Radiography 25 (2019) e45–e51

Contents lists available at ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

Evaluation of unenhanced axial T1W and T2W liver MR images acquired from institutions within the Republic of Ireland and the Kingdom of Saudi Arabia

S. Al-Dahery [a, *], A. McGee [a], L. Rainford [a], K. Khashoggi [b], N. Misha [b]

[a] Radiography and Diagnostic Imaging, School of Medicine, University College Dublin, Dublin, Ireland
[b] Radiology Department, King Abdulaziz University Hospital, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 7 April 2018
Received in revised form
8 October 2018
Accepted 20 October 2018
Available online 22 November 2018

Keywords:
Liver MRI
Image quality
T2W-FSE sequence
T1W-3D-GRE-FS sequence

ABSTRACT

Introduction: This multi-site study evaluated two breath-hold sequences commonly utilised for liver MRI; non-enhanced T1W-3D-FS-GRE-TRA and T2W-2D-FSE-TRA sequences, using physical measurements of SNR and CNR, and observer perceptions' (Visual Grading Analysis: VGA).
Methods: Liver MR image datasets (n = 168) from nine hospitals in the Kingdom of Saudi Arabia (KSA) and 11 hospitals in the Republic of Ireland were evaluated. Images were categorised into two groups per sequence, defined by slice thickness (T2W-2D-FSE, ≤5 mm vs ≥ 6 mm and T1W-3D-GRE-FS, ≤3 mm vs 4 mm). Images were evaluated using visual grading analysis VGA and physical measurements: SNR/CNR. Account was taken of varying patient sizes based on AP/transverse diameter measurements.
Results: Physical image quality measurements (SNR/CNR) returned no significant findings across Irish and KSA hospitals, for both sequences, despite variations in acquisition parameters. Statistically significant differences were found for some scoring criteria based on the observers' perceptions including spleen parenchyma, and spatial resolution for the non-enhanced T1W-3D-FS-GRE-TRA images, with a preference for images acquired using thin slices (≤3 mm). In addition, statistically significant difference was found for the scoring criteria motion artefact for the axial T2W-2D-FSE-TRA images, with a preference for images acquired using thick slices (≥5 mm). Negligible correlation was noted between SNR/CNR and measured abdominal AP/transverse diameters.
Conclusion: Whilst variations in sequences rendered no statistical differences in SNR/CNR findings, significant differences in observer image criteria scores was noted. The importance of both physical measurements and observers' perceptions evaluation methods for quality assessment of MR images was demonstrated and optimisation of liver sequence parameters is warranted.

Introduction

Differentiation of normal tissue from fluid- and/or fat-based pathologies is important in MR imaging.[1,2] In addition, Breath-hold acquisition technique is usually combined in liver MR imaging to minimise motion artefacts.[3] Breath-hold technique is usually combined with the axial two-dimensional T2-weighted fast-spin echo (T2W-2D-FSE-TRA) imaging of the liver which provides strong T2 contrast within short breath-hold times that are within the capabilities of most patients.[4] In addition, breath-hold technique is usually combined with the non-enhanced axial three-dimensional T1-weighted gradient-echo with fat suppression (T1W-3D-FS-GRE-TRA), which provides thin-slices and high spatial resolution images of the liver within a single breath-hold, without compromising on SNR.[5] The rationale for this study was based on literature evidence of variation in image quality for the non-enhanced T1W-3D-FS-GRE-TRA and T2W-2D-FSE-TRA sequences combined with breath-hold technique of the liver when trying to balance SNR, spatial and temporal resolution.[6] Whilst there has been investigation of liver MR image quality comparing standard T2W-2D-FSE relative to T2W-FSE with BLADE or with the HASTE technique[7,8] this study aimed to evaluate two commonly utilised breath-hold sequences; non-enhanced T1W-3D-FS-GRE-TRA and T2W-2D-FSE-TRA sequences. Literature evidence supports

* Corresponding author. Diagnostic Imaging Department, School of Medicine and Medical Science, University College Dublin, Belfield, Dublin 4, Dublin, Ireland.
E-mail address: shrooq.al-dahery@ucdconnect.ie (S. Al-Dahery).

https://doi.org/10.1016/j.radi.2018.10.005
1078-8174/

the clinical utility of these sequences for liver lesion detection and characterisation.[4,5]

Image interpretation accuracy is dependent upon image quality and observer competency. Physical image quality measures such as signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) are common and applicable to MR.[9–14] Literature reports that physical objective measurement of image quality is not sufficient to describe clinical image quality and observer perceptions of image quality are required.[15–17] Likewise, MR radiographers need to be aware of these quality indices and the interrelationship between the factors to be able to produce, recognise and evaluate high-quality MR images.

MRI departments in the Republic of Ireland and the Kingdom of Saudi Arabia (KSA) participated by providing image datasets for analysis. Objective physical measurement of image quality in terms of SNR and CNR combined with observer perceptions', using visual grading analysis (VGA), as an indicator of clinical performance were recorded.[16,17]

Materials and methods

Subjects

Exemption from full ethical approval was granted by the Irish Institutional Human Research Ethics Committee and by the relevant Institutional Review Board (IRB) in Ministry of Health (MOH) for Saudi hospitals.

MR clinical specialists from participating hospitals each provided Digital Imaging and Communication in Medicine (DICOM) format MR images of anonymised liver MRI examinations (n = 10). Inclusion criteria for retrospective image collection were normal anatomy or images containing minimal pathology such as small cysts or small haemangiomas. Exclusion criteria included images with diffuse liver pathology. Information was sought from each participant centre to ensure that centres had broadly similar MRI scanners. The images were imported and coded into a McKesson PACS system (MDCC-6130, v1.13, monitor resolution of 3280 × 2048 pixels, 400 cd/m^2). Due to the retrospective data collection, patient BMI measurements were not available. Therefore, anterior-posterior (AP) abdominal and transverse diameters were measured and patients were grouped into two cohorts based on these measurements. Patients within the lower two thirds were categorised as small/medium size, those in the upper third of diameter values were categorised as 'large' (small/medium transverse diameter range: 121–247 cm, AP diameter range: 248–327 cm; large transverse diameter range: 332–342 cm, small/medium AP diameter range: 182–331 cm). Correlation coefficients for patients' AP and transverse diameters with SNR and CNR values were calculated for each sequence.[18]

Parameters for the T2W-2D-FSE-TRA and non-enhanced T1W-3D-GRE-FS-TRA sequences acquired at 1.5T are specified in Table 1.

MR image interpretation

Image evaluation involved two phases:

Phase one:

Signal intensities of the liver, spleen and standard deviation of the background noise were obtained for images acquired with both sequences using a region-of-interest (ROI) consistently placed on the right lobe of the liver, avoiding hepatic vessels and artefacts. A ROI was also placed on the spleen in order to calculate the CNR. A ROI representative of image background noise was drawn away from skin surface margins, clear of anatomy and any phase ghosting artefact.[8,19–21] In each case, a 1 cm ROI diameter was placed within the liver, spleen, and image background[11] (Fig. 1). All measurements were performed by the same reviewer. Liver SNR values were calculated as follows:

$$SNR_{Liver} = 0.655 \times Signal_{Liver\ parenchyma}/noise;$$

where the signal is the mean pixel intensity within the selected ROI, and the noise is the standard deviation of background air within the FOV.[22] It should be acknowledged that there are many complexities in MRI scanners, such that classical SNR measurements do not hold true, and a correction factor of 0.655 was applied to account for the Rayleigh distribution statistics.[23–25] The following equation was used to calculate the CNR:

$$CNR_{Liver} = SNR_{Liver\ parenchyma} - SNR_{Spleen\ parenchyma}/noise;$$

where the CNR is the difference in mean signal intensity of the selected ROIs (S_A and S_B).[22]

Statistical analysis

SNR and liver-spleen CNR values were compared by using the Mann Whitney U-test. A p-value <0.05 was considered to be statistically significant, analyses were performed using SPSS (IBM SPSS, version 20). G*Power 3.1.9.2 software calculated power and a range of 80–90% was confirmed for the analysis comparing Irish and Saudi centres.

Phase two:

The image data sets were categorised into two groups defined by slice thickness for each sequence; as this was the parameter that varied the most across participant centres (25 cases per group). The power calculation for the sample size was based on personal communication directly derived from an expert in VGA techniques.[26] Therefore, for the T2W-2D-FSE-TRA sequence, 25 images performed using slice thickness of ≤5 mm, and 25 images performed by using slice thickness of ≥6 mm were compared. Whilst for the T1W-3D-FS-GRE-TRA sequence, 25 images with a slice thickness of ≤3 mm and 25 images with a slice thickness of ≥4 mm were compared. These 100 liver MR images were randomly presented to a review panel of MRI experts: six MR Radiographers with 6–15 years scanning experience/M.Sc. qualification level and two radiologists experienced in MRI, one with 5 and one with 10 years' experience. Observers were blinded to the clinical information, demographic data and sequence parameters. To support consistency of image review, all observers completed a short training *PowerPoint* presentation prior to independently evaluating images.

Ergonomic image review conditions adhered to the American Association of Physicists in Medicine (AAPM) recommendations. Reading room lighting was controlled to avoid the monitor reflections and ambient lighting levels were measured with a calibrated Unfors Light-O-Meter photometer (Billdal, Sweden) to ensure; 15–60 lux in the MRI diagnostic reading station.[27,28] The observers were allowed to modify the window width and centre, at their own discretion.

Images were evaluated for: (i) severity of motion and other artefact using 4-point scale as follows: grade 1, "severe" (non-diagnostic); grade 2, "moderate"; grade 3, "minimal", or grade 4, "no artefact"; (ii) visualisation of the signal intensity of liver parenchyma, spleen parenchyma and vascular pattern of the liver using 4-point scale as follows: grade 1, "poor" (non-diagnostic); grade 2, "reasonable"; or grade 3, "good"; (iii) image quality (SNR, CNR, and spatial resolution) using 3-point scale as follows: grade 1; "poor" (non-diagnostic); grade 2, "good"; or grade 3, "excellent";

Table 1
Parameters for breath-hold T1W-3D-FS-GRE-TRA and T2W-2D-FSE-TRA sequences of the liver protocol.

Seq	Hosp	ST	No. Slices	TR (ms)	TE (ms)	NEX	Matrix	FOV (mm)	VS (mm)	ETL	BW Hz/Px	FA °
T1W-3D-FS-GRE-TRA	1	4	40	5.43	2.36	1	512 × 230	430 × 268	0.8 × 1.1 × 4	1	300	10
	2	2.5	72	5.6	2.38	1	256 × 168	400 × 250	1.5 × 1.4 × 2.5	1	250	10
	3	3	48	5.2	2.45	1	320 × 320	380 × 380	1.1 × 1.1 × 3	1	345	10
	4	2	80	4.24	2.38	1	256 × 176	400 × 275	1.5 × 1.5 × 2	1	725	10
	5	2.6	48	6.1	2.79	1	256 × 192	350 × 262	1.3 × 1.3 × 2.6	1	250	10
	6	3	64	4.4	2.3	1	256 × 144	350 × 262	1.3 × 1.8 × 3	1	490	15
	7	4	72	5.77	2.63	1	256 × 240	350 × 262	1.3 × 1 × 4	1	250	10
	8	3	44	5.58	2.38	1	512 × 512	400 × 400	0.7 × 0.7 × 3	1	260	10
	9	3.6	64	5.43	2.47	1	512 × 352	350 × 262	0.6 × 0.7 × 3.6	1	300	10
	10	3	48	5.11	2.25	1	512 × 384	360 × 270	0.7 × 0.7 × 3	1	300	10
	11	3	48	3–6	2.2	1	288 × 216	380 × 285	1.3 × 1.3 × 3	1	445	10
	12	5	72	4.3	1.88	1	320 × 256	340 × 255	1 × 0.9 × 5	1	625	12
	13	3	72	4.74	2.38	1	320 × 250	280 × 280	0.8 × 1.1 × 3	1	404	10
	14	4	44	3.5	1.68	1	512 × 512	340 × 280	0.6 × 0.5 × 4	1	325	12
	15	3.5	64	4.8	2.38	1	320 × 230	350 × 262	1 × 1.1 × 3.5	1	401	10
	16	4	48	3.9	1.8	1	192 × 192	395 × 395	2 × 2 × 4	1	434	10
	17	4	64	3.9	1.8	1	192 × 192	380 × 380	1.9 × 1.9 × 4	1	431	10
	18	4.4	72	3.6	1.7	1	512 × 512	440 × 440	0.8 × 0.8 × 4.4	1	244	10
	19	4	72	3.9	1.9	0.7	512 × 512	440 × 440	0.8 × 0.8 × 4	1	244	10
	20	4	64	4.6	2.03	1	576 × 400	380 × 380	0.6 × 0.9 × 4	1	404	10
T2W-2D-FSE-TRA	1	6	20	4500	107	1	320 × 320	380 × 380	0.8 × 0.8 × 6	25	260	140
	2	5	20	3000	90	1	560 × 560	400 × 281	0.7 × 0.5 × 5	23	344	140
	3	6	24	2900	97	1	384 × 288	360 × 360	0.9 × 1.2 × 6	21	195	153
	4	4	26	3500	105	1	320 × 320	310 × 310	0.9 × 0.9 × 4	19	200	150
	5	4	20	4000	103	1	512 × 384	370 × 277	0.7 × 0.7 × 4	29	260	150
	6	4/6	30	5000	90	3	512 × 320	350 × 207	0.6 × 0.6 × 4	29	260	130
	7	5	30	5000	108	1	320 × 320	360 × 360	1.1 × 1.1 × 5	25	260	140
	8	7	19	4410	97	1	256 × 144	450 × 256	1.7 × 1.7 × 7	29	300	140
	9	6	30	3000	100	1	256 × 256	400 × 262	1.5 × 1 × 6	30	501	150
	10	5	30	3000	77	1	512 × 384	360 × 288	0.7 × 0.7 × 6	17	310	133
	11	6	20	3200	67	1	384 × 288	350 × 262	0.9 × 0.9 × 6	29	260	133
	12	5	20	4500	80	1	384 × 288	340 × 340	0.8 × 1.1 × 5	9	260	150
	13	5	20	4900	97	2	352 × 352	375 × 375	1 × 1 × 5	30	300	150
	14	7	20	4500	80	2	384 × 384	280 × 280	0.7 × 0.7 × 7	30	300	150
	15	5	20	4000	109	1	512 × 512	400 × 400	0.7 × 0.7 × 5	23	122	150
	16	4	26	3500	105	1	320 × 320	310 × 310	0.9 × 0.9 × 4	19	200	150
	17	7	19	4410	97	1	256 × 144	450 × 256	1.7 × 1.7 × 7	29	300	150
	18	6	20	4500	107	1	320 × 320	380 × 380	0.8 × 0.8 × 6	25	260	150
	19	5	20	4000	109	1	512 × 512	400 × 400	0.7 × 0.7 × 5	23	122	150
	20	6	30	3000	100	1	256 × 256	400 × 262	1.5 × 1 × 6	30	501	150

Key: Seq: sequence; Hosp: hospital; ST: slice thickness (mm); TR: time repetition (ms); TE: echo time (ms); NEX: number of averages; FOV: field-of-view (mm); VS: voxel size (mm); ETL: echo train length, BW: bandwidth (Hz/Px) and FA: flip angle (°).

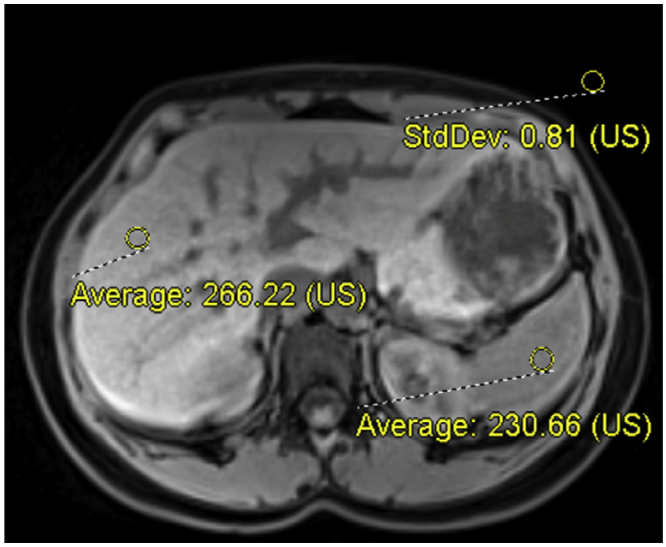

Figure 1. Non-enhanced T1W-3D-FS-GRE-TRA sequence; 1 cm ROI was placed on the right lobe of the liver, spleen and background noise.

(iv) overall image quality using 3-point scale as follows: grade 1, "unacceptable"; grade 2, "marginally acceptable"; or grade 3, "highly acceptable".

Statistical analysis

The Mann–Whitney U-test compared image quality findings across the different variations of the same sequence as employed across participating clinical sites. A Bonferroni p-value <0.005 was considered statistically significant for multiple testing errors.[29] Inter-observer agreement for the qualitative grading was calculated by using the R statistical analysis program (version 3.3.2: R Core Team, Bell Laboratory). Ordinal alpha reliability coefficient of 0.41–0.60, 0.61–0.80, and ≥0.80 were regarded as indicative of moderate, substantial, and excellent agreement, respectively.[30,31] Intra-observer reliability was measured by using the weighted kappa coefficient, which is suitable for measuring agreement between two sets of ordinal items.[31] Observers reviewed two datasets twice for each sequence.

Results

Of the 12 Irish and 12 Saudi hospitals who were invited to participate only 11 Irish hospitals and 9 Saudi hospitals were

undertaking liver MR imaging from which data sets were collected of patients who underwent liver MRI at 1.5T.

Quantitative analysis

The average gradient amplitude and slew rate of MR scanners across participating centres was 30 mT/m, and 125 mT/m/ms respectively. SNR and CNR values for images acquired with T2W-2D-FSE-TRA and non-enhanced T1W-3D-FS-GRE-TRA sequences are displayed in Table 2. The lowest median SNR and CNR obtained for T2W-2D-FSE-TRA images was 14.5 and 17.08 respectively and the highest median SNR and CNR obtained for T2W-2D-FSE-TRA images was 73.5 and 157.34 respectively. The lowest median SNR and CNR obtained for T1W-3D-FS-GRE-TRA images was 15.1 and 1 respectively and the highest median SNR and CNR obtained for T1W-3D-FS-GRE-TRA images was 256.5 and 72.9 respectively. Whilst no statistical significant findings were identified it is noted that a p value of 0.06 was recorded for the T1W-3D-FS-GRE-TRA images, Irish versus Saudi data, Recommendations for further research would be to increase data set numbers, particularly if evaluating subtle pathologies.

Across both countries negligible correlation between SNR/CNR - abdominal AP and transverse diameters was found for the non-enhanced T1W-3D-FS-GRE-TRA sequence. Similarly, for the T2W-2D-FSE-TRA sequence in KSA hospitals no correlation between SNR - abdominal AP and transverse diameters was noted, but moderate correlation was noted between CNR - abdominal AP and transverse diameters. Similarly, no correlation in the above variables was found for images acquired with both sequences in Irish hospitals (Table 3).

Perceptual VGA

Observer perception scores are displayed in Tables 4 and 5. Inter-observer agreement was high for the non-enhanced T1W-3D-FS-GRE-TRA and T2W-2D-FSE-TRA sequences (Table 5). In general, intra-observer agreement was high for the T2W-2D-FSE-TRA sequence, and perfect for non-enhanced T1W-3D-FS-GRE-TRA sequence for both study phases (Table 6).

Table 3
Pearson's Correlation coefficient findings for AP/transverse diameter measurements and SNR/CNR.

Variables	The Republic of Ireland		KSA	
	R	*P*	R	*P*
	T1W-3D-FS-GRE-TRA			
Transverse/SNR	−0.064	0.616	0.157	0.218
Transverse/CNR	−0.166	0.195	0.179	0.161
AP/SNR	−0.162	0.204	0.169	0.189
AP/CNR	−0.221	0.081	0.104	0.418
	T2W-2D-FSE-TRA			
Transverse/SNR	−0.169	0.221	0.163	0.234
Transverse/CNR	−0.09	0.519	0.12	0.384
AP/SNR	−0.92	0.508	0.061	0.656
AP/CNR	−0.117	0.4	0.174	0.204

Discussion

SNR and CNR values for the liver parenchyma for non-enhanced axial images acquired with the T1W-3D-FS-GRE-TRA sequence demonstrated no statistically significant differences across the participating centres (n = 20) despite variations in parameter values such as slice thickness. Literature suggests different optimal slice thicknesses for the non-enhanced T1W-3D-FS-GRE-TRA sequence, and for this study, the two image groups were defined by slice thickness (≤3 mm *vs.* ≥4 mm). A 4–5 mm slice thickness range was proposed by Manian & Szklaruk (2010)[4] the current study identified 7 hospitals utilising a slice thickness of 4 mm, while hospital 12 used a 5 mm slice thickness, the greatest across all participating sites (Table 1). Two hospitals, 15 & 9 used a slice thickness of 3.5 mm and 3.6 mm respectively as recommended in literature,[32–34] however nine centres applied slice thicknesses lower than 3 mm. The overall findings align with literature recommendations for thinner slice utilisation to increase spatial resolution while preserving SNR due to the volume acquisition technique for this sequence.[35]

Observers preferred images acquired with a slice thickness of ≤3 mm compared to thicker slice acquisition (≥4 mm). Statistically significant differences for the scoring criteria: spleen parenchyma, SNR, and spatial resolution were noted for ≤3 mm versus ≥4 mm slice thickness options for this non-enhanced T1W-3D-FS-GRE-TRA sequence (Table 4) as illustrated in Fig. 2. Motion artefact was noted on some data sets however the frequency of this effect was not consistent across both slice thickness and the principles of aligning sequence selection with observed patient breath hold capabilities remain.

The highest and lowest median SNR values aligned with liver-spleen CNR findings, confirming their interrelationship (Table 2). There is a paucity of literature recommending optimal SNR and CNR values for liver parenchyma on non-enhanced T1W-3D-FS-GRE-TRA images however these studies were performed almost a decade ago.[32,36]

No significant differences in SNR and liver-spleen CNR values for images acquired from the participant Saudi and Irish centres for the breath-hold T2W-2D-FSE-TRA sequence were noted. Despite this, variability in the median CNR values was noted, attributed to differences in the timing parameters (TR and TE) employed across the participant centres. Images generated with longer TRs (4000–5000 ms) together with matrix sizes ranging from 256 × 256 to 384 × 384 had higher measured SNR and liver-spleen CNR values, reflecting their stronger T2W, in line with the optimal TR range of 4000–6000 ms for this sequence proposed by Maniam & Szklaruk (2010).[4]

Observers preferred T2W-2D-FSE-TRA images acquired using a slice thickness of ≥6 mm based on the mean rank of scoring criteria such as the clarity of visualisation of the signal intensity of the

Table 2
Summary of SNR and CNR values by country and by sequence (T1W-3D-FS-GRE-TRA and T2W-2D-FSE-TRA).

Sequence	Parameters	Irish SNR	Saudi SNR	Irish CNR	Saudi CNR
T1W-3D-GRE-TRA	Median (Q1–Q3)	128.76 (78.2–191.2)	140.8 (70–211)	21.22 (7.5–55)	12.75 (5–29)
	Mean rank	63.32	63.68	68.99	55.94
	p-value	0.955		0.06	
T2W-2D-FSE-TRA	Median (Q1–Q3)	25.6 (15.2–46)	37.32 (11.5–75.4)	25.05 (13.85–41)	20.62 (10.76–48.7)
	Mean rank	39.28	45.8	40.87	39.45
	p-value	0.327		0.81	

Table 4
Descriptive statistics for the comparison of two different parameter settings for T1W-3D-FS-GRE-TRA & T2W-2D-FSE-TRA sequences.

Scoring Criteria		Groups	Mean Rank	P-value		Groups	Mean Rank	p-value
Motion Artefact	T1W-3D-FS-GRE-TRA	≤3 mm	196.96	0.504	T2W-2D-FSE-TRA	≤5 mm	172.62	<0.005
		≥4 mm	204.04			≥6 mm	221.38	
Other Artefact		≤3 mm	214.34	0.002		≤5 mm	200.51	0.400
		≥4 mm	186.66			≥6 mm	192.33	
Spleen Parenchyma		≤3 mm	219.77	<0.005		≤5 mm	198.57	0.684
		≥4 mm	181.23			≥6 mm	194.34	
Liver Parenchyma		≤3 mm	213.72	0.015		≤5 mm	198.01	0.773
		≥4 mm	187.28			≥6 mm	194.92	
Vascular Pattern		≤3 mm	212.07	0.035		≤5 mm	188.70	0.141
		≥4 mm	188.93			≥6 mm	204.63	
SNR		≤3 mm	223.17	<0.005		≤5 mm	193.26	0.536
		≥4 mm	177.84			≥6 mm	199.88	
CNR		≤3 mm	196.39	0.432		≤5 mm	195.26	0.797
		≥4 mm	204.61			≥6 mm	197.79	
Spatial Resolution		≤3 mm	215.24	0.005		≤5 mm	190.47	0.239
		≥4 mm	185.76			≥6 mm	202.78	
Overall Image Quality		≤3 mm	210.59	0.055		≤5 mm	194.67	0.715
		≥4 mm	190.41			≥6 mm	198.41	

Table 5
Inter-observer reliability findings: ordinal alpha reliability coefficients of 0.41–0.60, 0.61–0.80, and ≥0.80 as moderate, substantial, and excellent agreement, respectively.

Scoring Criteria	T2W-2D-FSE-TRA	T1W-3D-FS-GRE-TRA
Motion artefact	0.94	0.90
Other artefact	0.84	0.89
Spleen parenchyma	0.83	0.90
Liver parenchyma	0.87	0.85
Vascular pattern	0.9	0.86
SNR	0.89	0.92
CNR	0.87	0.89
Spatial resolution	0.74	0.88
Overall image quality	0.74	0.94

splenic and liver parenchyma (Table 4). Statistically significant difference was found in the frequency of motion artefact identified by the observers within images assigned to the thick-slice group (≥6 mm) (Fig. 3). As with the T1W-3D-FS-GRE-TRA sequence the higher frequency of motion artefact observed in thick-slice images may be attributable to a random cohort of patients who had compromised breath-hold capacity for those patients however the retrospective nature of data collection prohibited full investigation.

This study identified MR imaging variations across participating which in theory could impact on image quality however no significant differences in physical SNR and CNR values for the liver images analysed were identified. Contrary to these objective measures observers' opinion confirmed the benefits of using thicker slices for the T2W-2D-FSE-TRA sequence. The importance of including both physical and subjective image quality evaluation is demonstrated by these findings. MR radiographers need a comprehensive understanding of all parameters that impact upon image quality to fully understand image quality optimisation.

A potential limitation of this research was the inability to fully explore the impact of patient body mass index (BMI) on image quality. Yang et al., (2010) demonstrated larger patients could cause inhomogeneities in the magnetic field potentially impairing image quality.[37] Body size was not as strong a contributing factor to MR image quality in this work unlike published literature related to CT imaging.[18] It is also acknowledged that advanced liver imaging sequences such as DWI were not analysed as these sequences were not undertaken by all participant hospitals.

Table 6
Intra-observer reliability findings, using the weighted kappa coefficient, for the non-enhanced T1W-3D-FS-GRE-TRA and T2W-2D-FSE-TRA sequences.

Criteria	T1W-3D-FS-GRE-TRA					T2W-2D-FSE-TRA				
	Lower CI	Weighted kappa	Upper CI	Mean	SD	Lower CI	Weighted kappa	Upper CI	Mean	SD
MA.1	1.00	1.00	1.00	1.69	0.48	0.22	0.60	0.97	2.44	0.51
MA.2				1.69	0.48				2.25	0.45
OA.1	0.24	0.66	1.00	3.75	0.45	0.60	0.80	0.99	2.75	2.00
OA.2				3.75	0.45				3.00	3.00
SP.1	0.74	0.90	1.00	1.5	0.63	0.43	0.75	1.00	2.62	0.50
SP.2				1.56	0.63				2.5	0.52
LP.1	1.00	1.00	1.00	1.56	0.51	0.60	0.80	1.00	2.31	0.48
LP.2				1.56	0.51				2.38	0.50
VP.1	1.00	1.00	1.00	1.75	0.45	0.60	0.80	1.00	2.62	0.50
VP.2				1.75	0.45				2.56	0.51
SNR.1	1.00	1.00	1.00	1.38	0.50	1.00	1.00	1.00	2.44	0.51
SNR.2				1.38	0.50				2.44	0.51
CNR.1	1.00	1.00	1.00	1.62	0.50	0.22	0.60	1.00	2.62	0.50
CNR.2				1.62	0.50				2.56	0.51
SP.1	1.00	1.00	1.00	1.50	0.52	0.40	0.80	1.00	1.75	0.45
SP.2				1.50	0.52				1.81	0.40
IQ.1	0.60	0.80	1.00	1.38	0.50	0.270	0.67	1.00	2.31	0.48
IQ.2				1.44	0.51				2.19	0.40

Key: MA: motion artefact; OA: other artefact; SP: spleen parenchyma; LP: liver parenchyma; VP: vascular pattern; SNR: signal-to-noise ratio; CNR: contrast-to-noise ratio; SR: spatial resolution; and IQ: overall image quality.

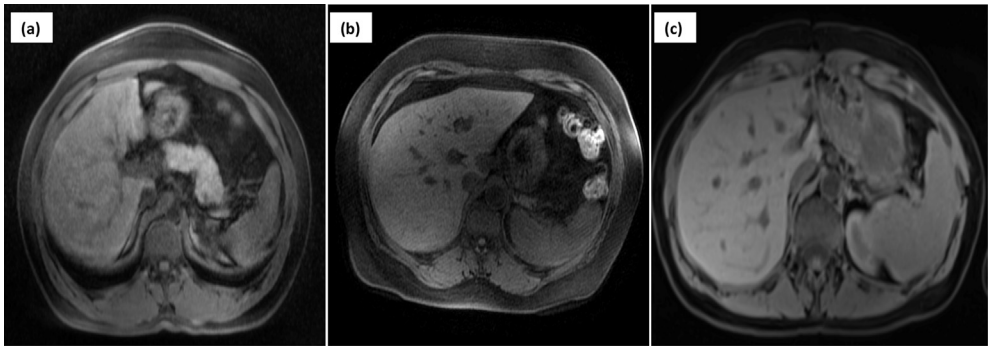

Figure 2. Standard breath-hold of non-enhanced breath-hold T1W-3D-FS-GRE-TRA images of the liver using: (a) a slice thickness of 5 mm (Hospital 12); (b) a slice thickness of 4 mm (Hospital 14); (c) a slice thickness of 3 mm (Hospital 8).

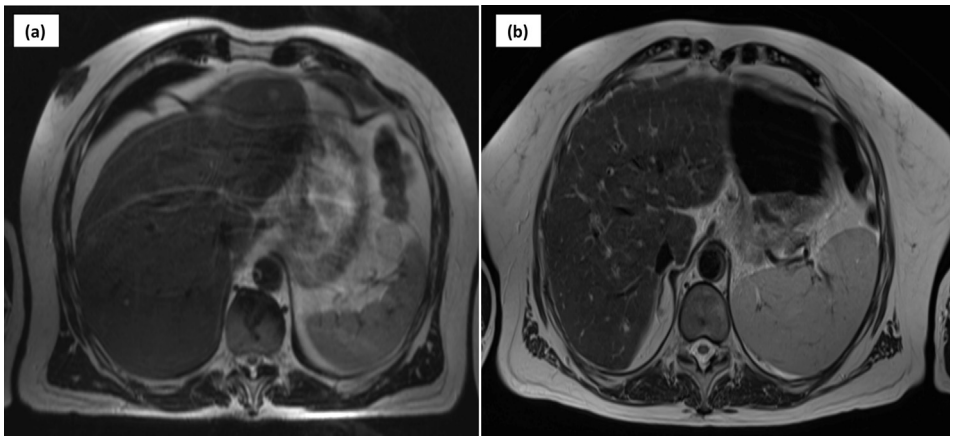

Figure 3. Standard breath-hold T2W-2D-FSE-TRA images of the liver using: (a) a slice thickness of 4/6 mm and TR of 5000 ms (Hospital 6); (b) a slice thickness of 4 mm and TR of 3500 ms (Hospital 4) to demonstrate variations in image quality levels.

Conclusion

Variations exist in pulse sequence parameters for liver imaging however no statistically significant differences were recorded in the physical measures of SNR and CNR. The importance of observers' perceptions in evaluating image quality was demonstrated as, contrary to physical measures, significant differences were identified for some of the image criteria, which may impact on the diagnostic efficacy of resultant images. The findings indicate optimisation of sequence parameters for liver imaging has the potential to enhance diagnostic efficacy of liver MR examinations. The evidence base for MR sequence optimisation is limited, most publications date back to the last decade. Further research investigating MR performance characteristics that reflect current practise is warranted. The inclusion of the visualisation of specific liver lesions, for a spectrum of pulse sequences currently employed clinically is recommended.

Conflict of interest statement

The authors declare no conflict of interest.

Acknowledgements

The authors greatly appreciate the efforts of the MR radiographers and radiologists in the participating clinical sites in KSA and the Republic of Ireland. Image evaluation panel members and Miss, and Marie Galligan for her support regarding quantitative data analysis. This study was funded by the Saudi Arabia Ministry of Higher Education.

References

1. Murakami T, Tsurusaki M. Hypervascular benign and malignant liver tumors that require differentiation from hepatocellular carcinoma: key points of imaging diagnosis. *Liver Cancer* 2014;**3**(2):85–96.
2. Kim BS, Kim JH, Choi GM, Kim SH, Park JK, Song B-C, et al. Comparison of three free-breathing T2-weighted MRI sequences in the evaluation of focal liver lesions. *Am J Roentgenol* 2008;**190**(1):W19–27.
3. Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: a complex problem with many partial solutions. *J Magn Reson Imag* 2015;**42**(4):887–901.
4. Maniam S, Szklaruk J. Magnetic resonance imaging: review of imaging techniques and overview of liver imaging. *World J Radiol* 2010;**2**(8):309–22.
5. Kim BS, Angthong W, Jeon Y-H, Semelka RC. Body MR imaging: fast, efficient, and comprehensive. *Radiol Clin* 2014;**52**(4):623–36.
6. Li T, Mirowitz SA. Fast T2-weighted MR imaging: impact of variation in pulse sequence parameters on image quality and artifacts. *Magn Reson Imag* 2003;**21**(7):745–53.
7. Haneder S, Dinter D, Gutfleisch A, Schoenberg S, Michaely H. Image quality of T2w-TSE of the abdomen and pelvis with Cartesian or BLADE-type k-space sampling: a retrospective interindividual comparison study. *Eur J Radiol* 2011;**79**(2):177–82.
8. Hirokawa Y, Isoda H, Maetani YS, Arizono S, Shimada K, Okada T, et al. Hepatic lesions: improved image quality and detection with the periodically rotated overlapping parallel lines with enhanced reconstruction technique—evaluation of SPIO-enhanced T2-weighted MR images 1. *Radiology* 2009;**251**(2):388–97.
9. Alsleem H, Davidson R. Quality parameters and assessment methods of digital radiography images. *Radiogr – Offic J Austral Inst Radiogr* 2012;**59**(2):46.
10. Båth M, Månsson LG. Visual grading characteristics (VGC) analysis: a non-parametric rank-invariant statistical method for image quality evaluation. *Br J Radiol* 2007;**80**:169–76.
11. Dietrich O, Raya JG, Reeder SB, Reiser MF, Schoenberg SO. Measurement of signal to noise ratios in MR images: influence of multichannel coils, parallel imaging, and reconstruction filters. *J Magn Reson Imag* 2007;**26**(2):375–85.

12. De Crop A, Smeets P, Van Hoof T, Vergauwen M, Dewaele T, Van Borsel M, et al. Correlation of clinical and physical-technical image quality in chest CT: a human cadaver study applied on iterative reconstruction. *BMC Med Imag* 2015;**15**(1):32.
13. Ludewig E, Richter A, Frame M. Diagnostic imaging—evaluating image quality using visual grading characteristic (VGC) analysis. *Vet Res Commun* 2010;**34**(5):473–9.
14. Zarb F, McEntee MF, Rainford L. Visual grading characteristics and ordinal regression analysis during optimisation of CT head examinations. *Insights Imag* 2015;**6**(3):393–401.
15. Zarb F, Rainford L, McEntee MF. Image quality assessment tools for optimization of CT images. *Radiography* 2010;**16**(2):147–53.
16. Månsson L. Methods for the evaluation of image quality: a review. *Radiat Protect Dosim* 2000;**90**(1–2):89–99.
17. Seeram E, Davidson R, Bushong S, Swan H. Image quality assessment tools for radiation dose optimization in digital radiography: an overview. *Radiol Technol* 2014;**85**(5):555–62.
18. Zarb F, Rainford L, McEntee MF. AP diameter shows the strongest correlation with CTDI and DLP in abdominal and chest CT. *Radiat Protect Dosim* 2010;**140**(3):266–73.
19. Dohan A, Gavini J-P, Placé V, Sebbag D, Vignaud A, Herbin C, et al. T2-weighted MR imaging of the liver: qualitative and quantitative comparison of SPACE MR imaging with turbo spin-echo MR imaging. *Eur J Radiol* 2013;**82**(11):e655–61.
20. Frydrychowicz A, Nagle SK, D'Souza SL, Vigen KK, Reeder SB. Optimized high-resolution contrast-enhanced hepatobiliary imaging at 3 tesla: a cross-over comparison of gadobenate dimeglumine and gadoxetic acid. *J Magn Reson Imag* 2011;**34**(3):585–94.
21. Pijl ME, Doornbos J, Wasser MN, van Houwelingen HC, Tollenaar RA, Bloem JL. Quantitative analysis of focal masses at MR imaging: a plea for standardization 1. *Radiology* 2004;**231**(3):737–44.
22. McRobbie DW, Moore EA, Graves MJ, Prince MR. *MRI from picture to proton*. Cambridge University Press; 2007.
23. Firbank M, Coulthard A, Harrison R, Williams E. A comparison of two methods for measuring the signal to noise ratio on MR images. *Phys Med Biol* 1999;**44**(12):N261.
24. Kaufman L, Kramer DM, Crooks LE, Ortendahl DA. Measuring signal-to-noise ratios in MR imaging. *Radiology* 1989;**173**(1):265–7.
25. Stecco A, Brambilla M, Puppi A, Lovisolo M, Boldorini R, Carriero A. Shoulder MR arthrography: in vitro determination of optimal gadolinium dilution as a function of field strength. *J Magn Reson Imag* 2007;**25**(1):200–7.
26. Båth M. *Personal communication between the author*. 2016.
27. Brennan PC, McEntee M, Evanoff M, Phillips P, O'Connor WT, Manning DJ. Ambient lighting: effect of illumination on soft-copy viewing of radiographs of the wrist. *Am J Roentgenol* 2007;**188**(2):W177–80.
28. Samei E, Badano A, Chakraborty D, Compton K, Cornelius C, Corrigan K, et al. Assessment of display performance for medical imaging systems: executive summary of AAPM TG18 report. *Med Phys* 2005;**32**(4):1205–25.
29. Noble WS. How does multiple testing correction work? *Nat Biotechnol* 2009;**27**(12):1135–7.
30. Altman D. *Practical statistics for medical research*. London: Chapman and Hall; 1991.
31. Watson P, Petrie A. Method agreement analysis: a review of correct methodology. *Theriogenology* 2010;**73**(9):1167–79.
32. Ramalho M, Heredia V, Tsurusaki M, Altun E, Semelka RC. Quantitative and qualitative comparison of 1.5 and 3.0 Tesla MRI in patients with chronic liver diseases. *J Magn Reson Imag* 2009;**29**(4):869–79.
33. Tsurusaki M, Semelka RC, Zapparoli M, Elias J, Altun E, Pamuklar E, et al. Quantitative and qualitative comparison of 3.0 T and 1.5 T MR imaging of the liver in patients with diffuse parenchymal liver disease. *Eur J Radiol* 2009;**72**(2):314–20.
34. Kim BS, Lee KR, Goh MJ. New imaging strategies using a motion-resistant liver sequence in uncooperative patients. *BioMed Res Int* 2014;**2014**. 142658.
35. Merkle EM, Dale BM. Abdominal MRI at 3.0 T: the basics revisited. *Am J Roentgenol* 2006;**186**(6):1524–32.
36. Ohno T, Isoda H, Furuta A, Togashi K. Non-contrast-enhanced MR portography and hepatic venography with time-spatial labeling inversion pulses: comparison at 1.5 Tesla and 3 Tesla. *Acta Radiol Short Rep* 2015;**4**(5). 2058460115584110.
37. Yang RK, Roth CG, Ward RJ, deJesus JO, Mitchell DG. Optimizing abdominal MR imaging: approaches to common problems 1. *Radiographics* 2010;**30**(1):185–99.