12-, 24-, 36-, 60-, 90-, and 120-month overall survival of the study population. Descriptive analyses were presented with mean (SD) for parametric variables, median (range) for non-parametric variables, and frequency (%) for categorical parameters. Differences between groups were tested for statistical significance with a chi-squared test, the Mann-Whitney U test, and linear regression, as appropriate. The univariate survival analysis was conducted using Kaplan-Meier curves and comparisons between prognostic groups were performed using the log-rank test. The area under the curve (AUC) of machine learning models was calculated and evaluated. All of the analyses were performed using IBM SPSS Statistics for Windows, Version 27.0 (IBM Corp., Armonk, NY, USA). A type-I error level of 5% ($p < 0.05$) was the threshold limit for statistical significance.

Model, Feature Selection, Training, and Outcomes

The Light Gradient Boosting Machine (LightGBM) (Microsoft Corp., Redmond, WA, USA) was used to test the performance of this approach to survival prediction. LightGBM is an ensemble of multiple decision trees that learn from each other to generate a more accurate final model.^[4] The technical details and method of use have been well defined elsewhere. Briefly, multiple train/test splits were stratified in order to estimate the generalizability error. The model was trained and re-evaluated 50 times with different training and test sets. In each iteration, the data were shuffled and then split into training and test sets with a 4:1 ratio. The mean AUC of all 50 models is presented as the estimated generalization error, and the SD of the scores represents the confidence interval. Model-agnostic feature importance techniques of permutation feature importance and leave-one-out feature importance were used. No method for missing data was employed, as it has the risk of bias. Variables that performed well in different models were selected.

Results

Population Characteristics

The final study population was 100 biopsy-proven or radiologically HCC cases diagnosed between January 2007 and July 2014. The mean age of the group was 62 ± 9 years, and 81% were male. The etiology of the liver disease was chronic hepatitis B in 47 patients, chronic hep-

Table 1. Characteristics of the patient population

	Mean (SD)/Number
Age (years)	62 (9)
Gender	
Male	82
Female	18
Etiology of liver disease	
Chronic hepatitis B	47
Chronic hepatitis C	25
Cryptogenic	12
Non-alcoholic steatohepatitis	6
Alcohol use	10
CTP class	
A	54
B	36
C	6

CTP: Child-Pugh-Turcotte.

Table 2. Features of hepatocellular carcinoma at the time of diagnosis and treatments used during follow-up

	Mean (SD)/Number
Barcelona clinic liver cancer	
Stage 0	7
Stage A	30
Stage B	21
Stage C	34
Stage D	6
Okuda	
Stage 1	51
Stage 2	40
Stage 3	7
Number of lesions	4 (1–15)
Diameter of biggest lesion	2.5 (0.73)
TACE	24
TARE	10
RFA	25
Resection	53
Transplantation	2
Chemotherapy	20

RFA: Radiofrequency ablation; TACE: Transarterial chemoembolization; TARE: Transarterial radioembolization.

atitis C in 25 patients, alcoholic liver disease in 10 patients, non-alcoholic steatohepatitis in 6 patients, and cryptogenic in 12 patients. At the time of diagnosis, 54 patients were classified as CTP class A, 36 as CTP class B and 6 patients as CTP class C (Table 1). Sixty-four cases had a single lesion, 24 had 2 or 3 lesions, and 12 cases had >3 lesions. The mean alpha-fetoprotein (AFP) level at the time of diagnosis was 204 ng/mL (Range: 671). Vascular invasion of the portal or hepatic veins was present in 48 patients. The BCLC and Okuda stages

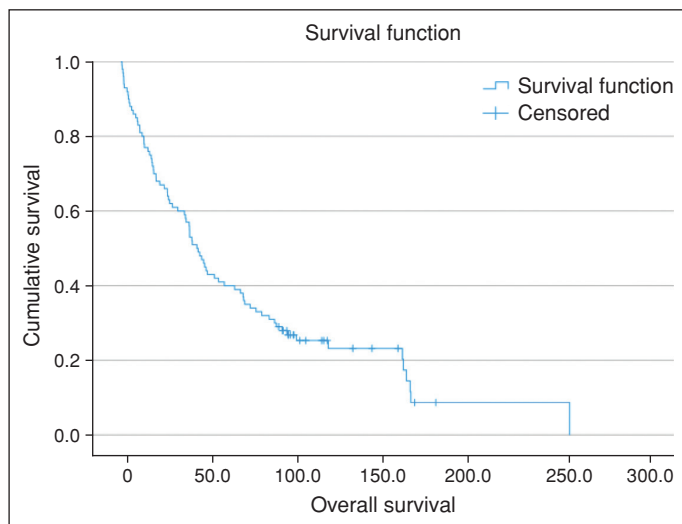


Figure 1. Kaplan-Meier survival curve of the study population.

Table 3. Survival characteristics of the study population

	6 months	1 year	2 years	3 years	5 years	10 years
Survival rate	88%	81%	67%	60%	40%	11%
Overall survival	43.9 (0.7–256) months					

were 7 cases of BCLC stage 0, 30 cases of BCLC stage A, 51 cases of Okuda stage 1, 21 cases of BCLC stage B, 40 cases of Okuda stage 1, 34 cases of BCLC stage C, 7 cases of Okuda stage 3, and 6 cases of BCLC stage D. A hepatic resection was performed in 53 cases, and 2 patients underwent liver transplantation. Transarterial treatment was performed in 34 patients: 24 chemoembolizations and 10 radioembolizations. Twenty-five patients had radiofrequency ablation and 16 had systemic treatment (Table 2).

Survival Characteristics and Analyses

The median overall survival of the study population was 43.0 months (interquartile range: 16.5–118.8 months) (Fig. 1). The survival rate for 6 months and 1, 2, 3, 5, and 10 years was 88%, 81%, 67%, 60%, 40%, and 11%, respectively (Table 3). Survival was statistically different across different BCLC (log rank: $p=0.000$) and Okuda (log rank: $p=0.000$) stages (Fig. 2, 3). A high AFP level, advanced age, and male gender were associated with lower survival in univariate Cox-regression analysis and remained significant in multivariable analysis after adjusting for BCLC stage.

Model Outputs

Deep learning models were designed to predict 6-, 12-, 36-, 60-, and 120-month overall survival in an HCC cohort. The features selected for the machine learning models were age, bilirubin, AFP, smoking status, alcoholic liver disease etiology, and GGT. The mean AUC of 50 machine learning models for overall survival prediction was 0.92 (0.061) for >6 months, 0.81 (0.107) for >1 year, 0.78 (0.11) for >2 years, 0.81 (0.083) for >3 years, 0.82 (0.079) for >5 years, 0.81 (0.96) for >8 years, and 0.66 (0.14) for >10 years (Fig. 3).

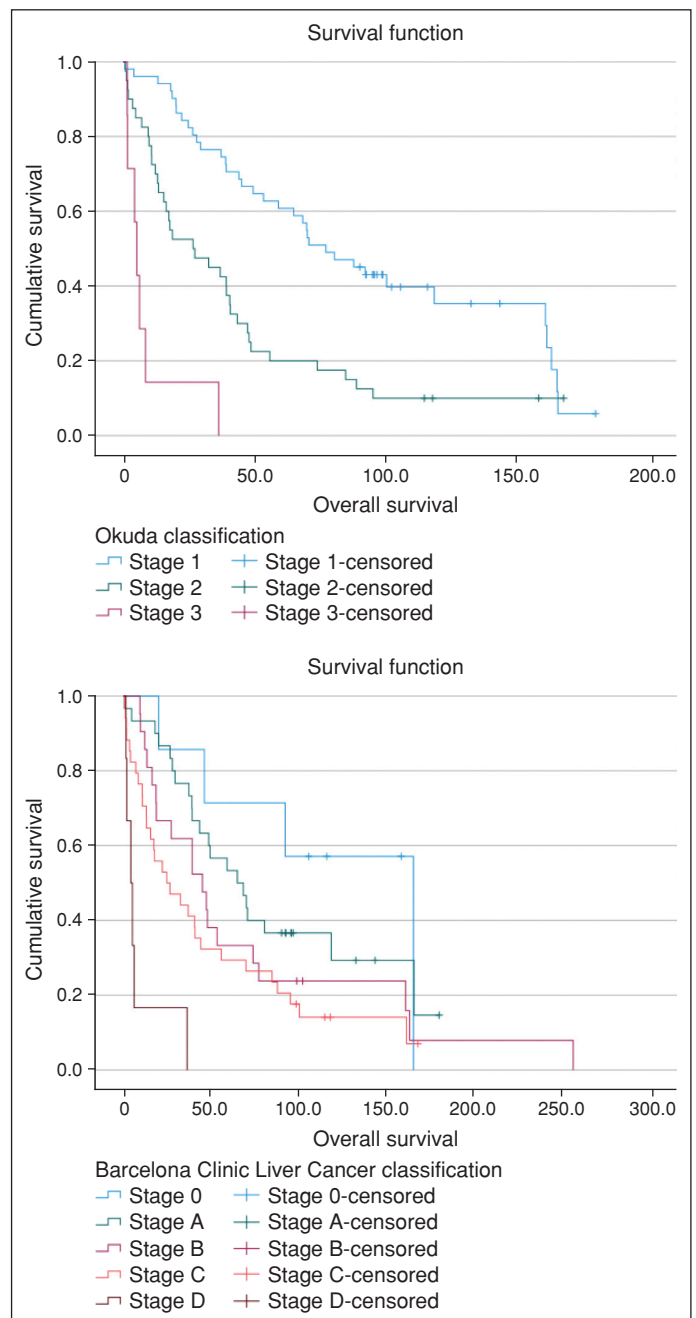


Figure 2. Survival of patients according to the Barcelona Clinic Liver Cancer and Okuda stages.

Discussion

We tested the performance of a machine learning algorithm to predict the overall survival probability of HCC patients at short- and long-term time intervals. The performance of the model forecast of 6-month, 1-, 2-, 3-, 5-, and 10-year overall survival ranged between 0.66 and 0.92. To best of our knowledge, this is the first study to evaluate such an approach to survival prediction in HCC.

The extent and reach of machine learning in HCC have been increasing in the past decade and some of that research is briefly described here to provide perspective. Machine learning models have been used to clas-

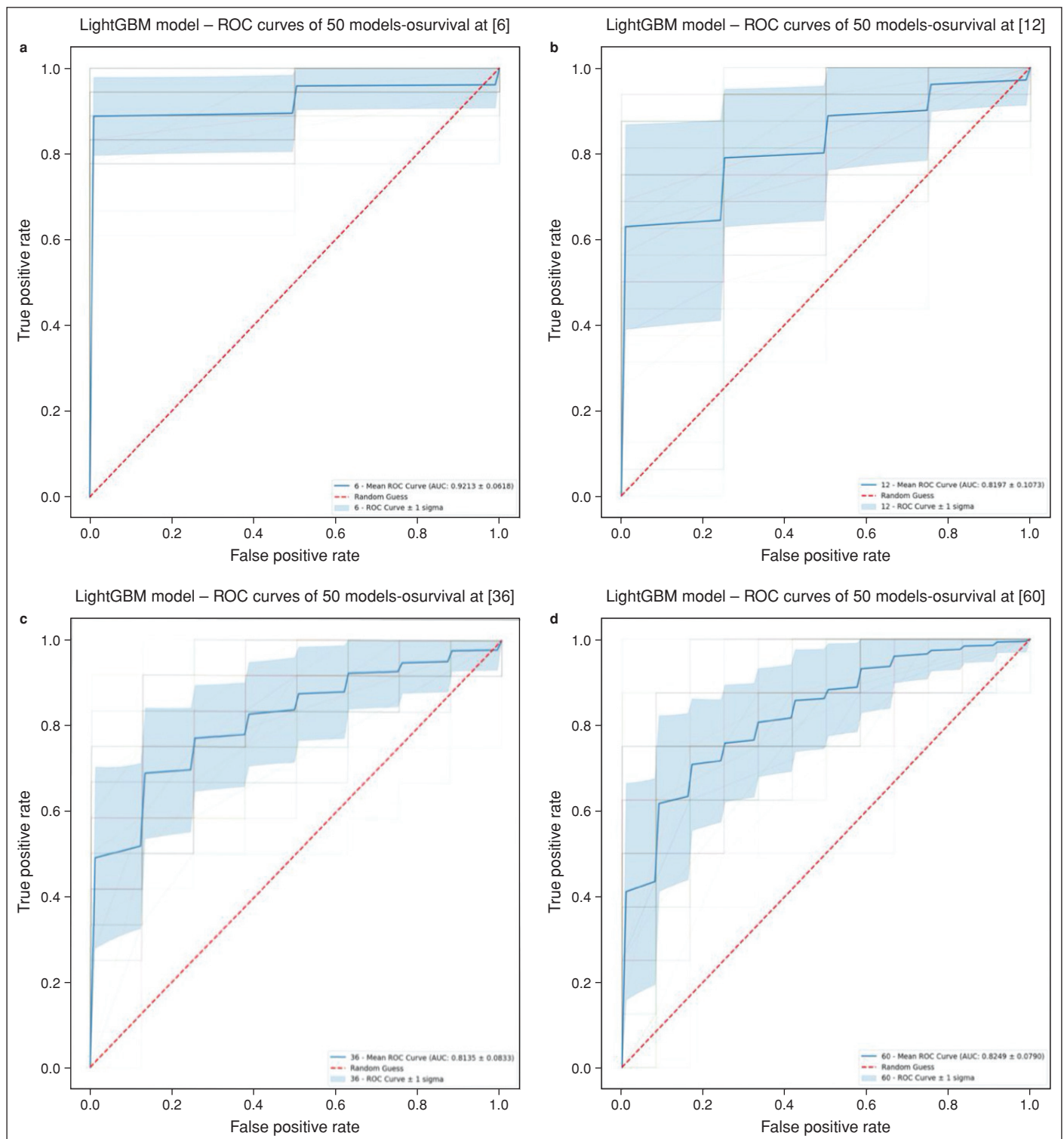


Figure 3. Outputs of the machine learning model for different time points. (a) 6 months, (b) 12 months, (c) 36 months, and (d) 60 months.

MAFLD: Metabolic-associated fatty liver disease.

sify 30 relevant genes with RNA signatures and cytokine signatures to classify HCC patients.^[5,6] RNA signatures were also used to predict 5-year overall survival in HCC and yielded an AUC of 0.69.^[7] HCC was also predicted using breath and urine metabolite analyses for screening

with an AUC approximating 0.9.^[8,9] Differential diagnosis of a hepatic lesion is another area that has been widely studied, and CT and MRI scans have been reported to have a classification accuracy of 99%.^[10] Machine learning has also been employed in treatment studies, and has

been used to predict progression or relapse after transarterial chemoembolization, resection, and transplantation.^[11–14]

Research of the potential of machine learning to predict the survival of HCC patients is limited. One methodological study in 2015 used previously collected, incomplete demographic and laboratory data from 165 patients to demonstrate the feasibility of this approach in a real-life setting. The sensitivity and specificity of the final model for the prediction of 1-year survival was 65% and 83%, respectively.^[15] However, studies that used artificial intelligence models to predict survival and outcomes for numerous cancers, including lung, breast, central nervous system, urogenital, and gastrointestinal cancer, with data from institutional and national clinical databases as well as multi-omic data, have been published and yielded better performance than traditional statistical methods.^[16]

The limitations of our study are mostly inherent to the small and heterogeneous population, the retrospective design, and machine learning methods. We included patients with long-term follow-up data from a single institution, which yielded smaller sized lesions and early-stage cases since we excluded patients without clinical, tumor-related, or survival data. Additional analyses of the different disease stages and treatments would increase the power of the findings. The drawbacks to using a small population, such as limited generalizability and statistical power, are acknowledged. In addition, due to this sampling bias, our population did not represent the full current spectrum of HCC, as they were primarily early-stage cases, resulting higher survival rates. The advantages of a robust model were prioritized, since the aim was to demonstrate the feasibility of machine learning models in this context. Finally, machine learning has specific limitations, such as model overfit to the study data, regardless of train-test splits and cross-validations. Therefore, further validation of our model or others using different and larger datasets will be required.

Machine learning is increasingly being used in hepatology. HCC is one area in which this approach will likely be useful due to the availability of extensive multi-omic and clinical data. Although radiology naturally spearheads machine learning efforts, we believe further research on other aspects of management in HCC -- as well as other aspects of hepatology -- will prove to be equally valuable.

Ethics Committee Approval: The Hacettepe University Clinical Research Ethics Committee granted approval for this study (date: 16.03.2021, number: GO 21/351).

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – CS, HYB, SY, DCG; Design – CS, HYB, SY, DCG; Supervision – HYB, SY; Data Collection and/or Processing – TKS, OS, IET, CS, DCG; Analysis and/or Interpretation – CS, DCG; Literature Search – CS, DCG; Writing – CS; Critical Reviews – HYB, SY, DCG.

Conflict of Interest: Cem Simsek is an equity holder in Algomedicus.

Financial Disclosure: This study was supported by Algomedicus Artificial Intelligence and Medical Simulation Company, Ankara, Turkey (A-2021.02).

References

1. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. *Int J Cancer* 2021 Apr 5. doi: 10.1002/ijc.33588. [Epub ahead of print]
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;127(12):2893-2917.
3. Petrick JL, Kelly SP, Altekruse SF, McGlynn KA, Rosenberg PS. Future of Hepatocellular Carcinoma Incidence in the United States Forecast Through 2030. *J Clin Oncol* 2016;34(15):1787-1794.
4. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*. 2017;30:3146-154.
5. Kim JW, Ye Q, Forgues M, Chen Y, Budhu A, Sime J, et al. Cancer-associated molecular signature in the tissue samples of patients with cirrhosis. *Hepatology* 2004;39(2):518-527.
6. Estevez J, Chen VL, Podlaha O, Li B, Le A, Vutien P, et al. Differential Serum Cytokine Profiles in Patients with Chronic Hepatitis B, C, and Hepatocellular Carcinoma. *Sci Rep* 2017;7(1):11867.
7. Jiang H, Zhao L, Chen Y, Sun L. A four-long noncoding RNA signature predicts survival of hepatocellular carcinoma patients. *Journal of Clinical Laboratory Analysis* 2020. doi: 10.1002/jcla.23377. [Epub ahead of print]
8. Miller-Atkins G, Acevedo-Moreno LA, Grove D, Dweik RA, Tonelli AR, Brown JM, et al. Breath metabolomics provides an accurate and noninvasive approach for screening cirrhosis, primary, and secondary liver tumors. *Hepatology Communications* 2020;4(7):1041-1055.
9. Wang J, Jain S, Chen D, Song W, Hu CT, Su YH. Development and evaluation of novel statistical methods in urine biomarker-based hepatocellular carcinoma screening. *Sci Rep* 2018;8(1):3799.
10. Naeem S, Ali A, Qadri S, Mashwani WK, Tairan N, Shah H, et al. Machine-learning based hybrid-feature analysis for liver cancer classification using fused (MR and CT) images. *Applied Sciences-Basel* 2020;10(9):3134.
11. Morshid A, Elsayes KM, Khalaf AM, Elmohr MM, Yu J, Kaseb AO, et al. A machine learning model to predict hepatocellular carcinoma response to transcatheter arterial chemoembolization. *Radiol Artif Intell* 2019;1(5):e180021.
12. Iwahashi S, Ghaibeh AA, Shimada M, Morine Y, Imura S, Ikemoto T, et al. Predictability of postoperative recurrence on hepatocellular carcinoma through data mining method. *Mol Clin Oncol* 2020;13(5):46.
13. Nam JY, Lee JH, Bae J, Chang Y, Cho Y, Sinn DH, et al. Novel Model to Predict HCC Recurrence after Liver Transplantation Obtained Using Deep Learning: A Multicenter Study. *Cancers (Basel)* 2020;12(10):2791.
14. Choi GH, Yun J, Choi J, Lee D, Shim JH, Lee HC, et al. Development of machine learning-based clinical decision support system for hepatocellular carcinoma. *Sci Rep* 2020;10(1):14855.
15. Santos MS, Abreu PH, Garcia-Laencina PJ, Simão A, Carvalho A. A new cluster-based oversampling method for improving survival prediction of hepatocellular carcinoma patients. *J Biomed Inform* 2015;58:49-59.
16. Huang S, Yang J, Fong S, Zhao Q. Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. *Cancer Lett* 2020;471:61-71.