Research Article
Innovative Diagnostics, Modelling and Digital Hepatology

JOURNAL OF HEPATOLOGY

AI model using CT-based imaging biomarkers to predict hepatocellular carcinoma in patients with chronic hepatitis B

Hyunjae Shin[1,†], **Moon Haeng Hur**[2,†], **Byeong Geun Song**[3,†], **Soo Young Park**[4,†], **Gi-Ae Kim**[5,†], Gwanghyeon Choi[6], Joon Yeul Nam[7], Minseok Albert Kim[8], Youngsu Park[2], Yunmi Ko[2], Jeayeon Park[2], Han Ah Lee[9], Sung Won Chung[10], Na Ryung Choi[11], Min Kyung Park[6], Yun Bin Lee[2], Dong Hyun Sinn[3], Seung Up Kim[12], Hwi Young Kim[11], Jong-Min Kim[13], Sang Joon Park[13,14], Hyung-Chul Lee[15], Dong Ho Lee[14], Jin Wook Chung[14], Yoon Jun Kim[2], Jung-Hwan Yoon[2], Jeong-Hoon Lee[2,16,*]

Journal of Hepatology **2025**. vol. ■ | 1–9

Background & Aims: Various hepatocellular carcinoma (HCC) prediction models have been proposed for patients with chronic hepatitis B (CHB) using clinical variables. We aimed to develop an artificial intelligence (AI)-based HCC prediction model by incorporating imaging biomarkers derived from abdominal computed tomography (CT) images along with clinical variables.

Methods: An AI prediction model employing a gradient-boosting machine algorithm was developed utilizing imaging biomarkers extracted by DeepFore, a deep learning-based CT auto-segmentation software. The derivation cohort (n = 5,585) was randomly divided into the training and internal validation sets at a 3:1 ratio. The external validation cohort included 2,883 patients. Six imaging biomarkers (*i.e*. abdominal visceral fat–total fat volume ratio, total fat–trunk volume ratio, spleen volume, liver volume, liver–spleen Hounsfield unit ratio, and muscle Hounsfield unit) and eight clinical variables were selected as the main variables of our model, PLAN-B-DF.

Results: In the internal validation set (median follow-up duration = 7.4 years), PLAN-B-DF demonstrated an excellent predictive performance with a c-index of 0.91 and good calibration function (p = 0.78 by the Hosmer-Lemeshow test). In the external validation cohort (median follow-up duration = 4.6 years), PLAN-B-DF showed a significantly better discrimination function compared to previous models, including PLAN-B, PAGE-B, modified PAGE-B, and CU-HCC (c-index, 0.89 *vs.* 0.65–0.78; all p <0.001), and maintained a good calibration function (p = 0.42 by the Hosmer-Lemeshow test). When patients were classified into four groups according to the risk probability calculated by PLAN-B-DF, the 10-year cumulative HCC incidence was 0.0%, 0.4%, 16.0%, and 46.2% in the minimal-, low-, intermediate-, and high-risk groups, respectively.

Conclusion: This AI prediction model, integrating deep learning-based auto-segmentation of CT images, offers improved performance in predicting HCC risk among patients with CHB compared to previous models.

Introduction

Hepatocellular carcinoma (HCC) is a major global health issue, currently ranked as the fourth leading cause of cancer-related deaths worldwide.[1] In East Asia, chronic hepatitis B (CHB) is the most prevalent cause of HCC, resulting in substantial healthcare burdens.[2] While nucleos(t)ide analogues have been shown to reduce the risk of HCC, there still remains a residual risk that requires ongoing surveillance for the early detection of HCC.[3,4]

The risk of HCC varies significantly among patients undergoing treatment, leading to the development of various risk prediction models.[5,6] However, predicting the incidence of HCC remains challenging, owing to the complexity of HCC development, which is influenced by various risk factors, ranging from well-known factors such as age or fibrotic burden to virus-related factors including hepatitis B envelope antigen (HBeAg) positivity or serum hepatitis B virus (HBV) DNA titer.[7,8] Moreover, the impact of metabolic risk factors such as obesity, diabetes mellitus, and hypertension, as well as the use of medications for these conditions (*e.g*. statins or aspirin), was investigated.[9,10] Abdominal visceral fat and subcutaneous fat have been shown to confer an additive risk for liver-related complications, either independently or through MASLD (metabolic dysfunction-associated steatotic liver disease) and other metabolic comorbidities.[11,12] It has been suggested that hepatic steatosis or steatohepatitis may increase the risk of HCC in patients with CHB.[13] In addition, myosteatosis, which

* Corresponding author. Address: Department of Internal Medicine, Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea; Tel.: +82-2-2072-2228, FAX: +82-2-743-6701.
E-mail addresses: pindra@empal.com, JHLeeMD@snu.ac.kr (J.-H. Lee).
† These five authors contributed equally to this study as co-first authors.
https://doi.org/10.1016/j.jhep.2024.12.029

indicates fat infiltration or fatty degeneration within muscles, has emerged as a notable risk factor.[12,14] Thus, a comprehensive system that integrates these various risk factors is essential for accurate HCC prediction.

Given the complexity of the risk factors that contribute to the development of HCC, machine learning-based modeling may be more effective. Recently, artificial intelligence (AI)-based prediction models, such as PLAN-B, have been introduced, exhibiting superior performance compared to traditional models based on regression analyses.[5,10] In addition, efforts are ongoing to enhance the accuracy of AI models by integrating imaging variables with clinical variables.[15] Recent advances in abdominal computed tomography (CT) auto-segmentation have made it possible to quantitatively measure visceral fat and spleen volume and assess the degree of myosteatosis.[16]

Therefore, we aimed to establish a highly accurate HCC prediction model for patients with CHB by utilizing variables extracted from baseline CT imaging in addition to clinical parameters. This model may enhance risk stratification and provide a more precise tool for optimizing HCC surveillance, leading to individualized surveillance and improved outcomes through timely and targeted interventions.

Patients and methods

Patients

Patients undergoing treatment for CHB who received abdominal CT imaging at seven tertiary referral hospitals in Korea between 2007 and 2021 were included in the study (Fig. 1). Patients received either entecavir or tenofovir according to the practice guidelines from the Korean Association for the Study of Liver.[17] Patients were excluded if either of the following criteria were met: i) development of HCC within the first year, and ii) missing variables for prediction. The derivation cohort comprised a total of 5,585 patients from five hospitals. The derivation cohort was randomly divided into training (n = 4,188) and internal validation (n = 1,397) sets at a ratio of 3:1. An external validation cohort (n = 2,883) was independently established from two hospitals. The institutional review board at each participating center granted approval for this study, and due to its retrospective nature, the requirement for informed consent was waived.

Development of an AI model using CT auto-segmentation

Patients were enrolled at the time of abdominal CT imaging (Table S1). Baseline demographics and comorbidities were evaluated within 1 month before or after enrollment. The series of CT images in the precontrast phase were analyzed with DeepFore (v1.0, MedicalIP Co. Ltd., Seoul, Korea). DeepFore enabled deep learning-based CT analysis and extracted imaging biomarkers through automated segmentation and allocation of organs and body compartments in stacked CT images (Fig. S1).[16,18–20] Our clinical trial on the prediction of HCC using DeepFore was approved by the Korean MFDS (approval number: 1713).

A gradient-boosting machine (GBM) algorithm was adopted to develop the model (supplementary methods).[21] The hyperparameters of the GBM algorithm were optimized using a randomized search method in the XGBoost library. The parameters included: learning rates (0.05, 0.07, and 0.09), maximum depth of a decision tree (4, 5, and 6), number of decision trees (50, 75, and 100), fraction of samples randomly selected (0.5, 0.7, and 1), and fraction of variables randomly selected (0.5, 0.7, and 1), all evaluated through five-fold cross-validation.

Selection of variables

Through a comprehensive review of potential risk factors for the development of HCC, 40 clinical variables and 9 imaging biomarkers were identified as potential risk factors supported by clinical evidence (Table S2). Among these variables, those

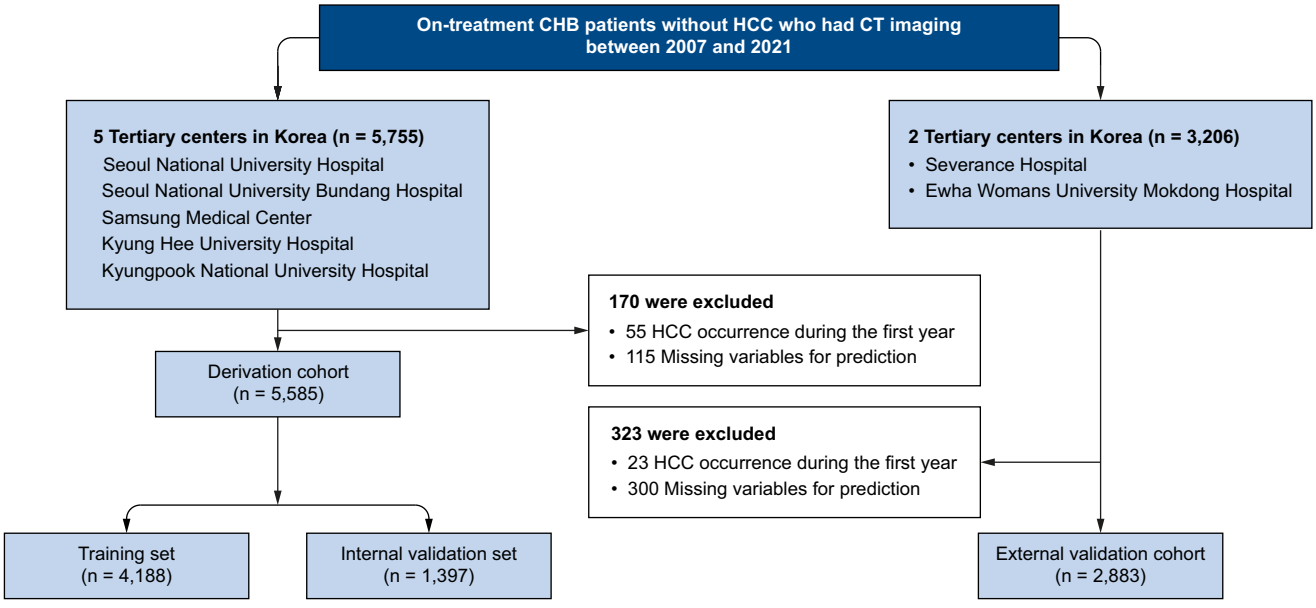


Fig. 1. Patient flow diagram. CHB, chronic hepatitis B; CT, computed tomography; HCC, hepatocellular carcinoma.

presumed to have similar clinical mechanisms influencing HCC occurrence (*e.g.*, the clinical diagnosis of cirrhosis, liver stiffness measurement [LSM], and spleen volume) were analyzed and selected one by one to develop multiple candidate models (supplementary methods). After testing a number of candidate models with various combinations of these variables, a model with the highest predictive accuracy using a minimal set of variables was selected as a final model, which was designated as PLAN-B-DF. As a result, from this pool of 40 candidate clinical variables, eight variables were utilized in the PLAN-B-DF model: age, type of antiviral treatment, platelet count, serum HBV DNA level, HBeAg status, alanine aminotransferase, and the presence of diabetes mellitus and hypertension. Similarly, six imaging biomarkers extracted from CT scans (*i.e.*, liver volume, spleen volume, total fat–trunk volume ratio, abdominal visceral fat–total fat volume ratio, liver–spleen Hounsfield unit [HU] ratio, and muscle HU) were utilized.

Follow-up and assessment

All patients were assessed during regular outpatient visits in each center after enrollment, adhering to the national HCC surveillance program in Korea.[22] HCC surveillance consisted of monitoring serum alpha-fetoprotein and liver imaging with ultrasonography, CT, or magnetic resonance imaging (MRI) every 6 months. The follow-up period ended at the date of HCC diagnosis, death of any cause, or the last follow-up, whichever occurred first. The primary outcome was the development of HCC. A diagnosis of HCC was established according to the international HCC guidelines.[23,24]

Statistical analysis

Non-parametric continuous variables are presented as median (IQR) unless otherwise noted. Categorical variables are presented as absolute numbers of cases and/or percentages. The chi-square test or Mann-Whitney *U* test was used to compare baseline characteristics between the cohorts. Survival analysis was performed to estimate the cumulative incidence of HCC, using Kaplan-Meier curves and the log-rank test.

Patients were stratified according to the output values calculated by the GBM algorithm, which were proportional to the risk of HCC in each patient. The output values that divide the training set into quartiles were rounded to two decimal places and selected as cut-off values. Consequently, the cut-off values were set at 0.03, 0.05, and 0.15, categorizing patients in the internal validation set and the external validation cohort into four groups: minimal-risk, low-risk, intermediate-risk, and high-risk. The discriminant function was assessed using Harrell's concordance index (c-index). The calibration function was assessed by plotting the risk of HCC development according to the risk group. The quality of calibration was assessed using the Hosmer-Lemeshow goodness-of-fit test. To minimize the impact of different observational periods between the cohorts, Gönen and Heller's c-index (GHCI), which utilizes inverse probability of censoring weighting, was also calculated. Two-sided *p* values were calculated for all analyses. *p* values less than 0.05 were considered statistically significant.

The prediction model was developed using Python version 3.7 (Python Software Foundation, Wilmington, DE, USA) with the XGBoost, scikit-learn, and SHAP libraries. The GBM model was visualized using the Graphviz library. All data were analyzed using R statistics version 4.2.0 (R Foundation, Vienna, Austria) with the compareC package for comparison of c-indices.

Results

Baseline characteristics

The baseline characteristics of the derivation and external validation cohorts, including CT-derived radiologic biomarkers, are outlined in Table 1. The percentage of patients with cirrhosis at baseline differed between the derivation and external validation cohorts (42.3% *vs.* 39.5%, respectively, $p < 0.001$). Meanwhile, spleen volume was not statistically different between the two cohorts (202.2 *vs.* 210.1 ml, respectively, $p = 0.06$). In the two cohorts, CT scans were performed as part of health check-ups in 12.4%, for HCC surveillance in 32.6%, and because of acute liver-related events in <5% of patients (Table S3).

In the derivation cohort, 582 patients (10.4%) developed HCC during a median follow-up of 7.4 (IQR, 5.0—9.7) years with an annual cumulative incidence of 1.5% (Table S4). The proportions of patients with more than 2 years and 5 years of follow-up were 97.8% and 74.7%, respectively. The training and internal validation sets were well balanced through random allocation (Table S5). Of these, 437 patients (10.3%) in the training set and 144 patients (10.4%) in the internal validation set developed HCC over 7.5 and 7.3 years, with annual incidences of 1.5% and 1.6%, respectively. In the external validation cohort, 263 patients (9.1%) developed HCC during a median follow-up of 4.6 (IQR, 2.4—7.3) years with an annual cumulative incidence of 1.9%. The proportions of patients with >2 years and >5 years of follow-up were 93.1% and 46.9%, respectively.

Development and validation of a novel AI model

A novel AI model named "**P**rediction of **L**iver cancer using **A**rtificial intelligence-driven model with **N**eural network-based software for hepatitis **B**–**D**eep**F**ore (PLAN-B-DF)" was developed in the derivation cohort using the GBM algorithm. The source code for PLAN-B-DF model is available upon request for research purposes. The hyperparameters of PLAN-B-DF were chosen with a learning rate of 0.07, a maximum depth of a decision tree of 5, a number of decision trees of 50, a fraction of samples randomly selected of 0.7, and a fraction of variables randomly selected of 0.7 (Fig. S2).

The Shapley values of PLAN-B-DF are presented in Fig. 2. PLAN-B-DF showed a c-index of 0.91 (95% CI 0.89–0.93) in the internal validation set. In the external validation cohort, PLAN-B-DF had a c-index of 0.89 (95% CI 0.87–0.90), and there was no significant difference in the discriminant function between the internal validation set and the external validation cohort ($p = 0.17$).

The discriminant functions of the candidate models were statistically lower than those of the PLAN-B-DF model (all

Table 1. Baseline characteristics of the derivation cohort and the external validation cohort.

	Total	Derivation cohort	External validation cohort		
	N = 8,468	n = 5,585	n = 2,883	*p* values	*SMD*
Age, years	50 [42–57]	49 [42–56]	50 [42–57]	0.07	0.031
Sex, male	5,179 (61.2%)	3,374 (60.4%)	1,805 (62.6%)	0.03	0.045
Tenofovir treatment, n (%)	3,784 (44.7%)	2,414 (43.2%)	1,370 (47.5%)	<0.001	0.350
Cirrhosis, n (%)	3,503 (41.4%)	2,363 (42.3%)	1,140 (39.5%)	<0.001	0.056
Hypertension, n (%)	1,174 (14.1%)	814 (15.0%)	360 (12.5%)	0.01	0.073
Diabetes mellitus, n (%)	816 (9.8%)	546 (10.1%)	270 (9.4%)	0.17	0.024
Obesity, n (%)	2,952 (37.2%)	2,117 (41.1%)	835 (29.9%)	<0.001	0.234
HBeAg positivity, n (%)	4,018 (47.9%)	2,592 (47.1%)	1,426 (49.5%)	<0.001	0.048
Platelet count, 10^9/L	158 [118–203]	159 [120–202]	158 [114–205]	0.32	0.012
Albumin, g/dl	4.1 [3.7–4.4]	4.2 [3.9–4.4]	4.1 [3.8–4.4]	<0.001	0.157
AST, U/L	64 [41–112]	69 [44–120]	56 [35–97]	<0.001	0.117
ALT, U/L	80 [41–150]	88 [47–166]	63 [33–129]	<0.001	0.194
HBV DNA, $\log_{10}$ IU/ml	6.1 [4.6–7.4]	5.6 [3.5–7.2]	6.2 [5.1–7.5]	<0.001	0.004
Total bilirubin, mg/dl	0.9 [0.7–1.2]	0.9 [0.7–1.3]	0.9 [0.6–1.2]	<0.001	0.036
Creatinine, mg/dl	0.8 [0.7–1.0]	0.8 [0.7–1.0]	0.8 [0.7–0.9]	<0.001	0.092
Liver volume, ml	1,102.0 [929.9–1317.0]	1,104.3 [937.8–1318.5]	1,092.9 [905.4–1309.9]	<0.001	0.092
Spleen volume, ml	203.9 [142.2–312.0]	202.2 [141.8–306.1]	210.1 [143.5–325.4]	0.06	0.036
Liver–spleen HU ratio	1.1 [0.9–1.2]	1.0 [0.9–1.2]	1.1 [1.0–1.2]	<0.001	0.178
Abdominal visceral fat–total fat volume ratio[1], (%)	38.0 [25.7–48.9]	37.3 [25.1–48.3]	39.1 [27.3–50.3]	0.001	0.109
Total fat–trunk volume ratio[2], (%)	61.8 [52.4–69.1]	61.8 [52.1–68.9]	61.9 [53.9–69.4]	0.18	0.109
Muscle HU	41.8 [35.3–47.3]	39.5 [33.1–44.5]	42.6 [36.1–48.1]	<0.001	0.025

AST, aspartate aminotransferase; ALT, alanine aminotransferase; HBeAg, hepatitis B envelope antigen; HU, Hounsfield unit; SMD, standardized mean difference.
Data are expressed as n (%) or median (interquartile range).
*p values for the comparison of the derivation and external validation cohorts were calculated using the chi-square test or Mann-Whitney U test.
[1]Calculated by abdominal visceral fat volume/total fat volume x 100 (%).
[2]Calculated by total fat volume/trunk volume x 100 (%).

p <0.05, supplementary results and Table S6). The discriminant function of the candidate model when spleen volume was replaced by LSM via transient elastography (PLAN-B-DF$_{\text{Spleen volume}\rightarrow\text{LSM}}$) did not show a statistically significant difference when compared to the PLAN-B-DF (p = 0.13, supplementary results). However, the discriminant function became statistically inferior when spleen volume was replaced by the presence of cirrhosis (PLAN-B-DF$_{\text{Spleen volume}\rightarrow\text{LC}}$, p = 0.03). Replacing

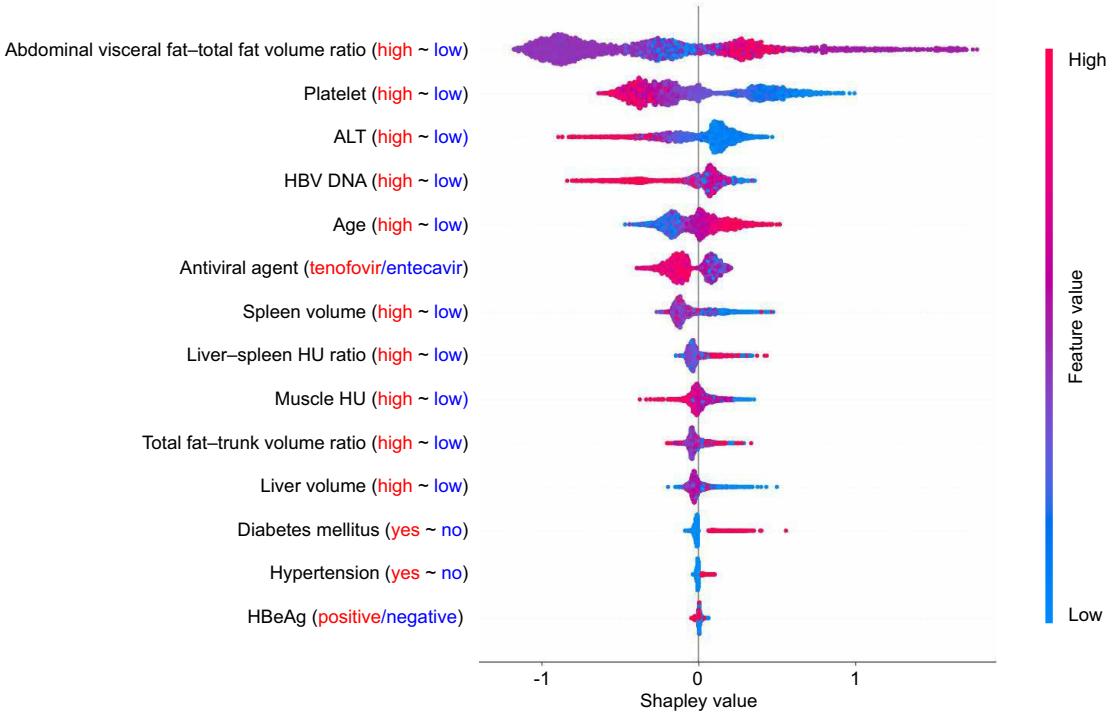


Fig. 2. Shapley plot of PLAN-B-DF in the derivation cohort. Variables higher on the figure are more important contributors to the predictive accuracy of PLAN-B-DF. High Shapley values for each variable are associated with an increased predicted risk of hepatocellular carcinoma. ALT, alanine aminotransferase; HBeAg, hepatitis B envelope antigen; HBV, hepatitis B virus; HU, Hounsfield unit.

the liver–spleen HU ratio with the controlled attenuation parameter score or sonographic fatty liver resulted in a decreased c-index (supplementary results). In addition, the Shapley plots for the candidate models revealed a variable hierarchy similar to that of PLAN-B-DF (Figs S3–S10).

In addition, PLAN-B-DF exhibited good calibration functions in both cohorts according to the Hosmer-Lemeshow test (p = 0.78 in the internal validation set, and p = 0.42 in the external validation cohort; Fig. S11). The Brier score of PLAN-B-DF in both cohorts also indicated a good calibration function (Table S7).

As a sensitivity analysis, GHCI was calculated and PLAN-B-DF exhibited superior performance compared to previous prediction models (supplementary results and Table S8). Additionally, we validated PLAN-B-DF after excluding patients who underwent CT scans due to suspected acute liver-related events, demonstrating similar results to our primary findings (supplementary results and Table S9).

Comparison of PLAN-B-DF to previous prediction models

PLAN-B-DF outperformed previous models based on conventional statistics (PAGE-B,[25] modified PAGE-B,[26] REACH-B,[27] CU-HCC,[6] GAG-HCC,[28] THRI,[29] and PAGED-B[8]) as well as machine learning-based models which used clinical variables only (PLAN-B[5]) in both the internal validation set (c-index, 0.91 *vs.* 0.65–0.82; all p <0.001; Table S10), and the external validation cohort (c-index, 0.89 *vs.* 0.63–0.78; all p <0.001; Table 2). Furthermore, PLAN-B-DF exhibited a significantly superior time-dependent AUROC (area under the receiver-operating characteristic curve) in comparison to the previous models during the 2-year, 4-year, 6-year, 8-year, and 10-year observation periods (0.87–0.91 *vs.* 0.63–0.82, all p <0.001; Table S11). In addition, PLAN-B-DF demonstrated a significantly higher AUPRC (area under the precision-recall curve) compared to previous models in both the internal validation set (0.66 *vs.* 0.16–0.32, all p <0.001) and the external validation cohort (0.44 *vs.* 0.13–0.24, all p <0.001; Fig. S12), as well as in other prediction performance metrics (Table S12).

HCC risk stratification by PLAN-B-DF

Each risk group, the minimal-risk, low-risk, intermediate-risk, and high-risk groups, showed a significantly different risk of developing HCC in the internal validation set (Fig. 3A) and the external validation cohort (Fig. 3B). In the external validation cohort, the minimal-risk group (788 out of 2,883, 27.3%) had a cumulative incidence of HCC of 0.0% through 10 years of follow-up. The low-risk group (662 out of 2,883, 23.0%) also showed a cumulative incidence of HCC of 0.2%, 0.4%, 0.4% at years 2, 6, and 10, respectively, whereas the intermediate-risk group (757 out of 2,883, 26.3%) showed a cumulative incidence of HCC of 1.3%, 11.9% and 16.0% at years 2, 6, and 10, respectively. The high-risk group (676 out of 2,883, 23.4%) exhibited a significantly increased cumulative incidence of HCC of 8.3%, 34.9% and 46.2% at years 2, 6, and 10, respectively.

Table 2. Discrimination function of PLAN-B-DF compared to previous prediction models in the external validation cohort.

	C-index (95% CI)	p values
PLAN-B-DF	0.89 (0.87–0.90)	Ref.
PLAN-B	0.78 (0.76–0.80)	<0.001
PAGE-B	0.72 (0.70–0.75)	<0.001
mPAGE-B	0.73 (0.70–0.75)	<0.001
REACH-B	0.67 (0.64–0.70)	<0.001
CU-HCC	0.65 (0.62–0.69)	<0.001
GAG-HCC	0.63 (0.59–0.67)	<0.001
THRI	0.73 (0.70–0.76)	<0.001
PAGED-B	0.71 (0.68–0.74)	<0.001

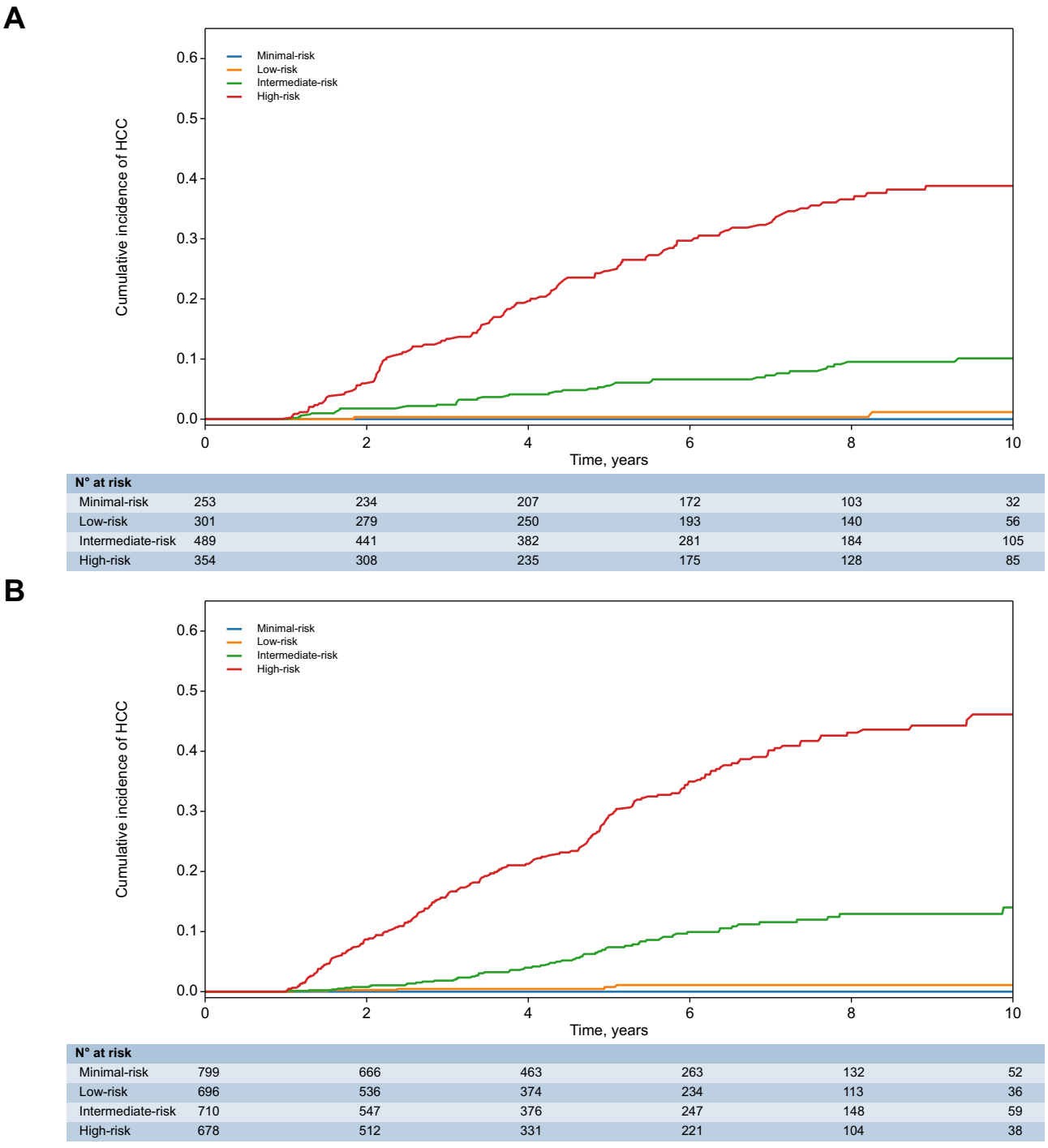
Subgroup analyses

The GBM algorithm was applied to subgroups of the external validation cohort to validate the robustness of PLAN-B-DF (Table S13). Across subgroups defined by the baseline HBeAg status (positive or negative), type of antiviral agent (entecavir or tenofovir), cirrhosis status (present or absent), and baseline HBV DNA status (undetectable or detectable), PLAN-B-DF maintained consistent predictive performance. Notably, in patients without cirrhosis, the application of the GBM algorithm yielded superior predictive accuracy compared to those with cirrhosis. For patients with early clinical phases of CHB, specifically the HBeAg-positive phase without cirrhosis, PLAN-B-DF showed an improved discriminant function over previous prediction models (c-index, 0.91 *vs.* 0.43–0.85, all p <0.05; Table S14), with the exception of the PLAN-B model (c-index, 0.91 *vs.* 0.90, p = 0.58), which also employed the GBM algorithm.

Discussion

In this study, we developed an AI prediction model, PLAN-B-DF, utilizing eight clinical variables and six CT-derived radiologic biomarkers, which exhibited outstanding performance in predicting HCC in patients with CHB. PLAN-B-DF achieved a high predictive accuracy (c-index of 0.89) in the independent external validation cohort. In addition, this model effectively classified patients with CHB into minimal-risk (with less than 0.1% HCC incidence over a 10-year period) and high-risk (with HCC incidence approaching 50% over the same period) groups. The robustness of this model, demonstrated by its reproducibility across various CT machines, was confirmed by various sensitivity and subgroup analyses. These findings demonstrate the potential role of PLAN-B-DF in improving the accuracy and efficiency of risk stratification for HCC development, thereby contributing to personalized management and surveillance of patients with CHB.

This model has significantly improved the prediction of HCC development by utilizing six quantifiable imaging biomarkers extracted from CT, which can be categorized into central obesity-related (abdominal visceral fat–total fat volume ratio, and total fat–trunk volume ratio), steatosis-related (liver–spleen HU ratio and muscle HU), and portal hypertension-related biomarkers (liver volume and spleen volume). Central obesity-related imaging biomarkers were key contributors to our model. Abdominal visceral fat–total fat volume ratio, which consistently ranked at the top in each Shapley plot of PLAN-B-DF and candidate models, was confirmed as a critical determinant in our model. To date, the impact of visceral fat on the cascade of MASLD, metabolic dysfunction-associated steatohepatitis, HCC, and subsequent mortality has been well established.[30,31] Visceral obesity is also thought to contribute not only through the cascade, but also directly through pro-inflammatory cytokines, which may play a carcinogenic role in



N° at risk						
Minimal-risk	253	234	207	172	103	32
Low-risk	301	279	250	193	140	56
Intermediate-risk	489	441	382	281	184	105
High-risk	354	308	235	175	128	85

N° at risk						
Minimal-risk	799	666	463	263	132	52
Low-risk	696	536	374	234	113	36
Intermediate-risk	710	547	376	247	148	59
High-risk	678	512	331	221	104	38

Fig. 3. Cumulative incidence of HCC according to risk groups stratified by PLAN-B-DF. In (A) internal validation set and (B) external validation cohort. HCC, hepatocellular carcinoma.

patients with CHB.[32] Total fat volume–trunk volume ratio, with subcutaneous fat as a major constituent, also exhibits negative impacts on HCC in patients with chronic liver disease.[33,34]

Among steatosis-related radiologic biomarkers, the liver–spleen HU ratio, indicative of hepatic steatosis, was identified as a risk factor for developing HCC in patients with CHB.[35] Given the strong positive correlation between MRI proton density fat fraction and liver HU, hepatic steatosis could be adequately assessed as a continuous variable by the liver–spleen HU value.[36,37] Additionally, myosteatosis – the fatty degeneration within muscles, known to reduce the quality of muscle – emerged as another significant risk factor for HCC.[14,38–40] Its importance in HCC risk assessment is well incorporated into our model by the muscle HU value. Portal

hypertension-related imaging biomarkers, such as liver and spleen volumes, were effective predictors of HCC development, as these are surrogate indicators for the diagnosis of cirrhosis. Our candidate model showed a decrease in the c-index when spleen volume was substituted for cirrhosis. As a binary variable based on clinical diagnosis, cirrhosis was typically inferred from the presence of splenomegaly or surface nodularity in imaging reports.[41] In contrast, assessing spleen volume as a continuous variable may be advantageous, as it is particularly indicative of portal hypertension, a key consequence of cirrhosis, especially in cases where liver biopsy is not available.[42,43] Comparison with our candidate model highlighted the utility of spleen volume compared to LSM via transient elastography, which was not suitable for testing in some patients with central obesity.[44]

In addition to imaging biomarkers, the clinical variables incorporated in PLAN-B-DF are crucial in enhancing its predictive accuracy for HCC. Notably, HBV DNA titer has recently been found to have a parabolic association with HCC risk, reinforcing the complexity of its impact on liver-related complications.[7,8] Given that both cohorts in our model include a substantial portion of patients with both low (<5 $\log_{10}$ IU/ml) and high (>8 $\log_{10}$ IU/ml) viral loads, PLAN-B-DF might have benefited from its non-linearity-based algorithm. Interestingly, our analysis indicates that the use of statin and aspirin did not markedly affect the predictive power of the model, which contrasts with several retrospective studies suggesting an association between HCC and metabolic drugs.[45–48] Similarly, while the type of antiviral treatment was identified as a factor related to HCC development, its impact, as shown in the Shapley plots, was comparatively less significant than that of other variables.[49,50] The inclusion of diabetes mellitus as a variable, parallel to its role in the PAGED-B model, showed minimal predictive power in our model.[51] This limitation may be due to the model reflecting only the diagnosis of diabetes or the use of antidiabetic drugs, rather than the disease control status, which might be indicated by hemoglobin A1c levels.[8,52] Hypertension, incorporated in a manner similar to its use in the SMART-HCC model, also contributed to our model, adding another layer to the multifactorial risk assessment for HCC in patients with CHB. However, the causality between hypertension and HCC risk remains underexplored.[10]

The clinical applicability of PLAN-B-DF relies on its ability to stratify HCC risk, which enables personalized surveillance by effectively classifying patients with CHB into minimal-to high-risk groups. By tailoring surveillance to risk groups, high-risk patients may benefit from more intensive surveillance, while less intensive surveillance in minimal- or low-risk patients may conserve healthcare resources. The threshold for cost-effectiveness of surveillance may vary depending on regional or national circumstances.[53,54] Considering that some international guidelines recommend surveillance when the HCC incidence is at least 1.0%/year,[54,55] it may be reasonable to limit surveillance to intermediate- or high-risk patients only. High-risk patients may also benefit from utilizing alternative CT or MRI, rather than relying exclusively on ultrasonography, to enable the early detection of HCC. For patients with modifiable risk factors for HCC, such as hypertension, diabetes, or excess visceral fat, physicians may recommend improving these factors. However, real-world evidence remains constrained, and further multinational studies are required to confirm whether tailored surveillance based on PLAN-B-DF results can improve clinical outcomes and cost-effectiveness of surveillance programs for patients with CHB.

There are several limitations in the current model. First, to utilize our PLAN-B-DF model, abdominal CT images are required, which may impose radiation exposure.[56] However, to obtain imaging biomarkers and apply PLAN-B-DF, only a single CT evaluation is needed. Additionally, in terms of early detection of HCC when it occurs, the occasional use of CT may be superior to surveillance using only ultrasonography.[57] Particularly in patients with cirrhosis or steatotic liver disease, where poor sonographic visualization or suboptimal assessment is expected, CT has been shown to be superior to ultrasonography for surveillance, according to previous studies.[58–60] In fact, approximately 40% of the included patients in our cohorts had cirrhosis, and the annual incidence of HCC was relatively high. Furthermore, the imaging biomarkers addressed in this study can be extracted from low-dose CT, which has recently been investigated to improve the efficiency of HCC surveillance.[61] In addition, the utilization of CT allows for operator-independent acquisition of imaging biomarkers compared to ultrasonography. Thus, future research is necessary to weigh the potential harms *vs*. the benefits of CT evaluation for individuals who are at risk of HCC. It would also be advantageous to investigate the possibility of substituting the imaging biomarkers utilized in the PLAN-B-DF model with other biomarkers extracted from non-radiation-based imaging modalities, such as MRI or bioimpedance analysis. Second, being a single-nation study predominantly involving genotype C HBV, the generalizability of our results may be limited and further validation of PLAN-B-DF is warranted in an independent international cohort. Third, the varying lengths of longitudinal follow-up in the retrospective cohorts may lead to right-censoring of the time-to-event data, potentially introducing bias in Harrell's c-index. To address this bias, we utilized additional performance metrics, including time-dependent AUROCs, AUPRCs, and another concordance index (*i.e*., GHCI), which were consistent with our main results.

In conclusion, our study establishes a highly accurate AI prediction model for HCC in patients with CHB, utilizing a combination of imaging biomarkers and clinical variables. This model demonstrates significant improvements in predictive accuracy and risk stratification compared to preexisting models and may offer valuable enhancements for HCC surveillance and management.

Affiliations

[1]*Center for Liver and Pancreatobiliary Cancer, National Cancer Center, Goyang, Gyeonggi-do, Korea;* [2]*Department of Internal Medicine and Liver Research Institute, Seoul National University College of Medicine, Seoul, Korea;* [3]*Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea;* [4]*Department of Internal Medicine, School of Medicine, Kyungpook National University, Daegu, Korea;* [5]*Divisions of Gastroenterology and Hepatology, Department of Internal Medicine, Kyung Hee University Hospital, College of Medicine, Kyung Hee University, Seoul, Korea;* [6]*Department of Internal Medicine, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Gyeonggi-do, Korea;* [7]*Seoul HIM Clinic, Seoul, Korea;* [8]*Department of Internal Medicine, ABC Hospital, Hwaseong, Gyeonggi-do, Korea;* [9]*Department of Internal Medicine, College of Medicine, Chung-Ang University, Seoul, Korea;* [10]*Division of Gastroenterology, Liver Center, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea;* [11]*Department of Internal*

Medicine, College of Medicine, Ewha Womans University, Seoul, Korea; [12]*Department of Internal Medicine, Yonsei University College of Medicine, Seoul, Korea;* [13]*AI Center, MedicalIP. Co. Ltd., Seoul, Korea;* [14]*Department of Radiology, Seoul National University College of Medicine, Seoul, Korea;* [15]*Department of Anesthesiology, Seoul National University College of Medicine, Seoul, Korea;* [16]*Inocras Inc., San Diego, CA, USA*

Abbreviations

AI, artificial intelligence; CHB, chronic hepatitis B; CT, computed tomography; GBM, gradient-boosting machine; GHCI, Gönen and Heller's c-index; HBeAg, hepatitis B envelope antigen; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HU, Hounsfield unit; LSM, liver stiffness measurement; MRI, magnetic resonance imaging.

Financial support

This work was supported by National IT Industry Promotion Agency grant funded by the Korea Ministry of Science and ICT (No. S0252-21-1001).

Conflict of interest

Dr. Yun Bin Lee reports receiving research grants from Samjin Pharmaceuticals and Yuhan Pharmaceuticals; Dr. Yoon Jun Kim receives research grants from BTG, Boston Scientific, AstraZeneca, Gilead Sciences, Samjin, BL&H, and Bayer, and lecture fees from Roche, Abbvie, Eisai, Boston Scientific, BMS, BTG, Bayer, MSD, Novo Nordisk, GC Cell, Boehringer Ingelheim, and Gilead Sciences; Dr. Jung-Hwan Yoon reports receiving research grants from Bayer HealthCare Pharmaceuticals, Daewoong Pharmaceuticals, and Bukwang Pharmaceuticals; Dr. Jeong-Hoon Lee reports receiving research grants from Yuhan Pharmaceuticals, and lecture fees from GC Cell, Daewoong Pharmaceuticals, Gilead Korea, and Samil Pharmaceuticals; No other potential conflict of interest relevant to this article was reported.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Jeong-Hoon Lee had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis.
Specific author contributions: Conceptualization and Methodology: Hyunjae Shin, Moon Haeng Hur, Jeong-Hoon Lee; Software: Hyunjae Shin, Moon Haeng Hur, Jong-Min Kim; Validation: Moon Haeng Hur, Jeong-Hoon Lee; Formal Analysis: Hyunjae Shin; Data Curation and Investigation: Hyunjae Shin, Moon Haeng Hur, Byeong Geun Song, Gi-Ae Kim, Soo Young Park; Resources: Gwanghyeon Choi, Joon Yeul Nam, Minseok Albert Kim, Youngsu Park, Yunmi Ko, Jeayeon Park, Han Ah Lee, Na Ryoung Choi, Min Kyung Park, Yun Bin Lee, Dong Hyun Sinn, Seung Up Kim, Hwi Young Kim, Jong-Min Kim, Hyung-Chul Lee, Dong Ho Lee, Jin Wook Chung, Yoon Jun Kim, Jung-Hwan Yoon, Moon Haeng Hur, Jeong-Hoon Lee; Writing – Original Draft Preparation: Hyunjae Shin; Writing – Review & Editing: Moon Haeng Hur, Jeong-Hoon Lee; Visualization: Hyunjae Shin, Moon Haeng Hur, Jong-Min Kim; Supervision: Jeong-Hoon Lee; Approval of the final article: all authors.

Data availability statement

Patient data in this study are not publicly available due to the IRB regulations of each center. However, data may be provided upon a reasonable request and following approval by each center.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2024.12.029.

References

Author names in bold designate shared co-first authorship

[1] Devarbhavi H, Asrani SK, Arab JP, et al. Global burden of liver disease: 2023 update. J Hepatol 2023;79:516–537.
[2] 2022 KLCA-NCC Korea practice guidelines for the management of hepatocellular carcinoma. Clin Mol Hepatol 2022;28:583–705.
[3] Papatheodoridis GV, Lampertico P, Manolakopoulos S, et al. Incidence of hepatocellular carcinoma in chronic hepatitis B patients receiving nucleos(t)ide therapy: a systematic review. J Hepatol 2010;53:348–356.
[4] Tseng CH, Hsu YC, Chen TH, et al. Hepatocellular carcinoma incidence with tenofovir versus entecavir in chronic hepatitis B: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2020;5:1039–1052.
[5] **Kim HY**, Lampertico P, Nam JY, et al. An artificial intelligence model to predict hepatocellular carcinoma risk in Korean and Caucasian patients with chronic hepatitis B. J Hepatol 2022;76:311–318.
[6] Wong VW, Chan SL, Mo F, et al. Clinical scoring system to predict hepatocellular carcinoma in chronic hepatitis B carriers. J Clin Oncol 2010;28:1660–1665.
[7] Choi WM, Kim GA, Choi J, et al. Non-linear association of baseline viral load with on-treatment hepatocellular carcinoma risk in chronic hepatitis B. Gut 2024;73:649–658.
[8] **Chun HS**, Papatheodoridis GV, Lee M, et al. PAGE-B incorporating moderate HBV DNA levels predicts risk of HCC among patients entering into HBeAg-positive chronic hepatitis B. J Hepatol 2024;80:20–30.
[9] Li W, Deng R, Liu S, et al. Hepatitis B virus-related hepatocellular carcinoma in the era of antiviral therapy: the emerging role of non-viral risk factors. Liver Int 2020;40:2316–2325.
[10] Lin H, Li G, Delamarre A, et al. A liver stiffness-based etiology-independent machine learning algorithm to predict hepatocellular carcinoma. Clin Gastroenterol Hepatol 2024;22:602–610.
[11] Feng H, Wang X, Zhao T, et al. Myopenic obesity determined by visceral fat area strongly predicts long-term mortality in cirrhosis. Clin Nutr 2021;40:1983–1989.
[12] Ko YH, Wong TC, Hsu YY, et al. The correlation between body fat, visceral fat, and nonalcoholic fatty liver disease. Metab Syndr Relat Disord 2017;15:304–311.
[13] Kim MN, Han K, Yoo J, et al. Increased risk of hepatocellular carcinoma and mortality in chronic viral hepatitis with concurrent fatty liver. Aliment Pharmacol Ther 2022;55:97–107.
[14] Cespiati A, Meroni M, Lombardi R, et al. Impact of sarcopenia and myosteatosis in non-cirrhotic stages of liver diseases: similarities and differences across aetiologies and possible therapeutic strategies. Biomedicines 2022;10:182.
[15] Zheng J, Xia Y, Xu A, et al. Combined model based on enhanced CT texture features in liver metastasis prediction of high-risk gastrointestinal stromal tumors. Abdom Radiol (NY) 2022;47:85–93.
[16] Jeon SK, Joo I, Park J, et al. Fully-automated multi-organ segmentation tool applicable to both non-contrast and post-contrast abdominal CT: deep learning algorithm developed using dual-energy CT images. Sci Rep 2024;14:4378.
[17] KASL clinical practice guidelines for management of chronic hepatitis B. Clin Mol Hepatol 2022;28:276–331.
[18] Lee YS, Hong N, Witanto JN, et al. Deep neural network for automatic volumetric segmentation of whole-body CT images for body composition assessment. Clin Nutr 2021;40:5038–5046.
[19] Chang Y, Yoon SH, Kwon R, et al. Automated comprehensive CT assessment of the risk of diabetes and associated cardiometabolic conditions. Radiology 2024;312:e233410.
[20] Park J, Joo I, Jeon SK, et al. Automated abdominal organ segmentation algorithms for non-enhanced CT for volumetry and 3D radiomics analysis. Abdom Radiol (NY) 2024. s00261-024-04581-5.
[21] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. San Francisco, California, USA: Association for Computing Machinery; 2016. p. 785–794.
[22] 2022 KLCA-NCC Korea practice guidelines for the management of hepatocellular carcinoma. J Liver Cancer 2023;23:1–120.
[23] European Association for the Study Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2018;69:182–236.
[24] Marrero JA, Kulik LM, Sirlin CB, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American association for the study of liver diseases. Hepatology 2018;68:723–750.
[25] Papatheodoridis G, Dalekos G, Sypsa V, et al. PAGE-B predicts the risk of developing hepatocellular carcinoma in Caucasians with chronic hepatitis B on 5-year antiviral therapy. J Hepatol 2016;64:800–806.
[26] **Kim JH**, Kim YD, Lee M, et al. Modified PAGE-B score predicts the risk of hepatocellular carcinoma in Asians with chronic hepatitis B on antiviral therapy. J Hepatol 2018;69:1066–1073.

[27] Yang HI, Yuen MF, Chan HL, et al. Risk estimation for hepatocellular carcinoma in chronic hepatitis B (REACH-B): development and validation of a predictive score. Lancet Oncol 2011;12:568–574.
[28] Yuen MF, Tanaka Y, Fong DY, et al. Independent risk factors and predictive score for the development of hepatocellular carcinoma in chronic hepatitis B. J Hepatol 2009;50:80–88.
[29] Sharma SA, Kowgier M, Hansen BE, et al. Toronto HCC risk index: a validated scoring system to predict 10-year risk of HCC in patients with cirrhosis. J Hepatol 2018;68:92–99.
[30] Ohki T, Tateishi R, Shiina S, et al. Visceral fat accumulation is an independent risk factor for hepatocellular carcinoma recurrence after curative treatment in patients with suspected NASH. Gut 2009;58:839–844.
[31] Fujiwara N, Nakagawa H, Kudo Y, et al. Sarcopenia, intramuscular fat deposition, and visceral adiposity independently predict the outcomes of hepatocellular carcinoma. J Hepatol 2015;63:131–140.
[32] Zhao J, Lawless MW. Stop feeding cancer: pro-inflammatory role of visceral adiposity in liver cancer. Cytokine 2013;64:626–637.
[33] Fan R, Niu J, Ma H, et al. Association of central obesity with hepatocellular carcinoma in patients with chronic hepatitis B receiving antiviral therapy. Aliment Pharmacol Ther 2021;54:329–338.
[34] von Hessen L, Roumet M, Maurer MH, et al. High subcutaneous adipose tissue density correlates negatively with survival in patients with hepatocellular carcinoma. Liver Int 2021;41:828–836.
[35] Mao X, Cheung KS, Peng C, et al. Steatosis, HBV-related HCC, cirrhosis, and HBsAg seroclearance: a systematic review and meta-analysis. Hepatology 2023;77:1735–1745.
[36] Pickhardt PJ, Blake GM, Graffy PM, et al. Liver steatosis categorization on contrast-enhanced CT using a fully automated deep learning volumetric segmentation tool: evaluation in 1204 healthy adults using unenhanced CT as a reference standard. AJR Am J Roentgenol 2021;217:359–367.
[37] Bae JS, Lee DH, Suh KS, et al. Noninvasive assessment of hepatic steatosis using a pathologic reference standard: comparison of CT, MRI, and US-based techniques. Ultrasonography 2022;41:344–354.
[38] Yoshikawa K, Shimada M, Morine Y, et al. Clinical impact of myosteatosis measured by magnetic resonance imaging on long-term outcomes of hepatocellular carcinoma after radical hepatectomy. BMC Surg 2023;23:281.
[39] Di Cola S, D'Amico G, Caraceni P, et al. Myosteatosis is closely associated with sarcopenia and significantly worse outcomes in patients with cirrhosis. J Hepatol 2024;81:614–650.
[40] **Kim H-K**, Bae S-J, Lee MJ, et al. Association of visceral fat obesity, sarcopenia, and myosteatosis with non-alcoholic fatty liver disease without obesity. Clin Mol Hepatol 2023;29:987–1001.
[41] Cacciottolo TM, Kumar A, Godfrey EM, et al. Spleen size does not correlate with histological stage of liver disease in people with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2023;21:535–537.e1.
[42] Patel M, Tann M, Liangpunsakul S. CT-Scan based liver and spleen volume measurement as a prognostic indicator for patients with cirrhosis. Am J Med Sci 2021;362:252–259.
[43] Geng Y, Shao WQ, Lin J. Spleen to non-cancerous liver volume ratio predicts liver cirrhosis in hepatocellular carcinoma patients. Abdom Radiol (NY) 2023;48:543–553.
[44] Myers RP, Pomier-Layrargues G, Kirsch R, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. Hepatology 2012;55:199–208.
[45] Jang H, Lee YB, Moon H, et al. Aspirin use and risk of hepatocellular carcinoma in patients with chronic hepatitis B with or without cirrhosis. Hepatology 2022;76:492–501.
[46] Lee TY, Hsu YC, Ho HJ, et al. Daily aspirin associated with a reduced risk of hepatocellular carcinoma in patients with non-alcoholic fatty liver disease: a population-based cohort study. EClinicalMedicine 2023;61:102065.
[47] Sharpton SR, Loomba R. Emerging role of statin therapy in the prevention and management of cirrhosis, portal hypertension, and HCC. Hepatology 2023;78:1896–1906.
[48] Zeng RW, Yong JN, Tan DJH, et al. Meta-analysis: chemoprevention of hepatocellular carcinoma with statins, aspirin and metformin. Aliment Pharmacol Ther 2023;57:600–609.
[49] **Choi WM**, Yip TC, Wong GL, et al. Hepatocellular carcinoma risk in patients with chronic hepatitis B receiving tenofovir- *vs.* entecavir-based regimens: individual patient data meta-analysis. J Hepatol 2023;78:534–542.
[50] **Hur MH**, Park MK, Yip TC, et al. Personalized antiviral drug selection in patients with chronic hepatitis B using a machine learning model: a multinational study. Am J Gastroenterol 2023;118:1963–1972.
[51] Nakatsuka T, Tateishi R. Development and prognosis of hepatocellular carcinoma in patients with diabetes. Clin Mol Hepatol 2023;29:51–64.
[52] Shin HS, Jun BG, Yi SW. Impact of diabetes, obesity, and dyslipidemia on the risk of hepatocellular carcinoma in patients with chronic liver diseases. Clin Mol Hepatol 2022;28:773–789.
[53] Xie D, Shi J, Zhou J, et al. Clinical practice guidelines and real-life practice in hepatocellular carcinoma: a Chinese perspective. Clin Mol Hepatol 2023;29:206–216.
[54] European Association for the Study of the Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2018;69:182–236.
[55] Singal AG, Llovet JM, Yarchoan M, et al. AASLD Practice Guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. Hepatology 2023;78.
[56] Berrington de González A, Mahesh M, Kim K-P, et al. Projected cancer risks from computed tomographic scans performed in the United States in 2007. Arch Intern Med 2009;169:2071–2077.
[57] Daher D, Seif El Dahan K, Cano A, et al. Hepatocellular carcinoma surveillance patterns and outcomes in patients with cirrhosis. Clin Gastroenterol Hepatol 2024;22:295–304.e2.
[58] Yoon JH, Lee JM, Lee DH, et al. A comparison of biannual two-phase low-dose liver CT and US for HCC surveillance in a group at high risk of HCC development. Liver Cancer 2020;9:503–517.
[59] Gupta P, Soundararajan R, Patel A, et al. Abbreviated MRI for hepatocellular carcinoma screening: a systematic review and meta-analysis. J Hepatol 2021;75:108–119.
[60] **Park MK**, Lee DH, Hur BY, et al. Effectiveness of US surveillance of hepatocellular carcinoma in chronic hepatitis B: US LI-RADS visualization score. Radiology 2023;307:e222106.
[61] Kang HJ, Lee JM, Ahn C, et al. Low dose of contrast agent and low radiation liver computed tomography with deep-learning-based contrast boosting model in participants at high-risk for hepatocellular carcinoma: prospective, randomized, double-blind study. Eur Radiol 2023;33:3660–3670.

Keywords: radiologic biomarker; deep learning; visceral fat; myosteatosis; segmentation.
Received 23 June 2024; received in revised form 12 November 2024; accepted 7 December 2024; Available online xxx