Abdominal Radiology (2020) 45:3716–3729
https://doi.org/10.1007/s00261-020-02538-y

HEPATOBILIARY

Evaluation of simultaneous-multislice diffusion-weighted imaging of liver at 3.0 T with different breathing schemes

Yigang Pei[1,2,3] · Siming Xie[1] · Wenzheng Li[1] · Xianjing Peng[1] · Qin Qin[4,5] · Qian Ye[1] · Mengsi Li[1] · Jiaxi Hu[1] · Jiale Hou[1] · Guijing Li[6] · Shuo Hu[2,7]

Published online: 30 April 2020

Abstract

Purpose To obtain the optimal simultaneous-multislice (SMS)—accelerated diffusion-weighted imaging (DWI) of the liver at 3.0 T MRI by systematically estimating the repeatability of apparent diffusion coefficient (ADC), signal-to-noise ratio (SNR) and image quality of different breathing schemes in comparison to standard DWI (STD) and other SMS sequences.

Methods In this institutional review board-approved prospective study, hepatic DWIs ($b = 50$, 300, 600 s/mm^2) were performed in 23 volunteers on 3.0 T MRI using SMS and STD with breath-hold (BH-SMS, BH-STD), free-breathing (FB-SMS, FB-STD) and respiratory-triggered (RT-SMS, RT-STD). Reduction of scan time with SMS-acceleration was calculated. ADC and SNR were measured in nine anatomic locations and image quality was assessed on all SMS and STD sequences. An optimal SMS-DWI was decided by systematically comparing the ADC repeatability, SNR and image quality among above DWIs.

Results SMS-DWI reduced scan time significantly by comparison with corresponding STD-DWI (27 vs. 42 s for BH, 54 vs. 78 s for FB and 42 vs. 97 s for RT). In all DWIs, BH-SMS had the greatest intraobserver agreement (intraclass correlation coefficient (ICC): 0.920–0.944) and good interobserver agreement (ICC: 0.831–0.886) for ADC measurements, and had the best ADC repeatability (mean ADC absolute differences: $0.046–0.058 \times 10^{-3}$mm^2/s, limits of agreement (LOA): $0.010–0.013 \times 10^{-3}$mm^2/s) in nine locations. BH-SMS had the highest SNR in three representative sections except for RT-STD. There were no significant differences in image quality between BH-SMS and other DWI sequences (median BH-SMS: 4.75, other DWI: 4.5–5.0; $P > 0.0.5$).

Conclusion BH-SMS provides considerable scan time reduction with good image quality, sufficient SNR and highest ADC repeatability on 3.0 T MRI, which is thus recommended as the optimal hepatic DWI sequence for those subjects with adequate breath-holding capability.

Keywords Optimal SMS-DWI · STD-DWI · Various breathing schemes · ADC repeatability · Liver DWIs

Abbreviations

SMS	Simultaneous-multislice
DWI	Diffusion-weighted imaging
STD	Standard DWI
ADC	Apparent diffusion coefficient
SNR	Signal-to-noise ratio
BH	Breath-hold
FB	Free-breathing
RT	Respiratory-triggered

✉ Shuo Hu
hushuo2019@163.com

Yigang Pei
xypyg0731@163.com

[1] Department of Radiology, Xiangya Hospital, Central South University, Changsha 410008, Hunan, People's Republic of China

[2] Key Laboratory of Biological, Nanotechnology of National Health Commission, Changsha 410008, Hunan, People's Republic of China

[3] Xiangya Hospital, Central South University, Changsha 410008, Hunan, People's Republic of China

[4] Department of Radiology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

[5] Kirby Center, Kennedy Krieger Institute, Baltimore, MD 21205, USA

[6] Simens Healthcare Ltd, Guangzhou Branch, Guangzhou 510620, People's Republic of China

[7] PET-CT Center, Xiangya Hospital, Central South University, Changsha 410008, Hunan, People's Republic of China

ICC Intraclass correlation coefficient
LOA Limits of agreement

Introduction

Diffusion-weighted imaging (DWI) has been widely performed to detect and characterize liver lesions, monitor and predict the treatment response of hepatic tumors [1–3].In DWI, the apparent diffusion coefficient (ADC), representing mobility of water molecules within the tissue, is a quantitative parameter as imaging and therapeutic biomarker for patients with liver diseases [4]. Unfortunately, ADCs can be interfered by image artifacts from breathing and cardiac motion [5]. Respiratory-triggering technique can reduce these artifacts in DWI protocols, but it is at the cost of rather long and also uncertain scan times, which in turn can markedly increase patient's uncomfortableness and decrease image quality [6]. Therefore, a DWI technique with scan time reduction and fewer artifacts is desired to obtain good image quality and reliable ADC quantification.

Recently, simultaneous-multislice (SMS) imaging was introduced using simultaneous multi-band radiofrequency excitation with echo-planar imaging readout and blipped-CAIPIRINHA technique that can reduce the scan time and increase patient's comfort. In SMS, the acceleration factor can be achieved using controlled aliasing introduced using the blipped-CAIPIRINHA technique, which was 2 in our study [7–9]. Compared to standard DWI (STD), SMS-DWI has been shown with similar image quality, but 40–70% shorter scan time under the same conditions [6, 10]. Currently, different breathing schemes have been used in SMS-DWI, including breath-hold (BH-SMS), free breathing (FB-SMS) and respiratory-triggered (RT-SMS). Among them, an optimal SMS-DWI would be able to improve workflow in daily routine, especially in the face of tight examination schedules, which presents a reliable ADC measurement, good image quality and sufficient SNR. To date, only a few studies on STD-DWI have reported the ADC repeatability with different breathing schemes [11–14]. Kandpal and Kwee et al. found that RT-STD could provide better image quality and SNR than BH-STD and FB-STD [11, 14]. LEE et al. demonstrated that echocardiography triggering STD was more effective in improving the repeatability of ADC and IVIM measurement than FB-STD and RT-STD [12]. Chen et al. considered FB-STD to have better repeatability and shorter acquisition time compared to that of BH, RT and NT-STD [13]. However, the DWI repeatability was just evaluated with different STD techniques.

With the growing clinical interests of hepatic SMS-DWI [15, 16], the aim of our study was to evaluate SMS-DWIs with three different breathing schemes (BH, FB and RT) at 3.0 T MRI and obtain an optimal respiration method for SMS sequences with respect to ADC repeatability, signal-to-noise ratio (SNR) and image quality, in comparison with STD and other SMS sequences.

Materials and methods

Subjects

All volunteers gave written informed consent to this prospective study, which was approved by the local Institutional Review Board. 23 healthy young volunteers with similar age (juniors in a medical college, mean 21 years, range 20–22 years) were enrolled (12 males, 11 females). In our study, the inclusion criteria included normal liver function tests, no history of alcohol and drug abuse, viral hepatitis, and prior abdominal surgery.

MR imaging protocol

This MRI study were carried out on a Siemens Prisma 3.0 T whole-body MRI system (80 mT/m maximum gradient strength, Siemens Healthcare, Erlangen, Germany) with an 18-channel anterior surface body coil in combination with 12 elements of a 32-channel spine coil. Each subject was scanned twice with a short break (5 min) and repositioned between two identical sessions. This allowed the quantification of intersession variability. Within each session, each subject underwent SMS- and STD-DWIs in the axial view with three different breathing acquisition schemes, with the order as RT-STD, RT-SMS, BH-STD, BH-SMS, FB-STD and FB-SMS. Three b-values of 50, 300, and 600 s/mm^2 were sampled in three orthogonal diffusion directions (three-scan trace). BH technique required the participants to hold the breath on the end-expiratory (BH-STD: twice breath-holds, BH-SMS: one end-expiration), FB and RT adopted an air-filled pressure sensor to measure respiratory-induced pressure changes with a respiration belt around the subjects. Their scan parameters were kept as close as possible and the detailed parameters of all SMS and STD sequences were summarized in Table 1. In our study, SMS was a commercially available sequence, which was achieved based on STD-DWI with the same acceleration factor. However, other acceleration factors = 2 were used in three SMS sequences (SMS AF = 2). For SMS and STD sequences, the acquisition volumes were large enough to completely cover the whole liver. The fat suppression was achieved with spectral adiabatic inversion recovery and the minimum time of echo (TE) was applied in all DWIs without any filter. Additionally, a minimum TR was applied in BH-SMS for saving scan time to meet the requirement of breath-hold ability. A *k*-space-based parallel imaging technique, generalized

Table 1 The summarized parameters of all DWI sequences

Parameters	RT-STD	FB-STD	BH-STD	RT-SMS	FB-SMS	BH-SMS
TR (ms)	1200	3000	1400	1500	2000	1800
TE (ms)	46	45	46	49	49	49
Slices	24					
Slice thickness (mm)	5					
Slice gap (mm)	1.5					
Bandwidth	2332					
Echo spacing (msec)	0.49					
Voxel size (mm^3)	1.5 × 1.5 × 5.0					
iPAT	GRAPPA 2					
ACS model	Separate					
ACS lines	28					
Reconstruction method	GRAPPA					
Concatenations	3	1	1	1	1	1
NSA	2/2/3	2/2/3	1/1/1	2/2/3	2/2/3	1/1/1
b-value (s/mm^2)	50,300,600					
SMS AF	–	–	–	2	2	2
Coil elements	30 (18 body, 12 spine)					
Net scan time	97 s	78 s	21 s	42 s	54 s	27 s

The parameters are the same for all DWI methods in those columns with one number. For BH-STD, the total acquisition time was 42 s due to the twice scans with breath-hold. However, for BH-SMS, just once scan was carried out on the end-expiratory for all volunteers

STD standard DWI sequence, *SMS* simultaneous-multislice DWI sequence, *RT* respiratory-triggered, *FB* free-breathing, *TR* time of repetition, *TE* echo time, *iPAT* integrated parallel acquisition technique, *ACS* accelerationrate auto-calibration signal, *GRAPPA* generalized autocalibrating partially parallel acquisition, *NSA* the numbers of signal acquisition for 50, 300, 600 s/mm^2, respectively, *SMS AF* simultaneous-multislice acceleration factor

autocalibrating partially parallel acquisition (GRAPPA), was used and the scan time was recorded in our study (Table 1).

Image analysis

All SMS- and STD-DWI images were transferred to a workstation and apparent diffusion coefficient (ADC) maps were calculated by image post processing (Syngo. via, VB10, Siemens Healthcare). In our study, ADC and SNR were independently measured by two radiologists (Y.P. and W.L., readers 1 and 2 with 5 and 10 years of clinical experience in liver MRI, respectively), and the image quality was assessed by another two radiologists (J.H. and S.X., reader 3 and 4 with 4 and 3 years of clinical experience in liver MRI). ADC maps were calculated by the log-linear fitting algorithm with three different b-values (b = 50, 300, 600 s/mm^2) according to the following equation:

$$I_{\text{Trace}} = I_{0e}^{-ADC*b} = I_{0e}^{-b*(D1+D2+D3)/3} = (I1 * I2 * I3)^{1/3} \quad (1)$$

where *I*1, *I*2, *I*3 are the measured diffusion-weighted images with three orthogonal gradient directions and *D*1, *D*2, *D*3 are the corresponding diffusion coefficients. Nine regions of interests (ROIs) were drawn in three representative sections (superior, central and inferior) on each ADC map. The central section was defined by the main stem of right portal vein. The superior and inferior sections were six consecutive section levels above or below the central section. ROIs were positioned only in right liver lobes because ADC and SNR values in left lobes were found unreliable due to cardiac motion artifacts [13, 17]. ROIs were placed with reference to the anatomic landmarks, such as portal, hepatic veins and their main branch, which were easy to discriminate on $b = 50$ s/mm^2 images. Thus, nine circular ROIs were manually positioned in three representative slices on the $b = 50$ s/mm^2 images and then copied them to ADC maps (ADC measurements) and $b = 600$ s/mm^2images (SNR measurements) (Fig. 1). Nine ROIs were kept the same size ($0.8mm^2$) on 12 DWIs. All intrahepatic ROIs were positioned with a distance of at least 5 mm to the organ capsule and kept away from macroscopically visible vessels and bile ducts. For each ROI, the maximum, minimum, mean values and their standard deviation were recorded. Thus, a total of 108 ADCs were collected for each volunteer (three ROIs per section, three sections, two repeated series and six techniques). ADCs were measured twice in a 2-week period for reader 1 and once in a week for reader 2. $b = 600$ s/mm^2 has been recommended as an optimal *b*-value in DWI for differentiation

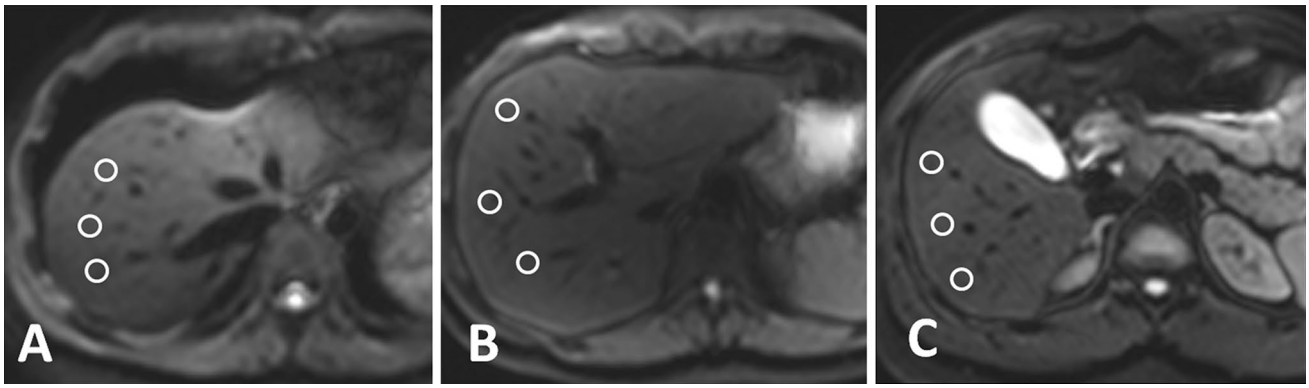


Fig. 1 A diagram of ROIs manually drawn in superior, central and inferior representative slices in the first session with BH-SMS sequence

of abdominal lesions [18]. Therefore, the highest $b = 600$ s/mm^2 images were used to calculate the estimated SNR of each representative section with the formula:

$$\mathrm{SNR} = \mathrm{SI}/\mathrm{SD} \quad (2)$$

where SI is the mean signal in the three ROIs of the same sections, and SD is the standard deviation of the signal from ROIs with the same size placed at nearby background in the corresponding section. In all subjects, SNRs were measured once in a 1-week period for reader 1and reader 2, which were recorded with the maximum, minimum, mean value and their standard deviation. Assessment of image quality was performed on the first DWI series (RT- STD, RT-SMS, BH-STD, BH-SMS, FB-STD and FB-SMS) for each participator. Image quality was rated on a five-point Likert scale with respect to following five aspects (L1–L5). L1: the sharpness of liver edge (5 = sharp and no blurring, 4 = mild blurring, 3 = moderate blurring, 2 = Severe blurring, 1 = non diagnostic), L2: severity of artifacts (5 = no artifacts, 4 = mild artifacts, 3 = moderate artifacts, 2 = severe artifacts, 1 = non diagnostic), L3: imaging quality of the liver dome (5 = excellent, 4 = good and not affecting interpretation, 3 = moderate and potentially affecting interpretation, 2 = poor and definitely affecting interpretation, 1 = non diagnostic), L4: image quality of segments caudal of the liver dome (5 = excellent, 4 = good and not affecting interpretation, 3 = moderate and potentially affecting interpretation, 2 = poor and definitely affecting interpretation, 1 = non diagnostic), L5: Overall image quality (5 = excellent, 4 = good and not affecting interpretation, 3 = moderate and potentially affecting interpretation, 2 = poor and definitely affecting interpretation, 1 = non diagnostic) [19]. Image quality scores were recorded in ADC maps and three DWI images with $b = 50$, 300 and 600 s/mm^2 for each volunteer, respectively.

Statistical analysis

All statistical analyses were performed using SPSS (SPSS for Windows, version 13.0; SPSS, Chicago, IL). ADC and SNR were expressed as mean and SD (minimum, maximum). Ratings were expressed as median (minimum, maximum). In our study, *P*-values of less than 0.05 were considered statistically significant. The interobserver agreement was analyzed by calculating weighted kappa coefficients (quadratic weighting) with kappa values of 0.01–0.20 representing slight agreement, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect agreement.

Intra- and interclass correlation coefficients (ICCs) were used to evaluate the intra- and interobserver agreement of ADC measurements [20, 21]. Intraobserver ICC was compared in two ADC measurements of the reader 1, and interobserver ICC was calculated between the first ADC measurements of reader 1 and the ADC measurements of reader 2. An ICC greater than 0.8 was indicative of good agreement. In addition, the repeatability of ADC measurements was evaluated with the Bland–Altman method [22].The mean absolute difference (bias) and the 95% confidence interval of the mean difference (limits of agreement, LOA) between the first and second DWI series were compared [16]. The differences of SNR between the six DWI techniques were assessed by using two-way classification analysis of variance. The Bonferroni method was used to adjust for multiple comparisons [23]. Image quality was compared between the six sequences using Friedman test. If the Friedman test showed a statistically significant P-value, the Dunn-Bonferroni post-hoc method for all pairwise comparisons were performed [6, 24].

Results

SMS and STD with different breathing schemes were successfully carried out in all volunteers. Acquisition time reduction with the SMS techniques was achieved with a good image quality.

Scan time

To obtain DWI of 24 slices, accelerated SMS acquisition reduced scan time significantly in comparison to STD-DWI: 36% for BH scheme (27 vs. 42 s), 31% for FB (54 vs. 78 s), and 55% for RT (42 vs. 97 s) (Table 1). BH-SMS enabled the shortest scan time (27 s) to obtain whole-liver DWI among six DWIs.

The ADC measurementsand their intra- and interobserver agreement

The average ADC values of three representative sections were gained with six DWI sequences. They are more reliable for intraobserver ADC measurements when using BH (almost all $P > 0.05$) than RT and FB techniques (almost all $P < 0.05$). For example, in reader 1's two measurements (first and second) with BH-SMS, the average ADC values were $(1.074 \pm 0.140) \times 10^{-3}$ mm^2/s vs.$(1.073 \pm 0.114) \times 10^{-3}$ mm^2/s for superior section ($P = 0.912$), which yielded less variation than those with RT-SMS ($(1.043 \pm 0.144) \times 10^{-3}$ mm^2/s vs. $(1.077 \pm 0.152) \times 10^{-3}$ mm^2/s, $P = 0.022$); and FB-SMS ($(1.011 \pm 0.138) \times 10^{-3}$ mm^2/s vs. $(1.049 \pm 0.132) \times 10^{-3}$ mm^2/s, $P = 0.005$) for superior slice (Table 2). Furthermore, BH-SMS had the highest intraobserver ICCs (0.920 ~ 0.941) in six DWI sequences (Table 2). For the interobserver average ADC values, they all were not significantly different between Reader1 and Reader 2 in six DWI sequences (all $P > 0.05$). For instance, the average ADC values with BH-SMS were $(1.074 \pm 0.140) \times 10^{-3}$ mm^2/s and $(1.075 \pm 0.134) \times 10^{-3}$ mm^2/s for superior section ($P = 0.562$), $(1.059 \pm 0.122) \times 10^{-3}$ mm^2/s and $(1.053 \pm 0.125) \times 10^{-3}$ mm^2/s for the central section ($P = 0.578$), and $(1.029 \pm 0.111) \times 10^{-3}$ mm^2/s and $(1.034 \pm 0.093) \times 10^{-3}$ mm^2/s for the inferior section ($P = 0.934$) (Table 2). Furthermore, BH-SMS had a good interobserver agreement (0.831–0.886) (Table 2).

Table 2 The ADC measurements in six DWI techniques and their intra- and interobserver agreement

		ADCs measured by reader 1			ADCs measured by reader 2		ADC measurement agreement	
		First	Second	*P*-value	First	*P*-value	Intraobserver	Interobserver
RT-STD	Superior	1.041 ± 0.094	1.076 ± 0.082	< 0.0001	1.037 ± 0.103	0.623*	0.842 (0.745–0.902)	0.898 (0.835–0.937)
	Central	1.069 ± 0.011	1.102 ± 0.094	< 0.0001	1.053 ± 0.103	0.053*	0.911 (0.857–0.945)	0.891 (0.824–0.932)
	Inferior	1.068 ± 0.100	1.092 ± 0.094	0.005	1.064 ± 0.104	0.585*	0.862 (0.777–0.914)	0.875 (0.797–0.922)
FB-STD	Superior	1.089 ± 0.158	1.112 ± 0.127	0.037	1.099 ± 0.142	0.228*	0.895 (0.830–0.935)	0.907 (0.850–0.942)
	Central	1.093 ± 0.151	1.106 ± 0.131	0.176*	1.090 ± 0.156	0.775*	0.910 (0.854–0.944)	0.905 (0.847–0.941)
	Inferior	1.095 ± 0.115	1.106 ± 0.119	0.249*	1.107 ± 0.123	0.377*	0.857 (0.768–0.911)	0.866 (0.784–0.917)
BH-STD	Superior	1.038 ± 0.143	1.024 ± 0.123	0.095*	1.059 ± 0.136	0.989*	0.934 (0.893–0.959)	0.828 (0.722–0.893)
	Central	1.041 ± 0.127	1.034 ± 0.111	0.348*	1.053 ± 0.154	0.376*	0.935 (0.895–0.960)	0.813 (0.697–0.884)
	Inferior	1.042 ± 0.115	1.029 ± 0.097	0.079*	1.042 ± 0.100	0.101*	0.915 (0.863–0.948)	0.741 (0.581–0.839)
RT-SMS	Superior	1.043 ± 0.144	1.077 ± 0.152	0.022	1.052 ± 0.141	0.653*	0.835 (0.734–0.898)	0.872 (0.793–0.921)
	Central	1.077 ± 0.150	1.103 ± 0.138	0.023	1.087 ± 0.142	0.347*	0.833 (0.730–0.897)	0.908 (0.851–0.943)
	Inferior	1.073 ± 0.138	1.102 ± 0.118	0.025	1.078 ± 0.129	0.464*	0.829 (0.724–0.894)	0.872 (0.794–0.921)
FB-SMS	Superior	1.011 ± 0.138	1.049 ± 0.132	0.005	1.011 ± 0.148	0.388*	0.804 (0.684–0.879)	0.899 (0.837–0.937)
	Central	1.030 ± 0.128	1.063 ± 0.119	0.007	1.031 ± 0.102	0.929*	0.806 (0.687–0.880)	0.848 (0.755–0.906)
	Inferior	1.057 ± 0.127	1.115 ± 0.129	< 0.0001	1.065 ± 0.143	0.976*	0.798 (0.673–0.875)	0.912 (0.858–0.946)
BH-SMS	Superior	1.074 ± 0.140	1.073 ± 0.114	0.912*	1.075 ± 0.134	0.562*	0.941 (0.904–0.963)	0.886 (0.816–0.929)
	Central	1.059 ± 0.122	1.079 ± 0.113	0.003	1.053 ± 0.125	0.578*	0.944 (0.910–0.966)	0.831 (0.727–0.895)
	Inferior	1.029 ± 0.111	1.037 ± 0.094	0.219*	1.034 ± 0.093	0.934*	0.920 (0.871–0.950)	0.833 (0.731–0.897)

ADC are given in $*10^{-3}$ mm^2/s. Mean ADC values measured in different anatomical regions. Data in parentheses are 95% confidence intervals

P values were gained by using the paired *t* test to compare differences for reader1 between the first and second ADC measurement, and to compare for the first ADC measurement between reader1 and reader2

$P > 0.05$ were considered no significant difference (*)

Repeatability of ADC measurements

In the right liver lobe, the repeatability of ADC measurements in nine different anatomic locations were varied for each technique. However, the mean ADC absolute differences (bias) with BH-SMS ($0.046–0.058 \times 10^{-3}$ mm^2/s) were lower than other five DWI sequences (FB-SMS: $0.063–0.101 \times 10^{-3}$ mm^2/s, RT-SMS: $0.060–0.109 \times 10^{-3}$ mm^2/s, BH-STD: $0.047–0.068 \times 10^{-3}$ mm^2/s, FB-STD: $0.064–0.090 \times 10^{-3}$ mm^2/s, RT-STD: $0.055–0.069 \times 10^{-3}$ mm^2/s) in nine locations (Table 3). Furthermore, BH-SMS had the highest ADC measurement repeatability with the lowest LOA ($0.010–0.013 \times 10^{-3}$ mm^2/s) in all six sequences (Table 3, Fig. 2). In addition, the ADC repeatability in the central middle position was better to that in other anatomic location for all DWI techniques, yielding mean absolute differences of ADCs ± LOA (BH-SMS: $(0.056 \pm 0.011) \times 10^{-3}$ mm^2/s, FB-SMS: $(0.084 \pm 0.025) \times 10^{-3}$ mm^2/s, RT-SMS: $(0.074 \pm 0.018) \times 10^{-3}$ mm^2/s, BH-STD: $(0.053 \pm 0.010) \times 10^{-3}$mm^2/s, FB-STD: $(0.064 \pm 0.018) \times 10^{-3}$mm^2/s, RT-STD: $(0.055 \pm 0.012) \times 10^{-3}$mm^2/s) (Table 3).

SNR quantitative analysis

Compared to various protocols, BH-SMS had the highest surrogate SNR in three representative sections except for RT-STD (Fig. 3). For example, in superior representative section, the surrogate SNR was 29.81 ± 20.49 with BH-SMS measured by reader1, which was obviously higher than FB-STD (9.57 ± 5.91, $P < 0.0001$) and BH-STD (11.26 ± 7.24, $P < 0.0001$), and was slightly higher than RT-SMS (20.18 ± 13.60, $P = 0.239$) and FB-SMS (19.19 ± 9.54, $P = 0.120$). In addition, an excellent agreement were found in superior ($r = 0.95$), middle ($r = 0.94$) and inferior section ($r = 0.86$) for BH-SMS between reader 1 and reader 2 (Table 4).

Qualitative image analysis

The interobserver agreements of two readers were very good for assessing image quality in five aspects (L1–L5) and were shown in Table 5. For instance, the kappa value was 0.951 ($P < 0.0001$) for the overall image quality assessment between Reader 3 and Reader 4. A detailed summary of all image quality criteria rankings as stratified by reader3 were also presented in Table 5. There were significant difference for the severity of artifacts (L2) and overall image quality (L5) in ADC map (L2: $P < 0.0001$, L5: $P < 0.0001$), $b = 50$ images (L2: $P < 0.0001$, L5: $P = 0.027$), $b = 300$ images (L2: $P < 0.0001$, L5: $P = 0.012$), b = 600 images (L2: $P < 0.0001$, L5: $P = 0.018$), which presented differences between STD and SMS sequences by further analysis ($P < 0.05$). However, it was not significantly different between BH-SMS and other 5 DWI sequences ($P > 0.05$) (Table 5, Figs. 4, 5).

Discussion

SMS-DWI technique has been proposed in clinical applications for detecting and characterizing liver lesions with shorter examination time and similar image quality in comparisons with STD-DWI sequence [6, 19]. In this study, the scan time has been saved about 36% for BH-SMS, 31% for

Table 3 The mean absolute differences of ADC measurements and their 95%confidence intervals in nine anatomic locations with six DWI techniques

Locations	RT-STD	FB-STD	BH-STD	RT-SMS	FB-SMS	BH-SMS
Superior left (1)	0.062 (0.017)	0.069 (0.020)	0.061 (0.012)	0.104 (0.023)	0.089 (0.035)	0.058 (0.011)
Superior middle (2)	0.065 (0.013)	0.075 (0.020)	0.056 (0.012)	0.095 (0.025)	0.102 (0.029)	0.056 (0.011)
Superior right (3)	0.067 (0.016)	0.090 (0.020)	0.068 (0.011)	0.096 (0.029)	0.089 (0.024)	0.046 (0.013)
Central left (4)	0.056 (0.016)	0.059 (0.020)	0.047 (0.012)	0.109 (0.032)	0.063 (0.026)	0.048 (0.012)
Central middle (5)	0.054 (0.012)	0.064 (0.018)	0.053 (0.010)	0.074 (0.018)	0.084 (0.025)	0.056 (0.011)
Central right (6)	0.061 (0.016)	0.084 (0.020)	0.053 (0.016)	0.097 (0.028)	0.098 (0.033)	0.048 (0.012)
Inferior left (7)	0.058 (0.016)	0.076 (0.015)	0.049 (0.014)	0.101 (0.028)	0.108 (0.033)	0.049 (0.011)
Inferior middle (8)	0.069 (0.015)	0.072 (0.016)	0.053 (0.011)	0.083 (0.031)	0.090 (0.034)	0.055 (0.011)
Inferior right (9)	0.063 (0.012)	0.076 (0.018)	0.0586 (0.011)	0.060 (0.015)	0.091 (0.026)	0.049 (0.011)

The mean absolute differences of ADC Measurement were given in *10^{-3}mm^2/s, which were calculated between the first and second DW imaging series and the differences represented ADC reproducibility

The 95% confidence interval of the mean absolute differences (limits of agreement [LOAs]) were shown in the parentheses

STD standard DWI sequence, *SMS* simultaneous-multislice DWI sequence with acceleration factor 2, *RT* respiratory-triggering, *FB* free-breathing, *BH* breath-hold

FB-SMS and 55% for RT-SMS, which cuts at least one-third of scan time of the corresponding STD-DWI without noticeable artifacts [6]. BH-SMS has used the shortest scan time to acquire whole-liver DWI with a good image quality on 3.0 T MRI, which is helpful to improve the work effectiveness and relieve the workflow tension.

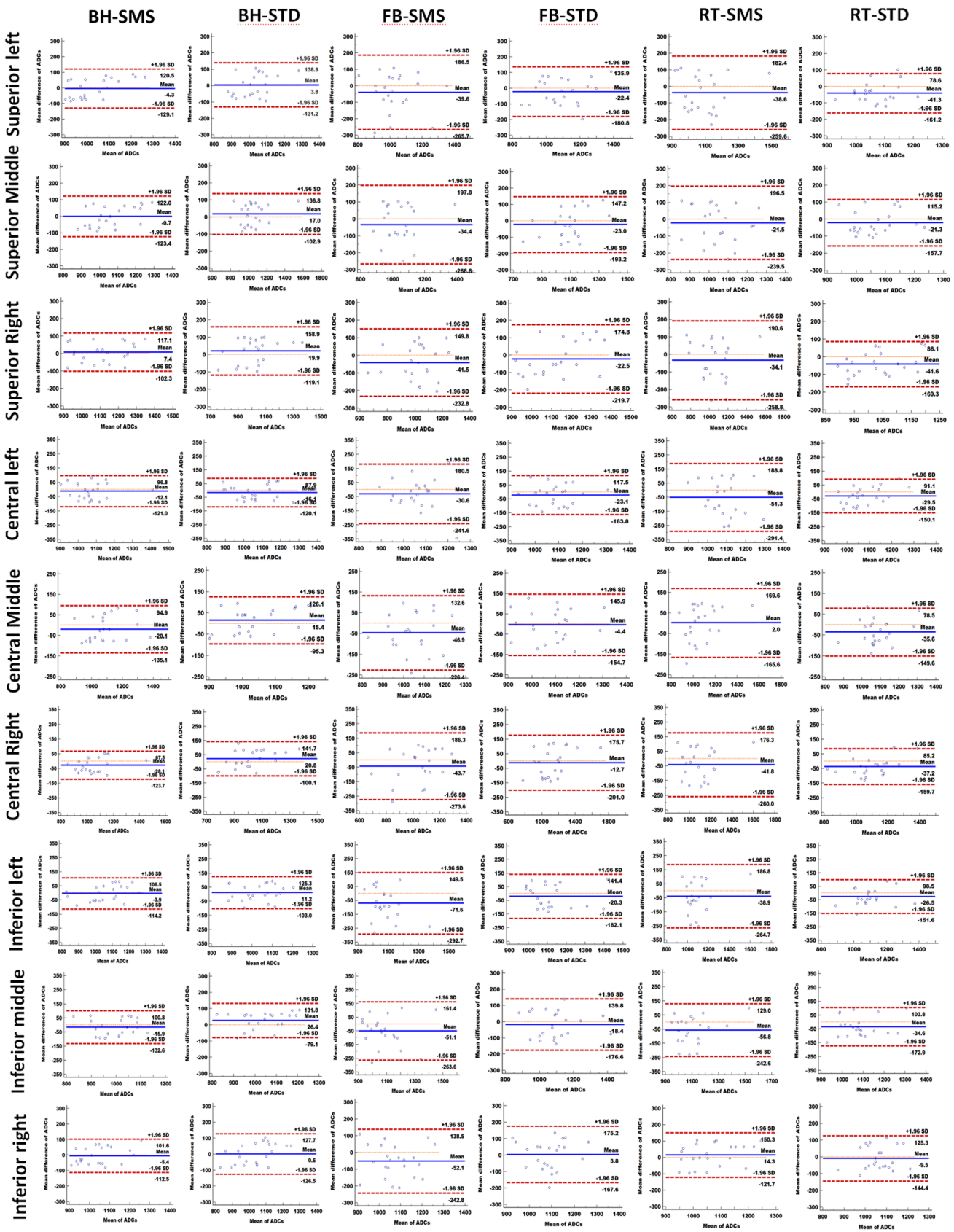
The mean ADC values were $0.9–1.1 \times 10^{-3}$ mm^2/s for three representative sections in all DWIs, which were near the low-end of the literature values [6, 12, 15]. The mean intraobserver ADC values with the BH technique were more reliable than RT and FB techniques. This might be associated with the mitigation of respiratory movement when using the BH scheme for those young volunteers who have the adequate breath-holding capability, which allows obtaining images with clear anatomic landmarks and placing ROIs in consistent locations. Therefore, BH-SMS offered higher repeatability for ADC measurements than RT-SMS and FB-SMS and the highest ICC (0.920–0.941) in intraobserver ADC measurements for three representative sections. For the interobserver average ADC measurements, BH-SMS had no significant difference between Reader 1 and Reader 2 and the interobserver agreement was good (0.831–0.886), which suggested that BH-SMS had sufficient reliability and repeatability for assessing ADC measurements.

In our results, the ADC repeatability in different locations was varied for each technique. However, it was the highest in the central middle location for all six DWI techniques. This finding was in agreement with the result of Chen et al. [13]. In addition, BH-SMS had the greatest ADC measurement repeatability in all six sequences yielded the smallest ADC mean differences and LOA. Taron. et. al reported that breathing schemes could affect the absolute ADC values (bias) [6]. In our study, the participators are young students with a good ability to hold their breaths and have fewer motion artifacts with BH technique than RT and FB techniques. Thus, the absolute differences of ADC values with the BH scheme were lower than those with RT and FB schemes. In addition, BH-SMS has a shorter scan time than BH-STD, which helps decrease the motion artifacts and achieve low ADC mean differences and LOA. Furthermore, our results found that all the LOAs were around 20–30% of the mean ADC value. This finding was in agreement with the findings of Chen and Kim et al. [13, 25]. Thus, for treatment response evaluation with ADC as a biomarker, we also recommend a confident threshold to be at least 20%, and the same DWI acquisition technique should be applied for all baseline and follow-up studies.

In our study, the greatest SNR was obtained for each DWI due to the application of a minimum TE in six DWIs, which could be affected by other factors such as the volume of the voxel, the bandwidth, and the numbers of signal acquisition (NSA). These were kept the same as much as possible in SMS and STD. In clinical work, NSA = 2, 2, 3 is usually used to acquire a signal for 50, 300, 600 s/mm^2, respectively, in FB and RT- DWI. However, NSA = 1, 1, 1 was performed in BH-DWI due to the maximum saving of scanning time without exceeding the upper limit of the breath-hold ability, which could yield a good DW image quality for the patients. In the quantitative analysis, the mean SNR of the liver in BH-SMS was significantly higher than almost all DWI sequences, except for RT-STD at $b = 600$ s/mm^2. RT-STD had approximately twice the SNR compared to BH-SMS, the reason may be due to the application of high NSA for improving SNR and 3 concatenations for gaining good DW image quality for RT-STD. But RT-STD takes at least twice the scan duration than BH-SMS, which is not preferred for clinical examinations. Interestingly, BH-SMS had distinctly higher SNR than FB-STD and BH-STD, and was slightly higher than RT-SMS and FB-SMS due to the fewer motion artifacts caused by the BH technique, which easily integrates with SMS. Also, an excellent agreement with BH-SMS was found for three representative sections. Therefore, BH-SMS can achieve a sufficient, reliable and reproducible SNR, which is a good choice for acquiring liver DWI to assess liver diseases at 3.0 T MRI.

Visual assessment is helpful for disease detection and characterization by observing the different signal attenuation on DWI. In our study, all SMS- and STD-DWI methods obtained good image quality. However, RT- and FB-based SMS-DWI were rated significantly lower than their corresponding STD-DWI techniques in terms of severity of artifacts and overall image quality, including all images ($b = 50$, 300, 600 s/ mm^2) and ADC maps ($P < 0.05$). Previous work conducted on 1.5 T reported that there were no significant differences in the overall image quality with RT- and FB-acquisitions between STD and SMS techniques [6], which was inconsistent with our findings. This may be due to the higher sensitivity of SMS technique to the breathing motion at 3.0 T than 1.5 T MRI. In addition, a good image quality was obtained with BH-SMS in our study, which was not significantly different with STD-DWI sequences (all P-value > 0.05). This may be attributed to that BH-DWI was performed during a breath-hold, which attempted to freeze motion, and the image quality can improve with less breath motion artifacts, especially for those young volunteers with a good respiration holding ability. Kartalis et al. [26] reported that the image quality of BH scheme was similar with RT scheme on DWI for detecting pancreatic ductal adenocarcinoma. Therefore, BH-SMS is preferable to the clinical routine when the patients have an excellent breath-holding ability.

There are several limitations in our study. First, the DWI assessments were performed only among young healthy volunteers. These subjects had a good breath-holding ability and gained the best ADC repeatability with BH-SMS. It suggested that BH-SMS can be an optimal hepatic DWI


BH-SMS
BH-STD
FB-SMS
FB-STD
RT-SMS
RT-STD
Superior left
Superior Middle
Superior Right
Central left
Central Middle
Central Right
Inferior left
Inferior middle
Inferior right
Mean difference of ADCs
Mean of ADCs
+1.96 SD
Mean
-1.96 SD

◂**Fig. 2** The repeatability of ADC measurements by Bland–Altman method with different DWI sequences in nine different anatomic locations. For BH technique, BH-SMS was slightly lower than BH-STD in mean ADC differences and LOA (The first and second lane). For FB and RT schemes, SMS sequence was higher than corresponding STD in mean ADC differences and LOA (from the third to sixth lane). However, BH-SMS had lower mean ADC differences and LOA than FB-STD and RT-STD (The first, fourth and sixth lane). They all suggested that BH-SMS had the highest ADC measurement repeatability in six DWI sequences

sequence. The diagnostic performance of the BH-SMS sequence warrant further studies for detecting hepatic lesions in clinical settings. Second, the TR were different among various DWI sequences, which may result into a reduced signal due to T1 saturation effects when use a smaller TR, especially in applying high *b*-values. Lastly, the DWI were assessed only in right lobes because the image quality of left lobes could be affected by the respiration motion, heart beat and gastrointestinal peristalsis. DWI study in liver left lobes should be carried out in the future.

In conclusion, the liver DWI using the BH-SMS technique provided considerable scan time reduction, comparable image quality, sufficient SNR and highest ADC repeatability comparable to STD and other two SMS techniques on 3.0 T MRI. Based on the results presented here, BH-SMS is recommended as the optimal hepatic DWI sequence for subjects with adequate breath-holding capability, which could improve work efficiency, especially in the condition of busy examination schedules in the daily routine.

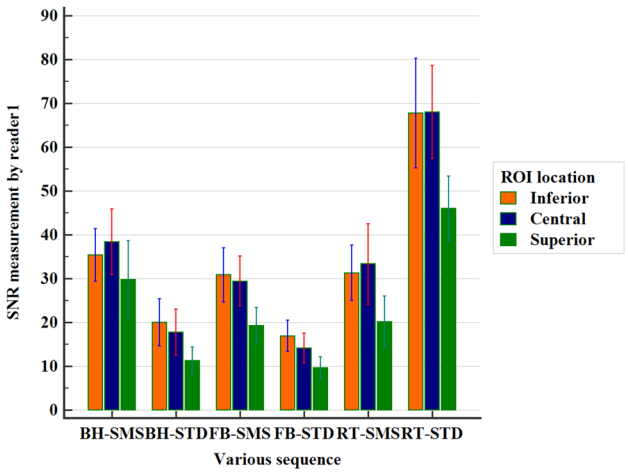


Fig. 3 SNR measurements of three representative sections by reader1with six DWI sequences. The SNR measurements from different ROI using BH-SMS were higher than the results using other sequences except RT-STD

Table 4 Signal-to-noise-ratio (SNR) of different representative sections on the right hepatic lobe

Sequences	Superior section			Central section			Inferior section		
	Reader 1	Reader 2	Agreement	Reader 1	Reader 2	Agreement	Reader 1	Reader 2	Agreement
RT-STD (1)	45.96 ± 17.00	48.98 ± 20.51	0.92 (0.81–0.97)	68.05 ± 24.59	64.72 ± 22.8	0.95 (0.88–0.98)	67.75 ± 29	71.73 ± 29.04	0.96 (0.90–0.98)
FB-STD (2)	9.57 ± 5.91	9.12 ± 6.18	0.98 (0.96–0.99)	14.14 ± 7.79	12.66 ± 6.48	0.93 (0.84–0.97)	16.93 ± 8.36	16.64 ± 10.8	0.95 (0.88–0.98)
BH-STD (3)	11.26 ± 7.24	10.12 ± 8.00	0.97 (0.93–0.99)	17.73 ± 12.16	16.77 ± 13.38	0.97 (0.94–0.99)	19.98 ± 12.52	19.39 ± 13.96	0.98 (0.95–0.99)
RT-SMS (4)	20.18 ± 13.60	18.17 ± 10.78	0.83 (0.59–0.93)	33.34 ± 21.27	31.57 ± 21.27	0.97 (0.92–0.99)	31.30 ± 14.72	31.80 ± 18.83	0.92 (0.91–0.97)
FB-SMS (5)	19.19 ± 9.54	19.02 ± 11.18	0.93 (0.83–0.97)	29.43 ± 13.15	26.67 ± 14.81	0.84 (0.63–0.93)	30.83 ± 14.26	29.35 ± 15.76	0.86 (0.67–0.94)
BH-SMS (6)	29.81 ± 20.49	29.14 ± 19.48	0.95 (0.88–0.98)	38.37 ± 17.36	37.13 ± 21.07	0.94 (0.85–0.97)	35.33 ± 13.84	34.90 ± 16.69	0. 86 (0.66–0.94)
P-value for comparisons in reader 1									
Overall	$P < 0.0001$			$P < 0.0001$			$P < 0.0001$		
Pairwise	6	1 $P = 0.001$ *		6	1 $P < 0.0001$*		6	1 $P < 0.0001$*	
		2 $P < 0.0001$*			2 $P < 0.0001$*			2 $P = 0.003$*	
		3 $P < 0.0001$*			3 $P = 0.001$ *			3 $P = 0.025$*	
		4 $P = 0.239$			4 $P = 1.000$			4 $P = 1.000$	
		5 $P = 0.120$			5 $P = 1.000$			5 $P = 1.000$	
	5	1 $P < 0.0001$*		5	1 P < 0.0001*		5	1 $P < 0.0001$*	
		2 $P = 0.239$			2 $P = 0.042$ *			2 $P = 0.122$	
		3 $P = 0.694$			3 $P = 0.319$			3 $P = 0.601$	
		4 $P = 1.000$			4 $P = 1.000$			4 $P = 1.000$	
	4	1 $P < 0.0001$*		4	1 $P < 0.0001$*		4	1 $P < 0.0001$*	
		2 $P = 0.120$			2 $P = 0.003$ *			2 $P = 0.062$	
		3 $P = 0.380$			3 $P = 0.034$			3 $P = 0.347$	
	3	1 $P < 0.0001$*		3	1 $P < 0.0001$*		3	1 $P < 0.0001$*	
		2 $P = 1.000$			2 P = 1.000			2 $P = 1.000$	
	2 *	1 $P < 0.0001$		2	1 $P < 0.0001$*		2	1 $P < 0.0001$*	

STD standard DWI sequence, *SMS* simultaneous-multislice DWI sequence with acceleration factor 2, *RT* respiratory-triggering, *FB*: free-breathing, *BH* breath-hold

*Statistical significance. good agreement: more than 0.8.SNR (signal-to-noise-ratio) was measured in different anatomical regions with $b = 600$ s/mm^{2}images.Mean values of SNR with the standard deviation and 95%confidence interval in parentheses are shown. The agreement betweenreader1 and reader 2 was given with 95% confidence interval in parentheses. At the bottom, *P*-values for the overall comparisons using Friedman test are given in reader1. If the Friedman test found a statistical significant *P*-value, additional *P*-value was presented with all pairwise comparisons by the Dunn–Bonferroni post-hoc test

Table 5 The evaluation of image quality by reader 3 and the agreement of image quality between reader 3 and reader 4

Criteria	RT-STD (1)	FB-STD (2)	BH-STD (3)	RT-SMS (4)	FB-SMS (5)	BH-SMS (6)	P-values for comparisons		Kappa value
							Overall	Pairwise	
L1: sharpness of liver									0.989 ($P<0.0001$)
ADC map	5 (4,5)	5 (3,5)	5 (4,5)	5 (3,5)	5 (3,5)	5 (4,5)	0.068		
b50	5 (4,5)	5 (4,5)	5 (4,5)	5 (4,5)	5 (4,5)	5 (4,5)	0.469		
b300	5 (4,5)	5 (4,5)	5 (4,5)	5 (4,5)	5 (4,5)	5 (4,5)	0.649		
b600	5 (4,5)	5 (4,5)	5 (5,5)	5 (4,5)	5 (4,5)	5 (4,5)	0.079		
L2: severity of artifacts									0.951 ($P<0.0001$)
ADC map	4 (3,5)	4 (4,5)	4 (4,5)	3 (3,4)	4 (3,4)	4 (3,4)	<0.0001*	2–4: 0.002; 3–4: 0.005;	
b50	5 (4,5)	5 (4,5)	5 (4,5)	4 (3,5)	4 (4,5)	4 (4,5)	<0.0001*	3–4: 0.049;	
b300	5 (4,5)	5 (4,5)	5 (4,5)	4 (3,5)	4 (4,5)	4 (4,5)	<0.0001*	2–5: 0.023; 3–4: 0.027;	
b600	5 (4,5)	5 (4,5)	5 (4,5)	4 (3,5)	4 (4,5)	4 (4,5)	<0.0001*	3–4: 0.011	
L3: imaging quality of the liver dome									0.989 ($P<0.0001$)
ADC map	5 (3,5)	5 (4,5)	5 (4,5)	5 (4,5)	5 (4,5)	5 (4,5)	0.057		
b50	5 (3,5)	5 (4,5)	5 (4,5)	5 (3,5)	5 (4,5)	5 (4,5)	0.156		
b300	5 (3,5)	5 (4,5)	5 (4,5)	5 (3,5)	5 (4,5)	5 (4,5)	0.318		
b600	5 (3,5)	5 (4,5)	5 (4,5)	5 (3,5)	5 (4,5)	5 (4,5)	0.071		
L4: Image quality of segments caudal of the liver dome									0.967 ($P<0.0001$)
ADC map	5 (3,5)	5 (4,5)	5 (4,5)	5 (4,5)	5 (3,5)	5 (3,5)	0.502		
b50	5 (4,5)	5 (4,5)	5 (5,5)	5 (4,5)	5 (3,5)	5 (4,5)	0.222		
b300	5 (4,5)	5 (4,5)	5 (5,5)	5 (4,5)	5 (3,5)	5 (4,5)	0.109		
b600	5 (4,5)	5 (5,5)	5 (5,5)	5 (4,5)	5 (3,5)	5 (4,5)	0.184		
L5: Overall image quality									0.951 ($P<0.0001$)
ADC map	4 (3,5)	5 (3,5)	5 (4,5)	4 (3,5)	4 (3,5)	5 (3,5)	<0.0001*	3–4:0.005; 3–5:0.001; 1–3:0.049	
b50	5 (4,5)	5 (4,5)	5 (4,5)	4.5 (3,5)	5 (4,5)	5 (4,5)	0.027*	3–4:0.018	
b300	5 (4,5)	5 (4,5)	5 (4,5)	4 (3,5)	4 (4,5)	5 (4,5)	0.012*	3–4:0.036;	
b600	5 (4,5)	5 (4,5)	5 (5,5)	5 (3,5)	5 (4,5)	5 (4,5)	0.018*	3–4: 0.029	

The findings of assessed image quality by reader 3 was presented and compared with RT-STD, FB-STD, BH-STD, RT-SMS, FB-SMS and BH-SMS. The agreement of image quality between reader 3 and reader 4 was shown with Kappa value

Median (min, max) values for the mean image quality ratings calculated are given separately for the three different b-values as well as for the ADC map. 5 = excellent, 4 = good, 3 = moderate, 2 = fair, 1 = nondiagnostic

STD standard DWI sequence, *SMS* simultaneous-multislice DWI sequence with acceleration factor 2, *RT* respiratory-triggering, *FB* free-breathing

*Statistical significance On the right, *P*-values for the overall comparisons using Friedman test are given. If the Friedman test showed a statistical significant *P*-value, the Dunn–Bonferroni post-hoc test for allpairwise comparisons was performed. *P*-values were presented only when they were less than 0.05

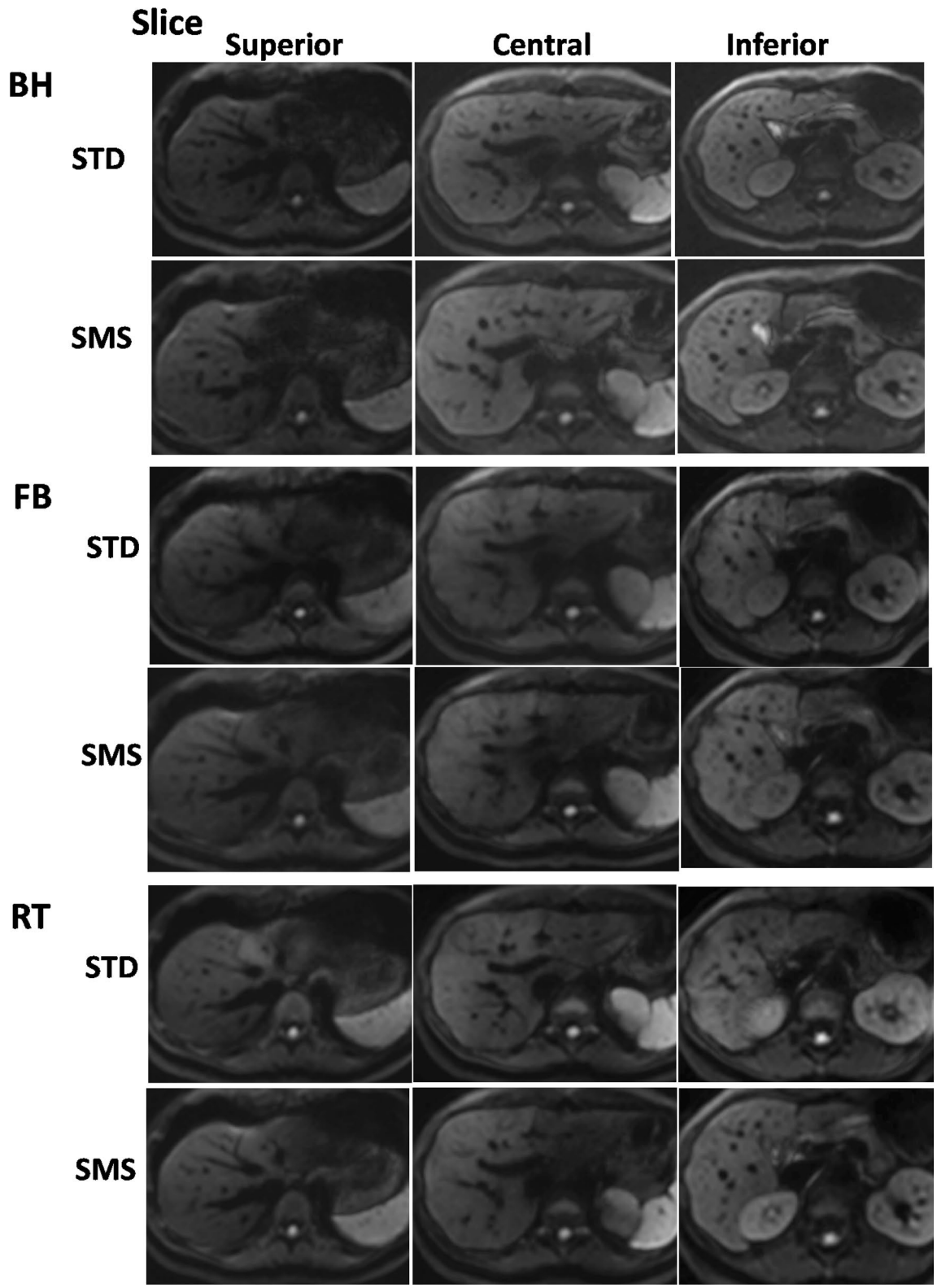


Fig. 4 Comparisons of image quality of STD- and SMS-DWI with $b=600$ s/mm^2 using different breathing schemes (*BH* breath-hold, *FB* free-breathing, *RT* respiratory-triggering). Image quality of BH-SMS did not significantly different with other five DWI acquisitions in three representative sections (superior, central and inferior), including the sharpness of liver edge (BH-SMS: 5 (4,5), other five DWIs: 5 (4,5)–5 (5,5), $P>0.05$), severity of artifacts (BH-SMS: 4 (4,5), other five DWIs: 4 (3,5)–5 (4,5), $P>0.05$) and Overall image quality (BH-SMS: 4 (4,5), other five DWIs: 5 (4,5), $P>0.05$)

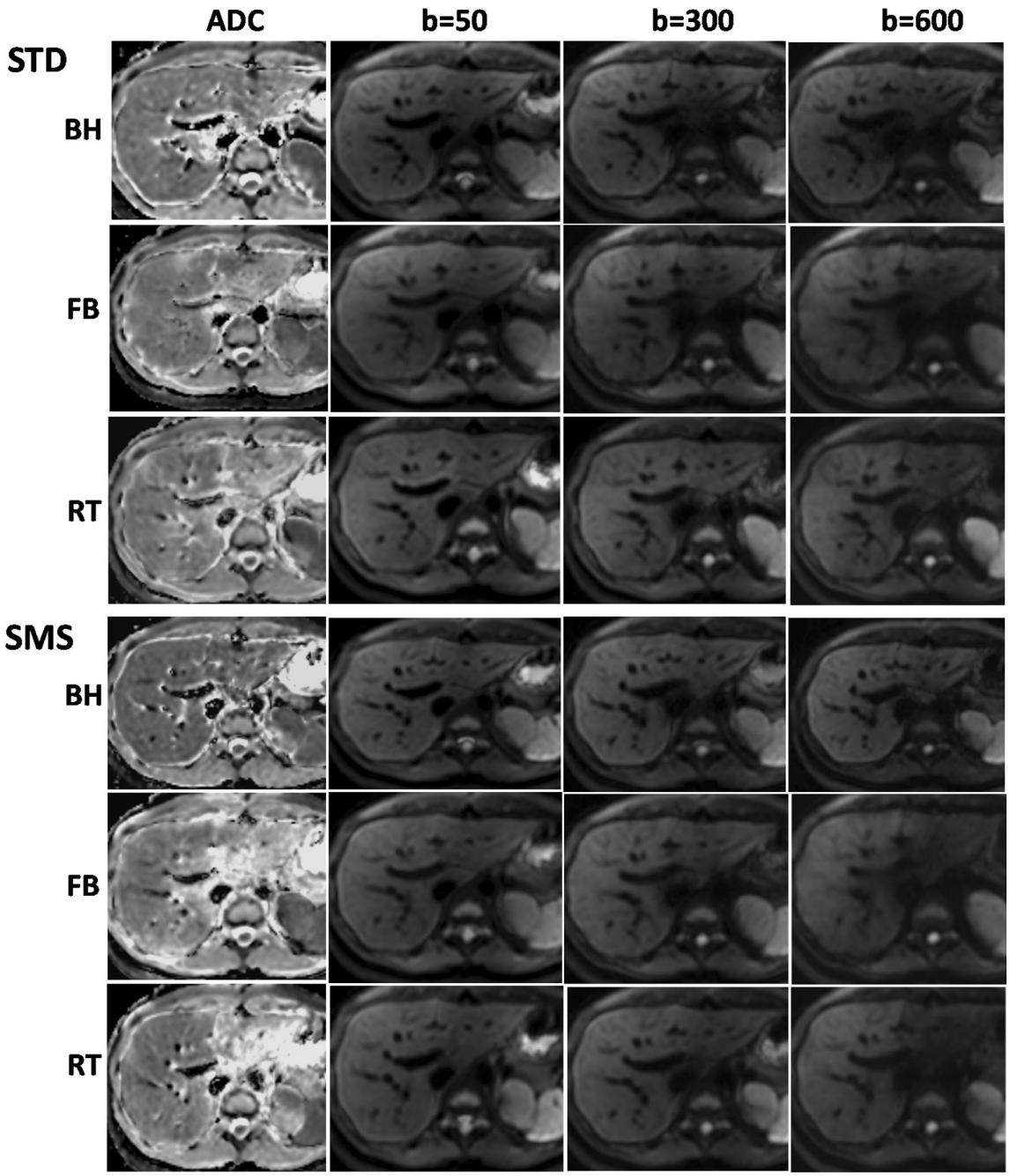


Fig. 5 Comparisons of image quality of STD-DWI (the upper three rows) and SMS-DWI (the lower three rows) using different breathing schemes (*BH* breath-hold, *FB* free-breathing, *RT* respiratory-triggering). Diffusion-weighted trace images with three different *b*-values (50, 300, 600 s/mm^2) with the corresponding ADC maps (left) are arrayed. Image quality with BH-SMS was slightly superior to other five DWI sequences in sharpness of liver edge and severity of artifacts, but they were not differ significantly between BH-SMS and other five DWI sequences ($P > 0.05$)

Funding This work was supported by the National Natural Science Foundation of China (Grant Number: 81371541, Beijing, China), Natural Science Foundation of HuNan Province (Grant Number: 2018JJ2656, Changsha, China), and China Postdoctoral Science Foundation(Grant Number: 2019M652807, Beijing, China).

Compliance with ethical standards

Conflict of interest There was no conflict of interest in this paper for all authors.

References

1. Shenoy-Bhangle A, Baliyan V, Kordbacheh H, Guimaraes AR, Kambadakone A. Diffusion weighted magnetic resonance imaging of liver: Principles, clinical applications and recent updates. World J Hepatol. 2017;9:1081-91. https://doi.org/10.4254/wjh.v9.i26.1081.
2. Galea N, Cantisani V, Taouli B. Liver lesion detection and characterization: role of diffusion-weighted imaging. J MAGN RESON IMAGING. 2013;37:1260-76. https://doi.org/10.1002/jmri.23947
3. Gong NJ, Wong CS, Chu YC, Gu J. Treatment response monitoring in patients with gastrointestinal stromal tumor using diffusion-weighted imaging: preliminary results in comparison with positron emission tomography/computed tomography. NMR BIOMED. 2013;26:185-92. https://doi.org/10.1002/nbm.2834
4. Bickel H, Pinker-Domenig K, Bogner W, Spick C, Bago-Horvath Z, Weber M, et al. Quantitative apparent diffusion coefficient as a noninvasive imaging biomarker for the differentiation of invasive breast cancer and ductal carcinoma in situ. INVEST RADIOL. 2015;50:95-100. https://doi.org/10.1097/RLI.0000000000000104
5. Zhuo J, Gullapalli RP. AAPM/RSNA physics tutorial for residents: MR artifacts, safety, and quality control. RADIOGRAPHICS. 2006;26:275-97. https://doi.org/10.1148/rg.261055134.
6. Taron J, Martirosian P, Erb M, Kuestner T, Schwenzer NF, Schmidt H, et al. Simultaneous multislice diffusion-weighted MRI of the liver: Analysis of different breathing schemes in comparison to standard sequences. J MAGN RESON IMAGING. 2016;44:865-79. https://doi.org/10.1002/jmri.25204.
7. Barth M, Breuer F, Koopmans PJ, Norris DG, Poser BA. Simultaneous multislice (SMS) imaging techniques. MAGN RESON MED. 2016;75:63-81. https://doi.org/10.1002/mrm.25897.
8. Taron J, Schraml C, Pfannenberg C, Reimold M, Schwenzer N, Nikolaou K, et al. Simultaneous multislice diffusion-weighted imaging in whole-body positron emission tomography/magnetic resonance imaging for multiparametric examination in oncological patients. EUR RADIOL. 2018;28:3372-83. https://doi.org/10.1007/s00330-017-5216-y.
9. Hsu YC, Chu YH, Tsai SY, Kuo WJ, Chang CY, Lin FH. Simultaneous multi-slice inverse imaging of the human brain. Sci Rep. 2017;7:17019. https://doi.org/10.1038/s41598-017-16976-0.
10. Weiss J, Martirosian P, Taron J, Othman AE, Kuestner T, Erb M, et al. Feasibility of accelerated simultaneous multislice diffusion-weighted MRI of the prostate. J MAGN RESON IMAGING. 2017;46:1507-15. https://doi.org/10.1002/jmri.25665.
11. Kwee TC, Takahara T, Koh DM, Nievelstein RA, Luijten PR. Comparison and reproducibility of ADC measurements in breathhold, respiratory triggered, and free-breathing diffusion-weighted MR imaging of the liver. J MAGN RESON IMAGING. 2008;28:1141-48. https://doi.org/10.1002/jmri.21569.
12. Lee Y, Lee SS, Kim N, Kim E, Kim YJ, Yun SC, et al. Intravoxel incoherent motion diffusion-weighted MR imaging of the liver: effect of triggering methods on regional variability and measurement repeatability of quantitative parameters. RADIOLOGY. 2015;274:405-15. https://doi.org/10.1148/radiol.14140759.
13. Chen X, Qin L, Pan D, Huang Y, Yan L, Wang G, et al. Liver diffusion-weighted MR imaging: reproducibility comparison of ADC measurements obtained with multiple breath-hold, free-breathing, respiratory-triggered, and navigator-triggered techniques. RADIOLOGY. 2014;271:113-25. https://doi.org/10.1148/radiol.13131572.
14. Kandpal H, Sharma R, Madhusudhan KS, Kapoor KS. Respiratory-triggered versus breath-hold diffusion-weighted MRI of liver lesions: comparison of image quality and apparent diffusion coefficient values. AJR Am J Roentgenol. 2009;192:915-22. https://doi.org/10.2214/AJR.08.1260.
15. Rosenkrantz AB, Oei M, Babb JS, Niver BE, Taouli B. 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. 2011;33:128-35. https://doi.org/10.1002/jmri.22395.
16. Larsen NE, Haack S, Larsen LP, Pedersen EM. Quantitative liver ADC measurements using diffusion-weighted MRI at 3 Tesla: evaluation of reproducibility and perfusion dependence using different techniques for respiratory compensation. MAGMA. 2013;26:431-42. https://doi.org/10.1007/s10334-013-0375-6.
17. Nasu K, Kuroki Y, Sekiguchi R, Kazama T, Nakajima H. Measurement of the apparent diffusion coefficient in the liver: is it a reliable index for hepatic disease diagnosis? Radiat Med. 2006;24:438-44. https://doi.org/10.1007/s11604-006-0053-y.
18. Koc Z, Erbay G. Optimal b value in diffusion-weighted imaging for differentiation of abdominal lesions. J Magn Reson Imaging. 2014;40:559-66. https://doi.org/10.1002/jmri.24403.
19. Boss A, Barth B, Filli L, Kenkel D, Wurnig MC, Piccirelli M, et al. Simultaneous multi-slice echo planar diffusion weighted imaging of the liver and the pancreas: Optimization of signal-to-noise ratio and acquisition time and application to intravoxel incoherent motion analysis. EUR J RADIOL. 2016;85:1948-55. https://doi.org/10.1016/j.ejrad.2016.09.002.
20. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. PSYCHOL BULL. 1979;86:420-28.
21. Busing KA, Kilian AK, Schaible T, Debus A, Weiss C, Neff KW. Reliability and validity of MR image lung volume measurement in fetuses with congenital diaphragmatic hernia and in vitro lung models. RADIOLOGY. 2008;246:553-61. https://doi.org/10.1148/radiol.2462062166.
22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. LANCET. 1986;1:307-10.
23. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ. 1995;310:170.
24. Dunn OJ. Multiple Comparisons Using Rank Sums. Technometrics. 1964;6:241-52. https://doi.org/10.1080/00401706.1964.10490181.
25. Kim SY, Lee SS, Byun JH, Park SH, Kim JK, Park B, et al. Malignant hepatic tumors: short-term reproducibility of apparent diffusion coefficients with breath-hold and respiratory-triggered diffusion-weighted MR imaging. RADIOLOGY. 2010;255:815-23. https://doi.org/10.1148/radiol.10091706.
26. Kartalis N, Loizou L, Edsborg N, Segersvard R, Albiin N. Optimising diffusion-weighted MR imaging for demonstrating pancreatic cancer: a comparison of respiratory-triggered, free-breathing and breath-hold techniques. EUR RADIOL. 2012;22:2186-92. https://doi.org/10.1007/s00330-012-2469-3.