RESEARCH ARTICLE

Quantitative liver ADC measurements using diffusion-weighted MRI at 3 Tesla: evaluation of reproducibility and perfusion dependence using different techniques for respiratory compensation

Nis Elbrønd Larsen · Søren Haack ·
Lars Peter Skovgaard Larsen · Erik Morre Pedersen

Abstract

Object Diffusion weighted imaging (DWI) of the liver suffers from low signal to noise making 3 Tesla (3 T) an attractive option, but 3 T data is scarce. It was the aim to study the influence of different b values and respiratory compensation methods (RCM) on the apparent diffusion coefficient (ADC) level and on ADC reproducibility at 3 T.

Materials and methods Ten healthy volunteers and 12 patients with malignant liver lesions underwent repeated (2–22 days) breathhold, free-breathing and respiratory triggered DWI at 3 T using b values between 0 and 1,000 s/mm^2.

Results The ADCs changed up to 150 % in healthy livers and up to 48 % in malignant lesions depending on b value combinations. Best ADC reproducibility in healthy livers were obtained with respiratory triggering (95 % limits of agreement: ±0.12) and free-breathing (±0.14). In malignant lesions equivalent reproducibility was obtained with less RCM dependence. The use of a lower maximum b value ($b = 500$) decreased reproducibility (±0.14 to ±0.32) in both normal liver and malignant lesions.

Conclusion Large differences in absolute ADC values and reproducibility caused by varying combinations of clinically realistic b values were demonstrated. Different RCMs caused smaller differences. Lowering maximum b value to 500 increased limits of agreement up to a factor of two. Serial ADC changes larger than approximately 15 % can be detected confidently on an individual basis in both malignant lesions and normal liver parenchyma at 3 T using appropriate b values and respiratory compensation.

Keywords Diffusion weighted MRI · Liver · Reproducibility of results

Introduction

Diffusion weighted magnetic resonance imaging (DWI) is a promising functional imaging technique with several liver imaging applications. The method is useful for liver fibrosis quantification [1–3], liver steatosis detection and characterization [4], lesion detection [5, 6] and lesion characterisation [1–3, 7–16]. The apparent diffusion coefficient (ADC), derived from DWI measurements, is a quantification of the random Brownian motion (diffusion) of water molecules in tissues on the μm scale [13, 17]. Since malignant tumors and areas of liver fibrosis are characterized by limited diffusion due to excessive extracellular space tortuosity, disrupted tissue organisation, and, in the case of malignant lesions, hypercellularity, DWI provides a high degree of contrast between these tissues and most benign tissues. The implementation of DWI in the clinical setting has several additional benefits: There is no use of radiation and no intravenous contrast administration. The method may potentially improve radio- and chemotherapy response evaluation using changes in the ADC [13, 15]. Furthermore, the method may aid in pre-treatment identification of patients, who, likely, will have a poor effect from chemotherapy [10, 13, 18].

Diffusion weighted imaging measurements in the liver are, however, still hampered by a number of technical challenges caused by motion of the liver during respiration,

N. E. Larsen (✉) · L. P. S. Larsen · E. M. Pedersen
Department of Radiology, NBG, Aarhus University Hospital,
Nørrebrogade 44, 8000 Århus C, Denmark
e-mail: niselbroend@hotmail.com

S. Haack
Department of Clinical Engineering, Aarhus University Hospital,
Nørrebrogade 44, 8000 Århus C, Denmark

rapid T2 signal decay [19] in the normal liver tissue causing low signal-to-noise ratio (SNR) and the single shot echo planar imaging (SS-EPI) MR sequence used for DWI in the body. The SS-EPI sequence is very sensitive to air/tissue transitions and motion resulting in shape distortions and susceptibility artefacts. The long echo trains cause blurring artefacts, and the large bandwidth makes the SNR of DW images poor. In addition, the large motion probing gradients (large b values with high sensitivity to diffusion) decreases the already low SNR. For this reason the use of higher magnetic field strengths is relevant to increase the SNR, and thereby the accuracy of ADC measurements, as well as the spatial resolution [13, 15]. Unfortunately, image artefacts as well as motion sensitivity increase with increasing field strength, highlighting the need for proper respiratory compensation when scanning at 3 Tesla (3 T). This might explain why a recent study [20] didn't show any significant differences in reproducibility between ADCs obtained at 1.5 and 3 T.

Thus, to establish DWI as a valid clinical biomarker in the liver, the respiratory compensation methods (RCMs) must ensure that reproducible ADC measurements can be acquired when following tumors over time, and studies concerning the reproducibility of DWI are therefore warranted [13]. At 1.5 T, conflicting evidence exists concerning the ability of respiratory triggered (RT) DWI to correctly estimate ADC [5, 6, 21–23] and to reproduce measurements [21–23]. It is, however, generally agreed that RT improves DWI image quality [5, 6, 24]. At 3 T a recently published study [25] found good agreement between ADCs obtained with BH, FB and RT using PACE. To our knowledge, only two 3 T studies [20, 26] have been published addressing reproducibility of ADC measurements in the liver. None of the two healthy volunteer studies do, however, cover RT, which is why further data assessing the various RCM influence on ADC reproducibility is lacking at 3 T.

It is known that the perfusion effect on signal intensity (SI), and thus on the calculated ADC, is almost exclusively related to the use of b values below 100–150 s/mm^2 [4, 13, 18, 27] due to the exponential nature of signal intensity at those lower b values. When the ADC is calculated using a b value combination employing $b = 0$ s/mm^2 and only one additional low maximum b value (for instance $b = 500$, s/mm^2, which is used in much of the corresponding literature, see Table 6), the ADCs can therefore theoretically be subjected to large changes in relation to alterations in the level of perfusion. A number of factors influence hepatic perfusion and the ADC calculation. In the context of liver MRI, it is important to secure that the measurement itself does not change the level of perfusion. It is known that large changes in caval blood flow velocity occurs during normal breathing due to intrathoracic pressure changes [28]. It can thus be hypothesized that inspiration and expiration can cause different levels of perfusion and, thus, different levels of signal intensity [29] on $b = 0$ s/mm^2 images, leading to differences in ADCs calculated from b values 0 and 500 s/mm^2.

Therefore, it was the aim of this 3 T study to:

1. Estimate the influence of different b values and RCMs on the ADC level in normal liver tissue and in malignant liver lesions.
2. Estimate the influence of perfusion sensitivity, different RCMs and b values on the ADC reproducibility in normal liver tissue and in malignant lesions.

Materials and methods

Study participants

The protocol was approved by the local ethics committee (The Regional Committee of Central Denmark, case number M-20100065). Written consent was obtained from all participants.

Inclusion criteria for healthy volunteers were: adults with no history of alcohol or drug abuse, no current medication and no prior liver disease.

Inclusion criteria for patients were: at least one lesion with a diameter larger than 1.5 cm, no current antitumor treatment and no liver surgery/biopsies performed within one month before first MR examination or between the two MR examinations.

Exclusion criteria for both groups were: general contraindications to MRI such as severe obesity, claustrophobia and implanted pacemakers or coils.

MR imaging

The MR imaging was performed on a 3.0 T MR scanner (Philips 3.0 T Achieva X-series system Quasar Dual gradients, Philips Medical Systems, Best, The Netherlands), software version 2.6.3, using a 16 channel XL-torso coil (Philips Medical Systems) for signal reception. Maximum gradient specifications were: amplitude (40 mT/m)/slew rate (200 mT/m/ms). The DWI of the liver was performed using a single-shot spin-echo echo-planar imaging (SS-EPI) sequence using tetrahedral encoding. Fat suppression was performed using chemical shift-selective fat-suppression. The RT was performed using a standard air-pressure driven respiratory monitoring device as supplied by the vendor which according to one cardiac MRI study [30] should be comparable to navigator triggering.

Diffusion weighted imaging was performed twice in each participant with a 2–22 day interval between the two

scans. A total of eight series (four series per session) for each healthy volunteer and 12 series (6 series per session) for each patient (in patients the breathhold sequences had to be divided into two separate sequences, because of the patients' reduced capacity for breathhold compared to the healthy volunteers, see Table 1) were acquired in the axial plane using the four RCMs: breathhold end inspiratory (BHI), breathhold end expiratory (BHE), free-breathing (FB) and RT (Fig. 1). The entire liver could not be covered with any of the employed RCMs using the settings listed in Table 1.

Participants were instructed to fast at least 4 h prior to each imaging session to minimize changes in hepatic perfusion [29].

In patients, care was taken to ensure that inspiratory and expiratory datasets contained the same focal pathology. Respiratory rate and pattern was documented, and total scan time recorded.

Complete acquisition parameters including scan times for the RCMs are included in Table 1.

Image analysis

Ten healthy volunteers [four women and six men; mean age (range), 24 (22–26) years] and 12 patients [six women and six men; mean age (range), 62 (54–80) years] with suspected or known malignant liver disease underwent abdominal MR-DWI with BHI, BHE, FB and RT twice with a 2 to 22 day interval between the two scans [Healthy volunteers; mean interval (range), 5 (3–7) days, Patients; mean interval (range), 9 (2–22) days] during a 12 months period, from September 2010 to August 2011.

After having performed the scans images were initially transferred to a Viewforum Workstation (Philips Medical Systems) where registration of the different excitations (NSA) was performed for each b value before averaging to reduce motion blurring. Subsequent image analysis was performed using Osirix software, v. 3.7.1 (Pixmeo, Geneva, Switzerland). ADC maps were calculated based on different sets of b values. All b value combinations used for ADC map calculations are listed in Table 2 according to the aims of the study. A threshold signal intensity value was calculated as "mean SI + 2 standard deviations (SD)" from a Non-Signal background ROI positioned in air outside the body. All pixel values below this calculated minimum SI value were registered as "not a value" and were excluded from the final ADC calculation [31].

ROIs were manually drawn on DW images and then copied to ADC maps for calculation of average ADC in the ROI. Parenchyma ROIs were drawn on DW images acquired with $b = 0$ s/mm^2 with maximal vessel to parenchyma contrast to avoid the inclusion of major hepatic vessels [13] (Fig. 2a).

Apparent diffusion coefficients of a large standardized ROI measuring 20 cm^2, approximately 320 pixels, was used to analyze liver parenchyma in healthy volunteers [22, 26, 32] (Fig. 2a).

The ROIs were placed laterally in the right liver lobe in the most homogenous parenchyma in three different slices: a reference slice at the level through the right portal vein, a slice located one cm (=two slices) above and a slice located one cm below the reference slice.

The three ROIs in the same cranio-caudal position were averaged and used as the representative ADC for the healthy liver parenchyma. All ROIs were placed minimally 0.5 cm from the liver border, and care was taken to avoid large hepatic vessels and artifacts.

The ROIs delineating lesions were drawn on DW images acquired with $b = 1,000$ s/mm^2 ensuring maximal contrast between lesion and parenchyma [13, 33] (Fig. 2b). The largest malignant lesion (if more than one was present) was used for evaluation in each patient to minimize partial volume averaging in small lesions. In small lesions (covering three successive slices or less) all available slices were used for analysis, while in larger lesions (i.e., lesions covering more than three successive slices) only the three slices with the largest lesion diameter were used for analysis [5, 21]. After copying ROIs from DW-images to calculated ADC-maps, ADCs from ROIs were exported to Microsoft Excel, where final averaged ADCs were calculated.

Statistical analysis

Mean ADCs were calculated as averages of the two repeated ADC measurements. Differences in mean ADCs between the RCMs were analyzed for each b value combination using the within-subject analysis of variance (ANOVA). If a significant difference was found using the within-subject ANOVA, paired t tests were used to compare the RCMs pairwise and establish which methods differed significantly.

Due to the low number of useful patient BHI sequences, only BHE, FB and RT were compared with the within-subject ANOVA to maintain statistical strength in the patient group. We additionally performed paired t tests to compare patient BHI and BHE sequences.

The within-subject ANOVAs and paired t tests were performed with the AnalystSoft, StatPlus:mac—statistical analysis program for Mac OS. Version 2009.

All statistical analysis of reproducibility was made as intrapersonal comparisons of ADCs obtained using the different RCMs and b values. Logarithmic transformed Bland–Altman plots presented as fractional values were constructed to determine the reproducibility of the employed RCMs, i.e., BHI, BHE, FB and RT [34]. The 95 % limits of agreement (95 % LOA) were used to establish the magnitude of changes

Table 1 Imaging parameters for breathhold (BHI and BHE), free-breathing (FB) and respiratory triggered (RT) protocols for healthy volunteers and patients

| Parameter	Patients			Healthy volunteers		
	BHI and BHE	FB	RT	BHI and BHE	FB	RT
Repetition time/echo time (ms)	1,600/72	2,000/72	2,000/71.3	1,600/72	2,000/72	2,000/71.3
Receiver bandwith (Hz/pixel)	2,833	3,028	3,028	2,833	3,028	3,028
Matrix	168 × 142	140 × 140	140 × 140	168 × 142	140 × 140	140 × 140
Slice thickness (mm)	5	5	5	5	5	5
No. of slices	15	20[b]	28[b]	15	20[b]	28[b]
No. of signals acquired (NSA)	2	7	5	2	7	5
Field of view (mm) RL/AP	400/282	350/282	350/282	400 × 286	350 × 280	350 × 280
Echo train length	49	59	59	47	59	59
Parallel Imaging factor	2, 2	2	2	2, 2	2	2
b values (s/mm^2)	0,150 + 500 and 0 + 150 + 1,000	0 + 150 + 500,1,000	0 + 150 + 500 + 1,000	0 + 150 + 750 + 1,000	0 + 150 + 500 + 750 + 1,000	0 + 150 + 500 + 750 + 1,000
Scan time	22 s	7:12 min	9–15 min[a]	32 s	9 :18 min	11–18 min[a]
Acquired voxel size (mm)	2.4 × 2, 9 × 5	2.5 × 2.5 × 5	2.5 × 2.5 × 5	2.4 × 2, 9 × 5	2.5 × 2.5 × 5	2.5 × 2.5 × 5
Phase-direction	AP	AP	AP	AP	AP	AP

[a] Approximated scan time, depending on respiratory frequency

[b] Interleaved slice acquisition

Fig. 1 DWI image examples from healthy volunteers ($b = 750$ s/mm^2, *top*) and patients ($b = 1{,}000$ s/mm^2, *bottom*) using each of the four examined respiratory compensation methods: breathhold end inspiratory (BHI), breathhold end expiratory (BHE), free-breathing (FB) and respiratory triggering (RT)

DWI image examples
Normal volunteers

BHI   BHE

FB   RT

Patients:

BHI   BHE

FB   RT
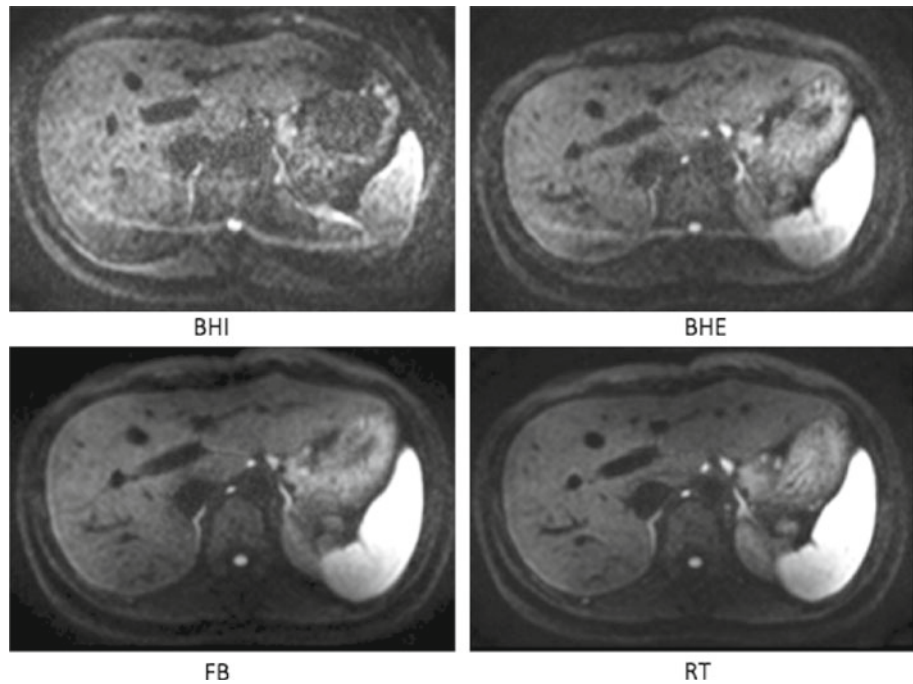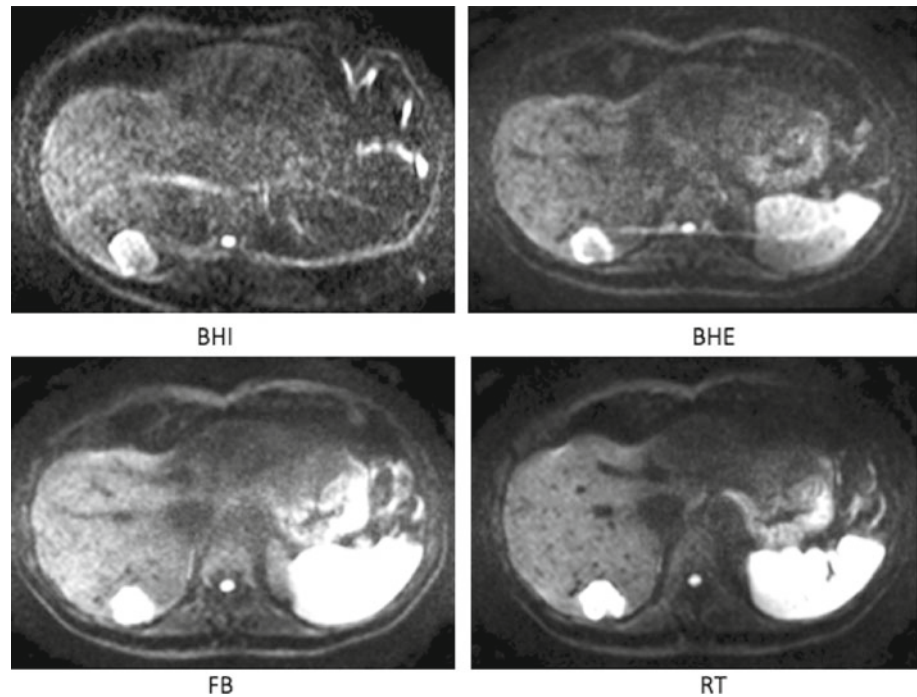
in ADC that can be confidently detected in a subject. For the sake of simplicity, all 95 % LOAs are presented as "± (upper limit + lower limit)/2" although it is realized that the upper and lower limits are not completely symmetrically distributed due to the logarithmic transformation. Paired t tests were used to test if any bias, i.e., systematic difference between the repeated scans were present.

Results

All healthy volunteer sequences obtained were sufficient for analysis. One patient was later excluded because of lacking a malignant diagnosis, leaving a total of 11 patients for further evaluation [five women and six men; mean age (range) 63 (54–80) years].

Table 2 b value combinations for ADC calculations according to study aims

Study aims	Healthy volunteers b values	Patients b values	All subjects RCMs
Influence of b value combinations on ADC	150 + 500	150 + 500	FB + RT
	150 + 750	150 + 1,000	FB + RT
	150 + 500 + 750 + 1,000	150 + 500 + 1,000	FB + RT
	0 + 500	0 + 500	FB + RT
	0 + 750	0 + 1,000	FB + RT
	0 + 1,000	0 + 150 + 500 + 1,000	FB + RT
	0 + 150 + 500 + 750 + 1,000		FB + RT
Influence of RCM on ADC	150 + 750	150 + 1,000	BHI + BHE + FB + RT
	0 + 750	0 + 1,000	BHI + BHE + FB + RT
Influence of perfusion and RCM on ADC reproducibility	150 + 750	150 + 1,000	BHI + BHE + FB + RT
	0 + 750	0 + 1,000	BHI + BHE + FB + RT
Influence of max b value on ADC reproducibility	0 + 500	0 + 500	FB + RT
	0 + 750	0 + 1,000	FB + RT

b values (s/mm^2) and RCM's (respiratory compensation methods) used

BHI breathhold end inspiratory, *BHE* breathhold end expiratory, *FB* free-breathing, *RT* respiratory triggered

Fig. 2 **a** Axial RT DW MR Image ($b = 0$ s/mm^2) demonstrating good contrast between parenchyma (*dark*) and vessels (*bright*). In the image the employed ROI can be seen. **b** Axial RT DW MR Image ($b = 1,000$ s/mm^2) showing good contrast between healthy parenchyma (*dark*) and lesion (*bright*) with surrounding ROI
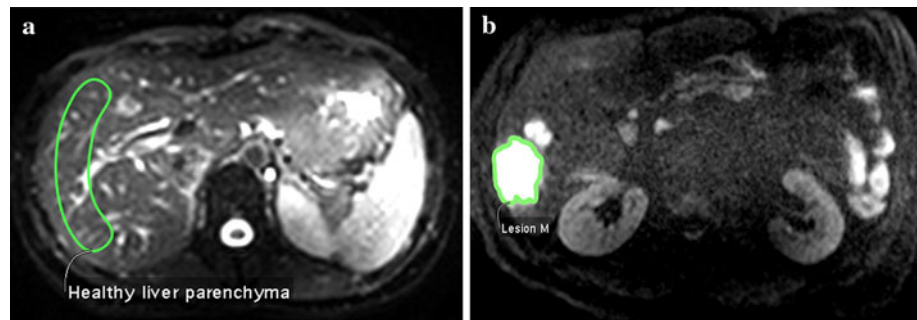
In the patients a total of 11 lesions (one lesion per included patient) was scanned and later analyzed (mean diameter 4.6 cm; range 1.7–12.8 cm). The lesions included one cholangiocarcinoma, five hepatocellular carcinomas and five metastases from colorectal carcinomas. In five patients, BHI sequences were insufficient because of: inability to hold breath ($n = 2$), sternal wire artifact ($n = 1$), lesion not included in field of view ($n = 1$) and coil malfunction ($n = 1$). In one patient, BHE sequences were insufficient (inability to hold breath). In conclusion, six BHI datasets, ten BHE datasets and 11 FB and RT datasets were used for further analysis.

ADC quantification: dependence of b values (Fig. 3)

Large ADC differences due to the use of different b value combinations were found in both healthy liver parenchyma and in malignant lesions. In the healthy liver parenchyma, the lowest ADCs of 0.71×10^{-3} mm^2/s (FB) and 0.725×10^{-3} mm^2/s (RT) were calculated using a b value combination of 150 + 500 + 750 + 1,000 s/mm^2 while

the highest ADCs of 1.766×10^{-3} mm^2/s (FB) and 1.811×10^{-3} mm^2/s (RT) were calculated using a b value combination of 0 + 500 s/mm^2. Thus, differences of approximately 150 % were found between the lowest and highest ADCs, depending on b values used.

A similar pattern was seen for the malignant lesions, although the measured ADC differences were somewhat smaller. The lowest ADCs of 0.807×10^{-3} mm^2/s (FB) and 0.808×10^{-3} mm^2/s (RT) were calculated using b values 150 + 500 + 1,000 s/mm^2 while the highest ADCs of 1.159×10^{-3} mm^2/s (FB) and 1.198×10^{-3} mm^2/s (RT) were calculated using b values 0 + 500 s/mm^2. Thus, differences of approximately 45 % were found between the lowest and highest ADCs in patients.

Influence of respiratory compensation methods on ADC level (Table 3)

In the healthy volunteers, no significant differences in ADC level were found between the four RCMs ($p > 0.05$) calculated with b values 0 + 750 s/mm^2.

Using the *b* value combination $150 + 750$ s/mm^2, significant differences in ADC levels were found between the four RCMs ($p < 0.001$). The subsequent students *t* tests revealed that BHI and BHE ADCs were significantly lower than the ADCs of FB and RT ($p < 0.001$). The ADC differences between BHI and BHE and between FB and RT were not significant ($p > 0.05$).
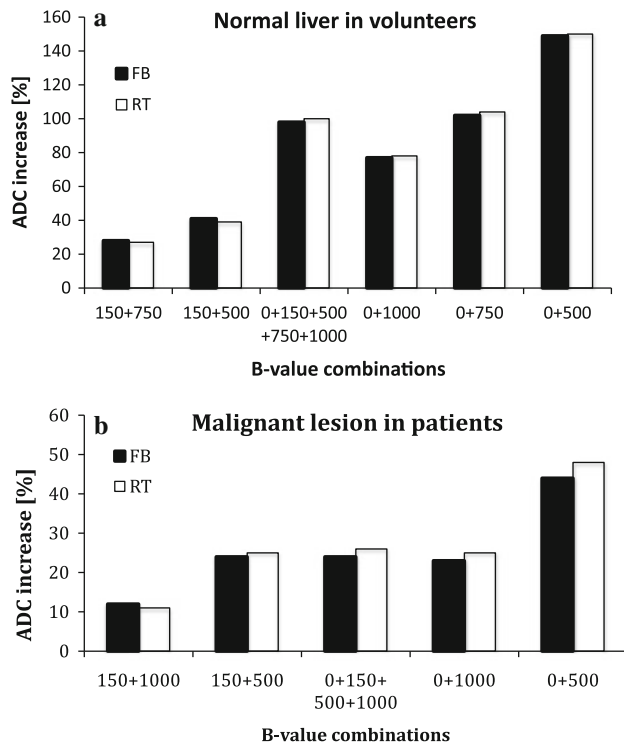

Fig. 3 Percentage difference (increase) in ADC-values obtained when using different *b* value combinations. In the right liver in normal volunteers (Fig. 3a), the reference ADC value is calculated from $b = 150 + 500 + 750 + 1,000$. In malignant lesions in patients (Fig. 3b), the reference ADC value is calculated from $b = 150 + 500 + 1,000$

Likewise, the ANOVA revealed significant differences in lesion ADCs between BHE, FB and RT calculated with *b* values $0 + 1,000$ s/mm^2 ($p < 0.01$). The subsequently performed *t* tests showed that the larger ADC acquired with BHE was significantly different from the FB and RT ADCs. Again, ADC differences between BHI and BHE and between FB and RT were not significant ($p > 0.05$).

Differences between BHE, FB and RT ADCs were non-significant using *b* values $150 + 750$ s/mm^2 ($p > 0.05$). This was also the case for the difference between BHI and BHE ADCs ($p > 0.05$).

Influence of lower perfusion sensitive *b* values and respiratory compensation methods on ADC reproducibility (Table 4)

Mean ADC bias between repeated sequences under all conditions ranged between 1 and 6 %. All bias were non-significant with p-values ranging between 0.06 and 0.79.

In healthy volunteers, we obtained the best reproducibility with an ADC calculation with *b* values $0 + 750$ s/mm^2 using RT and FB with 95 % LOA of ± 0.12 and ± 0.14, respectively. The BHI and BHE performed equally with 95 % LOAs of ± 0.25. With the ADC calculation using *b* values $150 + 750$ s/mm^2, the four RCMs performed very similarly with 95 % LOA ranging between ± 0.18 with BHE and ± 0.20 with RT. No obvious differences in terms of reproducibility could be found between ADC calculations using *b* values $150 + 750$ s/mm^2 and $0 + 750$ s/mm^2 with breathholding methods.

In the patients, the four RCMs were almost equally reproducible using the ADC calculation with *b* values $0 + 1,000$ s/mm^2 resulting in 95 % LOAs ranging between ± 0.10 with BHI and ± 0.15 with BHE and RT. A somewhat similar pattern was found for the ADC calculation with *b* values $150 + 1,000$ s/mm^2 with BHE, FB and RT being almost equally reproducible with 95 % LOA of

Table 3 ADC values for breathhold vs non-breathhold measurements

RCM	Normal liver		Malignant lesions	
	$b = 0 + 750$[§]	$b = 150 + 750$*	$b = 0 + 1,000$*	$b = 150 + 1,000$[§]
Breathhold insp (BHI)	1.52 ± 0.11	0.81 ± 0.09	1.22 ± 0.09	0.85 ± 0.12
Breathhold exp (BHE)	1.54 ± 0.1	0.83 ± 0.10	1.17 ± 0.09	0.86 ± 0.08
Free-breathing (FB)	1.44 ± 0.11	0.91 ± 0.08	0.99 ± 0.14	0.90 ± 0.12
Respiratory trig (RT)	1.48 ± 0.11	0.92 ± 0.07	1.01 ± 0.16	0.89 ± 0.15

ADC values (10^{-3} mm^2/s) $\pm$SD compared for different respiratory compensation methods (RCMs): BHE, BHI, FB and RT

* $p < 0.05$ for within-subject ANOVA. Subsequent paired *t* tests showed significant differences ($p < 0.05$) between all combinations of breathhold versus non-breathhold (BHI/BHE vs FB/RT) but no significant differences ($p > 0.05$) within breathhold (BHI vs BHE) or within non-breathhold (FB vs RT) groups, respectively

[§] $p > 0.05$ for within-subject ANOVA

Table 4 Reproducibility of ADC measurements: Influence of respiratory compensation methods and lower perfusion sensitive b values

RCM	Normal Liver		Malignant lesions	
	$b = 0 + 750$	$b = 150 + 750$	$b = 0 + 1{,}000$	$b = 150 + 1{,}000$
Breathhold insp (BHI)	0.99 ± 0.25	0.99 ± 0.18	1.04 ± 0.10	1.02 ± 0.23
Breathhold exp (BHE)	0.99 ± 0.25	1.01 ± 0.18	1.03 ± 0.15	1.04 ± 0.15
Free-breathing (FB)	1.05 ± 0.14	1.05 ± 0.18	0.97 ± 0.14	1.00 ± 0.16
Respiratory trig (RT)	1.04 ± 0.12	1.05 ± 0.20	1.00 ± 0.15	1.01 ± 0.14

Reproducibility of ADC measurements calculated as fractional mean bias $\pm 95~\%$ limits of agreement (Bland–Altman). Influence of different respiratory compensation methods (RCMs): BHI, BHE, FB and RT

± 0.15, ± 0.16 and ± 0.14, respectively. The BHI was less reproducible with 95 % LOA of ± 0.23.

Influence of maximum b value on ADC reproducibility (Table 5)

Reproducibility was found to increase with increasing maximum b value for both the healthy volunteers and the patients with FB and RT. In the healthy volunteers, the 95 % LOAs improved from ± 0.14 and ± 0.20 to ± 0.12 and ± 0.14 (RT and FB, respectively) when replacing $b = 500$ s/mm^2 with $b = 750$ s/mm^2 using an additional b value of 0 s/mm^2 for the ADC calculation. Similarly, reproducibility improved from ± 0.30 and ± 0.32 to ± 0.14 and ± 0.15 (FB and RT, respectively) in the patients when $b = 500$ s/mm^2 was replaced with $b = 1{,}000$ s/mm^2.

Discussion

Firstly, ADCs changed with the use of different b values. In the healthy volunteers, FB and RT ADCs varied up to approximately 150 %. In the patients, a similar pattern was found, although FB and RT ADCs only differed approximately 45 %.

Secondly, small but significant differences were found between breathhold and non-breathhold ADCs for some b value combinations. For the healthy volunteers, non-breathhold ADC calculation with b values 150 + 750 s/mm^2 gave significantly higher ADCs. Inversely, the breathhold methods gave significantly higher ADCs in patients using b values 0 + 1,000 s/mm^2 for ADC calculation. A non-significant trend of pairwise differences between breathhold and non-breathhold ADCs were also found with the other b value combinations.

Thirdly, we found ADC reproducibility values to be highly dependent on the selected b values. The use of a low maximum b value (500 s/mm^2) gave markedly reduced ADC reproducibility compared to ADC calculations using a higher maximum b value ($b = 750$ s/mm^2 in healthy volunteers and $b = 1{,}000$ s/mm^2 in patients).

Our findings of ADC level in relation to applied b values show that ADCs calculated on the basis of a monoexponential b value combination, including $b = 0$ s/mm^2, will be overestimated compared to ADCs calculated with a true "diffusion sensitive" b value of minimum 150 s/mm^2. This overestimation must be attributed to the perfusion effect on the $b = 0$ s/mm^2 images, which yields higher SI. Furthermore, our results emphasize that ADC results from different studies are not easily compared unless obtained with the same b values. To further complicate this problem, Sasaki et al. [35] showed that significant differences in grey and white matter ADCs can be attributed to the use of different coil systems, imagers, vendors and field strengths. It is likely that these findings also apply to DWI in the liver. Our results concerning the maximum b value's influence on ADC level are in good agreement with the results of 1.5 T MR studies [1, 11], a 3 T study in normal volunteers [20] (Table 6) and a 3 T rabbit study [36].

The influence of RCMs on ADCs showed a tendency of pairwise difference between ADCs for breathhold and non-breathhold RCMs in both normal liver and malignant lesions. The differences were not fully consistent and none of the methods should be disregarded due to deviant ADCs. Our results do, however, demonstrate that ADCs obtained with and without breathholding are not easily compared. These results are somewhat in contrast to the ones presented in other studies [22, 23] (Table 6). Kwee et al. [22] reported that healthy liver tissue ADCs acquired with RT

Table 5 Influence of maximum b value on ADC reproducibility

RCM	Normal liver		Malignant lesions	
	0 + 500	0 + 750	0 + 500	0 + 1,000
Free-breathing (FB)	1.04 ± 0.20	1.05 ± 0.14	0.98 ± 0.30	0.97 ± 0.14
Respiratory trig (RT)	1.03 ± 0.14	1.04 ± 0.12	1.00 ± 0.32	1.00 ± 0.15

Fractional reproducibility $\pm 95~\%$ limits of agreement (Bland–Altman) of ADC measurements in normal liver and malignant lesions showing increasing reproducibility with increasing maximum b value used for ADC calculation

Table 6 Comparison of papers regarding the influence of b values and respiratory compensation methods on ADC-values and reproducibility

Article	Field-strenght (T)	b values	RCMs	Patients/ healthy volunteers	ADC-values	Reproducibility ($\pm$95 % LOA)
Us	3.0	Healthy volunteers: 0 + 150 + 500 + 750 Patients: 0 + 150 + 500 + 1,000	BHI. BHE. FB and RT	Patients/ healthy volunteers	Healthy liver $b = 0 + 750/$ $150 + 750/(0 + 500)$: BHI: 1.52/0.81 BHE: 1.54/0.83 FB: 1.44/0.91/(1.77) RT: 1.48/0.92/(1.81) Malignant lesions $b = 0 + 1{,}000/150 +$ $1{,}000/(0 + 500)$: BHI: 1.22/0.85 BHE: 1.17/0.86 FB: 0.99/0.90/(1.16) RT: 1.01/0.89/(1.20)	Healthy liver $b = 0 + 750/150 +$ $750/(0 + 500)$: BHI: $\pm$0.25/0.18 BHE: $\pm$0.25/0.18 FB: $\pm$0.14/0.18/(0.20) RT: $\pm$0.12/0.20/(0.14) Malignant lesions $b = 0 + 1{,}000/150 +$ $1{,}000/(0 + 500)$: BHI: $\pm$0.10/0.23 BHE: $\pm$0.15/0.15 FB: $\pm$0.14/0.16/(0.30) RT: $\pm$0.15/0.14/(0.32)
Kim et al. [1]	1.5	3–846 (multiple)	BH	Patients/ Healthy volunteers	b values <410 s/mm^2 Healthy liver: 1.20 Malignant lesions: 1.24 b values <850 s/mm^2 Healthy liver: 1.02 Malignant lesions: 1.01	N/A
Koh et al. [11]	1.5	0 + 150 + 500	BH	Patients	0–500: Healthy liver: 1.48 Malignant lesions: 1.94 150 + 500: Healthy liver: 1.04 Malignant lesions: 1.36	N/A
Kandpal et al. [5]	1.5	0 + 500	BH and RT	Patients	Healthy liver: BH: 1.35 RT: 1.39 Metastasis: BH: 1.06 RT: 1.13	N/A
Nasu et al. [6]	1.5	0 + 500	FB and RT	Patient	FB: 1.47 RT: 1.36	N/A
Choi et al. [25]	3.0	50 + 800	BH. FB and PACE^b	Patients	BH: 0.93 FB: 0.87 PACE: 0.91	N/A
Rosenkrantz et al. [20]	1.5 and 3.0	BH: 0 + 300 + 400 and 0 + 700 + 800 FB: 0–800 (multiple)	BH and FB	Healthy volunteers	BH: 0–400: 1.63 (1.5 T) 1.49 (3.0 T) 0–800: 1.20 (1.5 T) 1.12 (3.0 T) FB: 0–800: 1.03 (1.5 T) 0.89 (3.0 T)	BH: 0–400: 11.8 % (1.5 T)^a 12.4 % (3.0 T)^a 0–800: 9.7 % (1.5 T)^a 13.5 % (3.0 T)^a FB: 0–800: 11.5 % (1.5 T)^a 15.3 % (3.0 T)^a
Kim et al. [21]	1.5	50 + 500	BH and RT	Patients	BH: 1.27–1.29 RT: 1.42–1.50	BH: 0.29–0.31 RT: 0.31
Kwee et al. [22]	1.5	0 + 500	BH. FB and RT	Healthy volunteers	BH: 1.57–1.62 FB: 1.62–1.65 RT: 2.07–2.27	BH: $\pm$0.16–0.24 FB: $\pm$0.20–0.29 RT: $\pm$0.59–0.60

Table 6 continued

Article	Field-strength (T)	b values	RCMs	Patients/ healthy volunteers	ADC-values	Reproducibility (±95 % LOA)
Kim et al. [23]	1.5	BH: 0 + 50 + 500 RT: 0 + 50 + 300 + 500 + 1,000	BH and RT	Patients	BH: 1.25–1.40 RT: 1.29–1.56	BH: ±0.19–0.24 RT: ±0.15–0.32
Braithwaite et al. [26]	3.0	0 + 400	FB	Healthy volunteers	Approx. 2.3	12.8 %[a]

[a] Reproducibility expressed as mean coefficient of variation

[b] PACE, navigator-triggered DWI with prospective acquisition correlation (PACE) technique

were significantly higher compared to ADCs acquired with BH and FB DWI. The study does, however, differ from our study in a number of ways, which makes direct comparison of the two studies difficult. Firstly, they used a 1.5 T scanner. Secondly, ADCs were calculated on coronally reformatted images. Thirdly, the repeated scans were performed within the same imaging session, and, finally, different ROI sizes and positioning were used. Similarly, Kim et al. [23] found a tendency of higher malignant lesions ADCs when comparing RT to BH. Again, several of the above mentioned factors make direct comparison of the studies difficult.

Braithwaite et al. [26] found higher ADC values at 3 T in normal volunteers than in the present study (Table 6). This difference might be explained by the maximum b value of 400 s/mm^2 in their study. Finally the use of imagers from different vendors could perhaps add to the differences in ADCs [35].

The best ADC reproducibility in healthy liver was obtained with a b value combination of $b = 0 + 750$ s/mm^2 (Table 4). This indicates that healthy liver parenchyma ADC calculations performed with FB and RT are more sensitive to changes in SI rather than to changes in perfusion. Probably due to a high SI in the $b = 0$ s/mm^2 DW images, relative to the $b = 150$ s/mm^2 images. The influence of perfusion is minimized, but the SI is lower and more prone to noise, which might influence the stability of the ADC calculation. Using a lower maximum b value in addition to a low perfusion sensitive b value (e.g. 0 s/mm^2) for ADC calculations decreased the reproducibility with a factor of approximately two, probably due to the higher sensitivity to changes in perfusion and the poorer mathematical stability of the mono-exponential ADC calculation [36].

In malignant lesions, a more consistent level of ADC reproducibility was seen (Table 4). This was probably due to a higher SI in the malignant lesions compared to healthy liver tissue. To our knowledge, we are the first to assess the influence of small perfusion sensitive b values and maximum b value on reproducibility of ADC-values in the liver

at 3 T. These findings have important implications for monitoring of parenchymal and tumor ADC changes over time, especially since the use of the less reproducible ADC calculation with b values $0 + 500$ s/mm^2 or similar are widely used in the present literature and probably in the clinical setting. Padhani et al. [13] recommend the use of b values between 100 and 750 s/mm^2 for tumor visualization and characterization in the liver. According to our results, these recommendations are appropriate for the assessment of healthy liver parenchyma.

In contrast, Kwee et al. [22] reported that healthy liver tissue ADCs acquired with RT were significantly less reproducible compared to ADCs acquired with BH and FB DWI at 1.5 T (Table 6). Likewise, Kim et al. [23] found a tendency of BH ADCs being more reproducible compared to RT ADCs when only using two b values for the malignant lesion ADC calculation (Table 6). They did, however, find RT to be equally reproducible to BH if employing multiple b values. The higher and less reproducible ADCs found in [22, 23] might be caused by the presence of hepatic pseudo-anisotropy, an artifact which in a study by Nasu et al. [37] was found to cause higher RT ADC values.

The reproducibility values presented by Braithwaite et al. [26] and Rosenkrantz et al. [20] could not be compared directly to our results since coefficients of variation (CV) were used as a measure of reproducibility instead of limits of agreement (Table 6).

Kim et al. [21] performed a reproducibility study at 1.5 T and found very similar 95 % LOAs compared to the ADC values calculated with b values $0 + 500$ s/mm^2 in our study (Table 6).

Because none of the ADCs achieved with the different RCMs were found to be consistently deviant, neither for the healthy volunteers nor for the patients, none of the methods should be disregarded in the clinical setting. We do, however, suggest that RT or FB with a b value combination including $b = 0$ s/mm^2 and a maximum b value larger than 500 s/mm^2 should be preferred for repeated healthy liver parenchyma measurement in clinical as well as research

settings due to superior reproducibility. The better RT and FB reproducibility enables detection of smaller changes in ADC than can be achieved with the breathholding methods. If a shorter examination time is preferred, FB should be chosen at the expense of a slightly reduced reproducibility. Furthermore, FB could be preferable in situations with patients with irregular breathing patterns. It is not advised to use breathhold and non-breathhold methods interchangeably in serial studies. It is uncertain if the use of $b = 0$ s/mm^2 will compromise the detection of changes in ADC when following pathological tissues over time. The large influence of perfusion on SI and thus on the calculated ADCs might conceal actual changes in diffusion, and, therefore, also conceal progression or regression of the pathological tissue. If this is the case, the alternative would be to employ a b value combination with a minimum b value of 150 s/mm^2 using a high number of signal averages (NSA) at the expense of prolonged acquisition time. In malignant lesions, ADC reproducibility was largely independent of employed RCM. The only exception being BHI using $b = 150 + 1,000$ s/mm^2. The BH methods can, therefore, potentially also be used for monitoring of malignant lesions. However, the limited resolution, limited coverage, limited SI and need for good breathhold compliance should be carefully considered before employing BH techniques.

Our study had some limitations. The small number of participants limits the strength of the study, especially for the patient BHI measurements. The intrapersonal study design does, however, still allow valid conclusions to be drawn.

Our study may be impaired by some degree of selection bias among the included participants. The healthy volunteers were recruited among a young, fit and homogenous population and the included patients had to accept participation in the study. This might cause inclusion of the "healthiest" patients. Our reproducibility results may, therefore, be overestimated to some degree if compared to a "true" patient population.

It can be considered a limitation to our study that we did not study subjects at both 1.5 and 3 T, but our ADC values and their reproducibility were in line with results from both 1.5 and 3 T (Table 6).

The lack of blinding when drawing ROIs is not considered to be a limitation because ADC results were unknown to the observer until the ROIs were copied automatically from the DW images to the ADC maps. Furthermore, one might argue that the chosen approach is closer to the actual clinical situation, where the same radiologist most likely will measure and compare ADCs from the repeated sequences.

When comparing breathhold and non-breathhold sequences it is a limitation that differences in, e.g., receiver bandwidth and NSA are necessary to comply with the breathholding constraints.

Future research should focus on further optimisation of the RCMs. This would lower time consumption of FB and RT and, thus, make the methods more attractive in the clinical setting. Likewise, acquisition time for BH DWI should be reduced to less than the 22 s used in our study in order for the BH methods to be consistently useful for scanning patients with a varying capacity for breathhold. Furthermore, it should be assessed if the limited spatial resolution achievable with the BH methods affects the evaluation of lesions with diameters smaller than the 1.5 cm used as lower limit in this study.

Conclusion

Apparent diffusion coefficients depend heavily on the b values chosen for the ADC calculation, especially in the normal right liver lobe. Small, but significant differences were found between breathhold and non-breathhold ADCs in both healthy volunteers and patients for some of the b value combinations.

Respiratory triggering and free-breathing using an ADC calculation with b values $0 + 750$ s/mm^2 gave the most reproducible ADCs when scanning healthy liver parenchyma. When scanning malignant lesions, fairly identical ADC reproducibility values were achieved with the different respiratory compensation methods. Using a low maximum b value of 500 s/mm^2 instead of 750 s/mm^2 (normal liver) and 1,000 s/mm^2 (malignant lesions) decreased ADC reproducibility with approximately a factor of two.

Apparent diffusion coefficient changes of approximately 15 % can be confidently detected in the right liver lobe of healthy volunteers and in malignant liver lesions at 3 T if appropriate b values and respiratory compensation methods are used.

References

1. Kim T, Murakami T, Takahashi S, Hori M, Tsuda K, Nakamura H (1999) Diffusion-weighted single-shot echoplanar MR imaging for liver disease. AJR Am J Roentgenol 173:393–398
2. Namimoto T, Yamashita Y, Sumi S, Tang Y, Takahashi M (1997) Focal liver masses: characterization with diffusion-weighted echo-planar MR imaging. Radiology 204:739–744
3. Taouli B, Vilgrain V, Dumont E, Daire JL, Fan B, Menu Y (2003) Evaluation of liver diffusion isotropy and characterization of focal hepatic lesions with two single-shot echo-planar MR imaging sequences: prospective study in 66 patients. Radiology 226:71–78
4. Guiu B, Petit JM, Capitan V, Aho S, Masson D, Lefevre PH, Favelier S, Loffroy R, Verges B, Hillon P, Krause D, Cercueil JP (2012) Intravoxel incoherent motion diffusion-weighted imaging in nonalcoholic fatty liver disease: a 3.0-T MR study. Radiology 265:96–103

5. 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. AJR Am J Roentgenol 192:915–922

6. Nasu K, Kuroki Y, Sekiguchi R, Nawano S (2006) The effect of simultaneous use of respiratory triggering in diffusion-weighted imaging of the liver. Magn Reson Med Sci 5:129–136

7. Bruegel M, Holzapfel K, Gaa J, Woertler K, Waldt S, Kiefer B, Stemmer A, Ganter C, Rummeny EJ (2008) Characterization of focal liver lesions by ADC measurements using a respiratory triggered diffusion-weighted single-shot echo-planar MR imaging technique. Eur Radiol 18:477–485

8. Gourtsoyianni S, Papanikolaou N, Yarmenitis S, Maris T, Karantanas A, Gourtsoyiannis N (2008) Respiratory gated diffusion-weighted imaging of the liver: value of apparent diffusion coefficient measurements in the differentiation between most commonly encountered benign and malignant focal liver lesions. Eur Radiol 18:486–492

9. Holzapfel K, Bruegel M, Eiber M, Ganter C, Schuster T, Heinrich P, Rummeny EJ, Gaa J (2010) Characterization of small (</=10 mm) focal liver lesions: value of respiratory-triggered echo-planar diffusion-weighted MR imaging. Eur J Radiol 76:89–95

10. Kele PG, van der Jagt EJ (2010) Diffusion weighted imaging in the liver. World J Gastroenterol 16:1567–1576

11. Koh DM, Scurr E, Collins DJ, Pirgon A, Kanber B, Karanjia N, Brown G, Leach MO, Husband JE (2006) Colorectal hepatic metastases: quantitative measurements using single-shot echo-planar diffusion-weighted MR imaging. Eur Radiol 16:1898–1905

12. Koike N, Cho A, Nasu K, Seto K, Nagaya S, Ohshima Y, Ohkohchi N (2009) Role of diffusion-weighted magnetic resonance imaging in the differential diagnosis of focal hepatic lesions. World J Gastroenterol 15:5805–5812

13. Padhani AR, Liu G, Koh DM, Chenevert TL, Thoeny HC, Takahara T, Dzik-Jurasz A, Ross BD, Van Cauteren M, Collins D, Hammoud DA, Rustin GJ, Taouli B, Choyke PL (2009) Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. Neoplasia 11:102–125

14. Parikh T, Drew SJ, Lee VS, Wong S, Hecht EM, Babb JS, Taouli B (2008) Focal liver lesion detection and characterization with diffusion-weighted MR imaging: comparison with standard breath-hold T2-weighted imaging. Radiology 246:812–822

15. Taouli B, Koh DM (2010) Diffusion-weighted MR imaging of the liver. Radiology 254:47–66

16. Agnello F, Ronot M, Valla DC, Sinkus R, Van Beers BE, Vilgrain V (2012) High-b-value diffusion-weighted MR imaging of benign hepatocellular lesions: quantitative and qualitative analysis. Radiology 262:511–519

17. Le Bihan D (1995) Molecular diffusion, tissue microdynamics and microstructure. NMR Biomed 8:375–386

18. Koh DM, Scurr E, Collins D, Kanber B, Norman A, Leach MO, Husband JE (2007) Predicting response of colorectal hepatic metastasis: value of pretreatment apparent diffusion coefficients. AJR Am J Roentgenol 188:1001–1008

19. de Bazelaire CM, Duhamel GD, Rofsky NM, Alsop DC (2004) MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results. Radiology 230:652–659

20. Rosenkrantz AB, Oei M, Babb JS, Niver BE, Taouli B (2011) Diffusion-weighted imaging of the abdomen at 3.0 Tesla: image quality and apparent diffusion coefficient reproducibility compared with 1.5 Tesla. J Magn Reson Imaging 33:128–135

21. Kim SY, Lee SS, Byun JH, Park SH, Kim JK, Park B, Kim N, Lee MG (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. Kwee TC, Takahara T, Koh DM, Nievelstein RA, 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

23. Kim SY, Soo Lee S, Bumwoo P, Kim N, Kim JK, Park SH, Byun JH, Song KJ, Koo JH, Kyung Choi E, Lee MG (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

24. Ivancevic MK, Kwee TC, Takahara T, Ogino T, Hussain HK, Liu PS, Chenevert TL (2009) Diffusion-weighted MR imaging of the liver at 3.0 Tesla using TRacking Only Navigator echo (TRON): a feasibility study. J Magn Reson Imaging 30:1027–1033

25. Choi JS, Kim MJ, Chung YE, Kim KA, Choi JY, Lim JS, Park MS, Kim KW (2012) Comparison of breathhold, navigator-triggered, and free-breathing diffusion-weighted MRI for focal hepatic lesions. J Magn Reson Imaging. [Epub ahead of print]

26. Braithwaite AC, Dale BM, Boll DT, Merkle EM (2009) Short- and midterm reproducibility of apparent diffusion coefficient measurements at 3.0-T diffusion-weighted imaging of the abdomen. Radiology 250:459–465

27. Moteki T, Horikoshi H (2011) Evaluation of noncirrhotic hepatic parenchyma with and without significant portal vein stenosis using diffusion-weighted echo-planar MR on the basis of multiple-perfusion-components theory. Magn Reson Imaging 29:64–73

28. Hjortdal VE, Christensen TD, Larsen SH, Emmertsen K, Pedersen EM (2008) Caval blood flow during supine exercise in normal and Fontan patients. Ann Thorac Surg 85:599–603

29. Hollingsworth KG, Lomas DJ (2006) Influence of perfusion on hepatic MR diffusion measurement. NMR Biomed 19:231–235

30. Santelli C, Nezafat R, Goddu B, Manning WJ, Smink J, Kozerke S, Peters DC (2011) Respiratory bellows revisited for motion compensation: preliminary experience for cardiovascular MR. Magn Reson Med 65:1097–1102

31. Koh DM, Thoeny HC (2010) Diffusion weighted MR imaging applications in the body. Springer, Berlin, Heidelberg

32. Colagrande S, Pasquinelli F, Mazzoni LN, Belli G, Virgili G (2010) MR-diffusion weighted imaging of healthy liver parenchyma: repeatability and reproducibility of apparent diffusion coefficient measurement. J Magn Reson Imaging 31:912–920

33. Vandecaveye V, De Keyzer F, Dirix P, Lambrecht M, Nuyts S, Hermans R (2010) Applications of diffusion-weighted magnetic resonance imaging in head and neck squamous cell carcinoma. Neuroradiology 52:773–784

34. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307–310

35. Sasaki M, Yamada K, Watanabe Y, Matsui M, Ida M, Fujiwara S, Shibata E (2008) Variability in absolute apparent diffusion coefficient values across different platforms may be substantial: a multivendor, multi-institutional comparison study. Radiology 249:624–630

36. Dong TM, Sun XJ, Xu X, Dai J, Zheng W, An NY, Wang XQ (2012) An experimental study on rabbit liver by 3.0 Tesla magnetic resonance diffusion-weighted imaging. Acta Acad Med Sin 34:51–57

37. 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 Phy 20:205–211