Can preoperative diffusion-weighted MRI predict postoperative hepatic insufficiency after curative resection of HBV-related hepatocellular carcinoma? A pilot study

Seung Up Kim\textsuperscript{a,g}, Young Chul Kim\textsuperscript{c,1}, Ji Soo Choi\textsuperscript{c,1}, Kyung Sik Kim\textsuperscript{d,g}, Gi Hong Choi\textsuperscript{d,g}, Jin Sub Choi\textsuperscript{d,g}, Jun Yong Park\textsuperscript{a,b,g}, Do Young Kim\textsuperscript{a,b,g}, Sang Hoon Ahn\textsuperscript{a,b,g}, Eun Hee Choi\textsuperscript{e}, Young Nyun Park\textsuperscript{f,g}, Chae Yoon Chon\textsuperscript{a,b,g}, Kwang-Hyub Han\textsuperscript{a,b,g}, Myeong-Jin Kim\textsuperscript{b,c,*}

\textsuperscript{a}Department of Internal Medicine, Yonsei University College of Medicine, Seoul 120-752, South Korea
\textsuperscript{b}Institute of Gastroenterology, Yonsei University College of Medicine, Seoul 120-752, South Korea
\textsuperscript{c}Department of Radiology, Yonsei University College of Medicine, Seoul 120-752, South Korea
\textsuperscript{d}Department of Surgery, Yonsei University College of Medicine, Seoul 120-752, South Korea
\textsuperscript{e}Department of Biostatistics, Yonsei University College of Medicine, Seoul 120-752, South Korea
\textsuperscript{f}Department of Pathology, Yonsei University College of Medicine, Seoul 120-752, South Korea
\textsuperscript{g}Liver Cirrhosis Clinical Research Center, Seoul 120-752, South Korea

Received 29 August 2009; revised 30 December 2009; accepted 5 March 2010

Abstract

Liver fibrosis determines the functional liver reserve. Several studies have reported that the apparent diffusion coefficient (ADC) values of diffusion-weighted magnetic resonance imaging (DW-MRI) can assess liver fibrosis. We investigated whether DW-MRI predicts postoperative hepatic insufficiency and liver fibrosis in patients with hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC). Twenty-six patients with HBV-related HCC who received preoperative DW-MRI on a 3-T MRI system were enrolled between July and December 2008. ADC values were measured twice by two observers. Three "b values" were used: 50, 400 and 800 s/mm\textsuperscript{2}. Postoperative hepatic insufficiency was defined as persistent hyperbilirubinemia (total bilirubin level >5 mg/dl for more than 5 days after surgery) or postoperative death without other causes. The mean age (21 men and 5 women) was 51.4 years. Three patients experienced postoperative hepatic insufficiency. Liver stiffness measurement predicted postoperative hepatic insufficiency, advanced fibrosis (F3–4), and cirrhosis significantly [area under the receiving operator characteristic curve (AUROC)=0.942, 0.771 and 0.818, respectively, with \(P=.047, .048\) and .006, respectively]; ADC values of DW-MRI, however, did not (AUROC=0.797, 0.648 and 0.491, respectively, with \(P=.100, .313\) and .938, respectively). Reliability of ADC values between right and left hepatic lobes (\(\rho=0.868\) and \(\rho=0.910\) in the first and second measures of Observer A; \(\rho=0.865\) and \(\rho=0.831\) in the first and second measures of Observer B) was high and the intra- and interobserver reliability (\(\rho=0.958\) in observer A and \(\rho=0.977\) in observer B; \(\rho=0.929\) in the first measure and \(\rho=0.978\) in the second measure between the two observers) were high. All reliability was significant (\(P<.001\)). Our results suggest that DW-MRI on a 3-T MRI system is not suitable for predicting postoperative hepatic insufficiency, advanced liver fibrosis, and cirrhosis in patients with HBV-related HCC, despite significantly high reliability.

© 2010 Elsevier Inc. All rights reserved.

Keywords: Fibrosis; Cirrhosis; Chronic hepatitis B; Hepatitis B virus; Diffusion-weighted magnetic resonance imaging; Apparent diffusion coefficient; Liver stiffness measurement; Transient elastography; Hepatectomy; Hepatocellular carcinoma
1. Introduction

Considerable advances in perioperative intensive care and surgical techniques have significantly decreased the rates of death and complications after liver resection [1–4]. Nevertheless, since many patients still have cirrhosis or other chronic liver disease (CLD), complications including death may follow liver resection. Therefore, preoperative evaluation of liver reserve remains important [5–7].

Since liver fibrosis is one of the most important factors determining the functional liver reserve, preoperative assessment of liver fibrosis is crucial and can help surgeons choose the optimal surgical strategy. Recently, liver stiffness measurement (LSM) using FibroScan has shown a considerable accuracy in assessing liver fibrosis in patients with CLD [8–10]. Accordingly, in our previous study, we hypothesized that preoperative assessment of liver fibrosis using LSM might predict postoperative hepatic insufficiency after liver resection and proved the potential usefulness of LSM as a non-invasive preoperative method for evaluating functional liver reserve [11].

To date, magnetic resonance imaging (MRI) has helped surgeons assess hepatic vascular and biliary anatomy preoperatively which is essential to ensure safe and successful liver resection [12,13]. Furthermore, recent improvements in MRI technology have made it feasible to apply diffusion-weighted MRI (DW-MRI) to abdominal imaging. Although several recent studies have reported that DW-MRI can assess liver fibrosis using apparent diffusion coefficient (ADC) values [14,15], conflicting results have been also reported and data on intra- and interobserver reliability of DW-MRI has not yet been evaluated [16].

Thus, we investigated both whether DW-MRI can predict postoperative hepatic insufficiency when compared to LSM in patients with hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC) undergoing curative liver resection, as well as how well DW-MRI can assess liver fibrosis preoperatively. The reliability of ADC values between the two hepatic lobes as well as the intra- and interobserver reliability of ADC values of DW-MRI were also evaluated.

2. Methods

2.1. Patients

From July and December 2008, 33 patients undergoing curative liver resection for HBV-related HCC at Severance Hospital, Yonsei University College of Medicine (Seoul, Korea) were enrolled in this prospective and blinded pilot study.

Patients were excluded if they met any of the following criteria: infection with other non-HBV hepatitis viruses; a greater than 5-year history of alcohol ingestion in excess of 40 g/day for men and 20g/day for women before surgery (n=1); comorbidities associated with HBV, such as nonalcoholic steatohepatitis, primary sclerosing cholangitis, and primary biliary cirrhosis (n=1); undergoing liver transplantation (n=2); Child-Pugh class C liver function; unreliable LSM value (n=1) or ongoing or previous history of antiviral treatment (n=2). With these criteria, seven patients were excluded and 26 patients were enrolled for statistical analysis. The preoperative diagnosis of HCC was made based on dynamic imaging studies, biopsy, and alpha-fetoprotein serology according to American Association for the Study of Liver Diseases guideline [17,18]. The typical vascular pattern on dynamic imaging studies meant that the lesion was hypervascular in the arterial phase, and washed out in the portal/venous phase [18].

Besides the demographic data, the following laboratory parameters were also collected from all the patients at the time of DW-MRI and LSM: hemoglobin, platelet count, aspartate aminotransferase (AST), alanine aminotransferase (ALT), albumin, total bilirubin and prothrombin time [international normalized ratio (INR)]. Hepatitis B surface antigen, hepatitis B e antigen (HBeAg) and antibodies were measured using standard enzyme-linked immunosorbent assays (Abbott Diagnostics, Abbott Park, IL, USA). HBV DNA levels were quantified with the COBAS TaqMan 48 analyzer (Roche Molecular Systems, Branchburg, NJ, USA) polymerase chain reaction assay with a 300 copies/ml lower detection limit.

The mean interval from the diagnosis of HCC to surgery and from DW-MRI and LSM examination to surgery was 7.2±1.2 and 4.2±1.1 days, respectively. The primary end point was the development of postoperative hepatic insufficiency. Postoperative hepatic insufficiency was defined as persistent hyperbilirubinemia (total bilirubin level >5 mg/dl for more than 5 days after surgery) or postoperative death without other causes [19,20]. The secondary end points were the performance of ADC values in the assessment of liver fibrosis, the reliability of ADC values between the two hepatic lobes, and the intra- and interobserver reliability of ADC values.

Written informed consent was obtained from all patients. This study was approved by the independent institutional review board of Severance Hospital and conformed to the ethical guidelines of the 1975 Helsinki declaration.

2.2. Measurement of indocyanine green retention at 15 min

After an overnight fast, 0.5 mg/kg of indocyanine green (ICG) was administered intravenously. Blood samples were drawn at 5, 10 and 15 minutes and the plasma ICG concentration was measured spectrophotometrically (710 nm). The plasma retention rate at 15 min (ICG R15, %) and the plasma disappearance rate (ICG−k, min−1) were calculated [11].

2.3. Liver stiffness measurement

LSM was performed preoperatively on the same days as DW-MRI according to the methods previously described [21,22]. Briefly, the tip of the transducer probe is placed on
the skin, between the ribs, at the level of the right liver lobe. Once the target area has been located, acquisition is triggered by pressing a button. The measurement depth is between 25 and 65 mm below the skin surface. Results were expressed in kilopascals. The interquartile range (IQR) was defined as an index of intrinsic variability in LSM corresponding to the interval around the LSM result containing 50% of the valid measurements between the 25th and 75th percentiles.

The median value was considered to be representative of the elastic modulus of the liver. The success rate was calculated as the number of validated measurements divided by the total number of measurements. In this study, only the procedures with at least ten validated measurements and a success rate of at least 60% were considered reliable. In addition, the median value of successful measurements was considered to be representative of the LSM value in a given patient only if the IQR of all validated measurements was less than 30% of the median value. Accordingly, any LSM value which did not satisfy the above conditions was considered unreliable and excluded from further analysis.

2.4. Non-invasive models for assessing liver fibrosis

The age–platelet index (API), AST to platelet ratio index (APRI), and age–spleen–platelet ratio index (ASPRI) were also evaluated preoperatively for comparison with DW-MRI and LSM in the assessment of liver fibrosis. The API, APRI and ASPRI were calculated according to formulas defined by Poynard et al., Wai et al. and Kim et al., respectively [23–25]. The upper normal limit of AST and ALT were 34 and 40 IU/ml, respectively. Spleen size was measured using ultrasonography and was defined as the greatest longitudinal dimension using electronic calipers on the image monitor [26].

2.5. Diffusion-weighted magnetic resonance imaging

MRI was performed preoperatively on the same day as LSM using a 3-T MRI system (MAGNETOM Trio Tim, Siemens, Erlangen, Germany) with a peak gradient amplitude of 45 mT/m and a slew rate of 200 T/m per second. A six-element body phased-array coil was placed on the anterior side of the patient with 6 elements of the spine coil integrated into the scanner table on the posterior side of the patient. Single-shot echoplanar diffusion-weighted imaging was synchronized with the patient’s respiration by means of the PACE (Prospective Acquisition Correction) navigator-triggered technique and was performed on all patients with three $b$ values to measure ADC values accurately: 50, 400 and 800 s/mm². Parallel imaging [the GRAPPA (Generalized Autocalibrating Partially Parallel Acquisition) technique with an iPAT (integrated Parallel Acquisition Techniques) factor of 2] was used to reduce both the susceptibility artifacts and acquisition time within the expiratory period. A frequency-selective fat saturation was additionally applied to cancel chemical shift artifacts. The MRI sequence parameters were as follows: a repetition time of one respiratory interval, echo time of 70 ms, 192×108 matrix size, two signal averages, 36–40-cm field of view, 25 slices with a 5-mm slice thickness and a band width of 1735 Hz/pixel. Although the DW-MRI protocol in this study was comparable to that of a previous study [14], imaging in this study was performed with a 3-T MRI system. The signal acquisition time was 3 min and 36 s, and the total scan time was varied depending on the respiratory period of the patients.

2.6. Measurement of apparent diffusion coefficient values by two observers

The diffusion gradients were applied in three orthogonal directions, and quantitative ADC maps were automatically calculated on a voxel-by-voxel basis with commercially available software on a workstation (Leonardo release VB15; Siemens, Erlangen, Germany). The mean ADC values (ADC-Rt reflects the right lobe and ADC-Lt, the left lobe) were calculated twice preoperatively and at the end of study by two observer from 5 values generated from 5 separate 0.5-cm diameter circular regions of interest (ROIs) placed randomly over the right and left hepatic lobes on 5 different slices from the ADC mapping images (Fig. 1). The R1-ADC value was defined as the representative mean ADC value generated from ten ADC values (five ADC values from the right hepatic lobe and another five from the left) from each observer at each measure and R2-ADC was defined as the representative mean ADC value generated from four R1-ADC values in a given patient (Fig. 1). Finally, R2-ADC was used for further analysis to compare the performance of DW-MRI to that of ICG R15, LSM, and other non-invasive models. Y.C. Kim, M.D. (second-year radiology fellow, Observer A) and J.S. Choi, M.D. (fourth-year radiology resident, Observer B) measured ADC values and were blind to this study protocol. Care was taken to exclude vessels and liver mass from the ROIs.

2.7. Liver resection

Each patient underwent abdominal ultrasonography, computed tomography and MRI including diffusion-weighted images, hepatic angiography and positron emission tomography to confirm the number, size, location and extent of the tumor as well as the existence of distant metastases. In addition to preoperative routine laboratory examinations and physical examination for determining Child-Pugh classification, ICG R15 was performed to evaluate liver reserve function.

All liver resection were performed by three surgeons (G. H. Choi, K.S. Kim and J.S. Choi). Indications for liver resection and the types of operative procedures were mainly determined based on the Makuuchi criteria, i.e., the presence or absence of ascites, the serum total bilirubin level, and ICG R15 [27]. During surgery, intraoperative ultrasonography was routinely performed for all patients to determine tumor localization and extent and to exclude the presence of
additional lesions in the residual liver. Operation time and bleeding amount were also recorded.

2.8. Extracted liver specimens

All extracted liver specimens were evaluated by an experienced hepatopathologist (YN Park) who was blinded to patients’ clinical histories and the study protocol. Liver histology was evaluated semiquantitatively according to the Metavir scoring system. Fibrosis was staged on a 0–4 scale: F0, no fibrosis; F1, portal fibrosis without septa; F2, portal fibrosis and a few septa; F3, numerous septa without cirrhosis and F4, cirrhosis. Activity was graded as A0, none; A1, mild; A2, moderate and A3, severe activity.

2.9. Statistical analysis

Patient characteristics are described as mean±S.D. or median (range). The independent t test, Mann–Whitney U test and Fisher’s exact test were used to compare baseline characteristics. The Pearson correlation coefficient (ρ) was used for evaluating the reliability of ADC values. The ability of
ADC values, LSM and ICG R15 to predict postoperative hepatic insufficiency and liver fibrosis was assessed by means of the receiver operating characteristic (ROC) curve and the area under the ROC curve (AUROC) for each. Two-sided P<.05 was considered significant. All statistical analyses were performed with SPSS 12.0 (SPSS, Chicago, IL, USA).

3. Results

3.1. Baseline characteristics

The baseline clinical and histological characteristics for all the patients at the time of DW-MRI and LSM examination are shown in Table 1. The mean age of all patients (21 men and five women) was 51.4±8.9 years. Twenty-six surgical specimens revealed a background of F1 fibrosis in one patient (3.8%), F2 in four (15.4%), F3 in six (23.1%) and F4 in 15 (57.7%). Seven patients (46.6%) out of 15 with cirrhosis (F4) showed mixed micro- and macro-nodular cirrhosis and eight patients (53.4%) showed macro-nodular cirrhosis (data not shown). All patients with F4 fibrosis showed Child-Pugh A liver function before surgery. A1 activity was noted in two (7.7%), and A2 was noted in 24 patients (92.3%). HBeAg was positive in 9 patients (34.6%) and HBV DNA level was more than 20,000 IU/ml in seven patients (26.9%) (data not shown). The mean ICG R15 and LSM values were 8.9±4.9% (median 8.4%, range 3.2–22.1%) and 14.4±9.5 kPa (median 11.2 kPa, range 4.3–35.0 kPa), respectively.

3.2. Variables associated with surgery and tumor characteristics

Variables associated with surgery and tumor characteristics are shown in Table 1. Five patients (19.2%) had segmentectomy only, one patient (3.8%) had bisegmentectomy, and 20 patients (76.9%) had a lobectomy. According to the Tumor-Node-Metastasis (TNM) staging of Liver Cancer Study Group of Japan in 2000 (modified UICC [Union Internationale Contre le Cancer] staging system), six patients (23.1%) had Stage I disease, 18 patients (69.2%) had Stage II disease, and two patients (7.7%) had Stage III disease.
3.3. Comparison between patients with and without postoperative hepatic insufficiency

The comparison between patients with and without postoperative hepatic insufficiency is shown in Table 1. Only three patients experienced postoperative hepatic insufficiency and there was no mortality associated with surgery. Extracted liver specimens of the three patients with postoperative hepatic insufficiency revealed evidence of liver cirrhosis (F4) and all of these patients experienced ascitic decompensation after surgery.

Comparing patients with postoperative hepatic insufficiency to those without, only LSM value was significantly different (P=.001). Noninvasive models such as API, APRI and ASPRI failed to distinguish patients with postoperative hepatic insufficiency (all P values were not significant). The three patients with postoperative hepatic insufficiency had a significantly higher LSM value before surgery than those without (median 33.3 vs. 10.5 kPa, respectively, P=.001). When we selected 12 patients with F4 fibrosis out of 23 patients without postoperative hepatic insufficiency and compared their LSM values to those with postoperative hepatic insufficiency, LSM values were still significantly higher in patients with postoperative hepatic insufficiency than in those without (median 33.3 vs. 13.2 kPa, P=.003, data not shown). Although some variables associated with liver function were favorable for patients without postoperative hepatic insufficiency, these differences did not reach statistical significances (P=.100 in albumin and P=.083 in prothrombin time, respectively). Preoperative body mass index, hemoglobin, AST, HBeAg positivity, HBV DNA level, spleen size, operation time, and blood loss during operation were not significantly different between two groups (data not shown).

3.4. Reliability of apparent diffusion coefficient values

ADC-Rt and ADC-Lt were calculated twice preoperatively and also at the end of study by two observers. Analysis of the reliability of the ADC values is shown in Table 2. The reliability between ADC-Rt and ADC-Lt was significantly high regardless of the measure order and observer (ρ=0.868 and ρ=0.910 in the first and second measures of observer A, respectively; ρ=0.865 and ρ=0.831 in the first and second measures of observer B, respectively, all P<.001).

Because of the high reliability between ADC-Rt and ADC-Lt, we could define the R1-ADC value as the representative ADC value for a certain observer and measure. Using the R1-ADC value, intraobserver reliability (ρ=0.958, P<.001 in observer A and ρ=0.977, P<.001 in observer B) and interobserver reliability (ρ=0.929, P<.001 in the first measure and ρ=0.978, P<.001 in the second measure between observers A and B, respectively) were also proven to be significantly high. This high intra- and interobserver reliability did not change when controlling the effect of tumor location (ρ=0.927 in observer A; ρ=0.977 in observer B; ρ=0.953 in the first measure; ρ=0.929 in the second measure; all P<.001; data not shown). Because reliability between the two hepatic lobes as well as intra- and interobserver reliability were significantly high in our study, we could again define the mean ADC value (R2-ADC) — calculated from four R1-ADC values for each observer at

Table 2
The reliability of apparent diffusion coefficient values

<table>
<thead>
<tr>
<th>ADC values</th>
<th>Measure order</th>
<th>Observer</th>
<th>Reliability (ρ)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC-Rt vs. ADC-Lt</td>
<td>1st measure</td>
<td>Observer A</td>
<td>0.868</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>ADC-Rt vs. ADC-Lt</td>
<td>2nd measure</td>
<td>Observer A</td>
<td>0.910</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>ADC-Rt vs. ADC-Lt</td>
<td>1st measure</td>
<td>Observer B</td>
<td>0.865</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>ADC-Rt vs. ADC-Lt</td>
<td>2nd measure</td>
<td>Observer B</td>
<td>0.831</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>R1-ADC</td>
<td>1st vs. 2nd measure</td>
<td>Observer A</td>
<td>0.958</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>R1-ADC</td>
<td>1st vs. 2nd measure</td>
<td>Observer B</td>
<td>0.977</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>R1-ADC</td>
<td>1st measure vs. 2nd measure</td>
<td>Observer A vs. B</td>
<td>0.929</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>R1-ADC</td>
<td>2nd measure vs. 1st measure</td>
<td>Observer A vs. B</td>
<td>0.978</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Table 3
Diagnostic performances in the prediction of postoperative hepatic insufficiency

<table>
<thead>
<tr>
<th>Predictor</th>
<th>AUROC</th>
<th>Standard error</th>
<th>P value</th>
<th>95% CI Low bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSM</td>
<td>0.942</td>
<td>0.047</td>
<td>.014</td>
<td>0.850</td>
<td>1.035</td>
</tr>
<tr>
<td>R2-ADC</td>
<td>0.797</td>
<td>0.116</td>
<td>.100</td>
<td>0.571</td>
<td>1.024</td>
</tr>
<tr>
<td>ICG R15</td>
<td>0.439</td>
<td>0.204</td>
<td>.968</td>
<td>0.092</td>
<td>0.893</td>
</tr>
</tbody>
</table>

CI, confidence interval.
each measure — as the final representative ADC value for DW-MRI in a given patient. Using the R2-ADC value, we performed a further analysis to compare the performance of DW-MRI to other noninvasive models.

The R2-ADC value did not show a tendency to increase or decrease according to the severity of liver fibrosis stage (0.62 for one patient with F1, 0.95±0.31 for 4 patients with F2, 1.05±0.25 for 6 patients with F3, and 0.96±0.31 10^3 mm^3/s for 15 patients with F4, P=.128 by one-way analysis of variance). Comparing patients with postoperative hepatic insufficiency to those without, R2-ADC values were not significantly different (1.18±0.08 vs. 0.92±0.32 10^3 mm^3/s, P=.180).

3.5. Diagnostic performance in the prediction of postoperative hepatic insufficiency

The diagnostic performances of LSM, R2-ADC values of DW-MRI, and ICG R15 in the prediction of postoperative hepatic insufficiency are shown in Table 3. Only LSM predicted postoperative hepatic insufficiency significantly (AUROC=0.942, P=.014), whereas R2-ADC values and ICG R15 did not (AUROC=0.797, P=.100 and AUROC=0.439, P=.968, respectively). The cutoff LSM value was set at 22.4 kPa, which gave the best statistical accuracy (sensitivity, 60.0%; specificity, 100.0%; positive predictive value, 100.0%; and negative predictive value, 91.3%). Corresponding ROC curves are depicted in Fig. 2.

3.6. Diagnostic performance of LSM, DW-MRI, and other non-invasive models in the prediction of advanced fibrosis (F3–4) and cirrhosis (F4)

The diagnostic performances of LSM, R2-ADC values of DW-MRI, and other non-invasive models in the prediction of advanced fibrosis (F3–4) and cirrhosis (F4) are shown in Table 4. Only LSM showed a statistically significant ability to predict advanced fibrosis and cirrhosis (AUROC=0.771, P=.048 and AUROC=0.818, P=.006, respectively). The R2-ADC and other non-invasive models failed to predict advanced fibrosis and cirrhosis. The cutoff LSM values for advanced liver fibrosis and cirrhosis were set at 7.4 and 10.7 kPa, respectively, which gave the best statistical accuracy (sensitivity for advanced liver fibrosis, 80.9%; specificity, 80.0%; positive predictive value, 94.4%; and negative predictive value, 50.0%; sensitivity for cirrhosis, 85.7%; specificity, 75.0%; positive predictive value, 80.0%; and negative predictive value, 81.8%). Corresponding ROC curves are depicted in Figs. 3 and 4.

4. Discussion

Recently, Lewin et al. reported that the performance of DW-MRI was comparable to that of other noninvasive tests for assessing liver fibrosis in patients with chronic hepatitis C (CHC) (AUROC=0.79 for F2–4 and 0.92 for F3–4) and concluded that DW-MRI could be used as an adjunctive imaging modality to assess liver fibrosis in patients for whom liver MRI is performed [14]. In their study, the performance of LSM was similar or slightly superior to that of DW-MRI (AUROC=0.87 for F2–4 and 0.92 for F3–4).

Inspired by that study, we planned to test the applicability of DW-MRI to patients with chronic hepatitis B (CHB). To investigate the ability of DW-MRI to both predict postoperative hepatic insufficiency and to assess liver fibrosis...
simultaneously, we prospectively and blindly recruited patients with HBV-related HCC undergoing curative liver resection after their preoperative evaluation with MRI. Additionally, we tested intra- and interobserver reliability, which have not been evaluated in prior studies. A significant difference between our study and that of Lewin et al. is that we used a 3-T MRI system. The overall quality of diffusion-weighted images with a 3-T MRI system is known to be superior to that of a 1.5-T MRI system, which was used in the study of Lewin et al. [28,29].

Of the various baseline characteristics, only LSM values were significantly different between patients with postoperative hepatic insufficiency and those without. However, other noninvasive models for assessing liver fibrosis — such as API, APRI and ASPRI — did not generate significantly different results between patients with and without postoperative hepatic insufficiency. Among LSM, DW-MRI using the R2-ADC value and ICG R15, only LSM was able to significantly predict postoperative hepatic insufficiency (AUROC=0.942, P=0.014). This confirmed the results of our previous study indicating that LSM could predict postoperative hepatic insufficiency in patients with HBV-related HCC undergoing curative liver resection with similar cutoff LSM values (22.4 vs. 25.6 kPa) [11]. Additionally, LSM in our current study significantly predicted advanced liver fibrosis (F3–4) and cirrhosis (F4), although our AUROC values were lower than those in previous studies [30]. Better performance of LSM in prediction of cirrhosis compared to advanced fibrosis (AUROC=0.818, P=0.006 vs. AUROC=0.771, P=0.048, respectively) is consistent with the results of a previous meta-analysis [31].

DW-MRI using R2-ADC values also failed to predict postoperative hepatic insufficiency (P=1.000) and the performance of DW-MRI in the prediction of advanced liver fibrosis (F3–4) and cirrhosis (F4) was disappointing (AUROC=0.648, P=0.313 and AUROC=0.491, P=0.938, respectively), despite our use of a 3-T MRI system which could have potentially provided better sensitivity to diffusion effects [14]. Furthermore, the R2-ADC values for each fibrosis stage did not show a tendency to increase or decrease. This result was different from those of Lewin et al. and Koinuma et al. [14,15] showing that ADC values for fibrotic liver decreased as the severity of liver fibrosis increased; on the contrary, our results were similar to those of other investigations indicating that ADC values do not decrease as the fibrosis scores increase [16,32,33].

Several studies have reported that the inferior performance of LSM in patients with CHB compared to those with CHC could be explained by the histopathologic characteristics that HBV infection has a tendency to make heterogeneous macronodular fibrotic structures in the liver [21,34,35]. In this study, most patients with cirrhosis (n=15) had a macronodular component to their fibrosis (seven patients with mixed micro- and macronodular cirrhosis and 8 patients with macronodular cirrhosis). We believe that these interfering histopathologic effects of HBV infection might influence the performance of DW-MRI similarly. Accordingly, although DW-MRI is potentially useful when other noninvasive tools have failed to assess liver fibrosis in patients with hepatitis C virus infection (which makes mostly micronodular structures in the liver) [14], it seems that DW-MRI with our current setting of 3-T MRI system is not applicable to patients with CHB. However, because the intra- and interobserver reliability of DW-MRI were proved to be significantly high in this study and various techniques are under investigation to improve the image quality [36–40], the usefulness of DW-MRI can be investigated again in the future for patients with CHB, especially for obese patients in whom LSM is often inadequate.

Magnetic resonance (MR) elastography measures the elastic parameters of the liver and is another promising imaging modality under investigation [41,42]. Because MR elastography can assess the total amount of fibrosis regardless of obesity and ascites, it could provide more exact predictions of postoperative hepatic insufficiency and liver fibrosis assessments in patients with CHB, if its related cost issues could be solved.

There are several limitations to our current study. First, although this study was designed as a pilot, the total number of recruited patients was small. Therefore, the results should be interpreted with cautions due to the possibility of type II statistical error. Second, because our study population had HBV-related HCC, the proportion of cirrhosis was inevitably high. Therefore, this spectrum bias might be involved [43]. These limitations could be solved with a large-scaled prospective study in the future.

In conclusion, DW-MRI on 3-T MRI system is neither suitable for predicting postoperative hepatic insufficiency nor evaluating the severity of liver fibrosis in patients with
HBV-related HCC undergoing curative liver resection irrespective of the significantly high intra- and interobserver reliability.

Acknowledgment

This study was supported by a grant from the Good Health R & D Project from the Ministry of Health, Welfare and Family Affairs, Republic of Korea (A050021) and in part by Brain Korea 21 Project for Medical Science. The authors wish to thank Joon Seong Kim, Ji Won Kim and Jeong Min Cho for their critical comments and support.

References


