Deep-Learning Based Analysis of Preoperative MRI Predicts Microvascular Invasion and Outcome in Hepatocellular Carcinoma

Bao-Ye sun  
Zhongshan Hospital Fudan University Department of Liver Surgery

Pei-Yi Gu  
Tongji University

Ruo-Yu Guan  
Zhongshan Hospital Fudan University Department of Liver Surgery

Cheng Zhou  
Zhongshan Hospital Fudan University Department of Liver Surgery

Jian-Wei Lu  
Tongji University

Zhang-Fu Yang  
Zhongshan Hospital Fudan University Department of Liver Surgery

Chao Pan  
Tongji University

Pei-Yun Zhou  
Zhongshan Hospital Fudan University Department of Liver Surgery

Ya-Ping Zhu  
Tongji University

Jia-Rui Li  
Tongji University

Zhu-Tao Wang  
Zhongshan Hospital Fudan University Department of Liver Surgery

Shan-Shan Gao  
Zhongshan Hospital Fudan University

Yong Yi  
Zhongshan Hospital Fudan University Department of Liver Surgery

Ye Luo  
Tongji University

Shuang-Jian Qiu  
Zhongshan Hospital Fudan University Department of Liver Surgery  
https://orcid.org/0000-0003-2056-6127
Abstract

Background & Aims:
Preoperative prediction of microvascular invasion (MVI) is critical for treatment strategy making in patients with hepatocellular carcinoma (HCC). We aimed to develop a deep learning (DL) model based on preoperative dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) to predict the MVI status and clinical outcomes in patients with HCC.

Methods
We retrospectively included a total of 321 HCC patients with pathologically confirmed MVI status. Preoperative DCE-MRI of these patients were collected, annotated and further analyzed by DL in this study. A predictive model for MVI integrating DL-predicted MVI status (DL-MVI) and clinical parameters was constructed with multivariate logistic regression.

Results
Of 321 HCC patients, 136 patients were pathologically MVI absent and 185 patients were MVI present. Recurrence-free survival (RFS) and overall survival (OS) were significantly different between the DL-predicted MVI-absent and MVI-present. Among all clinical variables, only DL-predicted MVI status and AFP were independently associated with MVI: DL-MVI (odds ratio [OR]=35.738; 95% confidence interval [CI]: 14.027-91.056; p<0.001), AFP (OR=4.634, 95% CI: 2.576-8.336; p<0.001). To predict the presence of MVI, DL-MVI combined with AFP achieved an area under the curve (AUC) of 0.824.

Conclusions
Our predictive model combining DL-MVI and AFP achieved good performance for predicting MVI and clinical outcomes in patients with HCC.

Introduction
Hepatocellular carcinoma (HCC) ranks the sixth most common malignancies worldwide and its incidence is increasing annually\(^1\). Surgical resection, liver transplantation and locoregional therapies may be potentially curative modality for HCC patients, whereas post-operative recurrence rate remains high, mainly due to the presence of vascular invasion\(^2,3\).

Microvascular invasion (MVI) is among the most vital prognostic factors for HCC and is a major risk indicator for early recurrence during the first 2 years after surgical resection\(^4-6\). MVI is defined as microscopic invasion of tumor cells within a vascular space lined by endothelium like smaller
intrahepatic vessels, including micro-vessels of portal vein or hepatic artery and small lymphatic vessels.

Although macrovascular invasion can be detected with diagnostic imaging, MVI is a histologic finding that can rarely be determined preoperatively. Currently, preoperative prediction of MVI remains challenging, despite several studies claimed that imaging features extracted from computed tomography (CT) and magnetic resonance imaging (MRI) were predictive of MVI. Gd-EOB-DTPA-enhanced MRI was reported to have a high value in predicting presence of MVI in HCC. MR imaging features, including arterial peritumoral enhancement, tumor margins, tumor size were independently associated with MVI, while these imaging features were extracted visually by experienced radiologists, limiting its clinical use. Additionally, it has been reported that a radio-genomic venous invasion (RVI) predictor, combining imaging features with gene expression, predicting accurately MVI in HCC. Through radiomic analysis of contrast-enhanced CT, Xun Xu et al. developed a computational approach integrating large scale clinic-radiologic and radiomic features to predict MVI and long-term clinical outcomes of patients with HCC. However, these criteria for a preoperative imaging diagnosis of MVI in HCC have not yet been widely recognized.

In recent years, with the continuous advancements achieved in computer science, deep learning (DL), with artificial intelligence as its core, has been paid more and more attention in the medical field. Compared with traditional empirical medicine, medical intelligence can integrate a large scale of existing data and experience to facilitate medical diagnosis and treatment. Image recognition is a now mature field in deep learning, whose research has gone into the analysis of medical images, such as the discrimination between benign occupancy and malignant nodules, the location of organs and lesions, the division of organs and its substructures, etc. Deep learning analysis of H-E scan slices (convolutional neural network, Resnet18) achieved an area under the curve (AUC) of 0.81-0.84 in predicting gastrointestinal tumor microsatellite instability (MSI). Deep learning also outperformed many experienced dermatologists in melanoma image classification. Deep learning neural networks based on magnetic resonance imaging (MRI), X-ray computer tomography (CT) and PET/CT have had great achievements in the characterization of prostate cancer, pulmonary nodules, hepatocellular carcinoma or benign occupancy.

Moreover, the accuracy could be further enhanced by the ability of deep learning to quickly compute high-dimensional data, based on real-time disease location and subsequent analysis of dynamic video such as endoscopy. Deep learning, combined with molecular expression information, high-throughput sequencing, and multi-group data is also a research area of concern in this field.

However, to the best of our knowledge, there have been few attempts to evaluate the diagnostic performance of deep learning in mining MR imaging features for predicting MVI of HCC and long-term clinical outcomes. This study aimed to investigate whether deep learning analysis of preoperative MR
imaging could be used to predict MVI, to determine its diagnostic performance and to evaluate whether it is associated with outcome in HCC patients.

**Material And Methods**

**Study design and patient population**

This retrospective study was approved by Zhongshan hospital Ethics committee and the requirement for written informed consent was waived. All procedures involving human participants were performed in accordance with the 1975 Helsinki declaration and its later amendments.

We queried our institution’s medical records to derive data from patients who underwent hepatic resection for HCC in year 2015 and year 2018 respectively. The key inclusion criteria for our study were as follows: (1) resectable HCC lesion without macroscopic vascular invasion; (2) underwent preoperative gadoxetic acid–enhanced and DW liver MR imaging within 1 month before surgery; (3) without a history of preoperative anti-cancer treatment; (4) pathological confirmation of HCC (5) MR imaging quality adequate for analysis. Exclusion criteria included: (1) received other anti-tumor therapies before surgery; (2) incomplete clinical or pathological information.

A total of 321 confirmed cases of HCC were identified, with 149 HCC patients forming the 2015 cohort and 172 patients forming 2018 cohort, according to the inclusion and exclusion criteria. Data of some preoperative laboratory examinations were collected, including liver function tests, hepatitis B and C immunology, serum a-fetoprotein (AFP) level, serum alanine aminotransferase (ALT), aspartate aminotransferase (AST), γ-glutamyl transpeptidase (GGT), serum total bilirubin (TB), conjugated bilirubin (CB), serum albumin (ALB), platelet count (PLT), prothrombin time (PT), international normalized ratio (INR). The diagnosis of HCC was histologically or clinically confirmed based on the criteria of the American Association for the Study of Liver Diseases (AASLD)³⁶.

**MR imaging acquisition**

All HCC patients underwent preoperative Gadoxetic acid-enhanced MR imaging examination by a 1.5T scanner (Siemens Healthcare, Erlangen, Germany). Image acquisition procedures were performed as previously reported³⁷. Namely, MR Imaging sequences included axial T2-weighted imaging(T2), diffusion-weighted imaging (DWI), in-phase and opposed-phase T1-weighted imaging(T1), and pre-contrast and post-contrast dynamic three-dimensional T1-weighted volumetric-interpolated breath-hold examination (VIBE) at arterial phase (T1A) , portal venous phase (T1V), delayed phase (T1D) after injection of 0.025 mmol/kg of gadoxetic acid (Primovist, Bayer Schering Pharma, Berlin, Germany) into the cubital vein, followed by a 20-mL saline flush.

**Deep learning network architecture and workflow**
Considering that different modalities of MRI contain different features to characterize MVI, analyzing the effect of different modalities and combinations of modalities for MVI prediction is necessary and important. To analyze MR Images from different modality, we first performed the feature extraction individually and then fused the extracted features to predict the status of MVI. Specifically, the whole procedure was divided into three steps: single modal image feature extraction, feature fusion and feature normalization (Fig.1).

**Step1: feature extraction from a single modality**

Given a 2D slice image from a single modality, a region of interest (ROI) in a rectangle shape from the original image was cropped by two experienced doctors. In order to focus on the boundary region of the tumor, we enlarged the ROI by 5-10 pixels at every boundary of the rectangular ROI. In other words, the inputs of our deep learning MR analysis model were the enlarged ROIs.

Considering that the large amount of training data could improve the performance of the model, we used data augmentation method to increase number of ROIs including image flipping, image scaling, adding gaussian noise. In order to make all the ROIs have the same size, we resized them into 320 320, then the processed ROI were input into the conventional neural network (CNN) network. The detailed network architecture of the CNN is as following. The first part of the CNN includes one $64 \times 7 @ 7$ convolution layer, a normalization module and a max pooling layer. After going through all these layers, we obtained a feature map of the input ROI roughly. The following structure of the CNN network contains six $64 \times 3 @ 3$ convolution bottleneck modules and six $128 \times 3 @ 3$ convolution bottleneck modules, in which all the bottleneck modules are employed from Resnet. The raw feature map of the ROI obtained at the first part was processed by these bottleneck modules sequentially, thus the feature of the ROI at different scales could be learned accordingly. The motivation behind this design of the network is that different scales of convolution bottleneck can help the CNN network learn feature with diverse scales. Features learned at shallow layers pay much attention to the detailed structure information while features learned at deep layers care more about the global information. Considering the target to analyze MR image is to predict the MVI level which is a global characteristic of an image, we directly use the output feature of the last $128 \times 3 @ 3$ convolution bottleneck module which is a 512-dimensional feature vector and feed into the next fully connected layer (FC).

**Step2: Feature fusion from multiple modalities**

For each patient in 2015 HCC cohort, 3 continuous slices showing the maximal diameter of the tumor were first exported from six modalities (i.e. T1, T1A, T1V, T1D, T2, DWI), and then analyzed in our model. As mentioned above, the feature of an image from each modality is extracted beforehand. Theoretically, we could use a CNN network with six branches sharing same weights to extract six modality images simultaneously, and then concatenated the learned six features as a fused one to feed into the FC layer. After passing the FC layer and the SoftMax, a confidence score identifying the level of MVI of the patient was obtained. However, the prediction performance of the CNN network by combining all six modalities
was not as well as expected, and three of six modalities (i.e. T1A, T1V, T1D) after empirically evaluation provides the best performance. More details could be referred to the experiment Section.

**Step3: Feature normalization**

After concatenating the features from different modalities, the fused feature was feed into a FC layer combined with a SoftMax classifier, which helped to normalize the feature into $n$ 1 bin vector. Here $n$ is the number of MVI levels.

**Step4: Network Training and Testing**

Considering that we had 3 continuous slices for each modality, anyone of the three slices could be used to characterize the MVI information from this modality. Thus, for each patient, 27 different slice combinations from three modalities (i.e. T1A, T1V, T1D) could be used as the input of the CNN network. In other words, for each patient, 27 training samples sharing the same label are given. Different from any data augmentation techniques, data shuffling like above is a unique way employed in our model thus the performance of our network could be further improved.

Given a testing sample, in the inference step, any one of the 27 different slice combinations could be used as the input of the trained CNN network. Without loss of the generality, we simply proceed to the first slice of 3 continuous slices for each modality.

**Histopathology**

All surgical specimens were examined by 2 experienced pathologists, particularly to detect the presence of MVI. MVI was defined as the presence of tumor invasion in smaller intrahepatic vessels including a portal vein, hepatic vein, or a large capsular vessel of the surrounding hepatic tissue lined by endothelium that was visible only on microscopy. MVI grade is classified as M1: the number of MVI < 5 and the distance of MVI ≤ 1 cm away from the tumor tissues, and M2: the number of MVI > 5 or the distance of MVI > 1 cm away from the tumor tissues, according to the practice guidelines for the Pathological Diagnosis of Primary Liver Cancer of China. The histologic parameters ordinarily included Edmondson-Steiner grade, size, surgical margin and MVI status of the tumor.

**Statistical analysis**

Statistical analysis was performed using SPSS v.25 (IBM Inc., Armonk, NY, USA) and R software (R software version 3.5.2, R Project for Statistical Computing, http://www.r-project.org). The discrimination performance of the DL predictive model was measured by the area under the ROC curve (AUC) value in the primary training/validation set. Calibration curves were plotted to analyze the diagnostic performance of the predictive model in the overall cohort. Decision curve analysis was conducted to determine the clinical usefulness and net benefits of the developed predictive model.
Patients were consistently followed up since the date of surgical resection at intervals of 2 to 3 months. Recurrence-free survival (RFS) and overall survival (OS) were defined as the interval between surgery and detection of first recurrence or death. Patients were censored in case of emigration, or on 31 Dec 2020, whichever came first. Survival curves were plotted using the Kaplan-Meier method and compared by log-rank test. A two-tailed p value <0.05 was considered statistically significant.

**Results**

**Baseline clinical characteristics**

Among the 321 patients enrolled in our study, histologic MVI was diagnosed in explanted tissue of 185 patients (57.6%). Patients with MVI had higher ALT, AST, GGT and AFP levels than those without MVI. Patients with MVI and patients without were similar in their distribution of sex, hepatic virus infection, cirrhosis, Child-Pugh stage, TB, CB, ALB, and PT. Risk coefficient estimated by univariate analysis is summarized in Table 1.

**Deep learning analysis of MR Images**

Before conducting the analysis, we used 50% random flipping, 0.75-1 random cropping and normalization as data augmentation. By this way to generate more data and make the deep neural network generalized well. The network is trained and optimized by Adam optimizer with 200 epochs and batch size of 2. We set initial learning rate to 0.1 and the whole experiments were conducted on a GTX1080 GPU card. Our training dataset was collected in 2015. We split the 2015 HCC cohort data into training set and validation set at a ratio of 9:1.

**Effectiveness of single modality:** To validate the effectiveness of the single modality, we used the same model structure to extract features, but input classifier without feature fusion. We found that compared to the best result by modal combination among T1, T1D, and T1V, single modality didn’t perform well. We got 63.19% accuracy for T1V modal, 58.91% accuracy for T1D model and 66.66% for T1 modal.

**Effectiveness of Multi-modalities:** Since we had 6 modalities of MRI image data, we’ve done the ablation study to figure out which kind of modality combination can lead to the best classification result. As shown in Table 2, the combination of T1, T1D and T1V resulted in the highest accuracy 92.11%. Meanwhile, we noticed that the modality of DWI was not a proper modal for MVI classification.

**Generalization between different cohorts:** We’ve tested our model on external validation dataset which was collected in 2018. And our model was trained on the dataset collected in 2015. The result showed our model did not work well on 2018 dataset. We obtained the accuracy of 68.69%, the precision of 76.92%, the recall of 75.76% and the F1-score of 76.34%.

In order to analyze the reasons for the performance drop when the network was trained on HCC 2015 cohort and tested on 2018 HCC cohort data, we used t-SNE algorithm to reduce the dimensionality of the data into 2D such that we could display and analyze the difference between the two cohort data sets.
visually as shown in Figure 2. The first row of Figure 2 showed the distribution difference for M0 group with modalities T1, T1V and T1D, respectively (Fig.2A). From the first row of Figure 2, we can see that the two datasets can be obviously separated on the modalities of T1 and T1D, but are blended on the modality of T1V. Similar results were shown in the second and the third rows for the M1 and the M2 groups simultaneously (Fig.2B-C). In other words, no matter how to select the combinations of the modalities among T1, T1D, T1V, there at least one modality cannot achieve satisfied classification results among M0, M1 and M2. The feature distribution inconsistency led to the bad classification performance of our model on 2018 HCC cohort data. Finally, we used the deep learning model constructed above to predict MVI status of the overall cohort and the results were denoted as DL-predicted MVI status (DL-MVI).

Predictors of survival

As of Dec. 2020, all the patients had completed the OS follow-up and PFS follow-up. The overall recurrence rate was 31.5% (101/321) and the overall death rate was 19.9% (64/321). The median OS of the patients was 59.5 months and patients with MVI had a median OS of 54.7 months (Fig. 3A). The median OS was 54.7 months for those with DL-predicted MVI presence and was not reached for those with DL-predicted MVI absence (Fig. 3C). The median PFS of the patients was 50.4 months, particularly 32.5 for patients with MVI and it was not reached for those without MVI (Fig. 3B). The median PFS was 36.3 months for patients with DL-predicted MVI presence and not reached for those with DL-predicted MVI absence (Fig. 3D).

Construction of DL-based predictive model for MVI

Among all clinical parameters, 4 clinical variables (ALT, AST, GGT and AFP) were identified by univariate logistic analysis. In the multivariate regression model, only 2 predictors were independent prognostic factors of histologic MVI: higher AFP (>20 ng/mL), DL-predicted MVI presence (Table 3). These independently associated risk factors were furthered enrolled to form the predictive model (Fig.4A), described by the formula: Y=-3.51+1.53×AFP+3.58×DL-MVI. The resulting DL-based predictive model demonstrated good accuracy in predicting the risk of MVI, with an AUC of 0.824 (Fig.4B). The calibration curve of the model demonstrated good agreement between predicted and observed MVI in the primary cohort (Fig.4C). The Decision curve for the predictive model is demonstrated in Fig.4D, the net benefit of the decision curve for the predictive nomogram is higher than that for assuming all patients have MVI when the threshold probability >4%.

Discussion

Recurrence and metastasis are the main reasons for poor prognosis in post-operative HCC patients. Approximately 70% of HCC patients treated with surgical resection develop a recurrence within 5 years\(^9\). Early recurrences, within 2 years after tumor resection, are frequently attributed to residual intrahepatic metastases. MVI is frequently present in HCC, and highly associates with several adverse biological...
markers, such as high grade, large tumor size, and elevated serum AFP. The presence of MVI more accurately predicted higher recurrence risk and poor clinical outcomes than factors included in the Milan criteria. Moreover, MVI determines the risk for intrahepatic or distant dissemination of malignant cells, and MVI-positive HCCs should potentially be treated with a wider resection margin compared with MVI-negative tumors.

The aim of this study was to investigate whether the DL-assisted model derived from large-scale clinical and imaging data, especially imaging features from DCE-MRI could be able to preoperatively predict MVI status and clinical outcomes in a cohort of 321 patients with HCC. Preoperative MVI status prediction is principle for clinicians to adopt appropriate therapeutic strategies, contributing to improve HCC patients’ overall survival. Histologic MVI has been claimed to be associated with poor HCC prognosis in many studies. Similar results were obtained in our primary cohort, patients with different DL-predicted MVI status, histologic MVI had different clinical outcomes.

Recently, there have been several studies attempting to predict MVI using only clinical parameters. Radiomics has been recently viewed as a vital imaging technology in medical oncology. Combining radiomics based on CT or MRI with clinical variables achieve the AUC from 0.796 to 0.906. However, the challenge of radiomics method is based on manually-defined precise boundary of the tumor, resulting in poor inter-reader reliability, and the results may not truly reflect the edge features of the target tumor.

The emerging DL method represents a new choice, due to its ability to integrate a large scale of clinical and imaging data. A recent study using DL based on preoperative CT showed a considerable efficacy (AUC: 0.906) in predicting MVI. Two other independent studies using DCE-MRI and 3D Convolutional Neural Networks instead of CT images to predict MVI achieved an AUC of 0.931 and 0.926 respectively.

In our study, higher serum AFP level (>20 ng/mL) and DL-predicted MVI presence were independently associated with histologic MVI by both univariate and multivariate logistic analysis, thus they were furthered included in the predictive model. The resulting DL-based predictive model demonstrated good accuracy in predicting the risk of MVI, with an AUC of 0.824. Similarly, a previous study achieved an AUC of 0.81 by combining DL With 3D Convolutional Neural Network for noninvasive prediction of MVI in HCC.

Algorithmically, the DL model applied in this study is a multi-input network. As we used six image sequences for MVI prediction, we first evaluated the effectiveness of single modality and then multi-modalities. It turned out that single modality didn't perform well as compared to multi-modalities. Furthermore, the combination of T1, T1D and T1V resulted in the highest accuracy (92.11%) in the training cohort. However, this model did not perform well on the validation dataset. In order to analyze the reasons for this apparent discrepancy between the 2015 training cohort and the 2018 validation cohort, we used t-SNE algorithm to reduce the dimensionality of the data into 2D such that we could display and
analyze the difference between the two cohort data sets visually. The results suggested that feature distribution inconsistency led to the bad classification performance of our model on 2018 HCC cohort data.

Several limitations of this study should be noted. First, because of the inherent character of a retrospective study, potential bias is possible. Second, this study was a single-center experience limited to our medical center, and the study results should be validated and reproduced by external medical centers.

In conclusion, we systematically investigated the large-scale clinical and MR imaging data of patients with HCC undergoing surgical resection with the assistance of DL for noninvasive prediction of MVI. Our predictive model integrating deep learning and serum AFP level demonstrated good performance for predicting MVI and clinical outcomes in patients with HCC.

**Abbreviations**

**MVI:** Microvascular invasion  
**HCC:** Hepatocellular carcinoma  
**DCE-MRI:** Dynamic contrast-enhanced magnetic resonance imaging  
**DL:** Deep learning  
**CNN:** Conventional neural network  
**AUC:** Area under curve  
**OS:** Overall survival  
**RFS:** Recurrence-free survival  
**CI:** Confidence interval  
**OR:** Odds ratio

**Declarations**

**Ethics approval and consent to participate**

This study was approved by the Research Ethics Committee of Zhongshan Hospital (Approval No: B2021-444) and informed consent was obtained from all patients. All procedures performed in this study were in accordance with the ethical standards of Zhongshan Hospital, Fudan University and with the 1964 Helsinki declaration.

**Funding**
This project is supported by National Natural Science Foundation of China (NO.81772510, NO.82072677, NO. 61806147 and NO.82072672), Clinical Research Project of Zhongshan Hospital (NO. 2020ZSLC62 and 2020ZHZS17).

**Competing interests**

The authors declare no competing interest of this study.

**Consent for publication**

All authors have read the final version and agreed on the publication.

**Availability of data and materials**

All data generated during this study are listed in this article. The data are available from the corresponding author upon reasonable request.

**Acknowledgement**

We acknowledged the Department of Pathology from Zhongshan Hospital for assisting in pathologic diagnosis.

**Author contributions**

Study concept, design, and supervision: Shuang-Jian Qiu, Ye Luo, Yong Yi;

MRI image annotation: Bao-Ye Sun, Zhang-Fu Yang, Zhu-Tao Wang, Shan-Shan Gao;

Deep learning network analysis: Pei-Yi Gu, Jian-Wei Lu, Chao Pan, Ya-Ping Zhu, Jia-Rui Li;

Analysis and interpretation of data: Bao-Ye Sun, Pei-Yi Gu, Ruo-Yu Guan;

Drafting of the manuscript: Bao-Ye Sun, Ruo-Yu Guan, Cheng Zhou, Pei-Yi Gu;

Preparation of figures and tables: Bao-Ye Sun, Pei-Yi Gu, Ruo-Yu Guan, Pei-Yun Zhou, Cheng Zhou;

Critical revision: Shuang-Jian Qiu, Ye Luo, Yong Yi;

**References**

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.

2. Villanueva A. Hepatocellular Carcinoma. N Engl J Med. 2019;380(15):1450-1462.

3. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet. 2018;391(10127):1301-1314.
4. Lim KC, Chow PK, Allen JC, et al. Microvascular invasion is a better predictor of tumor recurrence and overall survival following surgical resection for hepatocellular carcinoma compared to the Milan criteria. Ann Surg. 2011;254(1):108-113.

5. Zhou YM, Yang JM, Li B, et al. Risk factors for early recurrence of small hepatocellular carcinoma after curative resection. Hepatobiliary Pancreat Dis Int. 2010;9(1):33-37.

6. Imamura H, Matsuyama Y, Tanaka E, et al. Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. J Hepatol. 2003;38(2):200-207.

7. Roayaie S, Blume IN, Thung SN, et al. A system of classifying microvascular invasion to predict outcome after resection in patients with hepatocellular carcinoma. Gastroenterology. 2009;137(3):850-855.

8. Rodriguez-Peralvarez M, Luong TV, Andreana L, Meyer T, Dhillon AP, Burroughs AK. A systematic review of microvascular invasion in hepatocellular carcinoma: diagnostic and prognostic variability. Ann Surg Oncol. 2013;20(1):325-339.

9. Lee S, Kim SH, Lee JE, Sinn DH, Park CK. Preoperative gadoxetic acid-enhanced MRI for predicting microvascular invasion in patients with single hepatocellular carcinoma. J Hepatol. 2017;67(3):526-534.

10. Banerjee S, Wang DS, Kim HJ, et al. A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma. Hepatology. 2015;62(3):792-800.

11. Huang M, Liao B, Xu P, et al. Prediction of Microvascular Invasion in Hepatocellular Carcinoma: Preoperative Gd-EOB-DTPA-Dynamic Enhanced MRI and Histopathological Correlation. Contrast Media Mol Imaging. 2018;2018:9674565.

12. Zhang S, Xu G, Duan C, et al. Radiomics Analysis of MR Imaging with Gd-EOB-DTPA for Preoperative Prediction of Microvascular Invasion in Hepatocellular Carcinoma: Investigation and Comparison of Different Hepatobiliary Phase Delay Times. Biomed Res Int. 2021;2021:6685723.

13. Lei Z, Li J, Wu D, et al. Nomogram for Preoperative Estimation of Microvascular Invasion Risk in Hepatitis B Virus-Related Hepatocellular Carcinoma Within the Milan Criteria. JAMA Surg. 2016;151(4):356-363.

14. Xu X, Zhang HL, Liu QP, et al. Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. J Hepatol. 2019;70(6):1133-1144.

15. Qaiser T, Tsang YW, Taniyama D, et al. Fast and accurate tumor segmentation of histology images using persistent homology and deep convolutional features. Med Image Anal. 2019;55:1-14.

16. Kather JN, Pearson AT, Halama N, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med. 2019;25(7):1054-1056.

17. Hekler A, Utikal JS, Enk AH, et al. Pathologist-level classification of histopathological melanoma images with deep neural networks. Eur J Cancer. 2019;115:79-83.
18. Brinker TJ, Hekler A, Haushild A, et al. Comparing artificial intelligence algorithms to 157 German dermatologists: the melanoma classification benchmark. Eur J Cancer. 2019;111:30-37.

19. Brinker TJ, Hekler A, Enk AH, et al. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. Eur J Cancer, 2019;113:47-54.

20. Brinker TJ, Hekler A, Enk AH, et al. A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. Eur J Cancer, 2019;111:148-154.

21. Dascalu A, David EO. Skin cancer detection by deep learning and sound analysis algorithms: A prospective clinical study of an elementary dermoscope. EBioMedicine. 2019;43:107-113.

22. Yasaka K, Akai H, Abe O, Kiryu S. Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. Radiology. 2018;286(3):887-896.

23. Weston AD, Korfiatis P, Kline TL, et al. Automated Abdominal Segmentation of CT Scans for Body Composition Analysis Using Deep Learning. Radiology. 2019;290(3):669-679.

24. Choe J, Lee SM, Do KH, et al. Deep Learning-based Image Conversion of CT Reconstruction Kernels Improves Radiomics Reproducibility for Pulmonary Nodules or Masses. Radiology. 2019:181960.

25. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med. 2019;25(6):954-961.

26. Rubinstein E, Salhov M, Nidam-Leshem M, et al. Unsupervised tumor detection in Dynamic PET/CT imaging of the prostate. Med Image Anal. 2019;55:27-40.

27. Ouhmich F, Agnus V, Noblet V, Heitz F, Pessaux P. Liver tissue segmentation in multiphase CT scans using cascaded convolutional neural networks. Int J Comput Assist Radiol Surg. 2019.

28. Hamm CA, Wang CJ, Savic LJ, et al. Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. Eur Radiol. 2019;29(7):3338-3347.

29. Xu Y, Hosny A, Zeleznik R, et al. Deep Learning Predicts Lung Cancer Treatment Response from Serial Medical Imaging. Clinical Cancer Research. 2019;25(11):3266-3275.

30. Wang P, Xiao X, Glissen Brown JR, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng. 2018;2(10):741-748.

31. Nakagawa K, Ishihara R, Aoyama K, et al. Classification for invasion depth of esophageal squamous cell carcinoma using a deep neural network compared with experienced endoscopists. Gastrointest Endosc. 2019.

32. Horie Y, Yoshio T, Aoyama K, et al. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. Gastrointest Endosc. 2019;89(1):25-32.

33. Misawa M, Kudo S-e, Mori Y, et al. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience. Gastroenterology. 2018;154(8):2027-2029.e2023.

34. Xiao Y, Wu J, Lin Z, Zhao X. A deep learning-based multi-model ensemble method for cancer prediction. Comput Methods Programs Biomed. 2018;153:1-9.
35. Zhang Y, Hamada M. DeepM6ASeq: prediction and characterization of m6A-containing sequences using deep learning. BMC Bioinformatics. 2018;19(Suppl 19):524.

36. 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(2):723-750.

37. Yang L, Gu D, Wei J, et al. A Radiomics Nomogram for Preoperative Prediction of Microvascular Invasion in Hepatocellular Carcinoma. Liver Cancer. 2019;8(5):373-386.

38. Cong WM, Bu H, Chen J, et al. Practice guidelines for the pathological diagnosis of primary liver cancer: 2015 update. World J Gastroenterol. 2016;22(42):9279-9287.

39. Tabrizian P, Jibara G, Shrager B, Schwartz M, Roayaie S. Recurrence of hepatocellular cancer after resection: patterns, treatments, and prognosis. Ann Surg. 2015;261(5):947-955.

40. Feng LH, Dong H, Lau WY, et al. Novel microvascular invasion-based prognostic nomograms to predict survival outcomes in patients after R0 resection for hepatocellular carcinoma. J Cancer Res Clin Oncol. 2017;143(2):293-303.

41. Lee S, Kang TW, Song KD, et al. Effect of Microvascular Invasion Risk on Early Recurrence of Hepatocellular Carcinoma After Surgery and Radiofrequency Ablation. Ann Surg. 2021;273(3):564-571.

42. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.

43. Pote N, Cauchy F, Albuquerque M, et al. Performance of PIVKA-II for early hepatocellular carcinoma diagnosis and prediction of microvascular invasion. J Hepatol. 2015;62(4):848-854.

44. Zhu Y, Xu D, Zhang Z, et al. A new laboratory-based algorithm to predict microvascular invasion and survival in patients with hepatocellular carcinoma. Int J Surg. 2018;57:45-53.

45. Jiang YQ, Cao SE, Cao S, et al. Preoperative identification of microvascular invasion in hepatocellular carcinoma by XGBoost and deep learning. J Cancer Res Clin Oncol, 2021;147(3):821-833.

46. Song D, Wang Y, Wang W, et al. Using deep learning to predict microvascular invasion in hepatocellular carcinoma based on dynamic contrast-enhanced MRI combined with clinical parameters. J Cancer Res Clin Oncol, 2021.

47. Zhou W, Jian W, Cen X, et al. Prediction of Microvascular Invasion of Hepatocellular Carcinoma Based on Contrast-Enhanced MR and 3D Convolutional Neural Networks. Front Oncol. 2021;11:588010.

48. Zhang Y, Lv X, Qiu J, et al. Deep Learning With 3D Convolutional Neural Network for Noninvasive Prediction of Microvascular Invasion in Hepatocellular Carcinoma. J Magn Reson Imaging. 2021;54(1):134-143.

Tables

Table 1. The clinical and histologic characteristics of primary cohort.
| Variable                      | No. of patients (n=321) | Absent (n = 136) | Present (n = 185) | Estimate risk | p value |
|-------------------------------|-------------------------|------------------|------------------|---------------|---------|
| **Age**                      |                         |                  |                  |               |         |
| 0, ≤50 yr                    | 91                      | 33               | 58               | 1             |         |
| 1, >50 yr                    | 230                     | 103              | 127              | 0.702 (0.425-1.157) | 0.164  |
| **Sex**                      |                         |                  |                  |               |         |
| 0, Male                      | 273                     | 115              | 158              | 1             |         |
| 1, Female                    | 48                      | 21               | 27               | 0.936 (0.504-1.738) | 0.834  |
| **AFP**                      |                         |                  |                  |               |         |
| 0, ≤ 20 ng/mL                | 144                     | 84               | 60               | 1             |         |
| 1, >20 ng/mL                 | 177                     | 52               | 125              | 3.365 (2.118-5.347) | <0.001 |
| **Ascites**                  |                         |                  |                  |               |         |
| 0, absent                     | 301                     | 130              | 171              | 1             |         |
| 1, present                   | 20                      | 6                | 14               | 1.774 (0.664-4.741) | 0.248  |
| **Hepatic virus infection**  |                         |                  |                  |               |         |
| 0, absent                     | 77                      | 33               | 44               | 1             |         |
| 1, present (HBV/HCV)         | 244                     | 103              | 141              | 1.027 (0.612-1.723) | 0.921  |
| **Cirrhosis**                |                         |                  |                  |               |         |
| 0, absent                     | 84                      | 41               | 43               | 1             |         |
| 1, present                   | 237                     | 95               | 142              | 1.425 (0.864-2.351) | 0.164  |
| **ALT**                      |                         |                  |                  |               |         |
| 0, ≤40 U/L                   | 229                     | 108              | 121              | 1             |         |
| 1, >40 U/L                   | 92                      | 28               | 64               | 2.040 (1.220-3.412) | 0.006  |
| **AST**                      |                         |                  |                  |               |         |
| 0, ≤40 U/L                   | 226                     | 111              | 115              | 1             |         |
|                | Cases    | >40 U/L | 25 | 70 | 2.703 (1.597-4.573) | 0.001 |
|----------------|----------|---------|----|----|-------------------|-------|
| GGT            |          |         |    |    |                   |       |
| 0, ≤50 U/L     | 151      | 73      | 78 | 1  |                   | 1     |
| 1, >50 U/L     | 170      | 63      | 107| 1.590 (1.018-2.482)| 0.041 |
| TB             |          |         |    |    |                   |       |
| 0, ≤19 µmol/L  | 274      | 116     | 158| 1  |                   | 1     |
| 1, >19 µmol/L  | 47       | 20      | 27 | 0.991 (0.530-1.853)| 0.978 |
| CB             |          |         |    |    |                   |       |
| 0, ≤6.8 µmol/L | 244      | 107     | 137| 1  |                   | 1     |
| 1, >6.8 µmol/L | 77       | 29      | 48 | 1.293 (0.764-2.187)| 0.338 |
| ALB            |          |         |    |    |                   |       |
| 0, ≤40 g/L     | 91       | 40      | 51 | 1  |                   | 1     |
| 1, >40 g/L     | 230      | 96      | 134| 1.095 (0.671-1.787)| 0.717 |
| PLT            |          |         |    |    |                   |       |
| 0, ≤100*10⁹/L | 40       | 19      | 21 | 1  |                   | 1     |
| 1, >100*10⁹/L | 281      | 117     | 164| 1.268 (0.653-2.464)| 0.483 |
| INR            |          |         |    |    |                   |       |
| 0, ≤1.0        | 180      | 84      | 96 | 1  |                   | 1     |
| 1, >1.0        | 141      | 52      | 89 | 1.498 (0.955-2.349)| 0.078 |
| PT             |          |         |    |    |                   |       |
| 0, ≤12 s       | 241      | 105     | 136| 1  |                   | 1     |
| 1, >12 s       | 80       | 31      | 49 | 1.220 (0.728-2.046)| 0.45  |
| Surgical size (cm) | 321   | 4 (3-6) | 6.3 (4.5-9.25) | <0.001 |
| Tumor encapsulation |     |         |    |    |                   |       |
| 0, incomplete  | 194      | 88      | 106| 1  |                   | 1     |
Edmondson–Steiner grade

| Grade | Count | Complete | Accuracy |
|-------|-------|----------|----------|
| I-II  | 141   | 77       | 64       | 1        |
| III-IV| 180   | 59       | 121      | 2.467 (1.566-3.888) <0.001 |

AFP, a-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, γ-glutamyl transpeptidase; TB, total bilirubin; CB, conjugated bilirubin; ALB, albumin; PLT, platelet count; PT, prothrombin time; INR, international normalized ratio.

Table 2. MVI classification accuracy comparisons for different combinations of modalities.

| Combination | Accuracy |
|-------------|----------|
| T1 T1D T1V  | 92.11%   |
| T1 T1A T2   | 66.67%   |
| T1 T1D T2   | 84.21%   |
| DWI T1 T2   | 74.56%   |
| DWI T1A T2  | 77.19%   |
| T1V T1D     | 85.96%   |
| T1V T1D T1  | 87.72%   |
| T1V T1D T2  | 81.58%   |

T1: T1-weighted imaging; DWI: Diffusion-weighted imaging; T2: T2-weighted imaging; T1A: T1-weighted imaging at arterial phase; T1V: T1-weighted imaging at portal venous phase; T1D: T1-weighted imaging at arterial phase.

Table 3. Multivariate logistic regression analysis of factors associated with MVI.
### Variables

| Variables                                      | MVI                          | p value  |
|-----------------------------------------------|------------------------------|----------|
| ALT (>40 U/L versus ≤40 U/L)                  | 1.617 [0.763-3.427]          | 0.21     |
| AST (>40 U/L versus ≤40 U/L)                  | 1.939 [0.932-4.036]          | 0.077    |
| GGT (>50 U/L versus ≤50 U/L)                  | 0.737 [0.395-1.376]          | 0.338    |
| AFP (> 20 ng/mL versus ≤ 20 ng/mL)            | 4.634 [2.576-8.336]          | <0.001   |
| DL-MVI (present versus absent)                | 35.738 [14.027-91.056]       | <0.001   |

ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, γ-glutamyl transpeptidase; AFP, α-fetoprotein; OR: Odds ratio; CI: Confidence interval; DL-MVI: Deep learning predicted Microvascular invasion.

### Figures

#### Figure 1

Flowchart of DL model architecture

#### Figure 2

The distribution difference between the 2015 HCC cohort and the 2018 HCC cohort data from different modalities. (A) The distribution difference for M0 group with modalities T1, T1V and T1D, respectively. The difference for M1 group (B), and for M2 group (C) on T1, T1V and T1D, respectively.

#### Figure 3

Survival curves according to histological and DL-predicted MVI status. (A, C) OS and (B, D) RFS curves scaled by pathologic MVI status and DL-predicted MVI status with Kaplan-Meier analysis. MVI, microvascular invasion; DL, deep learning; OS, overall survival; RFS, recurrence-free survival.

#### Figure 4

Nomogram for predicting microvascular invasion (MVI) probabilities, receiver operating characteristics (ROC) curves, calibration of the nomogram and decision curve in the overall patients. (A) A nomogram
integrated DL-MVI and serum AFP level. (B) Receiver operating characteristic analysis of the nomogram. (C) Calibration curves of the nomogram in the overall datasets; X-axis is predicted probability of MVI. Y-axis is actual MVI. The diagonal dashed line indicates the ideal prediction by a perfect model. (D) Decision curve for the nomogram predicting the MVI in the overall patients.