Using Deep Learning to Predict Microvascular Invasion in Hepatocellular Carcinoma Based on Dynamic Contrast-enhanced MRI Combined with Clinical Parameters

Danjun Song

Zhongshan Hospital Fudan University

Yueyue Wang

Fudan University School of Basic Medical Sciences

Wentao Wang

Zhongshan Hospital Fudan University

Yining Wang

Zhongshan Hospital Fudan University

Jiabin Cai

Zhongshan Hospital Fudan University

Kai Zhu

Zhongshan Hospital Fudan University

Minzhi Lv

Zhongshan Hospital Fudan University

Qiang Gao

Zhongshan Hospital Fudan University

Jian Zhou

Zhongshan Hospital Fudan University

Jia Fan

Zhongshan Hospital Fudan University

Shengxiang Rao

Zhongshan Hospital Fudan University

Manning Wang

Fudan University School of Basic Medical Sciences

Xiaoying Wang (✉ xiaoyingwang@fudan.edu.cn)

Liver Cancer Institute, Zhongshan Hospital, Fudan University   https://orcid.org/0000-0001-5989-1539

Research Article

Abstract

Purpose Microvascular invasion (MVI) is a critical determinant of the early recurrence and poor prognosis of patients with hepatocellular carcinoma (HCC). Prediction of MVI status is clinically significant for the decision of treatment strategies and the assessment of patient's prognosis. A deep learning (DL) model was developed to predict the MVI status and grade in HCC patients based on preoperative dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) and clinical parameters.

Methods HCC patients with pathologically confirmed MVI status from January to December 2016 were enrolled and preoperative DCE-MRI of these patients were collected in this study. Then they were randomly divided into the training and testing cohorts. A DL model with eight conventional neural network (CNN) branches for eight MRI sequences was built to predict the presence of MVI, and further combined with clinical parameters for better prediction.

Results Among 601 HCC patients, 376 patients were pathologically MVI-absent, and 225 patients were MVI-present. To predict the presence of MVI, the DL model based only on images achieved an area under curve (AUC) of 0.915 in the testing cohort as compared to the radiomics model with an AUC of 0.731. The DL combined with clinical parameters (DLC) model yielded the best predictive performance with an AUC of 0.931. For the MVI grade stratification, the DLC models achieved an overall accuracy of 0.793. Survival analysis demonstrated that the patients with DLC-predicted MVI status were associated with the poor overall survival (OS) and recurrence-free survival (RFS). Further investigation showed that hepatectomy with the wide resection margin contributes to better OS and RFS in the DLC-predicted MVI-present patients.

Conclusion The proposed DLC model can provide a non-invasive approach to evaluate MVI before surgery, which can help surgeons make decisions of surgical strategies and assess patient's prognosis.

Introduction

Hepatocellular carcinoma (HCC) is ranked as the sixth in terms of incidence cases and the fourth with respect to the cause of cancer-related death worldwide (Villanueva 2019). Surgical treatment is regarded as the major curative treatment for patients with HCC (Ishizawa et al. 2008; Zhou et al. 2018). However, postoperative recurrence and metastasis remain major obstacles for the prognosis of HCC patients (Han et al. 2015; Hasegawa et al. 2013).

Microvascular invasion (MVI) is one of the most important biological characteristics of HCC and its prevalence in patients with HCC ranges from 15–57.1% (Rodríguez-Perálvarez et al. 2013). The presence of MVI is considered as an independent prognostic factor associated with the early recurrence and poor survival after both resection and transplantation (Lim et al. 2011; Mazzaferro et al. 2009a; Mazzaferro et al. 2009b; Vitale et al. 2014). For the MVI-positive patients, hepatectomy with expanding resection margins can significantly improve patient survival through eradicating micro-metastases (Han et al. 2019; Shindoh et al. 2013; Tsilimigras et al. 2020). Surgical resection rather than radiofrequency ablation (RFA)

may be suitable for the patient with small-size HCC displayed MVI (Imai et al. 2018; Lee et al. 2021) because it has been reported that among the patients with high risk of MVI, the cumulative 2-year recurrence rate was 90.0% in the surgical resection group, while 30.6% in the RFA group (Lee et al. 2021). Furthermore, additional adjuvant or targeted therapies were reported to improve survival for HCC patients with MVI compared to surgical resection alone (Sun et al. 2016; Wang et al. 2020b; Zhang et al. 2019). So preoperative prediction of MVI status is clinically significant for surgical decision making and other treatment strategies to improve patient prognosis.

However, assessment of MVI preoperatively is still challenging because it can only be confirmed by pathological diagnosis. Currently, a number of studies have attempted to predict MVI in HCC utilizing some clinical parameters, such as tumor number and size, alpha fetoprotein (AFP), protein induced by vitamin K absence or antagonist (PIVKA-II) and some serum component index (Nitta et al. 2019; Poté et al. 2015; Wang et al. 2020a; Zeng et al. 2019; Zhang et al. 2020a; Zheng et al. 2017; Zhu et al. 2018). The AUC of these predictors vary 0.529 from 0.803 and still required to be improved (Poté et al. 2015; Zhu et al. 2018). It is worth noting that medical imaging plays an irreplaceable role in preoperative diagnosis, assessment and clinical decision-making of HCC because it can provide much valuable information (Lee et al. 2017). Recent studies using radiomics combined with clinical parameters demonstrate impressive results in predicting MVI status, ranging from 0.796−0.899 of AUC (Ma et al. 2019; Xu et al. 2019; Zhang et al. 2020b). Nonetheless, extraction of "hand-engineered" features is still complicated, time-consuming, and lack of stable reproducibility (Liu et al. 2019).

Deep learning (DL) uses convolutional neural networks with a number of hidden layers to learn features associated with predefined task. It demonstrates superior performance than radiomics and was proved to be powerful in various challenging clinical tasks (Ardila et al. 2019; Coudray et al. 2018; Ting et al. 2017; Wang et al. 2018; Wang et al. 2019). A recent study using DL method based on CT images to predict MVI, achieved a good performance with an area under curve (AUC) of 0.906 (Jiang et al. 2021). Compared with traditional radiomics method, DL method needs less manpower and time for MVI prediction and it can learn features from the raw images without precise annotations (Wang et al. 2019). Magnetic resonance imaging (MRI) captures both the anatomical and functional information of the tumor from a series of imaging sequences, so it offers more valuable information than computerized tomography (CT) (Roberts et al. 2018). Therefore, we decided to utilize DL method based on MRI to predict MVI status. To the best of our acknowledge, few studies use the DL approach to predict MVI status based on dynamic contrast-enhanced MRI (DCE-MRI).

The aim of the current study is to improve the prediction performance of MVI. The proposed model using DL method integrating clinical parameters with DCE-MRI can predict MVI and assess patient oncological outcomes preoperatively. In addition, the potential significance of prediction MVI status in guiding the surgical resection range was also investigated based on our model.

Materials And Methods

Study population

This retrospective study was approved by the institutional review board of Zhongshan Hospital of Fudan University (Shanghai, China). The written informed consent was obtained from each patient. Considering follow-up time, 1829 consecutive patients, who underwent surgical resection for HCC from January to December 2016 at Zhongshan Hospital, Fudan University, were included. The pathology of HCC was confirmed by two independent pathologists. Patients who met the following inclusion criteria were enrolled in this study: (1) Pathologically confirmed HCC patients after surgical resection; (2) Available for the pathological assessment of MVI; (3) Available for effective DCE-MRI and clinical data within 1 month preoperatively. Patients who met the following exclusion criteria were excluded (1) With prior antitumoral therapies and hepatectomy; (2) With multiple liver tumors; (3) With macrovascular invasion and extrahepatic metastasis. The flow-chart of this study is shown in Supplementary Figure 1.

The final population consisted of 601 patients in this study. All patients were randomly divided into the training cohort (n=461, 76.7%) and the testing cohort (n=140, 23.3%).

Pathological evaluation and clinical parameters

The definition of MVI is the presence of tumor in the portal, hepatic venous system, or the branches surrounding the hepatic tissue lined by endothelium, which is visible only by microscopy (Cong et al. 2016; Roayaie et al. 2009). The MVI status is defined as MVI-absent and MVI-present. MVI grade is classified as MVI-grade 1: the number of MVI <5 and the distance of MVI ≤1 cm away from the tumor tissues, and MVI-grade 2: 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 (Cong et al. 2016). Both MVI status and grade were examined by two independent pathologists in the Zhongshan hospital. Clinical parameters were collected, including sex, age, routine blood test, blood biochemical test, blood coagulation function test, markers of hepatic fibrosis, hepatitis virus B carriers, AFP, and tumor size. Serum component index, such as platelet-lymphocyte ratio (PLR), neutrophil-lymphocyte ratio (NLR), lymphocyte-to-monocyte ratio (LMR), prognostic nutritional index (PNI), aspartate aminotransferase-to-platelet ratio index (APRI), aspartate aminotransferase-to-neutrophil ratio index (ANRI) and aspartate aminotransferase-lymphocyte ratio (ALR), were calculated as previous reported (Zheng et al. 2017). Details of clinical parameters can be found in the Supplementary Table 1.

Image annotation

All HCC patients underwent preoperative MRI with gadopentetate dimeglumine (Magnevist; Bayer Schering Pharma AG, Berlin, Germany) on a 1.5-T (Avanto, Siemens, Erlangen, Germany and Aera, Siemens, Erlangen, Germany) or 3.0 T scanners (Magnetom Verio, Siemens Medical Solution, Erlangen, Germany and Signa HDx, GE Healthcare, Milwaukee, WI, USA). Eight routine abdominal DCE-MRI sequences included turbo spin-echo T2-weighted (T2WI) with fat suppression, diffusion-weighted images (DWI) of b=0 and 500 s/mm^2, automatically generated apparent diffusion coefficient (ADC) maps under free-breathing, three-dimensional (3D) T1-weighted volumetric interpolated breath-hold examination of

pre-contrast phase, arterial phase (20–30 s), portal venous phase (about 80 s), and delayed phase (3 min). Details of MRI parameters were listed in the Supplementary Table 2. Tumor regions were drawn manually slice-by-slice by an abdominal radiologist and a hepatic surgeon. The annotation only requires the manual selection of the tumor region in each slice without precise tumor boundary. Before feeding to the models, these regions were confirmed and corrected by a senior radiologist and a senior hepatic surgeon with more than 10 years of experience in reading abdominal MRI.

Developing a DL model

A DL model was constructed to predict the MVI status and grade of MVI. The architecture of the DL model is shown in Figure 1. The DL model has eight inputs, which are 3D volumes of interest (VOIs) of the eight MRI sequences. This model has eight separate conventional neural network (CNN) branches, which are used for feature extraction from each of the eight VOIs. Then, the features extracted from the eight branches were fused, and the fused features were fed into fully connected (FC) layers and a SoftMax layer to obtain the predicted results. Notably, instead of using Resnet or other deep neural networks as the branches, we designed specific architecture as CNN branches for feature extraction in this task, as described in Supplementary Materials.

The input VOI of the DL model was cropped from raw MRI sequences according to the tumor annotation mask and resized to 64×64×16 pixels by third-order spline interpolation [0, 1]. The label of input VOIs was MVI grade encoded to one-hot. The DL model has three outputs, corresponding to the predicted probabilities of MVI-absent, MVI-grade 1, and MVI-grade 2, respectively. In the MVI grade prediction task, the category corresponded to the maximum value among the three outputs was recorded as the predicted MVI grade. Based on this result, MVI-grade 1 and 2 categories were both recorded as MVI-present in the MVI status prediction task.

In the training stage, the DL model was trained using a SoftMax cross-entropy loss to learn the MVI grade of the inputs. In order to account for class imbalance while calculating cross-entropy loss, each class was weighed according to its frequency, such that cases of rare conditions would contribute highly to the loss of function (Liu et al. 2020). Data augmentation was applied to improve generalization, which included horizontal flipping, vertical flipping, cropping, and zoom transformation for the training cohort. Next, we set learning rate at 10^{-4} and applied the Adam optimizer to update the model parameters on a batch size of 2. Early stopping was used with the patience parameter set at 50. Finally, the DL model was subjected to a maximum of 1000 epochs of training.

Developing a DL combined with clinical parameter (DLC) model

To improve the predictive performance, we integrated clinical information with DL model. The backward stepwise method was applied to select the significant clinical parameters. The minimum Akaike information criterion (AIC) index was used as the stop criterion to determine the optimal characteristics. Then, the selected parameters were incorporated into the DL model to form the DLC model (Figure 1).

Notably, for the selected parameters, categorical variables were encoded by one digit (i.e., -1 or 1 for each state), and continuous variables were normalized to [-0.5, 0.5].

During the DLC model training, eight VOIs were entered into the model as input, and the selected clinical parameters were input into the FC layer instead of extracting deep features and combining them offline. The training procedure and hyperparameter setting of the DLC model were similar to those of the DL model.

Developing a radiomics model

To compare the predictive performance of the DL and DLC models, we also constructed a radiomics model (Aerts 2016), which is described in the Supplementary Materials.

Correlation analysis between resection margins and oncological outcomes in the DLC-predicted MVI absent/present population

Further investigation between resection margins and oncological outcomes was performed in the DLC-predicted MVI-absent/-present population. The definition of resection margin is the minimum distance between the tumor and the cutting edge in the formalin fixed tissues. All resection margins were evaluated by two independent pathologists in our center. A maximal Youden index in receiver operating characteristic (ROC) curve analysis was calculated to determine the cutoff point of the resection margins. In this study, the margins at or less than cutoff points were defined as narrow resection margins, while the others were defined as wide resection margins.

Statistical analysis

Continuous data were expressed as using the mean ± standard deviation (SD) or median (interquartile range (IQR)) as appropriate. Continuous variables were analyzed using Student's t-test or Mann–Whitney U test appropriately. Categorical variables were analyzed using the χ^2 test. The recurrence-free survival (RFS) was recorded from the date of surgery to the date of radiographic detection of recurrence. Overall survival (OS) was defined as the interval between surgery and death or censor at the date of the last follow-up. A ROC curve was used to assess diagnostic performance. 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 proposed DL and DLC models were implemented using Python (version 3.5, https://www.python.org) based on the *Pytorch* package (https://pytorch.org/). A two-sided P-value <0.05 was considered statistically significant.

Results

Demographic comparison of baseline clinical features

A total of 601 patients who met the inclusion criteria were enrolled in this study. The flowchart of the study selection process is shown in Supplementary Fig. 1. Among the 601 patients, 225/601 (37.4%)

were pathologically identified as MVI-present, and 371/601 (62.6%) were determined as MVI-absent. Among the MVI-present patients, 140/225 (62.2%) was MVI-grade 1 and 85/225 (37.8%) was MVI-grade 2. After random distribution, patients were divided into the training cohort (n = 461) and the testing cohort (n = 140). The number of patients in three categories was 287/461 (62.3%) (MVI-absent), 106/461 (23.0%) (MVI-grade 1) and 68/461 (14.7%) (MVI-grade 2) in the training cohort, and 89/140 (63.6%) (MVI-absent), 34/140 (24.3%) (MVI-grade 1) and 17/140 (12.1%) (MVI-grade 2) in the testing cohort. No significant differences in baseline characteristics were observed between the training and the testing cohorts (Table 1).

Table 1
Patients and baseline characteristics

Characteristic	All patients	Training cohort	Testing cohort	P-values
Patient number	601	461	140	NA
Age, mean (± SD), year	56.45 (± 11.48)	56.39 (± 11.35)	56.66 (± 11.92)	0.803
Sex (Male/Female)	492/109	385/76	107/33	0.075
HBV carriers (%)	472 (78.5%)	358 (77.7%)	114 (81.4%)	0.404
Liver function				
TB, mean (± SD), μmol/L	12.64 (± 6.63)	12.77 (± 7.04)	12.23 (± 5.05)	0.405
ALB, mean (± SD), g/L	41.42 (± 3.49)	41.46 (± 3.52)	41.30 (± 3.42)	0.646
ALT, mean (± SD), U/L	34.74 (± 29.95)	34.97 (± 31.21)	33.98 (± 25.43)	0.733
GGT, mean (± SD), U/L	80.50 (± 124.68)	89.54 (± 136.62)	72.19 (± 71.47)	0.149
PT, mean (± SD), s	11.62 (± 0.80)	11.63 (± 0.84)	11.61 (± 0.67)	0.804
AFP, median (IQR), ng/mL	25.20 [4.40, 381.30]	27.40 [4.50, 504.20]	15.05 [4.10, 184.95]	0.233
Log_{10}AFP, mean (± SD)	1.70 (± 1.23)	1.75 (± 1.27)	1.57 (± 1.12)	0.145
Tumor size, median (IQR), cm	4.99 (± 3.59)	5.04 (± 3.77)	4.80 (± 2.92)	0.475
MVI status				0.856
MVI-absent (%)	376 (62.6%)	287 (62.3%)	89 (63.6%)	
MVI-present	225 (37.4%)	174 (37.7%)	51 (36.4%)	
MVI-grade 1/2	140/85	106/68	34/17	
HBV, hepatitis B virus; TB, total bilirubin; ALB, albumin; ALT, alanine aminotransferase; GGT, glutamyl transpeptidase; PT, prothrombin time; AFP, alpha fetoprotein; MVI, microvascular invasion.				

MVI status prediction

In the DL model, the AUC for prediction MVI achieved 0.909 (95% confidence interval (CI) 0.881−0.936) in the training cohort and 0.915 (95% CI 0.868−0.963) in the testing cohort (Fig. 2A, 2B and Table 2). To improve the predictive performance, five significant clinical parameters (the percentage of neutrophil, lymphocyte counts, international normalized ratio (INR), the logarithmic conversion ratio of the AFP (log_{10}AFP) and tumor size) were selected by backward stepwise analysis to establish the DLC model (Supplementary Table 3). The DLC models achieved the highest AUC of 0.934 (95% CI 0.912−0.956) in the training cohort and 0.931 (95% CI 0.884−0.979) in the testing cohort (Fig. 2A, 2B and Table 2). In comparison, the traditional radiomics model achieved the AUC of 0.764 (95% CI 0.719−0.809) in the training cohort and 0.731 (0.645−0.817) in the testing cohort (Fig. 2A, 2B and Table 2). The ROC curves and scatter plots of the prediction score are displayed in Fig. 2C, 2D and Supplementary Fig. 3. Furthermore, the prediction performance of our DLC model is better as compared to those of the previous published models in distinguishing MVI status (Supplementary Table 4).

Table 2
Performance comparison of different models in the training and the testing cohort.

Modality		AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Radiomics model	TR	0.764 (0.719−0.809)	0.727 (0.686−0.767)	0.667 (0.591−0.735)	0.763 (0.709−0.810)
	TE	0.731 (0.645−0.817)	0.671 (0.594−0.749)	0.647 (0.500-0.772)	0.685 (0.577−0.777)
DL model	TR	0.909 (0.881−0.936)	0.842 (0.808−0.875)	0.845 (0.780−0.894)	0.840 (0.791−0.879)
	TE	0.915 (0.868−0.963)	0.871 (0.816−0.927)	0.843 (0.709−0.925)	0.888 (0.799−0.942)
DLC model	TR	0.934 (0.912−0.956)	0.871 (0.805−0.937)	0.839 (0.774−0.889)	0.861 (0.814−0.897)
	TE	0.931 (0.884−0.979)	0.886 (0.824−0.948)	0.882 (0.754−0.951)	0.888 (0.799−0.942)

AUC, area under curve; CI, confidence interval; DL, deep learning; DLC, deep learning combined with clinical parameters; TR, training cohort; TE, testing cohort.

MVI grade prediction

In the training cohort, the DLC model exhibited a similar overall accuracy than the DL model (0.779 *vs.* 0.768) in MVI grade prediction. In the testing cohort, DLC model also had a comparable overall accuracy

as compared to the DL model (0.793 *vs.* 0.800). Figure 3 shows the confusion matrix of DL and DLC models between the training and testing cohorts. These results showed that DL and DLC models could distinguish the MVI-grade 1 group from MVI-grade 2 group.

Prognosis assessment based on DLC-predicted MVI status

The DLC model achieved the best result in MVI prediction, so the patient prognosis was analyzed based on the DLC predicted-MVI status and grade. The median follow-up time for all patients was 38.3 months: 36.9 months for those with MVI and 39.0 months for those without MVI. The 1-, 2-, and 3-year OS rates were 97.5%, 94.4%, 91.6% in those with DLC-predicted MVI-absent, 91.7%, 82.5%, 74.9% in those with DLC-predicted MVI-present, 93.9%, 86.5%, 77.6% in those with DLC-predicted MVI-grade 1, and 88.1%, 76.1%, 70.5% in those with DLC-predicted MVI-grade 2 (Fig. 4A&C). The 1-, 2-, and 3-year RFS rates were 88.3%, 79.0%, 73.5% in those with DLC-predicted MVI-absent, 68.0%, 56.8%, 52.1% in those with DLC-predicted MVI-present, 73.6%, 63.4%, 57.1% in those with DLC-predicted MVI-grade 1, and 59.1%, 46.2%, 44.1% in those with DLC-predicted MVI-grade 2 (Fig. 4B&D). The OS and RFS stratified by DLC-predicted MVI status and grade were similar to the trend of survival analysis in the patients with pathological MVI status and grade stratification (Fig. 4E-H). The results suggested the potential value of our model in assessment poor OS and early recurrence of patients based on the DLC-predicted MVI status and grade.

Correlation analysis between prognosis and wide/narrow resection margins in DLC-predicted MVI absent/present population

A 5 mm cutoff value of resection margin was selected through calculating the maximal Youden index. Therefore, margins of ≤ 5 mm were defined as narrow resection margins and > 5 mm were defined as wide resection margins.

Among the population with DLC-predicted MVI-present, the 1-, 2-, and 3-year OS rates were higher in the wide resection margin group than that in the narrow resection margin group (93.8%, 85.7%, and 80.2% *vs.* 89.8%, 79.7%, and 70.2%, respectively) (Fig. 5A) (P = 0.050). Additionally, the 1-, 2-, and 3-year RFS rates were also higher in the wide resection margin group than those in the narrow resection margin group (72.6%, 67.2%, and 63.6% *vs.* 64.1%, 47.7%, and 42.1%, respectively) (Fig. 5B) (P = 0.003). Among the patients with DLC-predicted MVI-absent, no statistical differences were observed in the OS and RFS rates between the wide resection margin and narrow resection margin groups (P = 0.094 and P = 0.350 respectively) (Fig. 5C&D). This is consistent with the trend of survival analysis in patients with pathological MVI absent/present after hepatectomy with different surgical margins (Fig. 5E-H). These findings demonstrated that hepatectomy with wide resection margin was associated with better OS and RFS in the DLC-predicted MVI present patients.

Discussion

MVI prediction is significant for clinicians to assess patient prognosis and adopt appropriate therapeutic strategies, contributing to improve patient survival (Han et al. 2019; Lee et al. 2021; Mazzaferro et al.

2009a; Rodríguez-Perálvarez et al. 2013; Sun et al. 2016; Wang et al. 2020b; Zhang et al. 2019). Currently, a number of studies have tried to predict MVI using the clinical parameters, which achieve the AUC from 0.529 to 0.803. Integrating clinical parameters with radiomics methods based on CT or MRI achieve the AUC from 0.796 to 0.906 (Poté et al. 2015; Xu et al. 2019; Zhang et al. 2020b; Zhu et al. 2018). However, predictive performance of clinical parameters still required to be improved. The shortcoming of radiomics method is based on manually-defined precise contours of the tumor and the analysis is conducted based on hand-crafted feature extractor, which do not represent the most optimal option. In contrast, the DL method can analyze the patterns and regions with less manpower and time (Park 2019; Wang et al. 2019). It can leverage a box containing the tumor and peripheral tissues as VOI without precise boundary, which is consistent with prior studies in exploring radiomics of the tumor and peripheral regions (Xu et al. 2019). A recent study using DL showed a considerable efficacy in identifying MVI based on preoperative CT (AUC 0.906, 95% CI 0.821−0.960) (Jiang et al. 2021). Therefore, the DL method was utilized to assess MVI preoperatively in this study.

Different from the previous DL study (Jiang et al. 2021), we chose DCE-MRI instead of CT images to predict MVI in this study. According to the clinical practice guidelines of National Comprehensive Cancer Network (NCCN, version 2019) (Network) and the American Association for the Study of Liver Diseases (AASLD, version 2018) (Marrero et al. 2018), DCE-MRI shows higher sensitivity than dynamic CT for the diagnosis of HCC. MRI also could capture both the anatomical and functional information of a tumor and provides more information than CT theoretically. Therefore, the DL method based on DCE-MRI may learn some different and meaningful features contributing to the improvement of predictive performance. Algorithmically, the DL model applied in this study is a multi-input network. As we used eight image sequences for MVI prediction, the DL both have eight CNN branches for each image input. In each CNN branch, to reduces the number of parameters, we proposed to use depthwise separable convolution instead of the ordinary convolution operation. Besides, to extract image features in different scales, we used the two kinds of convolution kernel size of 3×3×3 and 5×5×5. In this way, the model can extract richer features from the input image with fewer parameters.

To improve predictive performance, significant clinical characteristics (neutrophil percentage, lymphocyte counts, INR, $\log_{10}$AFP and tumor size) were also selected by backward stepwise with the lowest AIC value and then integrated with the DL model to develop the DLC model. As a result, the DLC model showed a robust performance in predicting the MVI status compared with the DL model (AUC: 0.931 *vs.* 0.915) in the testing cohort, which suggests the potential value of clinical parameters for MVI prediction. Attributed to these improvements mentioned above, the DLC model achieved a stronger predictive performance compared with previous predictive models based on CT or MRI (AUC: 0.931 *vs.* 0.796−0.906) (Supplementary Table 4). These findings validated the robustness of our DLC model and emphasized its suitability for clinical application. To the best of our knowledge, this is the first study using DL method combined with clinical parameters to predict MVI based on multi-parametric DCE-MRI in HCC patients.

In the current study, patients with solitary tumors were included. The main reason is that patients with multiple tumors are frequently associated with higher incidence of MVI (Reginelli et al. 2018; Xu et al.

2019), while those with solitary tumor displayed MVI are not easy to be identified. For patients with multiple tumors, previous studies only assessed the largest tumor to predict MVI status (Xu et al. 2019; Zhang et al. 2020b) but it is not accurate to identify which lesion exists MVI. The other point is that multi-carcinogenesis may not be controlled by extended liver resection due to the limited liver volume. Therefore, prediction of MVI status among these patients is more valuable to improve their survival.

In addition, our model identified different patient prognosis through stratification of DLC-predicted MVI status and grade, suggesting the strong performance of DLC model in preoperatively identifying HCC patients with high risk of early recurrence and poor prognosis.

The prediction of MVI also could provide potential values in guiding the resection range. The retrospective study showed hepatectomy with wide resection margin provides better prognosis compared to narrow resection margin (median OS: 93.5 *vs.* 69.2 months; median RFS: 53.1 *vs.* 37.5 months, respectively) (Han et al. 2019). Further investigation of our study showed that wide resection margins were closely related to improved oncological outcomes among the patients with DLC-predicted MVI-present. These findings suggested that wide resection margins should be the target of resection at the time of initial surgical strategy decisions for the patients with DLC-predicted MVI-present. So far, this study is first to demonstrate patient prognosis after hepatectomy with wide/narrow resection margins based on DLC-predicted MVI status.

Nevertheless, the present study also has some limitations. First, there is the possibility of potential bias because of the nature of the retrospective study. The DCE-MRI sequences were acquired from different scanners with different magnetic field strength. Standardized processing such as gray-level normalization and voxel-size resampling was performed, which reduced the dependency of differences in scanners and image parameters (Kim et al. 2019). However, some potential differences, such as scanning technique, still existed. Second, testing cohort in this study was from the same centers, and hence, an additional cohort from other centers using different scanners is conductive to confirm the reliability and generalizability of our DLC model. Considering the low percentage of patients with anatomical hepatectomy in this study, so we did not compare the difference of the survival in patients with anatomical hepatectomy and non-anatomical hepatectomy. The prognostic values of anatomical resection in the DLC-predicted MVI-present patients should be investigated in the future. Finally, the DL method is regarded as "black box" (Price 2018), and it is still challenging as well as necessary to explain the correlation between relevant features and results (Price 2018).

In conclusion, the DLC model displays robust ability in the prediction of MVI status and grade, and also can assess the oncological outcomes in HCC patients. Due to a high risk of tumor recurrence in patient with MVI, our findings suggest that hepatectomy with wide resection margin should be a target for the patients with DLC-predicted MVI-present. Therefore, our model not only makes it possible to assess MVI status before surgery, but also can help surgeons to make decisions regarding surgical strategies.

Abbreviation

MVI: microvascular invasion; HCC: hepatocellular carcinoma; DL: deep learning; DCE-MRI: dynamic contrast-enhanced magnetic resonance imaging; CNN: conventional neural network; AUC: area under curve; DLC: deep learning combined with clinical parameters; OS: overall survival; RFS: recurrence-free survival; RFA: radiofrequency ablation; AFP: alpha fetoprotein; PIVKA-II: prothrombin induced by vitamin K absence-II; MRI: magnetic resonance imaging; CT: computed tomography; PLR: platelet-lymphocyte ratio; NLR: neutrophil-lymphocyte ratio; LMR: lymphocyte-to-monocyte ratio; PNI: prognostic nutritional index; APRI: aspartate aminotransferase-to-platelet ratio index; ANRI: aspartate aminotransferase-to-neutrophil ratio index; ALR: aspartate aminotransferase-lymphocyte ratio; T2WI: turbo spin-echo T2-weighted; DWI: diffusion-weighted imaging; ADC: apparent diffusion coefficient; 3D: three-dimensional; VOI: volume of interest; FC; fully connected; AIC: akaike information criterion; ROC: receiver operating characteristic; INR: international normalized ratio; SD: standard deviation; IQR: interquartile range; CI: confidence interval; NCCN: National Comprehensive Cancer Network; AASLD: American Association for the Study of Liver Diseases.

Declarations

Funding

Availability of data and material

Data are available in the Article and Supplementary Information. All other data can be provided upon reasonable request to the corresponding authors.

Code availability

The code of the proposed method can be provided upon reasonable request to the corresponding authors.

Authors' contributions

Guarantors of integrity of entire study, W.X.Y., W.M.N., R.S.X.; Concepts and design, W.X.Y.; Administrative support, all authors; Data acquisition (including radiological and clinical data), all authors; Radiological images delineation, S.D.J., W.W.T., R.S.X., W.X.Y.; Development of methodology, W.Y.Y., W.M.N.; Statistical analysis, S.D.J., W.Y.Y., L.M.Z.; Manuscript drafting or revision, S.D.J., W.Y.Y..

Acknowledgement

Compliance with ethical standards

Conflict of interest

The authors of this manuscript declare no conflict of interest.

Ethical approval

This retrospective study was approved by IRB and the requirement for written informed consent is waived by IRB.

References

1. Aerts HJWL (2016) The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review JAMA Oncol 2:1636-1642 doi:10.1001/jamaoncol.2016.2631

2. Ardila D et al. (2019) End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography Nat Med 25:954-961 doi:10.1038/s41591-019-0447-x

3. Cong W-M, Bu H, Chen J, Dong H, Zhu Y-Y, Feng L-H, Chen J (2016) Practice guidelines for the pathological diagnosis of primary liver cancer: 2015 update World J Gastroentero 22:9279-9287

4. Coudray N et al. (2018) Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning Nat Med 24:1559-1567 doi:10.1038/s41591-018-0177-5

5. Han H-S, Shehta A, Ahn S, Yoon Y-S, Cho JY, Choi Y (2015) Laparoscopic versus open liver resection for hepatocellular carcinoma: Case-matched study with propensity score matching Journal of hepatology 63:643-650 doi:10.1016/j.jhep.2015.04.005

6. Han J et al. (2019) The impact of resection margin and microvascular invasion on long-term prognosis after curative resection of hepatocellular carcinoma: a multi-institutional study HPB (Oxford) 21:962-971 doi:10.1016/j.hpb.2018.11.005

7. Hasegawa K et al. (2013) Comparison of resection and ablation for hepatocellular carcinoma: a cohort study based on a Japanese nationwide survey J Hepatol 58:724-729 doi:10.1016/j.jhep.2012.11.009

8. Imai K et al. (2018) Microvascular Invasion in Small-sized Hepatocellular Carcinoma: Significance for Outcomes Following Hepatectomy and Radiofrequency Ablation Anticancer Res 38:1053-1060 doi:10.21873/anticanres.12322

9. Ishizawa T et al. (2008) Neither multiple tumors nor portal hypertension are surgical contraindications for hepatocellular carcinoma Gastroenterology 134:1908-1916 doi:10.1053/j.gastro.2008.02.091

10. Jiang YQ et al. (2021) Preoperative identification of microvascular invasion in hepatocellular carcinoma by XGBoost and deep learning J Cancer Res Clin Oncol 147:821-833 doi:10.1007/s00432-020-03366-9

11. Kim S, Shin J, Kim DY, Choi GH, Kim MJ, Choi JY (2019) Radiomics on Gadoxetic Acid-Enhanced Magnetic Resonance Imaging for Prediction of Postoperative Early and Late Recurrence of Single Hepatocellular Carcinoma Clin Cancer Res 25:3847-3855 doi:10.1158/1078-0432.Ccr-18-2861

12. Lee S et al. (2021) Effect of Microvascular Invasion Risk on Early Recurrence of Hepatocellular Carcinoma After Surgery and Radiofrequency Ablation Ann Surg 273:564-571 doi:10.1097/sla.0000000000003268

13. Lee S, Kim SH, Lee JE, Sinn DH, Park CK (2017) Preoperative gadoxetic acid-enhanced MRI for predicting microvascular invasion in patients with single hepatocellular carcinoma J Hepatol 67:526-534 doi:10.1016/j.jhep.2017.04.024

14. Lim K-C et al. (2011) 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 254:108-113 doi:10.1097/SLA.0b013e31821ad884

15. Liu Y et al. (2020) A deep learning system for differential diagnosis of skin diseases Nat Med 26:900-908 doi:10.1038/s41591-020-0842-3

16. Liu Z et al. (2019) The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges Theranostics 9:1303-1322 doi:10.7150/thno.30309

17. Ma X et al. (2019) Preoperative radiomics nomogram for microvascular invasion prediction in hepatocellular carcinoma using contrast-enhanced CT Eur Radiol 29:3595-3605 doi:10.1007/s00330-018-5985-y

18. Marrero JA et al. (2018) Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases Hepatology 68:723-750 doi:10.1002/hep.29913

19. Mazzaferro V et al. (2009a) Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis Lancet Oncol 10:35-43 doi:10.1016/S1470-2045(08)70284-5

20. Mazzaferro V et al. (2009b) Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis Lancet Oncol 10:35-43 doi:10.1016/S1470-2045(08)70284-5

21. Network NCC NCCN guidelines version 2.2019,Hepatobiliary cancers. https://www.nccn.org/professionals/physician_gls/pdf/hepatobiliary.pdf. Accessed April 2019

22. Nitta H et al. (2019) Prognostic Value and Prediction of Extratumoral Microvascular Invasion for Hepatocellular Carcinoma Ann Surg Oncol 26:2568-2576 doi:10.1245/s10434-019-07365-0

23. Park CM (2019) Can Artificial Intelligence Fix the Reproducibility Problem of Radiomics? Radiology 292:374-375 doi:10.1148/radiol.2019191154

24. Poté N et al. (2015) Performance of PIVKA-II for early hepatocellular carcinoma diagnosis and prediction of microvascular invasion J Hepatol 62:848-854 doi:10.1016/j.jhep.2014.11.005

25. Price WN (2018) Big data and black-box medical algorithms Sci Transl Med 10 doi:10.1126/scitranslmed.aao5333

26. Reginelli A et al. (2018) Can microvascular invasion in hepatocellular carcinoma be predicted by diagnostic imaging? A critical review Future Oncol 14:2985-2994 doi:10.2217/fon-2018-0175

27. Roayaie S et al. (2009) A system of classifying microvascular invasion to predict outcome after resection in patients with hepatocellular carcinoma Gastroenterology 137:850-855 doi:10.1053/j.gastro.2009.06.003

28. Roberts LR et al. (2018) Imaging for the diagnosis of hepatocellular carcinoma: A systematic review and meta-analysis Hepatology 67:401-421 doi:10.1002/hep.29487

29. Rodríguez-Perálvarez M, Luong TV, Andreana L, Meyer T, Dhillon AP, Burroughs AK (2013) A systematic review of microvascular invasion in hepatocellular carcinoma: diagnostic and prognostic variability Ann Surg Oncol 20:325-339 doi:10.1245/s10434-012-2513-1

30. Shindoh J et al. (2013) Risk factors of post-operative recurrence and adequate surgical approach to improve long-term outcomes of hepatocellular carcinoma HPB (Oxford) 15:31-39 doi:10.1111/j.1477-2574.2012.00552.x

31. Sun JJ et al. (2016) Postoperative Adjuvant Transcatheter Arterial Chemoembolization After R0 Hepatectomy Improves Outcomes of Patients Who have Hepatocellular Carcinoma with Microvascular Invasion Ann Surg Oncol 23:1344-1351 doi:10.1245/s10434-015-5008-z

32. Ting DSW et al. (2017) Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes JAMA 318:2211-2223 doi:10.1001/jama.2017.18152

33. Tsilimigras DI et al. (2020) Effect of Surgical Margin Width on Patterns of Recurrence among Patients Undergoing R0 Hepatectomy for T1 Hepatocellular Carcinoma: An International Multi-Institutional Analysis J Gastrointest Surg 24:1552-1560 doi:10.1007/s11605-019-04275-0

34. Villanueva A (2019) Hepatocellular Carcinoma N Engl J Med 380:1450-1462 doi:10.1056/NEJMra1713263

35. Vitale A et al. (2014) Is resectable hepatocellular carcinoma a contraindication to liver transplantation? A novel decision model based on "number of patients needed to transplant" as measure of transplant benefit Journal of hepatology 60:1165-1171 doi:10.1016/j.jhep.2014.01.022

36. Wang L, Jin YX, Ji YZ, Mu Y, Zhang SC, Pan SY (2020a) Development and validation of a prediction model for microvascular invasion in hepatocellular carcinoma World J Gastroenterol 26:1647-1659 doi:10.3748/wjg.v26.i14.1647

37. Wang L et al. (2020b) Postoperative adjuvant treatment strategy for hepatocellular carcinoma with microvascular invasion: a non-randomized interventional clinical study BMC Cancer 20:614 doi:10.1186/s12885-020-07087-7

38. Wang P et al. (2018) Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy Nat Biomed Eng 2:741-748 doi:10.1038/s41551-018-0301-3

39. Wang S et al. (2019) Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning Eur Respir J 53:1800986 doi:10.1183/13993003.00986-2018

40. Xu X et al. (2019) Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma J Hepatol 70:1133-1144 doi:10.1016/j.jhep.2019.02.023

41. Zeng F, Chen B, Zeng J, Wang Z, Xiao L, Deng G (2019) Preoperative neutrophil-lymphocyte ratio predicts the risk of microvascular invasion in hepatocellular carcinoma: A meta-analysis Int J Biol Markers 34:213-220 doi:10.1177/1724600819874487

42. Zhang H et al. (2020a) Predictive value of gamma-glutamyl transpeptidase to lymphocyte count ratio in hepatocellular carcinoma patients with microvascular invasion BMC Cancer 20:132 doi:10.1186/s12885-020-6628-7

43. Zhang X et al. (2020b) Contrast-enhanced CT radiomics for preoperative evaluation of microvascular invasion in hepatocellular carcinoma: A two-center study Clin Transl Med 10:e111 doi:10.1002/ctm2.111

44. Zhang XP et al. (2019) Postoperative adjuvant sorafenib improves survival outcomes in hepatocellular carcinoma patients with microvascular invasion after R0 liver resection: a propensity score matching analysis HPB (Oxford) 21:1687-1696 doi:10.1016/j.hpb.2019.04.014

45. Zheng J et al. (2017) Utility of Serum Inflammatory Markers for Predicting Microvascular Invasion and Survival for Patients with Hepatocellular Carcinoma Ann Surg Oncol 24:3706-3714 doi:10.1245/s10434-017-6060-7

46. Zhou J et al. (2018) Guidelines for Diagnosis and Treatment of Primary Liver Cancer in China (2017 Edition) Liver cancer 7:235-260 doi:10.1159/000488035

47. Zhu Y et al. (2018) A new laboratory-based algorithm to predict microvascular invasion and survival in patients with hepatocellular carcinoma Int J Surg 57:45-53 doi:10.1016/j.ijsu.2018.07.011
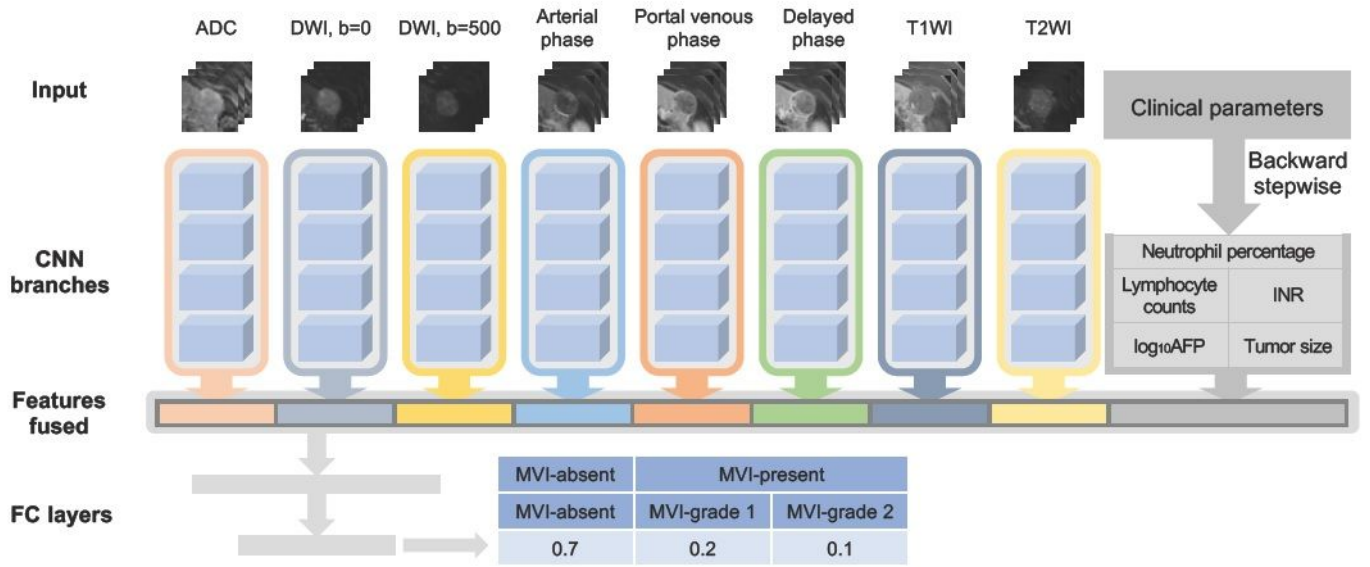
Figures

Figure 1

Figure 1

An illustration of the DL and DLC model architecture.
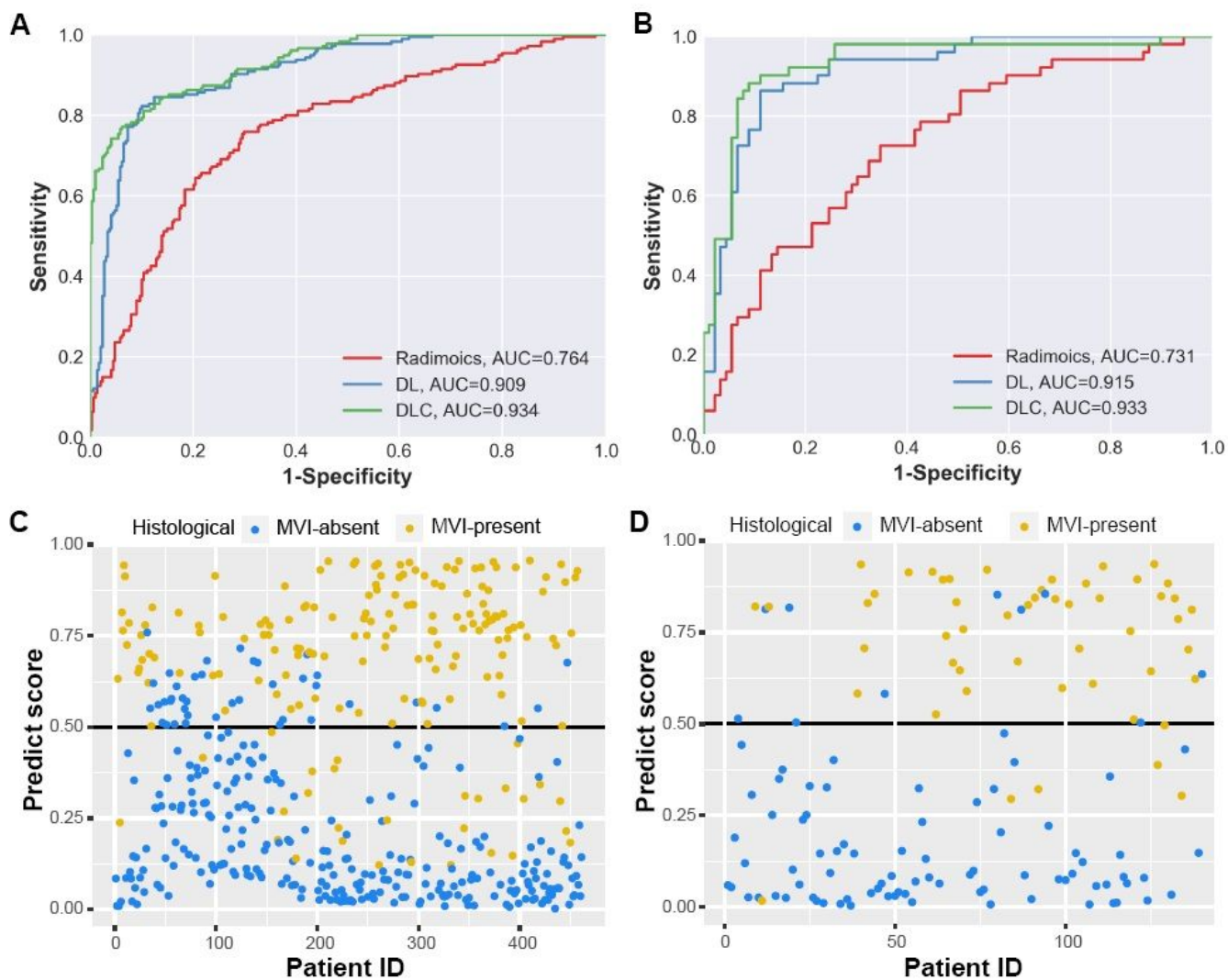
Figure 2

Figure 2

The ROC curves of the three models in the training cohort (A) and testing cohort (B). The scatter plots for the predictive score of DLC model in the training cohort (C) and the testing cohort (D).
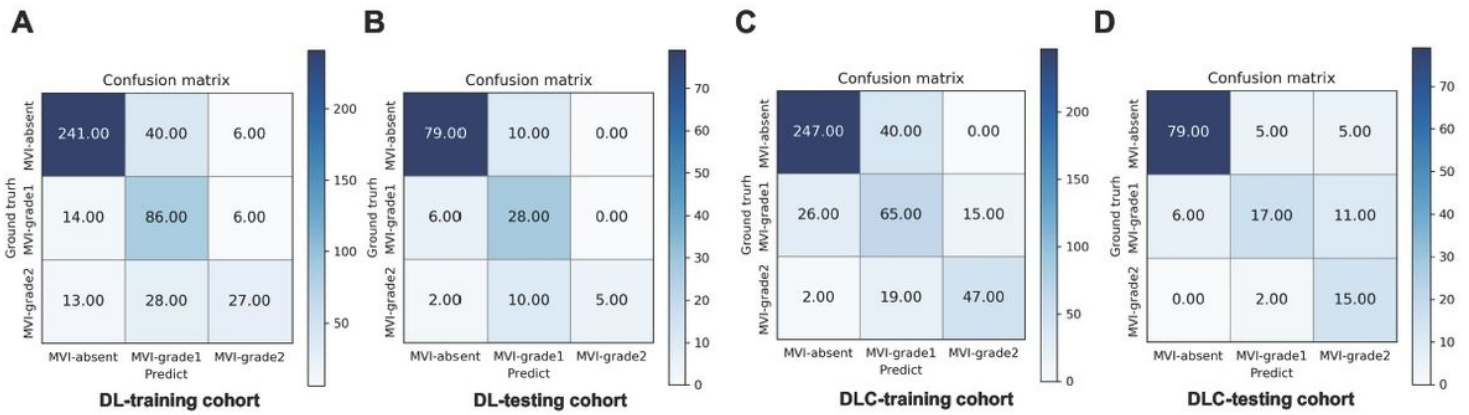
Figure 3

Figure 3

The confusion matrix of predicting MVI grade among MVI-absent, MVI-grade 1 and MVI-grade 2. (A) Confusion matrix of DL model in the training cohort. (B) Confusion matrix of DL model in the testing cohort. (C) Confusion matrix of DLC model in the training cohort. (D) Confusion matrix of DLC model in the testing cohort.
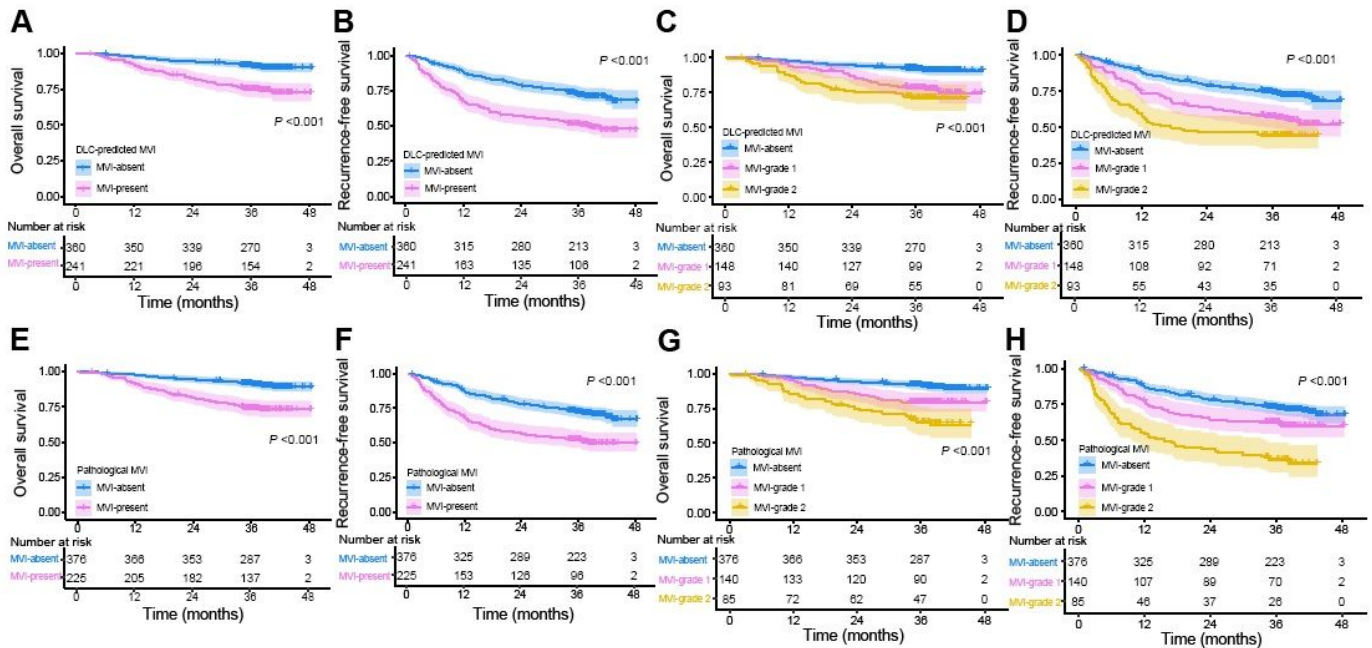
Figure 4

Figure 4

Overall survival (OS) and recurrence-free survival (RFS) curves scaled by DLC-predicted MVI status (A and B) and DLC-predicted MVI grade (C and D) with Kaplan-Meier. OS and RFS curves scaled by histological MVI status (E and F) and histological MVI grade (G and H) with Kaplan-Meier analysis.
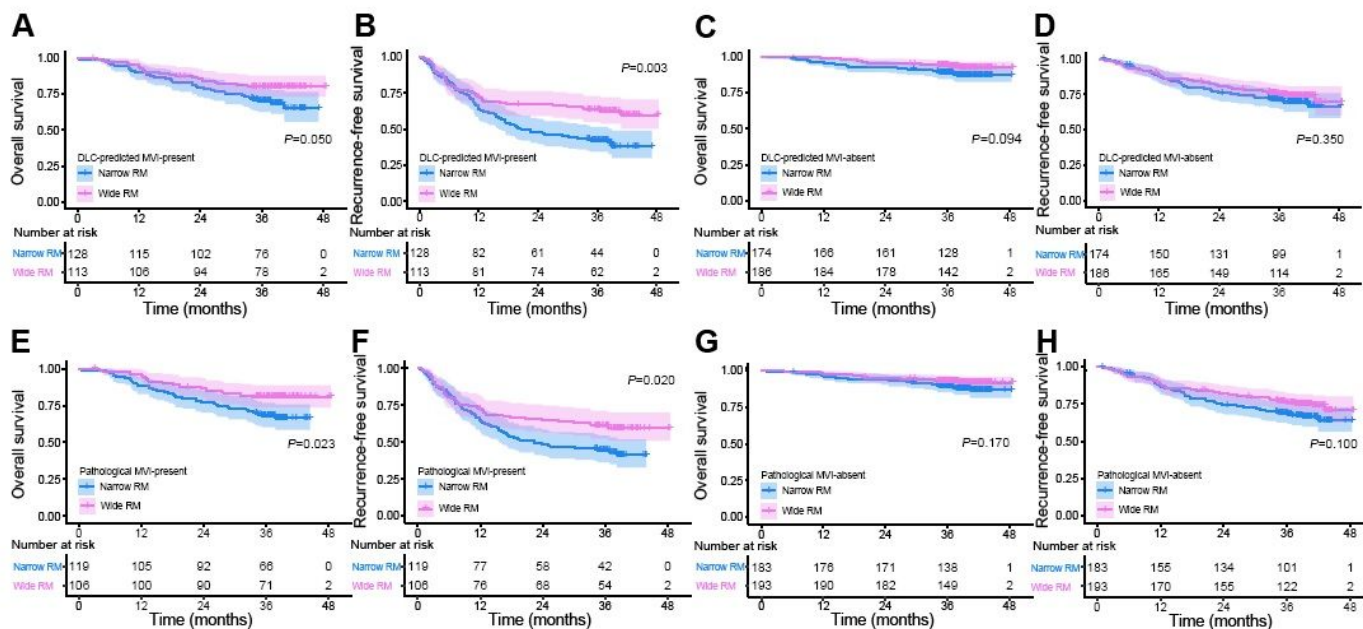
Figure 5

Figure 5

Overall survival (OS) and recurrence-free survival (RFS) curves scaled by resection margins (RM) in the patients with DLC-predicted MVI-present (A and B) and those with DLC-predicted MVI-absent (C and D). OS and RFS curves scaled by RM in the patients with histological MVI-present (E and F) and those with histological MVI-absent (G and H).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures.pdf
- Revisedatriclesupplementary.docx