Evaluating histologic differentiation of hepatitis B virus-related hepatocellular carcinoma using intravoxel incoherent motion and AFP levels alone and in combination

Qungang Shan,[1] Jingbiao Chen,[1] Tianhui Zhang,[1] Ronghua Yan,[1] Jun Wu,[1] Yunhong Shu,[2] Zhuang Kang,[1] Bingjun He,[1] Zhongping Zhang,[3] Jin Wang[1]

[1]Department of Radiology, The Third Affiliated Hospital of Sun Yat-Sen University, 600 Tianhe Rd, Guangzhou 510630, People's Republic of China
[2]Department of Radiology, Mayo Clinic, Rochester, MN, USA
[3]GE Healthcare MR Research China, Beijing, China

Abstract

Purpose: To evaluate histologic differentiation of hepatitis B virus (HBV)-related hepatocellular carcinomas (HCCs) using apparent diffusion coefficient (ADC) and intravoxel incoherent motion (IVIM)-derived metrics and to compare findings with alpha-fetoprotein (AFP) levels alone and in combination.
Materials and method: One hundred and six chronic HBV-related HCC patients who underwent IVIM diffusion-weighted magnetic resonance imaging with eleven *b* values were enrolled. Mean ADC, diffusion coefficient (*D*), pseudodiffusion coefficient (*D**), and perfusion fraction (*f*) values were determined for all detected lesions. The metrics and AFP levels of different histologically differentiated groups were compared. Spearman's rank correlation was used to assess the statistical dependence among the histologically differentiated HCCs. Receiver operating characteristic (ROC) analysis was performed to evaluate diagnostic performance of these metrics and AFP levels alone and in combination.
Results: ADC, *D*, and *f* values and AFP levels were significantly different among well-, moderately, and poorly differentiated HCCs. The four metrics were significantly correlated with histologic differentiation. The area under the ROC curve (AUC-ROC) of ADC, *D*, *f*, and AFP for diagnosing well-differentiated HCCs was 0.903, 0.84, 0.782, and 0.806, respectively, and the AUC-ROC of above metrics for diagnosing poorly differenti-ated HCCs was 0.787, 0.726, 0.624, and 0.633, respectively. The combination of ADC and AFP provided an AUC-ROC of 0.945 for well-differentiated HCC. However, this did not provide better performance for diagnosing poorly differentiated HCC.
Conclusion: ADC, IVIM metrics, and AFP levels may be useful for evaluating histologic differentiation of HBV-related HCCs, and the combination of ADC and AFP provides better diagnostic performance for well-differentiated HCC.

Key words: IVIM—AFP—HBV—HCC—Histologic differentiation

Hepatocellular carcinoma (HCC) is the most common primary malignancy of the liver. Histologic differentiation is an important factor affecting patient prognosis. Large tumor size, capsule formation, satellite lesions, and microvascular invasion are more frequently found in poorly differentiated HCCs compared with well- and moderately differentiated HCCs. Surgical resection and liver transplantation have been considered the most effective treatments, but suffer from a high recurrence rate, particularly with poorly differentiated lesions. Therefore, evaluation of histologic differentiation is helpful for selecting those HCC lesions that are more amenable to treatment [1–3]. Alpha-fetoprotein (AFP) is considered an independent prognostic factor which is correlated with histologic differentiation [4, 5]. Diffusion-weighted magnetic resonance imaging (DWI) could

Correspondence to: Jin Wang; *email:* wangjin3@mail.sysu.edu.cn

be used to characterize the diffusion process of molecules in vivo. It has been used as a diagnostic tool for the detection of tumor and the differentiation of benign and malignant lesions. Several previous studies have used the DWI calculated apparent diffusion coefficient (ADC) to predict the histologic differentiation of HCCs before surgery but the results are controversial [6–11]. Although the field strength of MR scanners used in these studies was 1.5 T in all cases, different b values ranging from 0 to 1000 s/mm^2 were used. In addition, ADC could not reflect the pure molecular diffusion because of the "pseudodiffusion" effect generated by blood flow, which results in an overestimation of ADC values [12].

The concept of intravoxel incoherent motion (IVIM) was initially proposed by Le Bihan et al. [12]. Using the IVIM theory, the relationship between signal attenuation and higher b values could be described more accurately with quantitative metrics that separately reflect the molecular diffusion of water and the microcirculation of blood. Respiratory triggered (RT) IVIM-DWI has been shown to generate good and reproducible image quality [13]. According to a recent study [14], pseudodiffusion tended to be underestimated in the liver when too few low b values ($0 < b < 50$ s/mm^2) were included. Hence, it is recommended that at least two low b values should be used when performing liver IVIM-DWI. Woo et al. [15] reported that ADC and pure diffusion coefficient (D) were both inversely correlated with histologic grade and D showed better diagnostic performance when eight b values were used. In the aforementioned studies, HCC patients with various etiologies were included. Among the various causes of HCC, chronic hepatitis B virus (HBV) infection is the main risk factor in most East Asian countries. In China, chronic HBV infection is the principal cause of HCC [16–18]. The tumor characteristics and prognosis of HBV-related HCC are different from those of HCC caused by other factors. Multiple nodules, larger tumors, vascular invasion, and higher serum AFP levels are more frequent in HBV-related HCC compared with HCV-related HCC [19, 20]. Therefore, histological differentiation of HBV-related HCC is needed. To our knowledge, no IVIM studies have evaluated histologic differentiation of HCC groups with a single etiology. It remains unclear whether the ADC and IVIM metrics could discriminate well-, moderately-, and poorly- differentiated HBV-related HCCs. Therefore, this study aimed to investigate ADC and IVIM-derived metrics for histologic differentiation of HBV-related HCCs and to compare and combine the metrics with AFP.

Materials and methods

Study population

The retrospective study was approved by the Institutional Review Board, and written informed consent was waived. After computerized search of the medical data, a total of 156 patients with histologically-proven HCC were included. All patients underwent liver MRI between April 2014 and November 2015 at our institution. We retrospectively assessed the images and clinical data of these patients and 106 chronic HBV patients (94 males and 12 females; age range 28–78 years; mean 50.9 ± 1.1 years; BMI 22.2 ± 0.3) with 109 pathologically confirmed HCCs (size range 14–150 mm; mean 47.9 ± 2.9 mm) were included in the final assessment (Fig. 1). The time interval between MR exam and surgery ranged from 1 to 28 days (mean 8.7 days). Fifty patients were excluded for the following reasons: (a) the needle-biopsy results may not truly reflect the accurate histologic HCC grade because HCC is heterogeneous and often contains different histologic grades ($n = 5$) [10], (b) previous transcatheter arterial chemoembolization ($n = 18$), (c) underlying risk factors other than HBV infection alone ($n = 5$), (d) the time interval between the MR examination and surgery was greater than 1 month ($n = 1$), (e) the lesion size was below 1 cm
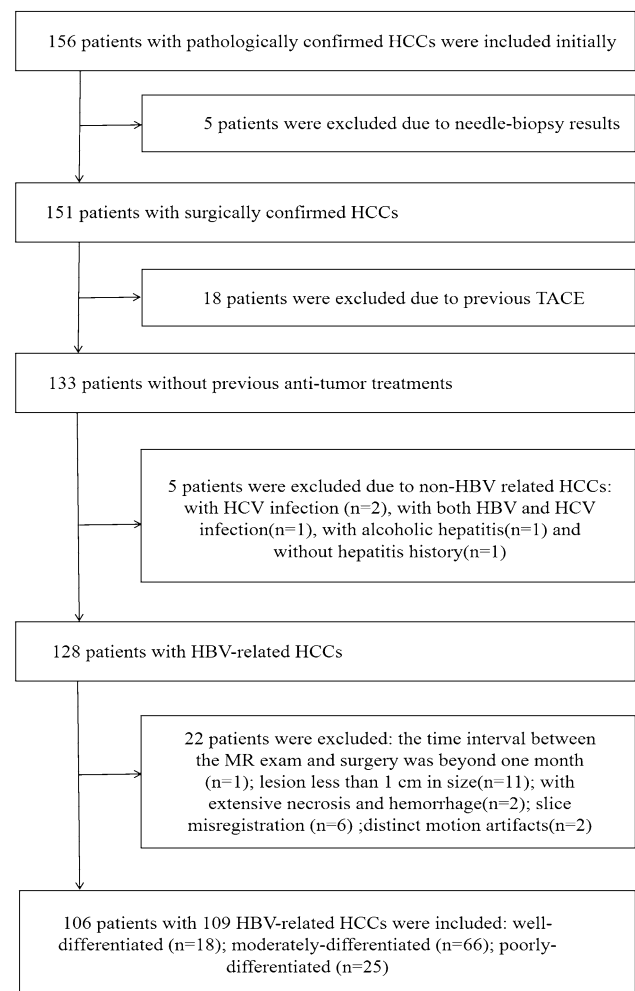

Fig. 1. Flowchart shows the patient selection process.

thereby increasing the possibility of partial volume effects affecting accurate tumor grading [21] ($n = 11$), (f) extensive necrosis and hemorrhage meant that no enough measurable area was present ($n = 2$), (g) slice misregistration or distinct motion artifacts which interfered with image interpretation ($n = 8$).

Liver transplantation was performed in 4 patients and surgical resection in the remaining 102 patients. Three patients had two lesions and 103 patients had one lesion. AFP was positive in 90 patients (ranging from 1 to 1210 ng/mL; mean 236.5 ± 32.5 ng/mL). Among the 106 patients, 94 were diagnosed with cirrhosis based on histological presentation, MRI features, or clinical symptoms. Liver function was assessed according to the Child–Pugh classification: 77 patients were grade A, 27 patients grade B, and 2 patients grade C. The differentiation of HCCs was classified into three groups as follows: well-differentiated ($n = 18$), moderately differentiated ($n = 66$) and poorly differentiated ($n = 25$). Table 1 showed the baseline characteristics of the included patients.

MRI and IVIM scan

All patients underwent MRI on a 3.0 T whole-body MR scanner (Discovery MR750, GE Healthcare, Milwaukee, WI) equipped with eight-channel phased array abdominal coil. All patients were required to fast for at least 4 h

Table 1. Patient' characteristics

Characteristics	Values
Mean age (years)	50.9 ± 11
Gender	
Men	94
Women	12
BMI	22.2 ± 3
Histological differentiation	
Well	18
Moderately	66
Poorly	25
Child–Pugh classification	
Grade A	77
Grade B	27
Grade C	2
Location of tumor	
Right lobe	75
Left lobe	31
Right and left lobe	3
Size of tumor (mm)	47.9 ± 2.9
Surgery	
Resection	102
Liver transplantation	4
Time interval between the MR exam and surgery (days)	8.7 ± 5.8
AFP	
(+)	90 (84.9%)
Well differentiated	12
Moderately differentiated	55
Poorly differentiated	23
(−)	16 (15.1%)
Well differentiated	4
Moderately differentiated	10
Poorly differentiated	2

BMI, body mass index; AFP, alpha fetal protein

before the examination. The patients were positioned in a supine and feet first position, with arms raised to the head. The scanning sequences used were as follows: breath-hold (BH) coronal fast imaging employing steady-state acquisition (FIESTA), BH coronal single-shot fast spin echo (SSFSE), and RT axial T2-weighted FSE. RT axial diffusion-weighted single-shot echo-planar imaging (EPI) was scanned with 11 b values ($b = 0, 30, 50, 100, 150, 200, 300, 500, 800, 1000, 1500$ s/mm^2). Afterwards, breath-holding axial liver acquisitions with volume acceleration (LAVA) were performed before and after the administration of gadobenate dimeglumine (MultiHance; Bracco Imaging, Milan, Italy) at a dose of 0.1 mmol/kg body weight. Contrast agent was injected antecubitally using a power injector (Spectris Solaris EP; Medrad, Warrendale, PA, USA) at a rate of 2.5 mL/s followed by a 20 mL saline flush. Dynamic images were acquired in four phases after contrast administration: dual arterial phase (15–20 s post contrast), dual portal phase (50–55 s post contrast), coronary enhancement phase (1 min and 30–40 s post contrast), and a delayed phase (3 min post contrast). The scanning parameters of all the pulse sequences are listed in Table 2.

Monoexponential and IVIM model of DWI

The ADC value was calculated by fitting all 11 b values to a simplified monoexponential model. The prototype software was provided by the scanner manufacturer (GE Healthcare).

IVIM data were analyzed with a nonlinear bi-exponential fitting model using all the b values as described by Le Bihan et al. [12]. The relation between diffusion-weighted signal and b values was expressed as follows:

$$S_b/S_0 = (1 - f) \exp(-bD) + f \exp(-bD^*),$$

where D (diffusion coefficient) is the diffusion parameter representing pure molecular diffusion, D^* (pseudodiffusion coefficient) is the perfusion-related diffusion parameter representing incoherent microcirculation within the voxel, and f (perfusion fraction) is the fraction of the diffusion linked to microcirculation.

Image analysis

All images were analyzed by an abdominal radiologist (J.W. with 23 years of experience of liver MRI) on a workstation of the picture archiving and communication system. HCC was diagnosed based on the diagnostic criterion proposed by the American Association for the Study of Liver Disease (AASLD) in 2010 [22], i.e., wash-in in the arterial phase and washout in the delayed phase. All measurements were performed in consensus using vendor provided software (Functool on GE Advantage Workstation 4.6, GE Healthcare) by two trained radi-

Table 2. Scanning parameters

Sequences	TR/TE (ms)	Flip angle (degree)	Matrix size	Field of view (cm)	Bandwidth (kHz)	Slice thickness (mm)	Slice gap (mm)	Acquisition time	NEX acceleration factor
FIESTA	3.1/1.1	45	224×256	38×38	125	8	1	10 s	1
SSFSE	1222/80.9	90	384×160	44×44	83.3	8	0	21 s	1
T2WI	6000/72.8	110	320×320	36×36	83.3	5	1	3 min 36 s	2
DW single-shot EPI	9231/56	90	128×128	38×30	250	5	1	3 min 52 s	1
LAVA-Flex plain scan	3.7/1.7	15	260×224	36×36	1200	5	1	14 s	1
LAVA-Flex dynamic	3.7/1.7	12	260×224	36×36	1200	5	1	–	1

TE, time of echo; TR, time of repetition; FIESTA, fast imaging employing steady-state acquisition; SSFSE, single-shot fast spin echo; DW, diffusion-weighted; EPI, echo-planar imaging; LAVA, liver acquisitions with volume acceleration

ologists (Q.G.S. with 3 years and R.H.Y. with 13 years of experience of liver MRI) who were blinded to the histological results, and further checked by the study coordinator J.W. After reviewing the DWI, T2WI and dynamic enhancement images, regions of interest (ROIs) were manually drawn on the axial b_{1000} images to encompass as much lesion body as possible on the slice with the maximum tumor cross-section. ROIs were placed at least 5 mm away from the margin of the tumor to minimize partial volume effects [21]. Necrosis and hemorrhage were avoided by referring to T2WI images and dynamic enhancement on T1WI images [15, 23]. All the ROIs were transferred to ADC and IVIM-derived maps for measurement (Fig. 2). The area of ROIs ranged from 73 to 11,720 mm^2 (mean area 1889 mm^2).

Histological analysis

All surgical specimens were analyzed by a pathologist with 10 years of liver experience. All HCCs were assigned to well-, moderately, and poorly differentiated groups. When different differentiations coexisted within a tumor, the most predominant differentiation of the tumor was selected [10].

Statistical analysis

ADC, D, D^*, f, and AFP values of different histologically differentiated groups were compared using Kruskal–Wallis test followed by post hoc multiple comparisons with Bonferroni correction. Spearman's rank correlation was used to assess the statistical dependence among the four metrics of HCCs with various histologic differentiations. Receiver operating characteristic (ROC) analysis was performed to evaluate diagnostic performance of ADC, D, D^*, f, and AFP in differentiating among HCCs with three histologic grades. The appropriate cut-off values were determined by choosing the point with the largest Yoden Index. The areas under the ROC curve were compared using the Delong et al. method [24]. For joint analysis of multiple parameters, we used a binary logistic regression model with stepwise regression to select an optimal combination. All the statistical analyses were

performed using SPSS 22.0 (IBM Corp., USA) and MedCalc13.0 (MedCalc Software bvba, Ostend, Belgium). $P < 0.05$ was considered statistically significant in all statistical analyses.

Results

Distribution of AFP_+ and AFP_- patients of different histologic grades

In our study, 90 patients were AFP_+ with a positive rate of 84.9%. The number of well-, moderately and poorly differentiated HCC patients was 12, 55, and 23, respectively, in the AFP_+ group and 4, 10, and 2, respectively, in the AFP_- group.

ADC, IVIM and AFP values among different groups

ADC, IVIM, and AFP values of well-, moderately, and poorly differentiated HCCs were summarized in Table 3. ADC, D, f and AFP values were significantly different among all three groups ($P < 0.001$). ADC, D, and f values of well-differentiated HCCs were significantly higher than those of moderately ($P < 0.001$, <0.001 and 0.002, respectively) and poorly differentiated HCCs ($P < 0.001$, <0.001 and 0.001, respectively). The ADC value of moderately differentiated HCCs was significantly higher than that of poorly differentiated HCCs ($P = 0.005$), but no significant difference was found in D and f values between moderately differentiated HCCs and poorly differentiated HCCs ($P = 0.05$ and 0.96, respectively). AFP levels of well-differentiated HCCs were significantly lower than those of moderately and poorly differentiated HCCs ($P < 0.001$), but no significant difference was found between moderately and poorly differentiated HCCs ($P = 0.755$). There was no significant difference in D^* value among the three groups ($P = 0.458$).

Correlation of histologic differentiation with ADC, IVIM metrics and AFP levels

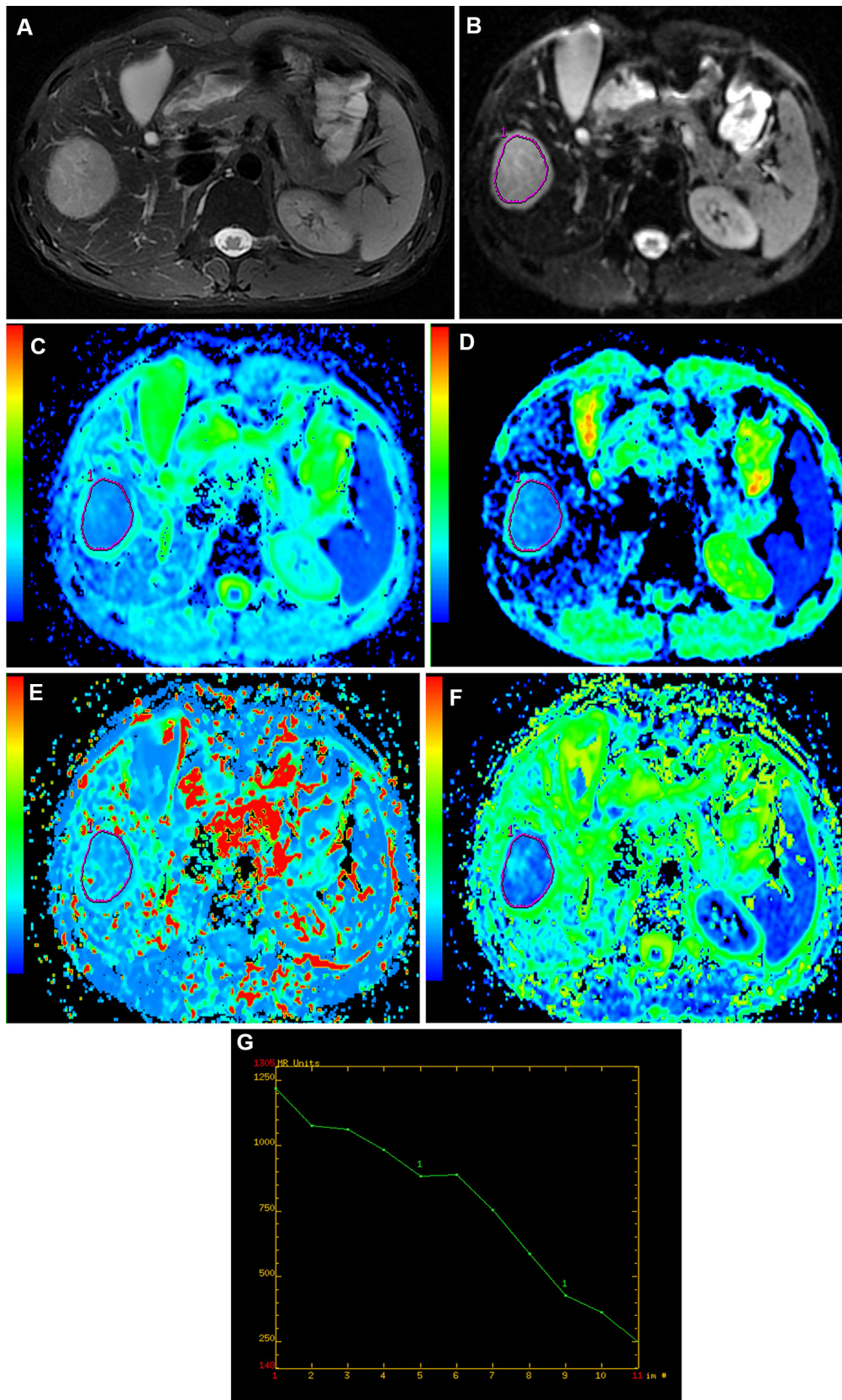ADC, D, and f all decreased and AFP increased with the grade of histologic differentiation. The four metrics were

Fig. 2. A surgically confirmed moderately differentiated HCC in a 31-year-old man. **A** T2WI image, **B** diffusion-weighted image with $b = 0$, **C–F** ADC, D, D^*, and f map, and **G** signal intensities vs. b-values. The mean values of ADC, D, D^*, and f of the tumor were $0.91 \times 10{-}3$ mm^2/s, $0.80 \times 10{-}3$ mm^2/s, $17.6 \times 10{-}3$ mm^2/s, and 21.6%, respectively.

Table 3. ADC, IVIM metrics and AFP levels of different histologically differentiated HBV-related HCCs

Parameters	Well-differentiated ($n = 18$)	Moderately differentiated ($n = 66$)	Poorly differentiated ($n = 25$)	P
ADC ($\times 10^{-3}$ mm^2/s)	1.06 (0.96–1.42)	0.88 (0.64–1.50)	0.80 (0.54–1.17)	<0.001
D ($\times 10^{-3}$ mm^2/s)	0.96 (0.76–1.26)	0.78 (0.44–1.38)	0.70 (0.50–1.14)	<0.001
D^* ($\times 10^{-3}$ mm^2/s)	14.0 (4.7–34.7)	8.46 (4.3–46.2)	9.35 (4.2–252.0)	0.458
f (%)	26.7 (12.1–39.9)	21.5 (9.9–53.2)	20.6 (10.3–28.4)	<0.001
AFP (ng/mL)	8.93 (1–466.9)	468.85 (1.05–1210)	486.30 (2.20–1210)	<0.001

Numbers in parentheses are ranges
* Data are median values

significantly correlated with histologic differentiation: $r = -0.586$ ($P < 0.001$), $r = -0.478$ ($P < 0.001$), $r = -0.332$ ($P < 0.001$) and $r = 0.36$ ($P < 0.001$), respectively. However, there was no correlation between histologic grade and D^* values ($r = -0.103$, $P = 0.287$). ADC and D were significantly correlated with AFP levels: $r = -0.339$ ($P < 0.001$) and $r = -0.224$ ($P = 0.019$), respectively.

## ROC analysis for discriminating different histologic subtypes

According to ROC analysis, an ADC value of 1.01×10^{-3} mm^2/s, a D value of 0.90×10^{-3} mm^2/s, a f value of 0.217, and an AFP value of 9.8 ng/mL were found to be the most accurate cut-off levels for discriminating well differentiated HCCs from moderately and poorly differentiated HCCs, with a sensitivity and specificity of 94.4% and 85.7% for ADC, 77.8% and 84.6% for D, 94.4% and 57.1% for f, and 72.2% and 83.5% for AFP, respectively. The area under the ROC curve (AUC-ROC) of ADC, D, D^*, f and AFP for diagnosing well-differentiated HCCs was 0.903 (range 0.832–0.952), 0.84 (range 0.758–0.903), 0.592 (range 0.493–0.685), 0.782 (range 0.693–0.856), and 0.806 (range 0.72–0.876), respectively. An ADC value of 0.839×10^{-3} mm^2/s and a D value of 0.793×10^{-3} mm^2/s were found to be the most accurate cut-off levels for discriminating poorly differentiated HCCs from well- and moderately differentiated ones, with a sensitivity and specificity of 76% and 73.8% for ADC and 84% and 56% for D, respectively. The AUC-ROC of the ADC, D, D^*, f, and AFP for diagnosing poorly differentiated HCCs was 0.787 (range 0.699–0.860), 0.726 (range 0.633–0.807), 0.535 (range 0.437–0.632), 0.624 (range 0.526–0.715), and 0.633 (range 0.536–0.724), respectively. The AUC-ROC of ADC was greater than that of D for discriminating well- from moderately, and poorly differentiated HCCs ($P = 0.03$), but there was no difference for discriminating poorly from well- and moderately differentiated HCCs ($P = 0.08$). The AUC-ROC of ADC was greater than that of f for discriminating poorly from well- and moderately differentiated HCCs ($P = 0.03$), but there was no difference for discriminating well- from moderately- and poorly-HCCs ($P = 0.07$). No difference was

found between the AUC-ROC of D and f for the diagnosis of well- and poorly differentiated HCCs ($P = 0.5$ and 0.28, respectively). ADC and AFP had a similar diagnostic performance ($P = 0.08$) for well-differentiated HCC, while ADC had a higher diagnostic value for poorly differentiated HCC ($P = 0.04$; Fig. 3). The logistic regression resulted in the selection of ADC and AFP as the optimal set for the diagnosis of well-differentiated HCC. Combining AFP, the AUC-ROC of ADC increased to 0.945 (range 0.898–0.992) which was higher than ADC alone ($P = 0.04$), with a sensitivity and specificity of 94.4% and 91.2%, respectively. Other combinations of parameters did not provide better performance compared with ADC. However, using the same method, the combination of ADC and other parameters did not display better performance in the diagnosis of poorly differentiated HCC compared with ADC alone. The number of HCC classified in each category for each parameter used alone and in combination is shown in Tables 4 and 5.

Discussion

We have shown that ADC, D, f, and AFP values were significantly different among well-, moderately, and poorly differentiated HCCs ($P < 0.001$). The four metrics were significantly correlated with histologic differentiation ($P < 0.001$), and ADC showed the strongest correlation compared with D, f, and AFP ($r = -0.586$, -0.478, -0.332, and 0.36, respectively). Histologic differentiation is one of the most important factors related to the recurrence of HCC [2, 5]. Therefore, evaluation of preoperative histologic differentiation could potentially help in selecting more proper treatment strategies and predicting the prognosis. In China, HBV infection is the most important etiology of HCC. HBV-related HCC is different from HCC caused by other etiologies in tumor characteristics and prognosis. Biomarkers which could predict histological differentiation of HBV-related HCC preoperatively would be clinically important.

In previous studies, AFP was reported to be correlated with histologic differentiation [4], and its levels were higher in HBV-related HCCs compared with HCV-related HCCs [20]. However, the diagnostic value of AFP for differentiation among HBV-related HCCs remains unknown. In our study, the positive rate of AFP was
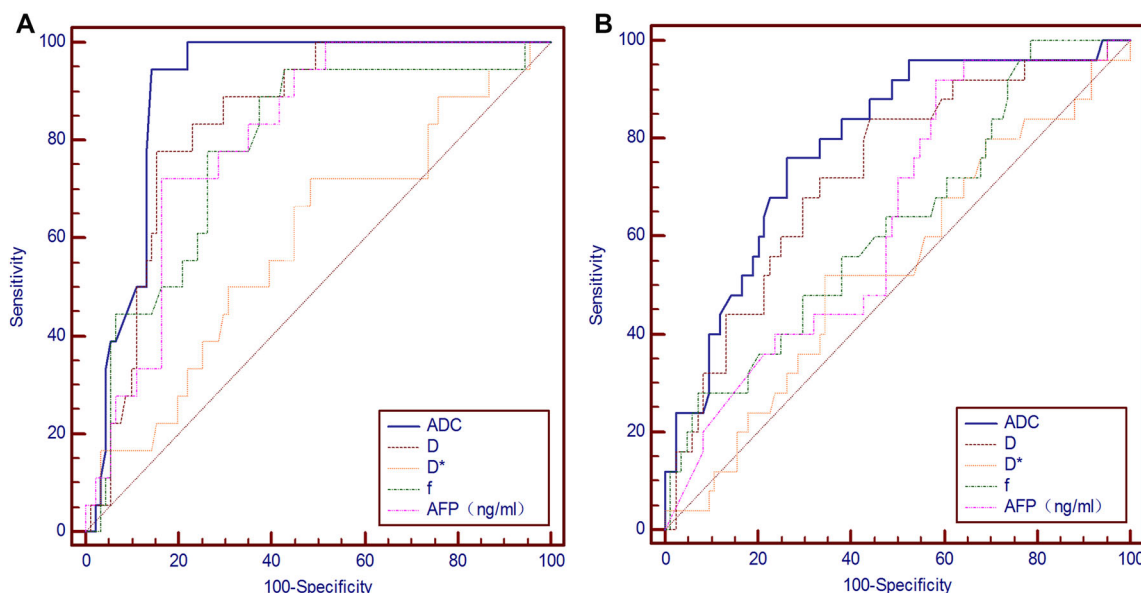
Fig. 3. The comparison of ROC curve analysis for diagnosing well: **A** and poorly differentiated and **B** HCC with ADC, D, f, and AFP values.

Table 4. The number of HCC classified in well and non-well differentiated group for the parameters used alone and in combination

Parameters	Pathological results		Sum
	Well differentiated ($n = 18$)	Non-well differentiated ($n = 91$)	
ADC_+	17	13	30
ADC_-	1	78	79
Sensitivity/specificity (%)	94.4	85.7	
D_+	14	14	28
D_-	4	77	81
Sensitivity/specificity (%)	77.8	84.6	
f_+	17	39	56
f_-	1	52	53
Sensitivity/specificity (%)	94.4	57.1	
AFP_+	13	15	28
AFP_-	5	76	81
Sensitivity/specificity (%)	72.2	83.5	
$(ADC + D)_+$	17	12	29
$(ADC + D)_-$	1	79	80
Sensitivity/specificity (%)	94.4	86.8	
$(ADC + f)_+$	18	9	27
$(ADC + f)_-$	0	82	82
Sensitivity/specificity (%)	100	90.1	
$(ADC + AFP)_+$	17	8	25
$(ADC + AFP)_-$	1	83	84
Sensitivity/specificity (%)	94.4	91.2	
$(ADC + D + f)_+$	17	9	26
$(ADC + D + f)_-$	1	82	83
Sensitivity/specificity (%)	94.4	90.1	
$(ADC + D + AFP)_+$	17	11	28
$(ADC + D + AFP)_-$	1	80	81
Sensitivity/specificity (%)	94.4	87.9	
$(ADC + f + AFP)_+$	17	9	26
$(ADC + f + AFP)_-$	1	82	83
Sensitivity/specificity (%)	94.4	90.1	
$(ADC + D + f + AFP)_+$	16	5	21
$(ADC + D + f + AFP)_-$	2	86	88
Sensitivity/specificity (%)	88.9	94.5	

Table 5. The number of HCC classified in poorly and non-poorly differentiated group for the parameters used alone and in combination

Parameters	Pathological results		Sum
	Poorly differentiated ($n = 25$)	Non-poorly differentiated ($n = 84$)	
ADC_+	19	22	41
ADC_-	6	62	68
Sensitivity/specificity (%)	76	73.8	
D_+	21	37	58
D_-	4	47	51
Sensitivity/specificity (%)	84	56	
$(ADC + D)_+$	18	18	36
$(ADC + D)_-$	7	66	73
Sensitivity/specificity (%)	72	78.6	

84.9%. We found significant correlation of AFP levels with histological differentiation of HBV-related HCCs. Furthermore, ADC showed a similar performance for diagnosing well-differentiated HCCs and a better performance for diagnosing poorly differentiated HCCs, compared with AFP levels.

Several previous studies have reported that ADC values showed an inverse correlation with histologic grade [6–9]. Woo et al. [15] reported that ADC and D values were significantly correlated with histologic grade: $r = -0.448$ ($P = 0.002$) and $r = -0.604$ ($P < 0.0001$), respectively, which was consistent with our results. The increased cellular density and nuclear/cytoplasmic ratio may restrict the diffusion process which causes ADC and D to decrease with histologic grade. Nasu et al. [10] found that the histologic grade of HCCs was not correlated with ADC values,

while Xu et al. [11] found that moderately, and poorly differentiated HCCs had significantly higher ADC values than well-differentiated HCCs. However, in these two studies, ROIs were placed to encompass as much of the tumor as possible which may have resulted in the overestimation of ADC values due to the inclusion of necrotic areas. In addition, only b values of 0 and 500 s/mm^2 were used in the study by Nasu et al. [10], so ADC values obtained might have been overestimated because of the "pseudodiffusion" effect.

Among all the IVIM metrics, f values decreased with increasing histologic grade in our study. Similarly, Lin et al. [25] reported that f values were lower in the high-grade gliomas (HGGs). On the other hand, Woo et al. [15] observed that f values did not show significant correlation with histologic grade and it decreased with the histologic grade ranging from grade 2 to grade 4, but the differences between f values among the four histologic grades were not significant ($P = 0.172$). Several possible reasons may explain these findings. Firstly, a previous study by Lemke et al. [26] found that longer echo time could result in a greater f value. In our study, the f values of HCCs might have been overestimated, and the difference in f values among HCCs with different histologic grades might be attributed to different T2 values. Secondly, with an increase in histologic grade, tumor capillaries characterized by irregular diameters, abnormal vascular branching pattern, incomplete basal membrane, and leaks begin to dominate, resulting in reduced blood circulation [27, 28]. Apart from the different b values used in DWI and the difference in the underlying diseases, the small sample size of grade 1 ($n = 4$) and grade 2 ($n = 4$) HCCs in Woo's study may have led to the discrepancy with our results.

According to the ROC analysis in our study, ADC values showed a better performance for diagnosing well-differentiated HCCs and a non-inferior performance for diagnosing poorly differentiated HCCs compared with D value, and it showed a better performance for diagnosing poorly differentiated HCCs and a non-inferior performance for diagnosing well-differentiated HCCs compared with f value. Lin et al. [25] reported that ADC values showed better efficacy in discriminating HGG from low-grade gliomas than D and f values, which was consistent with our findings. However, Woo et al. [15] showed that D values of HCCs showed significantly better diagnostic performance than ADC values in differentiating high-grade HCCs from low-grade HCCs. The discrepancy could have several possible causes. In this study, the significant decrease of both D and f values with the higher histologic grade indicated restricted diffusion and decreased microcirculation, which could both reduce ADC values. Therefore, ADC value as a composite metric which could reflect both pure diffusion and microcirculation may perform better in discriminating HCCs of different histologic grades than D and f values.

However, the efficacy of ADC and IVIM metrics requires further study. Using joint analysis of multiple parameters, we found that ADC values in combination with AFP levels could provide better diagnostic performance with increased specificity for well-differentiated HBV-related HCCs, which might potentially provide accurate prediction of prognosis. However, for the diagnosis of poorly differentiated HCC, the combination multiple parameters did not display better performance compared with ADC alone.

We found that there was no correlation between histologic differentiation and D^* values. This may be related to the poor reproducibility of D^* [13, 14, 29, 30]. In our study, the standard deviation of the D^* value was close to or even larger than its mean value, which indicated that it fluctuated greatly and may not be a reliable metric. Further improvement in the reproducibility of D^* is warranted.

No clear consensus on the optimal DWI protocol has been reached. As previous studies reported, RT technique has some advantages. Dyvorne et al. [13] found that RT IVIM-DWI could produce good image quality and reproducibility. In their study, RT sequences showed higher image quality than free-breathing (FB) sequences which were prone to respiratory motion artifact. Kandpal et al. [31] suggested that RT sequences yield better image quality and higher SNR compared with BH sequences. However, RT sequences also have some disadvantages. ADC values obtained from RT sequences are less reproducible than those from BH and FB sequences [32, 33]. RT acquisition takes more time, and there is a risk of pseudo-anisotropy artifact originating from respiratory motion which makes ADC values inaccurate [34].

The choice of number and distribution of b values could affect IVIM-derived metrics. For RT sequences, increasing the number of b values could improve reproducibility [32]. The estimation error of all IVIM metrics increases if fewer b values are used compared with the 16 b value acquisition as the reference standard [35]. According to a recent study [14], pseudodiffusion metrics tend to be underestimated in the liver when too few low b values ($0 < b < 50$ s/mm^2) are included, so it is recommended that at least two low b values should be used for liver IVIM-DWI. Therefore, we used 11 b values including 2 low b values in our study.

There are some limitations in our study. Firstly, the study was retrospective and there might be selection bias. Secondly, HCCs were inhomogeneous and in some cases various differentiations coexisted within a tumor. Although the ROIs were placed to encompass as much lesion body as possible to characterize the whole lesion, it was still an estimation and could not reflect the heterogeneity of HCCs. Histogram analysis would be helpful but was not performed because the software was not available when we started this study. Thirdly, the mea-

sured f values might be overestimated because of its dependency on time of echo (TE). We did not calculate T2 values in our study, and T2 correction, which may help correct f values, was not performed [26]. The effect of TE on the f value of HCCs with different histologic grades needs to be further investigated. Fourthly, in addition to pure diffusion and perfusion-related diffusion, a third diffusion compartment might exist, which might not be adequately described by a bi-exponential model [36]. Studies using a tri-exponential model may be needed. Fifthly, we did not assess the reproducibility of ADC and IVIM metrics. Lastly, RT sequences have limitations, as mentioned above.

In conclusion, our study showed that ADC, D, and f values derived from multiple b values DWI and serum AFP levels were significantly different among HBV-related HCCs with different histologic grades. ADC values demonstrate the best diagnostic efficiency, while the combination of ADC and AFP provide better diagnostic performance for the diagnosis of well-differentiated HCC.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. For this type of study, formal consent is not required.

Informed consent The Institutional Review Board approved the study and did not require additional informed consent for reviewing the patients' medical records and image.

References

1. Oishi K, Itamoto T, Amano H, et al. (2007) Clinicopathologic features of poorly differentiated hepatocellular carcinoma. J Surg Oncol 95:311–316
2. Kaibori M, Ishizaki M, Matsui K, Kwon A (2010) Predictors of microvascular invasion before hepatectomy for hepatocellular carcinoma. J Surg Oncol 102:462–468
3. Okusaka T, Okada S, Ueno H, et al. (2002) Satellite lesions in patients with small hepatocellular carcinoma with reference to clinicopathologic features. Cancer 95:1931–1937
4. Duvoux C, Roudot Thoraval F, Decaens T, et al. (2012) Liver transplantation for hepatocellular carcinoma: a model including α-fetoprotein improves the performance of Milan criteria. Gastroenterology 143:986–994
5. Yaprak O, Akyildiz M, Dayangac M, et al. (2012) AFP level and histologic differentiation predict the survival of patients with liver transplantation for hepatocellular carcinoma. Hepatobiliary Pancreat Dis Int 11:256–261
6. Muhi A, Ichikawa T, Motosugi U, et al. (2009) High-b-value diffusion-weighted MR imaging of hepatocellular lesions: estimation of grade of malignancy of hepatocellular carcinoma. J Magn Reson Imaging 30:1005–1011
7. Heo SH, Jeong YY, Shin SS, et al. (2010) Apparent diffusion coefficient value of diffusion-weighted imaging for hepatocellular carcinoma: correlation with the histologic differentiation and the expression of vascular endothelial growth factor. Korean J Radiol 11:295–303
8. Nishie A, Tajima T, Asayama Y, et al. (2011) Diagnostic performance of apparent diffusion coefficient for predicting histological grade of hepatocellular carcinoma. Eur J Radiol 80:e29–e33
9. Nakanishi M, Chuma M, Hige S, et al. (2012) Relationship between diffusion-weighted magnetic resonance imaging and histological tumor grading of hepatocellular carcinoma. Ann Surg Oncol 19:1302–1309
10. Nasu K, Kuroki Y, Tsukamoto T, et al. (2009) Diffusion-weighted imaging of surgically resected hepatocellular carcinoma: imaging characteristics and relationship among signal intensity, apparent diffusion coefficient, and histopathologic grade. Am J Roentgenol 193:438–444
11. Xu H, Li X, Xie J, Yang Z, Wang B (2007) Diffusion-weighted magnetic resonance imaging of focal hepatic nodules in an experimental hepatocellular carcinoma rat model. Acad Radiol 14:279–286
12. Le Bihan D, Breton E, Lallemand D, Aubin M (1988) Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology 168:497–505
13. Dyvorne HA, Galea N, Nevers T, Fiel MI (2013) Diffusion-weighted imaging of the liver with multiple b values: effect of diffusion gradient polarity and breathing acquisition on image quality and intravoxel incoherent motion parameters—a pilot study. Radiology 266:920–929
14. Cohen AD, Schieke MC, Hohenwalter MD, Schmainda KM (2015) The effect of low b-values on the intravoxel incoherent motion derived pseudodiffusion parameter in liver. Magn Reson Med 73:306–311
15. Woo S, Lee JM, Yoon JH, et al. (2014) Intravoxel incoherent motion diffusion-weighted MR imaging of hepatocellular carcinoma: correlation with enhancement degree and histologic grade. Radiology 270:758–767
16. de Martel C, Maucort-Boulch D, Plummer M, Franceschi S (2015) World-wide relative contribution of hepatitis B and C viruses in hepatocellular carcinoma. Hepatology 62:1190–1200
17. Ashtari S (2015) Hepatocellular carcinoma in Asia: prevention strategy and planning. World J Hepatol 7:1708
18. Yang JD, Roberts LR (2010) Epidemiology and management of hepatocellular carcinoma. Infect Dis Clin N Am 24:899–919
19. Zhou Y, Si X, Wu L, et al. (2011) Influence of viral hepatitis status on prognosis in patients undergoing hepatic resection for hepatocellular carcinoma: a meta-analysis of observational studies. World J Surg Oncol 9:108
20. Kao W, Su C, Chau G, et al. (2011) A comparison of prognosis between patients with hepatitis B and C virus-related hepatocellular carcinoma undergoing resection surgery. World J Surg 35:858–867
21. Kim SY, Lee SS, Byun JH, Park SH (2010) Malignant hepatic tumors: short-term reproducibility of apparent diffusion coefficients with breath-hold and respiratory-triggered diffusion weighted MR imaging. Radiology 255:815–823
22. Bruix J, Sherman M (2011) Management of hepatocellular carcinoma: an update. Hepatology 53:1020–1022
23. Wagner M, Doblas S, Daire J, Paradis V (2012) Diffusion-weighted MR imaging for the regional characterization of liver tumors. Radiology 264:464–472
24. Delong ER, Delong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845
25. Lin Y, Li J, Zhang Z, et al. (2015) Comparison of intravoxel incoherent motion diffusion-weighted MR imaging and arterial spin labeling MR imaging in gliomas. BioMed Res Int 2015:1–10
26. Lemke A, Laun FB, Simon D, Stieltjes B, Schad LR (2010) An in vivo verification of the intravoxel incoherent motion effect in diffusion-weighted imaging of the abdomen. Magn Reson Med 64:1580–1585
27. Semela D, Dufour J (2004) Angiogenesis and hepatocellular carcinoma. J Hepatol 41:864–880
28. Lewin M, Fartoux L, Vignaud A, et al. (2011) The diffusion-weighted imaging perfusion fraction f is a potential marker of sorafenib treatment in advanced hepatocellular carcinoma: a pilot study. Eur Radiol 21:281–290

29. Kakite S, Dyvorne H, Besa C, et al. (2015) Hepatocellular carcinoma: short-term reproducibility of apparent diffusion coefficient and intravoxel incoherent motion parameters at 3.0 T. J Magn Reson Imaging 41:149–156

30. Andreou A, Koh DM, Collins DJ, et al. (2013) Measurement reproducibility of perfusion fraction and pseudodiffusion coefficient derived by intravoxel incoherent motion diffusion-weighted MR imaging in normal liver and metastases. Eur Radiol 23:428–434

31. Kandpal H, Sharma R, Madhusudhan KS, Kapoor KS (2009) Respiratory-triggered versus breath-hold diffusion-weighted MRI of liver lesions: comparison of image quality and apparent diffusion coefficient values. Am J Roentgenol 192:915–922

32. Kim SY, Lee SS, Park B, Kim N (2012) Reproducibility of measurement of apparent diffusion coefficients of malignant hepatic tumors: effect of DWI techniques and calculation methods. J Magn Reson Imaging 36:1131–1138

33. Kwee TC, Takahara T, Koh D, Nievelstein RAJ, Luijten PR (2008) Comparison and reproducibility of ADC measurements in breathhold, respiratory triggered, and free-breathing diffusion-weighted MR imaging of the liver. J Magn Reson Imaging 28:1141–1148

34. Nasu K, Kuroki Y, Fujii H, Minami M (2007) Hepatic pseudo-anisotropy: a specific artifact in hepatic diffusion-weighted images obtained with respiratory triggering. Magn Reson Mater Phys 20:205–211

35. Dyvorne H, Jajamovich G, Kakite S, Kuehn B, Taouli B (2014) Intravoxel incoherent motion diffusion imaging of the liver: optimal b-value subsampling and impact on parameter precision and reproducibility. Eur J Radiol 83:2109–2113

36. Cercueil J, Petit J, Nougaret S, et al. (2015) Intravoxel incoherent motion diffusion-weighted imaging in the liver: comparison of mono-, bi- and tri-exponential modelling at 3.0-T. Eur Radiol 25:1541–1550