Original Research

Longitudinal radiomics algorithm of posttreatment computed tomography images for early detecting recurrence of hepatocellular carcinoma after resection or ablation

Jing-xian Shen [a],[1], Qian Zhou [b],[1], Zhi-hang Chen [c], Qiao-feng Chen [d], Shu-ling Chen [e], Shi-ting Feng [f], Xin Li [g], Ting-fan Wu [g], Sui Peng [h],*, Ming Kuang [c],[e],*

[a] State Key Laboratory of Oncology in Southern China, Department of Medical Imaging, Sun Yat-sen University Cancer Center, Guangzhou, China
[b] Department of Medical Statistics, Clinical Trials Unit, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China
[c] Department of Liver Surgery, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China
[d] Department of Gastroenterology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China
[e] Department of Medical Ultrasonics, Division of Interventional Ultrasound, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China
[f] Department of Radiology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China
[g] GE Healthcare, Shanghai, China
[h] Clinical Trials Unit, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

A R T I C L E   I N F O

A B S T R A C T

Objectives: To develop a radiomics algorithm, improving the performance of detecting recurrence, based on posttreatment CT images within one month and at suspicious time during follow-up.
Materials and methods: A total of 114 patients with 228 images were randomly split (7:3) into training and validation cohort. Radiomics algorithm was trained using machine learning, based on difference-in-difference (DD) features extracted from tumor and liver regions of interest on posttreatment CTs within one month after resection or ablation and when suspected recurrent lesion was observed but cannot be confirmed as HCC during follow-up. The performance was evaluated by area under the receiver operating characteristic curve (AUC) and was compared among radiomics algorithm, change of alpha-fetoprotein (AFP) and combined model of both. Five-folded cross validation (CV) was used to present the training error.
Results: A radiomics algorithm was established by 34 DD features selected by random forest and multivariable logistic models and showed a better AUC than that of change of AFP (0.89 [95% CI: 0.78, 1.00] vs 0.63 [95% CI: 0.42, 0.84], $P = .04$) and similar with the combined model in detecting recurrence in the validation set. Five-folded CV error in the validation cohort was 21% for the algorithm and 26% for the changes of AFP.
Conclusions: The algorithm integrated radiomic features of posttreatment CT showed superior performance to that of conventional AFP and may act as a potential marker in the early detecting recurrence of HCC.

Introduction

Hepatocellular carcinoma (HCC) is one of the most common cancers and the third leading cause of cancer-related mortality worldwide [1]. Surgical resection and ablation are widely considered as the important curative treatments for HCC [1]. However, over 70% of patients suffered from recurrence within 5 years after curative resection or ablation [2]. Early detection of recurrent HCC could make curative treatment available due to early stage of the recurrent carcinoma [3].

Currently, the surveillance after resection or ablation mainly depended on serological test and imaging examination. Serum alpha-fetoprotein (AFP) was one of the most common method for surveillance. However, variation of its cut-off value and low accuracy restricted the application of AFP in the follow-up examination [4,5]. Although regular imaging examination includes ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI), CT was still the most commonly used in the detection of recurrent HCC due to its high sensitivity, worldwide availability and easy interpretability for clinicians [6–9]. However, diagnostic performance of CT

for suspicious recurrence with small size or atypical image pattern was poor, leading to more frequent follow-up or invasive biopsy. It was difficult to confirm suspicious recurrence by eyes even for experienced radiologists. A prospective study showed that the accuracy of diagnosing recurrent HCC less than 1 cm after ablation only reached 72% [8].

Invisible changes existed in CT images of suspicious recurrence, which had potential to become markers of recurrence. Radiomics has been a rapidly growing discipline based on quantitative image analysis through computer algorithm, which can objectively extract far more features than manual extraction [10]. It has been used in the prediction of therapeutic response, survival and genetic features of tumors [11–13]. In addition, Segal et al. found that imaging features were associated with genetic features [14]. Thus, we hypothesized that radiomics features on CT images after resection or ablation could reflect biological behavior of tumor and might improve the performance of detection of recurrent HCC by detecting subtle imaging changes not visible for a radiologist. In this study, we aimed to identify potential radiomics features associated with recurrence of HCC patients after surgical resection or ablation by collecting a series of follow-up CT images to improve performance of detecting recurrence, thereby making earlier intervention available.

Materials and methods

Study design

This retrospective cohort study was approved by the institutional ethic review board of Sun Yat-sen University Cancer Center (Approval Number: GZR2019–153) in accordance with the Declaration of Helsinki. The requirement for informed consent was waived because of the retrospective nature of this study.

Study population

We reviewed all patients with HCC who underwent surgical resection or ablation in a single center from January 2009 to April 2018. The inclusion criteria were as follows: (1) underwent resection or ablation; (2) pathologic diagnosis as primary HCC; (3) underwent CT examination within one month after resection or ablation and suspected recurrent lesion detected by an experienced radiologist could be observed on CT images before confirmed recurrence. Suspected recurrent lesion was categorized as LR-2 or LR-3 according to the Liver Imaging Reporting and Data System (LI-RADS) [15]. Fig. 1 showed the flow diagram of patient selection. CT images included in the study were collected on two time points: (1) time point 1: within one month after resection or ablation (CT1); (2) time point 2: when suspected recurrent lesion was observed but cannot be confirmed as HCC during follow-up (CT2). Suspected lesions were found and reviewed retrospectively by a radiologist (JXS., with more than 20 years of experience in abdominal radiology), and confirmed by another radiologist (STF, with more than 20 years of experience in abdominal radiology). The radiologists were blinded when evaluating the images. Recurrence was finally confirmed by MRI (LR4/5) and a continuous follow-up imaging examination. Patients included were randomly split (7:3) into training and validation cohort. The clinical data were collected prospectively from the database of HCC.

CT image acquisition

Hepatic CT was performed with a 64-, 80- or 320-detector row scanner CT machine (Aquilion TSX-101A, Toshiba; iCT256, Philips; Discovery HD750, GE Healthcare). The scanning parameters were routinely set
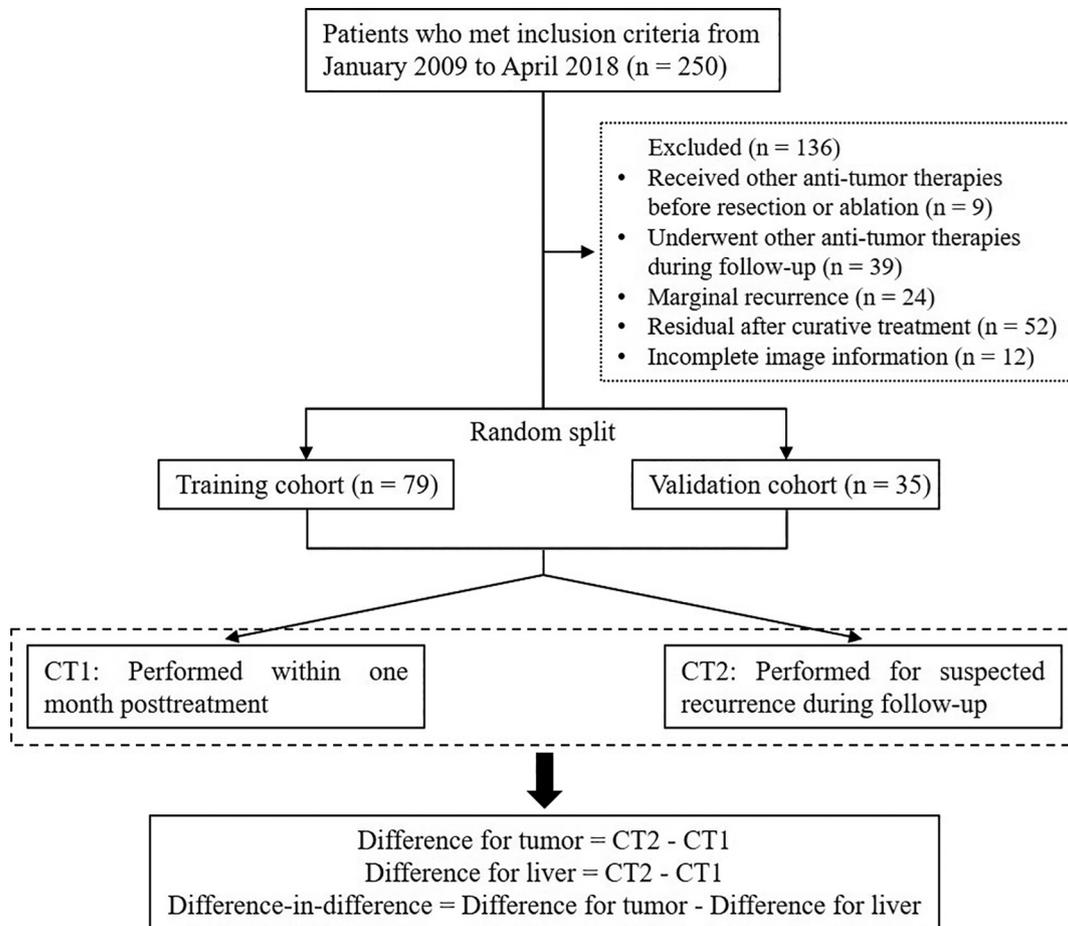
Fig. 1. Flow diagram of patient selection. CT1, CT images at time point 1 which was within one month after surgery. CT2, CT images at time point 2 which was the time that suspected recurrence lesion was found during regular follow-up.

regardless of the type of devices as follows: tube voltage, 120 kV; tube current, 240 mA; pitch factor, 1.0; slice thickness, 1 mm; and slice interval, 1 mm. Contrast-enhanced CT images were acquired after the injection of 1.0 ml/kg contrast agent (Ultravist, Bayer, Germany) into the antecubital vein at a rate of 2.0–3.0 ml/s followed by an immediate 20 ml saline flush delivered using a power injector (P3T abdomen module, Medrad Inc.). Hepatic arterial and portal venous phases were acquired at 27 s and 57 s after contrast agent injection, respectively. The slice thickness of the reconstructed arterial and portal venous phase images was 2 mm.

Region of interest identification, segmentation and radiomics features extraction

Portal vein phase of CT images were used. All regions of interest (ROIs) were delineated manually by two clinicians (Q.F.C. and Z.H.C., with more than 5 years of experience in abdominal radiology) under the guidance of J.X.S. using ITK-Snap software (open source software; www.itk-snap.org), and in the event of a disagreement, a consensus was reached by discussion. First, ROI was delineated around the liver lesion on the largest cross-sectional layer as indicated in the portal vein phase of CT2. Then, ROI delineated on CT2 was transferred to CT1 at the same location (Fig. 2a and Supplementary Fig. 1). To ensure that the area covered in CT1 was really the same area, we not only ensured that the tumor ROI in CT1 have the same number of pixels and make the shape as identical as possible but also evaluated the lesion position in CT2 through the blood vessels in the liver and delineated the tumor ROI in CT1 as far as possible. The CT image features of all ROIs were extracted and analyzed by the A.K. software (Analysis-Kit, GE healthcare) [16]. A total of 1044 radiomics features were extracted (Fig. 2b), including the following six categories (Supplementary Table 1): Histogram parameters, Texture parameters, Form factor parameters, Gray level co-occurrence matrix, Gray level run-length matrix, Gray level size zone matrix.

Difference-in-difference radiomics feature definition

Radiomics features extracted from tumor ROIs and liver ROIs on CT1 were defined tumor radiomics features at time point 1 and normal liver radiomics features at time point 1, respectively. Similarly, radiomics features extracted from tumor ROIs and liver ROIs on CT2 were defined tumor radiomics features at time point 2 and liver radiomics features at time point 2, respectively. The change of radiomics feature was defined as tumor radiomics feature at time point 2 minus tumor radiomics feature at time point 1. Considering the influence of background liver on tumorigenesis, we included normal liver radiomics features for analysis.

Difference-in-difference (DD) features were defined as difference for tumor minus difference for liver. The method of difference in difference was used to decrease the impact of different time points and locations of ROI on features [17]. Difference for tumor was tumor radiomics feature at time point 2 when suspected recurrent lesion was observed but cannot be confirmed as HCC minus tumor radiomics feature at time point 1 within one month after surgery, and difference for liver was liver radiomics feature at time point 2 minus liver radiomics feature at time point 1.

Statistical analysis

Continuous variables were presented as the mean $\pm$ standard deviation (SD) or median and interquartile change (IQR). Continuous variables between training and validation cohort, recurrence and no recurrence group were compared using Student's *t*-test or Wilcoxon rank sum test. Categorical variables were described as frequency (percentage) and were analyzed with chi-square test or Fisher exact test as appropriate. The inter-observer and intra-observer agreement for extracting radiomics features was evaluated by the intraclass correlation coefficient (ICC), and was graded as high consistency (ICC $\geq$ 0.8), middle (0.5–0.79), or low (<0.5).

Random forest algorithm was used to select important DD features associated with recurrence in the training cohort. Variable selection was not based on one random forest training but 100 iterations of random forest using the bootstrap method. Frequencies of top 100 important features per iteration were counted and ranked decreasingly. Subsequently, the top 40 DD features were selected initially, and six features were excluded to avoid overfitting according to model performance and 34 features were finally selected and combined into a radiomics algorithm (Fig. 2c and Supplementary Table 2). Logistic regression analysis was used to perform as binary outcome classifier for recurrence status instead of estimating the probability of recurrence in 1-, 3-, or 5-year follow-up time. The performance of the prediction model was evaluated with the area under the receiver operating characteristic curve (AUC) and its 95% confidence interval (*CI*) basing on the Youden index, and the accuracy, sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) of the optimal cutoff value were also calculated.

Internal evaluation in the training cohort was completed by 5-fold cross validation (CV) to determine the impact of different "training" and "validation" set sizes on prediction performance. External evaluation was performed
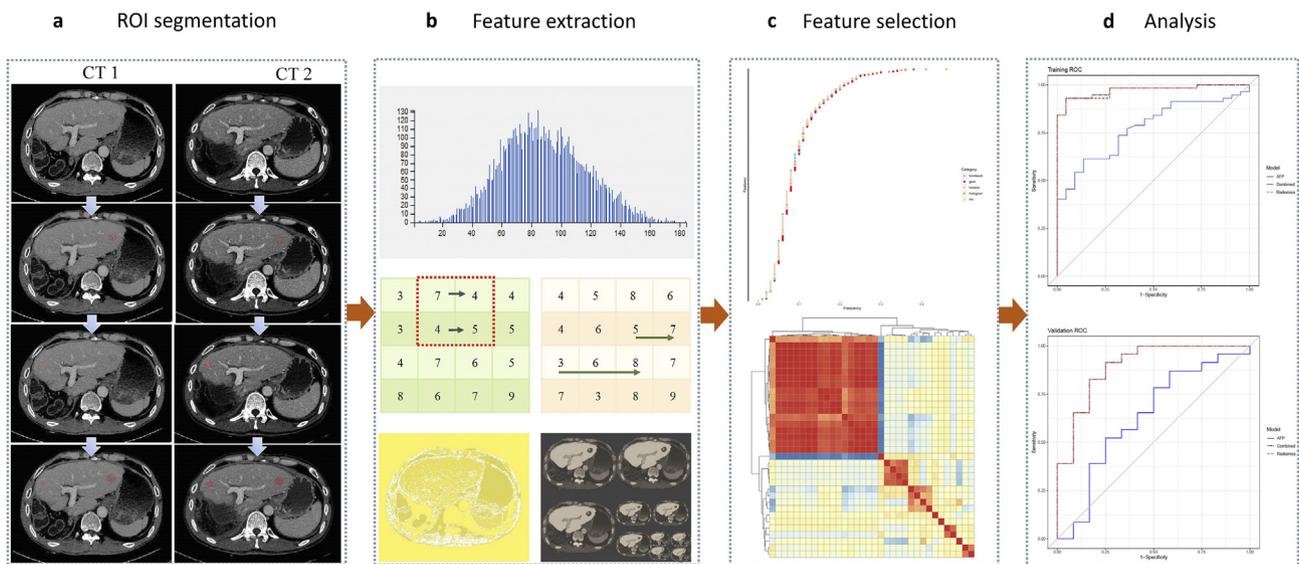
Fig. 2. Workflow of radiomics analysis in the study. (**a**) CT images acquisition and segmentation. The images of CT2 presented before that of CT1 was because ROI was first delineated around the liver lesion on the largest cross-sectional layer as indicated in the portal vein phase of CT2 and was then transferred to CT1 at the same location. (**b**) Radiomic features were extracted by an in-house software. (**c**) Features selection by random forest algorithm and logistic regression model. (**d**) Predictive model construction.

in the validation cohort. DeLong's test for two correlated ROC curves was performed to compare AUC of two logistic models (Fig. 2d). A two-sided *P*-value less than 0.05 was considered statistically significant. All analyses were performed by using R version 3.5.0.

Results

Baseline characteristics

In total, 250 patients fulfilling the inclusion criteria were considered, whereas the following patients were excluded from the study: receiving other anti-tumor therapies before resection or ablation ($n = 9$), undergoing other anti-tumor therapies during the follow-up ($n = 39$), with marginal recurrence ($n = 24$), with residual lesions after resection or ablation ($n = 52$) or incomplete imaging information ($n = 12$). Finally, 114 patients (228 images; mean age, 51 years ± 12; 103 men vs 11 women) with 70.2% recurrence were enrolled for the analysis, including 79 patients in the training cohort and 35 patients in the validation cohort (Fig. 1). The median time to recurrence of all the patients is 7.7 months. The negative lesions included transient hepatic perfusion disorder (13/34), regenerative nodule (4/34) and focal fatty liver (6/34). Other lesions (11/34) are unclassified in CT scans but confirmed by MRI or continuous follow-up. The baseline clinical characteristics of both cohorts were shown in Table 1. Baseline characteristics, except for higher serum AFP level in the training cohort at two time points, were not significantly different between the two cohorts. In the training cohort, higher serum AFP level at time point 2 ($P < .001$) were found in recurrent patients than those in non-recurrence group (Supplementary Table 3). In the validation cohort, all clinical characteristics were

Table 1
Baseline clinical characteristics of patients in the training and validation cohort.

Variables	Total	Training cohort	Validation cohort	P value
Age (years), mean ± SD	51 ± 12	50 ± 12	53 ± 12	0.23
Sex (man/women)	103/11	73/6	30/5	0.31
BMI	23.3 ± 3.4	23.4 ± 3.5	23.2 ± 3.2	0.98
HCV	2 (1.8)	1(1.3)	1(2.9)	0.52
Fatty liver	0(0%)	0(0%)	0(0%)	–
Drinking	19(16.7%)	12(15.2%)	7(20.0%)	0.53
Tumor size (cm)[a]	1.3 (1.1, 1.5)	1.3 (1.1, 1.5)	1.3 (1.0, 1.4)	0.30
HBsAg, n (%)				0.06
Unknown	2 (1.8)	0 (0.0)	2 (5.7)	
Negative	11 (9.6)	6 (7.6)	5 (14.3)	
Positive	101 (88.6)	73 (92.4)	28 (80.0)	
Primary treatment, n (%)				1.00
Surgical resection	102 (89.5)	71 (89.9)	31 (88.6)	
Ablation	12 (10.5)	8 (10.1)	4 (11.4)	
Child-Pugh class, n (%)				0.55
A	111 (97.4)	76 (96.2)	35 (100.0)	
B	3 (2.6)	3 (3.8)	0 (0.0)	
BCLC stage, n (%)				0.051
0	11 (9.6)	4 (5.1)	7 (20.0)	
A	89 (78.1)	63 (79.7)	26 (74.3)	
B	4 (3.5)	3 (3.8)	1 (2.9)	
C	10 (8.8)	9 (11.4)	1 (2.9)	
Baseline AFP (ug/L)	49.7 (6.0,666.1)	92.4 (8.22598)	6.7 (2.6287.5)	0.002
≤20	49(43.0%)	28(35.4%)	21(60.0%)	0.015
>20	65(57.0%)	51(64.6%)	14(40.0%)	
AFP 1 (ug/L)[a]	7.2 (3.6, 18.3)	9.2 (4.6, 21.7)	4.4 (2.5, 10.0)	0.001
AFP 2 (ug/L)[a,b]	5.1 (2.8, 16.0)	6.6 (2.9, 43.4)	3.7 (2.1, 7.5)	0.01
Change of AFP[a,b]	0.0 (−0.0, 0.0)	0.0 (−0.1, 0.1)	−0.0 (−0.0, 0.0)	0.45
Time to recurrence (months)	7.7 (3.8,20.8)	8.8 (3.4,22.1)	7.2 (4.4,17.6)	0.89
Recurrence, n (%)				0.49
Yes	80 (70.2)	57 (72.2)	23 (65.7)	
No	34 (29.8)	22 (27.8)	12 (34.3)	

Abbreviations: AFP 1, alpha-fetoprotein at time point 1. AFP 2, alpha-fetoprotein at time point 2. BCLC, Barcelona Clinic Liver Cancer. HBsAg, hepatitis B surface antigen.
[a] Date are presented as median (inter-quartile range, IQR).
[b] Change of AFP was defined as AFP 2 minus AFP 1.

found no significant differences between the two groups in terms of recurrence status (Supplementary Table 4).

Inter-observer and intra-observer agreement of feature extraction

The inter-observer ICC was ≥ 0.8, 0.5–0.79, < 0.5 for 88.1%, 10.5% and 1.3% of the features, respectively. The intra-observer ICC was ≥ 0.8, 0.5–0.79, < 0.5 for 82.9%, 15.8% and 1.2% of the features, respectively.

Development of radiomics model

A total of 1044 DD radiomics features were decreased to 34 potential risk predictors associated with recurrence by random forest in the training cohort (Fig. 3a). The heatmap which depicted correlation coefficients matrix of 34 selected DD radiomics features was shown in Fig. 3b. Correlation coefficient matrix showed these 34 features could be mainly clustered into three categories with positive correlation, indicating suitable for establishing a radiomics model.

Validation of radiomics model

The radiomics model which constructed by linearly combining the 34 DD features using multivariable logistic regression analysis, and yielded an AUC of 0.97 (95% CI: 0.94, 1.00) in the training cohort with a five-folded CV error of 21% and 0.89 (95% CI: 0.79, 1.00) in the validation cohort, demonstrating the good predictive performance of the model. In addition, the accuracy, sensitivity, specificity, NPV and PPV of the model in detecting recurrence were 0.86, 0.91, 0.75, 0.82and 0.88 in the validation cohort (Table 2).

Association of Radiomics Features and Change of AFP Level with Recurrence.

In the training cohort, the radiomics algorithm obtained from a multivariable logistic regression model was significantly associated with HCC recurrence (OR = 2.7; 95% CI: 1.5, 6.3; $P = .02$), suggesting that this radiomics algorithm had great potential to detect recurrence. However, no significant association was found between the changes of serum AFP level with recurrence (OR = 1.00; 95% CI: 0.97, 1.04; $P = .88$). In the combined model, the radiomics algorithm showed significant association with recurrence with an OR of 2.7 (95% CI: 1.5, 6.4; $P = .01$) and the changes of AFP level was still not statistically significant (OR = 0.99; 95% CI: 0.93, 1.03; $P = .72$).

Comparison of Predictive Performance of Radiomics, AFP and Combined Model.

The AUC, accuracy, sensitivity, specificity, NPV and PPV of the changes of serum AFP level in predicting recurrence were also calculated in the training and validation cohort, as shown in Table 2. Subsequently, a combined model which integrated the 34 DD radiomics features with the changes of serum AFP level was developed to predict recurrence in the two cohorts, and found that it yielded an AUC, accuracy, sensitivity, specificity, NPV and PPV of 0.97 (95% CI: 0.94, 1.00), 94%, 93%, 95%, 84% and 98% in the training cohort and 0.89 (95% CI: 0.78, 1.00), 86%, 91%, 75%, 82% and 88% in the validation cohort with a five-folded CV error of 14% (Table 2). As shown in Table 2, the predictive performance of the radiomics algorithm, the changes of AFP level and combined model for HCC recurrence were compared in the training and validation cohort. DeLong's test for two correlated ROC curves showed that the AUCs of the radiomics algorithm were both significantly larger than that of the changes of AFP level in the training cohort (0.97 vs 0.77, Fig. 4a) and the validation cohort (0.89 vs 0.63, $P = .04$, Fig. 4b), and demonstrated a higher predicting performance in detecting recurrence over the changes of AFP level alone. Furthermore, when combining the change of AFP and radiomics algorithm, the AUCs had no significant increase compared with the radiomics algorithm alone in both cohorts (Fig. 4).

Discussion

In the present study, we established a radiomics algorithm based on the change of radiomic features to detect recurrence, and yielded an area under
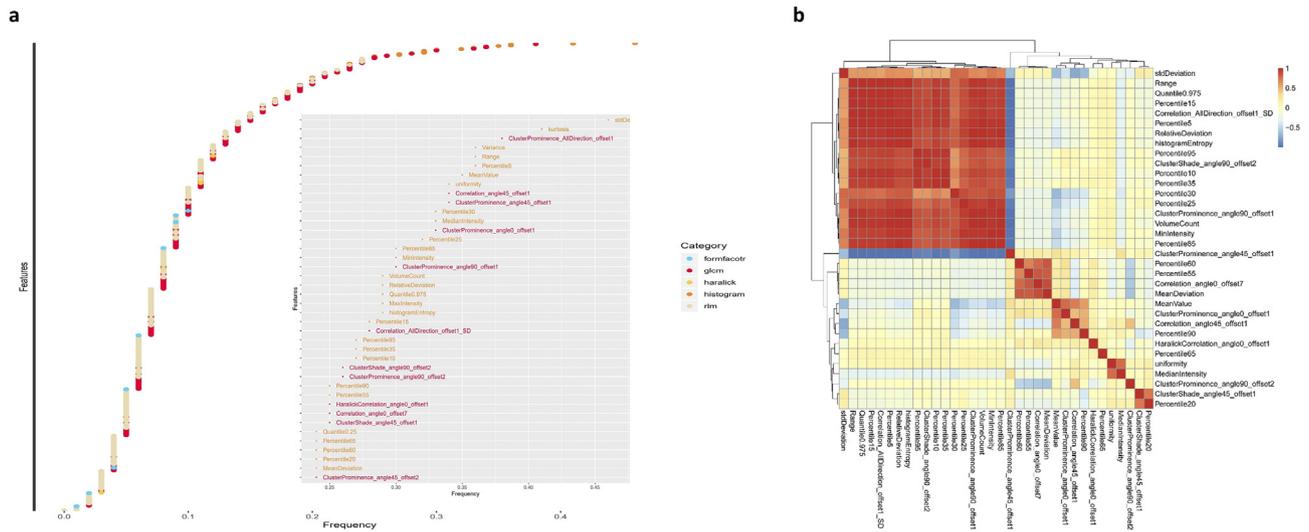
Fig. 3. Radiomics feature selection and radiomics heatmap. (**a**) Random forest algorithm was used to find the most important variables using 100 iteration. The number of trees used for the random forest were 100 and the number of variables tested at each node were five. The top 40 difference-in-difference features were selected, and six features were excluded to avoid overfitting according to model performance and 34 features were finally selected and combined into a radiomics signature. (**b**) The heatmap depicted correlation coefficients matrix of 34 selected difference-in-difference radiomics features in the training cohort. Unsupervised clustering analysis was used. Red color was proportional to the positive correlation, blue was negative correlation, and yellow was not correlated with each other. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the receiver operating characteristic curve of 0.97 and 0.89 in the training and validation cohort respectively. We concluded that longitudinal CT radiomic features with suspected recurrent lesion after treatment, integrated into a radiomic algorithm, could assist clinicians in diagnosis of HCC recurrence. Additionally, the radiomic algorithm was independently associated with hepatocellular carcinoma recurrence, suggesting that our finding play a critical role in the early detection and treatment of recurrent hepatocellular carcinoma.

Radiomics has been recognized as an important method in oncology. However, there were few studies based on radiomics to explore a feasible approach for surveillance of HCC patients after resection or ablation. Previous studies tried to use CT radiomics to predict the recurrence of HCC with preoperative images, for example, they combined the radiomics algorithm with clinical variables or used a single texture analysis parameter to predict early recurrence with an AUC or C-index less than 0.85 [18–22]. However, these studies shared the common flaws that the localization of lesions and specific time point of tumor recurrence remained unpredicted, thus making the precise intervention of recurrence impossible. Instead, our finding used a simple radiomics algorithm to analyze posttreatment images, and the results demonstrated not only early detection and intervention but also surveillance of recurrence could be achieved. In addition, we chose CT images to construct our model though MRI imaging can find more image features

and make smaller lesions obvious compared with CT. Because our method considered background liver information, MRI imaging might perform better than CT scan in our model. However, it is difficult for us to obtain a series of follow-up MRI images because MRI imaging is usually used as an assistant tool for the undermined in CT imaging. Due to easy access to follow-up CT images, we preferentially chose CT images to construct our model.

Previous study showed poor diagnostic performance for recurrent HCC less than 1 cm after ablation based on CT images with an AUC of 0.72 [8]. During follow-up period after resection or ablation, a suspected recurrent lesion was usually observed on CT images but could not clarify its characteristics, resulting in a wait of 1–2 months or further MRI examination if available. In our study, 79 suspected recurrent lesions less than 2 cm in size lacked typical CT dynamic features in the training cohort, and then applied to training our radiomics algorithm for the purpose of improving diagnostic performance of CT in detecting recurrent HCC less than 2 cm, which yielded an AUC of 0.89 in the validation cohort. This may imply that, we cannot observe recurrence status directly from these small lesions either through vision or radiomic features in CT imaging of one time point, but recurrence status could be detected by analyzing substantial change of radiomic features between CT images performed two time points, just as monitor changes in serum AFP level for predicting recurrence [23].

Table 2

Comparison of diagnostic performance in detecting recurrence of hepatocellular carcinoma.

Item	Training cohort			Validation cohort		
	Radiomics algorithm	Change of AFP	Combined model	Radiomics algorithm	Change of AFP	Combined model
AUC[c]	0.97 (0.94, 1.00)	0.77 (0.67, 0.88)[a]	0.97 (0.94, 1.00)[b]	0.89 (0.78, 1.00)	0.63 (0.42, 0.84)[a]	0.89 (0.78, 1.00)[b]
Accuracy (%)[c]	94 (86, 98)	68 (57, 78)	94 (86, 98)	86 (70, 95)	71 (54, 85)	86 (70, 95)
Sensitivity (%)[c]	93 (83, 98)	61 (48, 74)	93 (83, 98)	91 (72, 99)	87 (66, 97)	91 (72, 99)
Specificity (%)[c]	95 (77, 100)	86 (65, 97)	95 (77, 100)	75 (43, 95)	42 (15, 72)	75 (43, 95)
NPV (%)[c]	84 (64, 95)	46 (31, 63)	84 (64, 95)	82 (48, 98)	62 (24, 91)	82 (48, 98)
PPV (%)[c]	98 (90, 100)	92 (79, 98)	98 (90, 100)	88 (68, 97)	74 (54, 89)	88 (68, 97)
5-fold CV error	21%	26%	14%	–	–	–

Abbreviations: AUC, area under the operating characteristic curve; AFP, alpha-fetoprotein; *CI*, confidence interval; CV, cross-validation; DD, difference-in-difference; NPV, negative predictive value; PPV, positive predictive value.

[a] AUC of Radiomics algorithm vs Change of AFP level: $P < .05$.

[b] AUC of Combined model vs Radiomics algorithm: $P > .05$.

[c] Data in parentheses are 95% confidence interval.

a

b
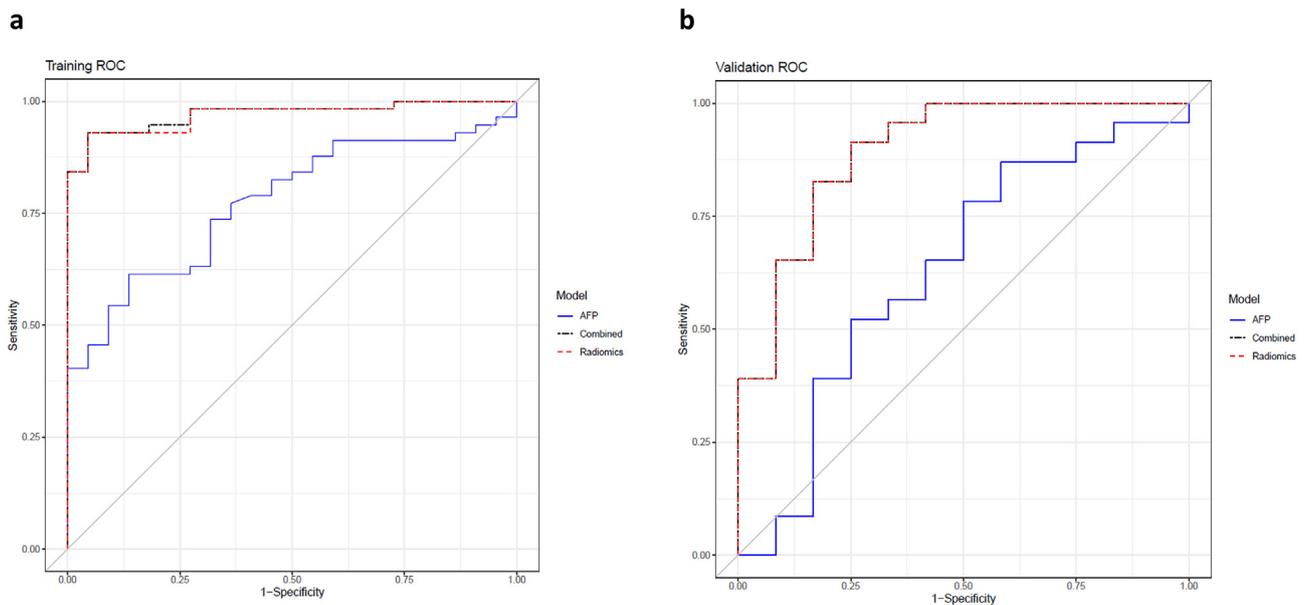
Fig. 4. Receiver operating characteristics curve (ROC) of the training and validation cohort. (a) Performance of radiomics signature, change of AFP level and combined model in the training cohort. (b) Performance of radiomics signature, change of AFP level and combined model in the validation cohort. AFP, alpha-fetoprotein. AUC, area under the receiver operating characteristic curve.

Mattonen et al. [24] showed that radiomics based on follow-up images could detect early changes associated with local recurrence after stereotactic ablative radiation therapy for lung cancer. In agreement with the previous study, our finding reported that radiomics algorithm based on the change of radiomic features could detect early changes associated with recurrence after resection or ablation.

Previous studies have found that radiomic features were closely related to tumor biological behavior and microscopic structure [25–27]. Our study discovered 34 DD radiomic features significantly associated with recurrence of HCC. These features could be categorized into two types, including GLCM and GLRLM. Among texture features, GLCM and GLRLM reflect signal mixing degree of the lesions by means of relative relationship between distribution and site of the gray level, and thus are important markers of intra-tumor homogeneity. DD radiomic features could detect the change of homogeneity of tumor region by eliminating the interference of background liver, and thus produced great performance of detecting recurrence with an AUC of 0.97. Previous studies have also demonstrated the ability of CT radiomics analysis for assessing nonalcoholic steatohepatitis [28] and microvascular invasion [29,30].

In addition, previous studies proposed that persistent changes of AFP level maybe a prognostic factor for HCC development, and AFP was the most frequently tested indicators in the detection of HCC [31]. For example, elevated AFP increases risk of tumor recurrence posttreatment [32]. However, AFP levels did not improve the detection rate of HCC though combining with ultrasound during follow-up, moreover, appropriate cutoff values may limit its sensitivity and specificity in clinical application [33]. By contrast, radiomic methods used CT images to quantify tumor features at the macroscopic level and may not meet these problems abovementioned. In our study, we compared the prediction performance of radiomic model with changes of AFP or the combined model, and the results suggest that a better predictive AUC of radiomic algorithm than that of change of AFP (0.89 vs 0.63, $P = .038$) and similar with the combined model in detecting recurrence. This indicated that radiomic features played a key role in early detecting HCC recurrence.

Major limitation of this study is the retrospective nature with selective bias and a single-center study with a limited sample size. Larger sample size is required to strengthen the power of the conclusion. Despite no external validation, cross validation would make up for this shortcoming and demonstrate robustness of our radiomics algorithm. Larger validation cohorts will be needed in future studies. In addition, definition of suspected

lesions was potentially subjective though using LI-RADS. If the radiologist was uncertain which the lesions belong to, another expert was consulted and agreement should be reached. In addition, almost all (97.4%) patients had a Child Score A in our study. For other patients who suffer from underlying health issues, our current model might not be applied. Further, to improve generalizability of our model, more patients with alcoholism, Hepatitis B or C should be included.

In conclusion, we found that difference-in-difference of CT radiomic features could detect early changes associated with recurrence and might assist clinicians in earlier intervention for recurrence of hepatocellular carcinoma.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tranon.2020.100866.

CRediT authorship contribution statement

Authors contribution: **J Shen**: Conceptualization, Data acquisition, Resources, Roles/Writing - original draft. **Q Zhou**: Conceptualization, Design, Methodology, Formal analysis, Roles/Writing - original draft. **Z Chen**: Data curation, Roles/Writing - original draft. **Q Chen**: Data curation, Writing - review & editing. **S Chen**: Writing - review & editing. **S Feng**: Administrative or material support. **X Li**: Software, Technical support. **T Wu**: Software, Technical support. **S Peng**: Design, Review and revision of draft manuscript. **M. Kuang**: Conceptualization, Design, Methodology, Interpretation, Supervision, Funding acquisition.

Funding

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Forner, M. Reig, J. Bruix, Hepatocellular carcinoma, Lancet 391 (2018) 1301–1314, https://doi.org/10.1016/S0140-6736(18)30010-2.

[2] P. Tabrizian, G. Jibara, B. Shrager, M. Schwartz, S. Roayaie, Recurrence of hepatocellular cancer after resection: patterns, treatments, and prognosis, Ann. Surg. 261 (2015) 947–955, https://doi.org/10.1097/SLA.0000000000000710.

[3] I. Hatzaras, D.A. Bischof, B. Fahy, D. Cosgrove, T.M. Pawlik, Treatment options and surveillance strategies after therapy for hepatocellular carcinoma, Ann. Surg. Oncol. 21 (2014) 758–766, https://doi.org/10.1245/s10434-013-3254-5.

[4] F. Trevisani, P.E. D'Intino, A.M. Morselli-Labate, G. Mazzella, E. Accogli, P. Caraceni, M. Domenicali, S. De Notariis, E. Roda, M. Bernardi, Serum alpha-fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: influence of HBsAg and anti-HCV status, J. Hepatol. 34 (2001) 570–575.

[5] S. Gupta, S. Bent, J. Kohlwes, Test characteristics of alpha-fetoprotein for detecting hepatocellular carcinoma in patients with hepatitis C - a systematic review and critical analysis, Ann. Intern. Med. 139 (2003) 46–50.

[6] A. Colli, M. Fraquelli, G. Casazza, S. Massironi, A. Colucci, D. Conte, P. Duca, Accuracy of ultrasonography, spiral CT, magnetic resonance, and alpha-fetoprotein in diagnosing hepatocellular carcinoma: a systematic review, Am. J. Gastroenterol. 101 (2006) 513–523.

[7] A. Furlan, D. Marin, P. Cabassa, A. Taibbi, E. Brunelli, F. Agnello, R. Lagalla, G. Brancatelli, Enhancement pattern of small hepatocellular carcinoma (HCC) at contrast-enhanced US (CEUS), MDCT, and MRI: Intermodality agreement and comparison of diagnostic sensitivity between 2005 and 2010 American Association for the Study of Liver Diseases (AASLD), Eur. J. Radiol. 81 (2012) 2099–2105, https://doi.org/10.1016/j.ejrad.2011.07.010.

[8] Y. Imai, K. Katayama, M. Hori, T. Yakushijin, K. Fujimoto, T. Itoh, T. Igura, M. Sakakibara, M. Takamura, M. Tsurusaki, H. Takahashi, K. Nakanishi, N. Usuki, K. Tsuji, H. Ohashi, T. Kim, T. Takehara, T. Murakami, Prospective comparison of Gd-EOB-DTPA enhanced MRI with dynamic CT for detecting recurrence of HCC after radiofrequency ablation, Liver Cancer 6 (2017) 349–359, https://doi.org/10.1159/000481416.

[9] V.M. Runge, Safety of the gadolinium-based contrast agents for magnetic resonance imaging, focusing in part on their accumulation in the brain and especially the dentate nucleus, Investig. Radiol. 51 (2016) 273–279, https://doi.org/10.1097/RLI.0000000000000273.

[10] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data, Radiology 278 (2016) 563–577, https://doi.org/10.1148/radiol.2015151169.

[11] Y. Li, X. Liu, K. Xu, Z. Qian, K. Wang, X. Fan, S. Li, Y. Wang, T. Jiang, MRI features can predict EGFR expression in lower grade gliomas: a voxel-based radiomic analysis, Eur. Radiol. 28 (2018) 356–362, https://doi.org/10.1007/s00330-017-4964-z.

[12] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (2017) 115–118, https://doi.org/10.1038/nature21056.

[13] N.M. Braman, M. Etesami, P. Prasanna, C. Dubchuk, H. Gilmore, P. Tiwari, D. Plecha, A. Madabhushi, Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI, Breast Cancer Res. 19 (2017) 57, https://doi.org/10.1186/s13058-017-0846-1.

[14] E. Segal, C.B. Sirlin, C. Ooi, A.S. Adler, J. Gollub, X. Chen, B.K. Chan, G.R. Matcuk, C.T. Barry, H.Y. Chang, M.D. Kuo, Decoding global gene expression programs in liver cancer by noninvasive imaging, Nat. Biotechnol. 25 (2007) 675–680.

[15] K.M. Elsayes, J.C. Hooker, M.M. Agrons, A.Z. Kielar, A. Tang, K.J. Fowler, V. Chernyak, M.R. Bashir, Y. Kono, R.K. Do, D.G. Mitchell, A. Kamaya, E.M. Hecht, C.B. Sirlin, 2017 version of LI-RADS for CT and MR imaging: an update, Radiographics 37 (2017) 1994–2017, https://doi.org/10.1148/rg.2017170098.

[16] J. Liang, X. Huang, H. Hu, Y. Liu, Q. Zhou, Q. Cao, W. Wang, B. Liu, Y. Zheng, X. Li, X. Xie, M. Lu, S. Peng, L. Liu, H. Xiao, Predicting malignancy in thyroid nodules: radiomics score versus 2017 American College of Radiology (ACR) thyroid imaging, reporting and data system (TI-RADS), Thyroid 28 (2018) 1024–1033, https://doi.org/10.1089/thy.2017.0525.

[17] C. Wing, K. Simon, R.A. Bello-Gomez, Designing difference in difference studies: best practices for public health policy research, Annu. Rev. Public Health 39 (2018) 453–469.

[18] T.C.H. Hui, T.K. Chuah, H.M. Low, C.H. Tan, Predicting early recurrence of hepatocellular carcinoma with texture analysis of preoperative MRI: a radiomics study, Clin. Radiol. 73 (2018) 1056.e11–1056.e16, https://doi.org/10.1016/j.crad.2018.07.109.

[19] Y. Zhou, L. He, Y. Huang, S. Chen, P. Wu, W. Ye, Z. Liu, C. Liang, CT-based radiomics signature: a potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma, Abdom. Radiol. (NY) 42 (2017) 1695–1704.

[20] W. Wang, Q. Chen, Y. Iwamoto, X. Han, Q. Zhang, H. Hu, L. Lin, Y.W. Chen, Deep learning-based radiomics models for early recurrence prediction of hepatocellular carcinoma with multi-phase CT images and clinical data, Conf. Proc. IEEE Eng. Med. Biol. Soc. 2019 (2019) 4881–4884, https://doi.org/10.1109/EMBC.2019.8856356.

[21] G.W. Ji, F.P. Zhu, Q. Xu, K. Wang, M.Y. Wu, W.W. Tang, X.C. Li, X.H. Wang, Radiomic features at contrast-enhanced CT predict recurrence in early stage hepatocellular carcinoma: a multi-institutional study, Radiology 191470 (2020)https://doi.org/10.1148/radiol.2020191470.

[22] G.W. Ji, F.P. Zhu, Q. Xu, K. Wang, M.Y. Wu, W.W. Tang, X.C. Li, X.H. Wang, Machine-learning analysis of contrast-enhanced CT radiomics predicts recurrence of hepatocellular carcinoma after resection: a multi-institutional study, EBioMedicine 50 (2019) 156–165, https://doi.org/10.1016/j.ebiom.2019.10.057.

[23] S. Siripongsakun, S.H. Wei, S. Lin, J. Chen, S.S. Raman, J. Sayre, M.J. Tong, D.S. Lu, Evaluation of alpha-fetoprotein in detecting hepatocellular carcinoma recurrence after radiofrequency ablation, J. Gastroenterol. Hepatol. 29 (2014) 157–164, https://doi.org/10.1111/jgh.12438.

[24] S.A. Mattonen, D.A. Palma, C. Johnson, A.V. Louie, M. Landis, G. Rodrigues, I. Chan, R. Etemad-Rezai, T.P. Yeung, S. Senan, A.D. Ward, Detection of local cancer recurrence after stereotactic ablative radiation therapy for lung cancer: physician performance versus radiomic assessment, Int. J. Radiat. Oncol. Biol. Phys. 94 (2016) 1121–1128, https://doi.org/10.1016/j.ijrobp.2015.12.369.

[25] P. Grossmann, O. Stringfield, N. El-Hachem, M.M. Bui, E. Rios Velazquez, C. Parmar, R.T. Leijenaar, B. Haibe-Kains, P. Lambin, R.J. Gillies, H.J. Aerts, Defining the biological basis of radiomic phenotypes in lung cancer, Elife 6 (2017)https://doi.org/10.7554/eLife.23421 pii:e23421.

[26] K. Hayano, F. Tian, A.R. Kambadakone, S.S. Yoon, D.G. Duda, B. Ganeshan, D.V. Sahani, Texture analysis of non-contrast-enhanced computed tomography for assessing angiogenesis and survival of soft tissue sarcoma, J. Comput. Assist. Tomogr. 39 (2015) 607–612, https://doi.org/10.1097/RCT.0000000000000239.

[27] B. Ganeshan, V. Goh, H.C. Mandeville, Q.S. Ng, P.J. Hoskin, K.A. Miles, Non-small cell lung cancer: histopathologic correlates for texture parameters at CT, Radiology 266 (2013) 326–336, https://doi.org/10.1148/radiol.12112428.

[28] S. Naganawa, K. Enooku, R. Tateishi, H. Akai, K. Yasaka, J. Shibahara, T. Ushiku, O. Abe, K. Ohtomo, S. Kiryu, Imaging prediction of nonalcoholic steatohepatitis using computed tomography texture analysis, Eur. Radiol. 28 (2018) 3050–3058, https://doi.org/10.1007/s00330-017-5270-5.

[29] J. Zheng, J. Chakraborty, W.C. Chapman, S. Gerst, M. Gonen, L.M. Pak, W.R. Jarnagin, R.P. DeMatteo, R.K.G. Do, A.L. Simpson, Hepatopancreatobiliary Service in the Department of Surgery of the Memorial Sloan Kettering Cancer Center, Research Staff in the Department of Surgery at Washington University School of Medicine, Preoperative prediction of microvascular invasion in hepatocellular carcinoma using quantitative image analysis, J. Am. Coll. Surg. 225 (2017) 778–788.e1, https://doi.org/10.1016/j.jamcollsurg.2017.09.003.

[30] X. Xu, H.L. Zhang, Q.P. Liu, S.W. Sun, J. Zhang, F.P. Zhu, G. Yang, X. Yan, Y.D. Zhang, X.S. Liu, Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma, J. Hepatol. 70 (2019) 1133–1144, https://doi.org/10.1016/j.jhep.2019.02.023.

[31] H. Imamura, Y. Matsuyama, E. Tanaka, T. Ohkubo, K. Hasegawa, S. Miyagawa, Y. Sugawara, M. Minagawa, T. Takayama, S. Kawasaki, M. Makuuchi, Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy, J. Hepatol. 38 (2003) 200–207, https://doi.org/10.1016/s0168-8278(02)00360-4.

[32] G.L. Wong, H.L. Chan, Y.K. Tse, H.Y. Chan, C.H. Tse, A.O. Lo, V.W. Wong, On-treatment alpha-fetoprotein is a specific tumor marker for hepatocellular carcinoma in patients with chronic hepatitis B receiving entecavir, Hepatology 59 (2014) 986–995, https://doi.org/10.1002/hep.26739.

[33] M. Biselli, F. Conti, A. Gramenzi, M. Frigerio, A. Cucchetti, G. Fatti, M. D'Angelo, M. Dall'Agata, E.G. Giannini, F. Farinati, F. Ciccarese, P. Andreone, M. Bernardi, F. Trevisani, A new approach to the use of α-fetoprotein as surveillance test for hepatocellular carcinoma in patients with cirrhosis, Br. J. Cancer 112 (2015) 69–76, https://doi.org/10.1038/bjc.2014.536.