Magnetic Resonance Imaging 63 (2019) 114–122

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

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

Original contribution

Susceptibility based multiparametric quantification of liver disease: Non-invasive evaluation of steatosis and iron overload

Monika Uhrig[a,1], Johannes Mueller[b,1], Thomas Longerich[c], Beate Katharina Straub[c,d], Lukas R. Buschle[a,e], Heinz-Peter Schlemmer[a], Sebastian Mueller[b], Christian H. Ziener[a,*]

[a] German Cancer Research Center (DKFZ), Department of Radiology, D-69120 Heidelberg, Germany
[b] Dept. of Medicine, Salem Medical Center and Center for Alcohol Research, University Hospital Heidelberg, D-69120 Heidelberg, Germany
[c] Dept. of Pathology, University Hospital Heidelberg, D-69120 Heidelberg, Germany
[d] Dept. of Pathology, University Hospital Mainz, D-55131 Mainz, Germany
[e] Faculty of Physics and Astronomy, University of Heidelberg, D-69120 Heidelberg, Germany

ARTICLE INFO

Keywords:
MR spectroscopy
Susceptometry
Room temperature susceptometer
Iron overload
CAP
Ferritin

ABSTRACT

Purpose: To evaluate if single-voxel MR spectroscopy (MRS) of iron and fat correlates with biopsy results of hepatic steatosis and iron overload, and to compare MR-measurements with room-temperature susceptometer (RTS), ultrasound, controlled attenuation parameter (CAP) and serum ferritin.

Material and methods: In this prospective study, a set of 42 patients out of 47 screened patients with several chronic liver diseases underwent MRI-examination at 1.5 T including R2-measurements by single-voxel high-speed T2-corrected multiecho spectroscopy, additional liver biopsy, abdominal ultrasound, CAP, and RTS. Routine blood and serum parameters were determined, including ferritin. Atomic absorption spectroscopy (AAS) and histologically confirmed extent of hepatic steatosis from liver biopsy were used as reference standard. For correlation of R2, RTS, CAP, ferritin, and ultrasound with results of AAS and histologically determined fat fraction of liver biopsy specimen, Spearman's and Pearson's correlation as well as receiver operating characteristics curve (ROC) analysis with cut-off values determined by maximizing Youden index was used.

Results: MRS iron assessment correlated best with AAS, with a Pearson correlation coefficient of 0.715 ($p < 0.001$), followed by RTS 0.520 ($p < 0.001$), and serum ferritin 0.213 ($p = 0.088$, not significant). MRS fat quantification correlated best with the histological confirmed extent of steatosis hepatis with a Spearman correlation coefficient of 0.836 ($p < 0.001$), followed by CAP 0.604 ($p < 0.001$) and sonographically diagnosed steatosis 0.358 ($p = 0.013$).

Conclusion: MRS by T2-corrected multiecho single-voxel spectroscopy correlated best with histological results of hepatic fat and iron content compared to RTS, CAP, abdominal ultrasound, and ferritin. Non-invasive methods to assess hepatic fat and iron are of clinical interest for follow-up examinations of patients with chronic liver diseases, where repeated biopsy is not indicated.

1. Introduction

Hepatic iron overload and steatosis can be caused by several chronic liver diseases [1]. Besides the alcoholic liver disease (ALD), the non-alcoholic fatty liver disease (NAFLD), viral hepatitis and several other diseases like M. Wilson and hereditary haemochromatosis can cause fatty degeneration and iron overload [2]. Furthermore, liver iron overload due to repeated blood transfusions is an important side effect that has to be monitored especially in children with acute leukemia [3]. Hepatic steatosis can develop into fibrosis, which is a risk factor of hepatocellular carcinoma (HCC). Iron overload is highly toxic and carcinogenic as well [4].

The gold standard for the diagnosis of liver iron and fat levels is the liver biopsy, where iron is typically analyzed by atomic absorption spectroscopy (AAS), and steatosis can be assessed by quantifying the amount of liver cells degenerated by fat [2,5,6]. However, due to potential complications invasive liver biopsy is not recommended for treatment monitoring.

There exist several methods to non-invasively follow-up patients. Serum markers such as ferritin and transferrin saturation are the

* Corresponding author at: Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany.
E-mail address: c.ziener@dkfz.de (C.H. Ziener).
[1] M. Uhrig and J. Mueller contributed equally to the manuscript and jointly share first authorship.

https://doi.org/10.1016/j.mri.2019.08.016
Received 10 May 2019; Received in revised form 11 July 2019; Accepted 15 August 2019

preferred screening tool for iron overload, even if there are many confounding variables like inflammation, cancer, and certain types of anemia, which can lead to over- or underestimation [7]. Abdominal sonography can be used to estimate the degree of steatosis by comparing echogenicity of the liver parenchyma to surrounding structures [8]. A more advanced technique is the measurement of the controlled attenuation parameter (CAP) based on the ultrasound wave attenuation, which is increasing with increasing hepatic fat content [9,10].

Techniques which determine non-invasively the magnetic susceptibility, e.g. using a Superconducting Quantum Interference Device (SQUID), are more precise but rarely available [11–13]. SQUID requires liquid helium, is expensive, and worldwide only three devices (Oakland (USA), Torino (Italy) and Hamburg (Germany)) for clinical use exist [14]. In contrast, room-temperature susceptometry (RTS) does not need helium and uses less expensive sensor technology [15,16]. Until now, RTS has not been established in clinical routine.

Imaging technologies to assess iron and fat content of the liver include dual-energy computed tomography, magnetic resonance imaging (MRI), and MR spectroscopy (MRS) [17–21]. Compared to SQUID and RTS, MRI is more widely available. T1-weigthed in- and opposed phase MRI-sequences as well as proton density fat fraction assessment allow estimating hepatic steatosis [19]. For iron overload estimation, different techniques are available, including T2 and T2* weighted images, quantitative relaxometry, and susceptometry [18]. The ability of MR spectroscopy to separate the water resonance from other confounding proton resonances, such as lipid, is a main difference to MRI, measuring a summed signal from each imaging voxel [21]. In all MR techniques, quantification of liver iron is based on the magnetic properties of ferritin and hemosiderin. These complexes agglomerate inside hepatic cells and create local magnetic field inhomogeneities in the form of a three-dimensional dipole field [22,23]. The strength of the local dipole field depends on the external magnetic field strength B_0 and the susceptibility difference $\Delta\chi$ between the iron agglomerations and the surrounding interstitial tissue [24]. These additional local dipole fields influence the dephasing process of the spin-bearing water and fat molecules in the tissue surrounding the iron agglomerations, resulting in an acceleration of the measured free induction decay [25]. Consequently, the Fourier transform of the signal decay, also known as the spectrum or the line-shape, is broadened. Since measurements in patients are performed at room temperature, magnetic resonance signal formation is not exclusively determined by the strength of the local dipole field, but is affected by the diffusion process of the spin-bearing water molecules as well. This diffusion process is characterized by the diffusions coefficient D of water molecules in the vicinity of the iron depositions. Due to the small radius R of the iron depositions, the dynamic frequency scale, D/R^2 is much smaller than the static frequency scale $\gamma \cdot \Delta\chi \cdot B_0$ [26]. Thus, the underlying diffusion regime is the motional narrowing regime where an increase of diffusion results in a narrowing of the spectral width and consequently to a slower signal decay. In contrast to the porous lung parenchyma, where the large size of the alveoli results in an asymmetric spectrum, the spectrum of liver tissue exhibits a symmetric Lorentzian shape [27].

The purpose of this study was to compare the susceptibility-based methods single- voxel MRS and RTS with ultrasound, CAP, and ferritin measurements using liver biopsy as gold standard.

2. Material and methods

2.1. Patient population

This prospective study was approved by the institutional review board and written informed consent was obtained from all patients.

Between May 2015 and October 2018, 47 patients were screened (Fig. 1). Five of the patients were excluded for the following reasons: One patient had an abnormally high AAS result (32 mg/g dw) but no stainable iron in histology, low ferritin, and low MRI and RTS signals; one patient denied liver biopsy; one patient stopped MRI examination because of claustrophobia; in one patient, MRI iron and fat assessment failed, and another patient had an unknown implant preventing MRI examination. In total, 42 patients with several diffuse liver diseases underwent the study MRI examination and had additional liver biopsy, abdominal ultrasound, CAP, and RTS measurements [6]. The diagnosed liver diseases included alcoholic liver disease (ALD, N = 12), non-alcoholic fatty liver disease (NAFLD, N = 7), haemochromatosis (N = 6), elevated ferritin or transaminase of unknown origin (N = 7), liver cirrhosis (N = 6), auto-immune hepatitis (N = 3), and M. Wilson (N = 1). For all patients, routine blood and serum parameters were determined, including ferritin (ng/mL). Liver biopsy, CAP, US, RTS, and MRI were performed within a one-week interval. Table 1 summarizes the patient characteristics.

2.2. MRI examination

For the quantification of the transverse relaxation rate of the liver, patients underwent MRI of the abdomen in a 1.5 Tesla MRI scanner (Siemens Aera) using a surface phased-array coil with 18 receiver channels. Standard morphological sequences including a transversal single shot T2 weighted echo-planar fast spin echo sequence (T2-HASTE: *H*alf-Fourier *A*cquisition *S*ingle-shot *T*urbo spin *E*cho imaging; TR 1400 ms, TE 95 ms) with and without spectral fat saturation, a triggered T2-turbo spin echo (TR 2793.8 ms, TE 100 ms), a gradient-echo T1 weighted sequence in breath hold technique (T1-VIBE: *V*olume *I*nterpolated *B*reathhold *E*xamination; TR 7 ms, TE 4.77 ms) transversal and coronal, and a diffusion weighted imaging (TR 6900 ms, b-values of 0, 50, 100, 200, 300, 400, 500, 600, 700, 800). For segmentation of the liver, a transversal gradient-echo T1 weighted sequence in breath hold and Dixon technique (T1-VIBE-Dixon; TR 7 ms, TE 4.77 ms and 2.39 ms) was used. For evaluation of fat and iron content, the relaxation rate R2 of fat and water from a high-speed T2-corrected multi-echo acquisition at 1H MR spectroscopy (HISTO-sequence), provided by the manufacturer (liver evaluation tool LiverLab, Siemens), was used (Fig. 2) [28]. This single-voxel spectroscopy uses a stimulated echo acquisition mode (STEAM) at five different echo times TE = 12, 24, 36, 48, 72 ms with repetition time TR = 3000 ms, resulting in a total acquisition time of 15 s acquired in one breath hold. A bandwidth of 1200 Hz with 1024 points was used, with one signal average per echo and with a voxel size of 3 cm × 3 cm × 3 cm [28]. The measurement was preceded by a 30s standard shimming. Simple linear baseline correction and low-pass sliding window filtering was applied for the acquired spectra. The algorithm estimates the integral areas of the water (3.6 to 5.8 ppm) and lipid (0 to 3.6 ppm) peaks for each measured echo. From the water and fat integrals, a monoexponential T2-decay fit with least-squares approximation was calculated with $S = M_0 \exp.(-R2 \cdot TE)$, leading to an estimation of the relaxation coefficient R2 and equilibrium magnetization M_0 for water and fat individually. Thus, HISTO provides T2-corrected lipid content by percentage of fat $= [M_{0,lipid}/(M_{0,lipid} + M_{0,water})] \cdot 100$.

The spectroscopic voxel was positioned in homogeneous parenchyma of the right liver lobe in distance to larger vessels and tissue boundaries to avoid susceptibility effects and static background gradients. The main difference between HISTO and other spectroscopy techniques based on one spectrum acquired at one single echo time is, that it allows the simultaneous measurement of R2 of water and R2 of fat, offering corrections for T2 effects. Spectroscopy results using only one echo time can vary depending on the chosen echo time. This effect is increasing with increasing susceptibility effects, e.g. in case of iron overload. Furthermore HISTO is designed for fast clinical application with only 15 s acquisition time in one breath hold [28].

2.3. Abdominal ultrasound, RTS, and CAP

For each patient, a clinical routine abdominal ultrasound (HI vision

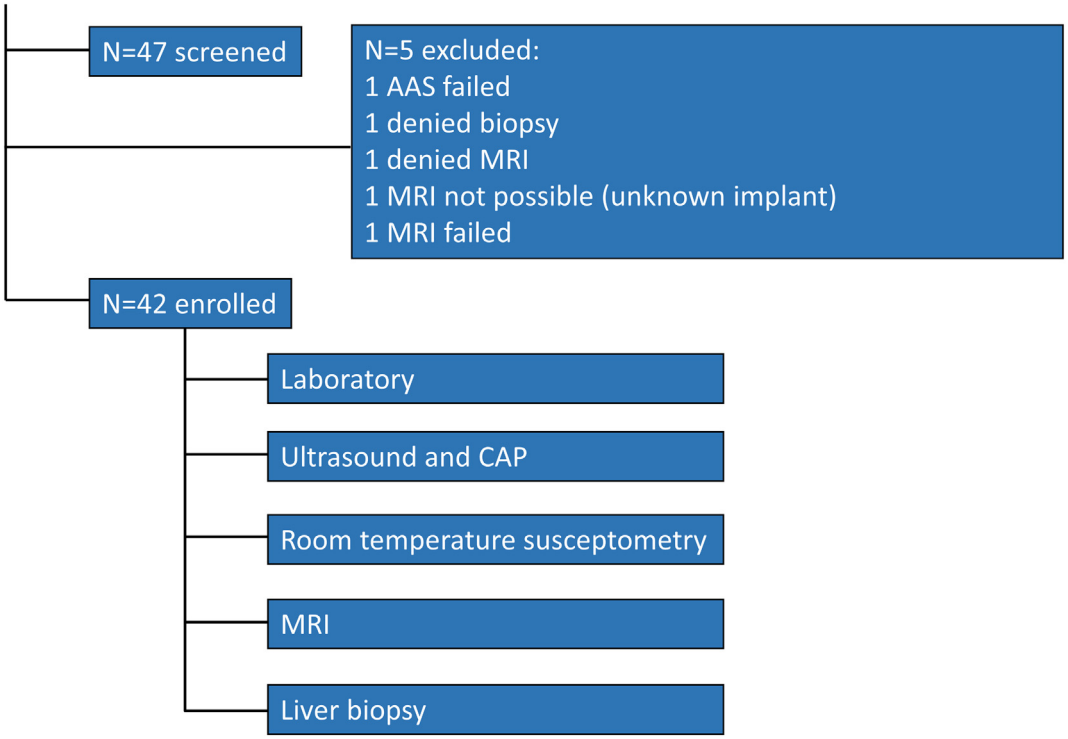

Fig. 1. Study Design. AAS: atomic absorption spectroscopy; CAP: controlled attenuation parameter.

Preirus, Hitachi) was performed and the degree of hepatic steatosis was assessed, with scores from 0 (no fatty infiltration), 1 (moderate: same echogenicity of portal vein branch walls and surrounding liver parenchyma), 2 (medium: hyperechogenic liver obscures the diaphragmatic outline) to 3 (severe fatty infiltration of the complete liver parenchyma). Liver size was measured in the median axillar line (MAL).

Furthermore, CAP was measured, which reflects the attenuation of the ultrasound wave [dB/m] [9,10]. This technique is coupled with a transient elastography measurement (Fibroscan, Echosense SA, Paris, France), using the M or XL probe [29,30]. Patients with ascites could successfully be examined with the XL probe when ascites lamina was < 30 mm.

RTS was performed after abdominal ultrasound, allowing to select an appropriate point for RTS measurements.

The basic idea behind RTS is, similar to SQUID, measuring the magnetic susceptibility relative to water. Comparable with MRI and SQUID, the liver iron content can be determined by the magnetic volume susceptibility of liver iron ($1600 * 10^{-6}$ ppm) [11–13]. The RTS is operated at Salem Medical Center, Heidelberg. The main difference to a SQUID is, that it does use copper instead of superconducting coils. It consists of two sensing coils of different diameter and a field generating coil, which produces a weak AC-magnetic field at 575 Hz. The signals of both channels are post-processed using a software-based lock-in amplifier. The magnetic susceptibility is measured relative to water. The sensor is placed in the abdominal region near the liver, at a point at least 5 cm away from the lung. For one measurement, the patient exhales and holds this state for 10 s. One RTS examination consisted of nine successive measurements with two times repositioning of the sensor. Before and after the patient measurement, a water reference measurement was performed in order to correct a possible background signal drift. The resulting liver iron concentration is the mean of the nine measurements and expressed in units of μg/g wet weight. A more detailed description of the method can be found elsewhere [15,16].

2.4. Liver biopsy and atomic absorption spectroscopy

Liver biopsy was performed in the center of the right liver lobe according to Menghini technique [31,32]. Ultrasound-guidance prevented damage of vessels and other critical structures. Samples were dried, weighed and analyzed by atomic absorption spectroscopy (AAS) (GTA120 AA 240-z, Agilent) in the institutional laboratory. Results of AAS were the gold standard for our study. We defined moderate iron overload by any liver iron concentration (LIC) in AAS ≥ 2 mg/g dry weight (dw) and severe iron overload by LIC in AAS ≥ 4 mg/g dw.

For histological analysis, samples were fixed in formalin and embedded in paraffin. Scoring of steatosis, lobular inflammation, hepatocellular ballooning and fibrosis staging was performed according to Kleiner et al. by two experienced pathologists, blinded to patients data [6].

2.5. Statistical analysis

Ordinal parameters (e.g. histological scores) were correlated using Spearman's rank correlation coefficient (correlation coefficient r, p). Metric parameters were assessed using Pearson's correlation coefficient (r, p). All p-values for correlations are one-sided, and Bonferroni-adjusted p-values corrected for multiple comparisons are shown separately. Differences were considered as significant at $p < 0.05$ for uncorrected values. In order to give a quantitative measure of precision of MRS, RTS and ferritin with respect to AAS, we converted MRS, RTS and ferritin into AAS units [mg/g dw] using a linear regression model. The deviation from AAS was then calculated as the standard deviation over the differences to AAS measurements for all three parameters in mg/g dw. Performance of the different methods (MRI, RTS, CAP, ultrasound, serum-ferritin), compared to the gold standard AAS, was determined by receiver operating characteristics curve (ROC) analysis, with cut-off values determined by maximizing the Youden index. Comparison of ROCs was performed using the Delong-test. By means of the defined cut-off values, the sensitivity, specificity, accuracy and positive predictive value (PPV) were calculated based on true and false negative

Table 1
Patient characteristics.

	Total N = 42	ALD N = 12	NAFLD N = 7	Haemochr. N = 6	Else N = 17
Morphological data					
Gender (male)	64%	58%	71%	100%	53%
Age (years)	53.7 ± 13.8	56.6 ± 11.0	52.1 ± 15.7	53.7 ± 14.3	52.5 ± 15.1
BMI (kg/m^2)	26.6 ± 4.3	25.9 ± 5.1	26.6 ± 4.6	27.5 ± 4.0	26.8 ± 4.1
Laboratory					
AST (U/L)	116 ± 154	151 ± 180	68 ± 28	35 ± 12	151 ± 188
ALT (U/L)	159 ± 294	97 ± 111	97 ± 46	51 ± 26	276 ± 441
GGT (U/L)	323 ± 573	667 ± 959	156 ± 114	50 ± 30	290 ± 369
Bilirubin total (mg/dL)	1.8 ± 3.1	3.0 ± 5.0	0.8 ± 0.4	0.8 ± 0.5	1.7 ± 2.5
Hemoglobin (g/dL)	14.2 ± 2	12.9 ± 2.4	15.9 ± 1.4	14.8 ± 1.9	14.0 ± 1.6
Ferritin (ng/mL)	1093 ± 741	1393 ± 703	754 ± 419	824 ± 713	1178 ± 832
Transferrin (g/L)	2.4 ± 0.7	1.9 ± 0.5	2.9 ± 0.6	2.2 ± 0.5	2.7 ± 0.7
Transferrin saturation (%)	41 ± 22	53 ± 26	28 ± 9	48 ± 25	34 ± 15
Ultrasound (US)					
Liver size (MAL, cm)	16.2 ± 3.3	17.9 ± 3.7	14.3 ± 3.6	14.6 ± 1.7	16.8 ± 2.8
Hepatic steatosis (US)	1.6 ± 1.0	2 ± 0.5	1.4 ± 1.4	1.3 ± 1.2	1.6 ± 0.9
Spleen size (cm)	10.8 ± 3.1	10.9 ± 3.1	10.6 ± 1.5	11.4 ± 3.0	11.0 ± 2.5
Ascites (0–1)	7.5%	20.0%	0.0%	0.0%	6.3%
Signs of cirrhosis (0–1)	16.1%	20.0%	0.0%	16.7%	18.2%
Histology & AAS					
Kleiner steatosis (0–3)	1.3 ± 1.0	1.5 ± 1.0	1.6 ± 1.1	1.4 ± 1	0.9 ± 1.1
Fibrosis stage (0–4)	2.3 ± 1.2	3.0 ± 1.2	2.0 ± 1.1	1.9 ± 1.2	2.2 ± 1.2
Iron-RES (0–3)	0.6 ± 0.9	0.8 ± 1.0	0.4 ± 0.7	1.0 ± 1.4	0.4 ± 0.5
Iron hepatocytes (0–3)	0.8 ± 1.1	0.9 ± 1.1	0.4 ± 0.5	1.7 ± 1.4	0.4 ± 0.8
LIC-AAS (mg/g dw)	2.5 ± 2.8	1.7 ± 1.5	1.7 ± 0.6	6.2 ± 5.9	2.1 ± 1.5
Transient elastography					
Liver stiffness (kPa)	15.0 ± 16.2	21.3 ± 17.6	10.9 ± 6.1	7.0 ± 3.5	16.6 ± 20.3
CAP (dB/m)	294 ± 64	291 ± 54	300 ± 55	300 ± 86	290 ± 66
MRI					
Fat fraction (%)	14.4 ± 11.5	15.8 ± 13.5	18.3 ± 11.0	14.6 ± 11.0	11.8 ± 10.9
R2 water	31.9 ± 7.7	30.2 ± 5.1	33.5 ± 5.5	37.7 ± 11.8	29.6 ± 6.2
RTS					
LIC-RTS (μg/g ww)	163 ± 323	122 ± 293	23 ± 295	437 ± 397	119 ± 260

+/−: standard deviation; ALD: alcoholic liver disease; NAFLD: non-alcoholic fatty liver disease; Haemochr.: haemochromatosis; Else: autoimmune hepatitis (N = 3), unknown ferritin or transaminase elevation (N = 7), M. Wilson (N = 1), Liver cirrhosis (N = 6); BMI: body mass index; AST: aspartate aminotransferase; ALT: alanine aminotransferase; GGT: γ-glutamyltransferase; MAL: median axillar line; ascites: percentage of patients with ascites; iron-RES: iron in reticuloendothelial system; LIC: liver iron concentration; AAS: atomic absorption spectroscopy; dw: dry weight; ww: wet weight; CAP: controlled attenuation parameter.

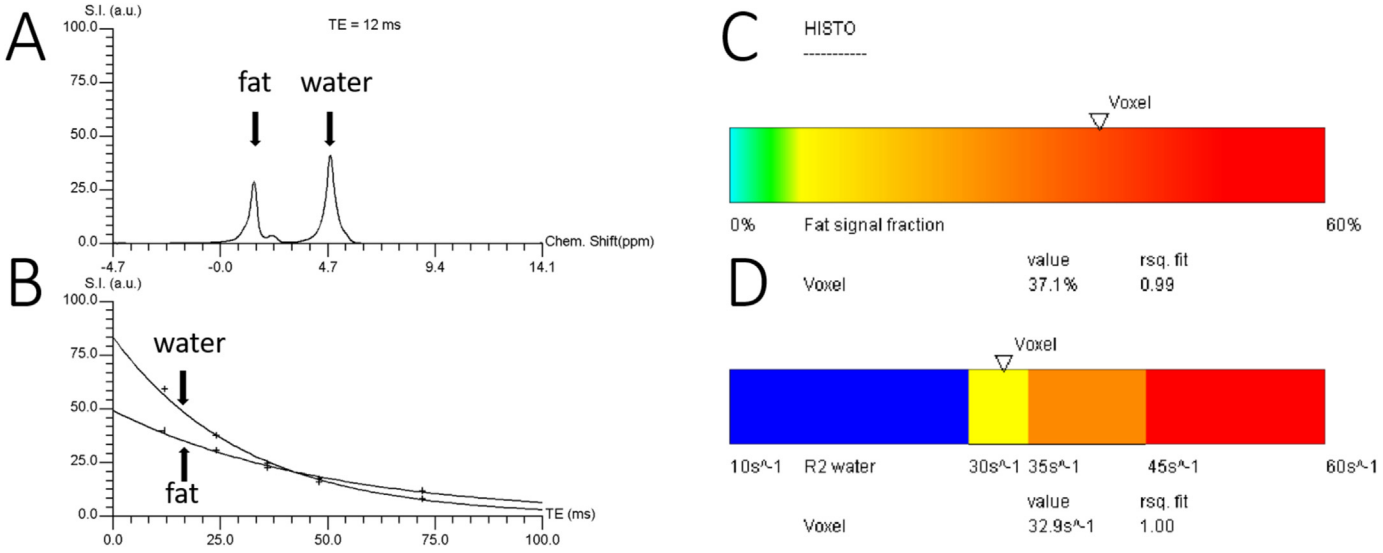

Fig. 2. Acquisition of high-speed T2-corrected multi-echo spectroscopy. Spectra at echo time TE = 12 ms (A) and signal decay (B) for water and fat signal. Evaluation report with fat signal fraction (C) and relaxation rate (D) within voxel.

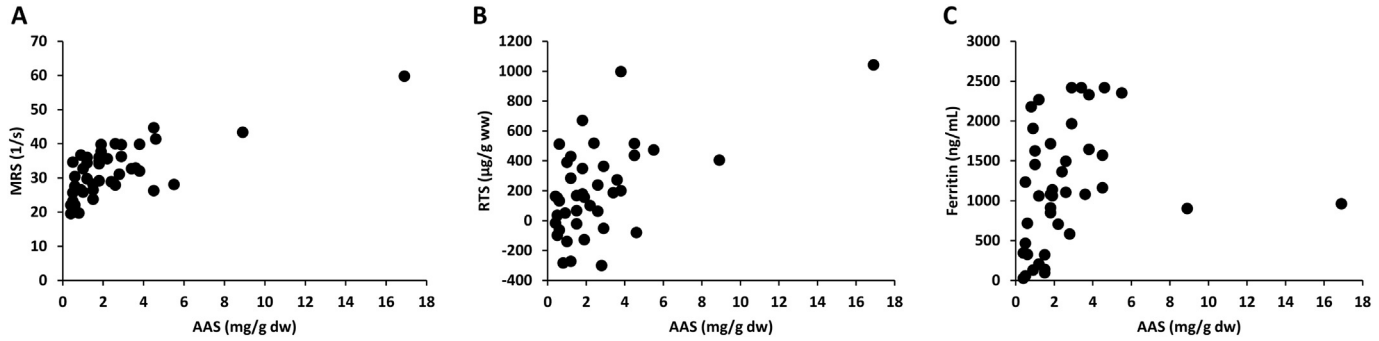

Fig. 3. Scatter plots for the correlations between the gold standard AAS (atomic absorption spectroscopy from liver biopsy probes) and the different techniques to assess liver iron content in vivo. MRS: Magnetic resonance spectroscopy; RTS: liver iron concentration measured by room temperature susceptometry.

and positive values. The corresponding confidence intervals (CI 95%) for sensitivity, specificity and accuracy were calculated using the Wilson score interval. Statistical analysis was performed using the software PASW Statistics 18, version 18.0.0 (SPSS, Inc., Munich, Germany).

3. Results

3.1. Iron

Among the 42 patients, N = 25 (59.5%) of the patients had normal liver iron with LIC in AAS of < 2 mg/g dw, N = 11 (26.2%) had moderate iron overload of LIC between 2 and 4 mg/g dw, and N = 6 (14.3%) % had LIC > 4 mg/g dw. The highest result was 16.9 mg/g dw in a haemochromatosis patient.

Measurements of R2 water correlated best with the gold standard AAS from liver biopsy, with a Pearson correlation coefficient of 0.715 ($p < 0.001$, adjusted $p < 0.001$), followed by RTS 0.520 ($p < 0.001$, adjusted $p < 0.001$), and serum ferritin 0.213 (0.088, adjusted $p = 0.262$, not significant) (see Fig. 3 for scatter plots). For patients with AAS-determined iron ≤2 mg/g dw, R2 correlated as well with the MRS-determined hepatic fat fraction with a Pearson correlation coefficient of 0.421 ($p < 0.05$). For AAS > 2 mg/g, the correlation coefficient between AAS and R2 was 0.74 ($p < 0.001$, adjusted $p < 0.001$), between AAS and RTS 0.58 ($p = 0.007$, adjusted $p = 0.021$), and between AAS and ferritin −0.20 ($p = 0.216$, adjusted $p = 0.646$, not significant). The precision, calculated as described in Section 2.5, was 2.0 mg/g dw for MRS, 2.4 mg/g dw for RTS and 7.5 mg/g dw for ferritin.

In the ROC-analysis, MRS and RTS performed better than serum ferritin for severe iron overload. ROC curves can be seen in Fig. 4. The area under the curve (AUROC) for severe iron overload (defined by LIC in AAS ≥ 4 mg/g dw) was 0.76 (95%CI 0.48–1.00), 0.80 (95%CI 0.57–1.00), and 0.69 (95%CI 0.48–0.89) for MRS, RTS, and ferritin, respectively, and 0.75 (95%CI 0.61–0.90), 0.72 (95%CI 0.56–0.88), and 0.77 (95%CI 0.64–0.91) for moderate iron overload (defined by LIC in AAS ≥ 2 mg/g dw). Using the Delong-test, the difference between the ROCs did not reach statistical significance.

The diagnostic accuracy for R2-measurements was 0.64 (95%CI 0.49–0.77) and 0.95 (95%CI 0.84–0.99) for moderate and severe iron overload, while the accuracy for RTS measurements was 0.74 (95%CI 0.59–0.85) and 0.86 (95%CI 0.72–0.93), respectively. Serum ferritin had accuracies of 0.67 (95%CI 0.52–0.79) and 0.5 (95%CI 0.36–0.65) (Table 2). The optimal cutoff value for moderate and severe iron overload was an R2-value of 27.8 s^{-1} and 40.7 s^{-1}, respectively, for MRS and 187 and 397 µg/g ww for RTS. The optimal cutoff value for ferritin was 878 ng/mL both for moderate and severe iron overload.

The box plots given in Fig. 5 underline the result of the positive correlation between iron content determined in liver biopsy and MRS-based determination of the relaxation rate R2.

3.2. Fat

N = 10 (23.8%) of the patients did not show any steatosis in histology, N = 14 (33.3%) had steatosis grade S1, N = 10 (23.8%) grade S2, and N = 8 (19.0%) grade S3. MRS-fat quantification correlated best with the histological confirmed extent of hepatic steatosis with a Spearman correlation coefficient of 0.836 ($p < 0.001$, adjusted $p < 0.001$), followed by CAP 0.604 ($p < 0.001$, adjusted $p < 0.001$) and sonographically diagnosed steatosis of 0.358 ($p = 0.013$, adjusted $p = 0.038$).

ROC analysis including optimal cutoff values are given in Table 3

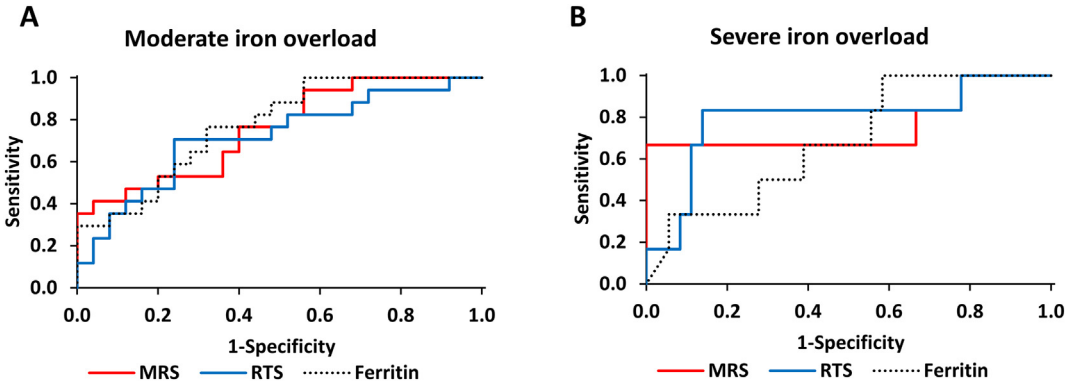

Fig. 4. Receiver operating curves (ROC) for iron measurements with MRS, RTS and serum ferritin in case of moderate (≥2 mg/g dw) iron overload (A) and severe (≥4 mg/g dw) iron overload (B). The area under the curve (AUROC) for moderate iron overload was 0.75 (95%CI 0.61–0.90), 0.72 (95%CI 0.56–0.88) and 0.77 (95%CI 0.64–0.91) and 0.76 (95%CI 0.48–1.00), 0.80 (95%CI 0.57–1.00) and 0.69 (95%CI 0.48–0.89) for severe iron overload for MRS, RTS and ferritin, respectively. Curves were not significantly different ($p > 0.10$, Delong-test). dw: dry weight; RTS: room temperature susceptometry.

Table 2
Diagnostic performance of iron measurements.

	Optimal cutoff	Sensitivity	Specificity	Accuracy	PPV
AAS ≥ 2 mg/g dw					
MRI (s^{-1})	27.8	0.94 (0.73–0.99)	0.44 (0.27–0.63)	0.64 (0.49–0.77)	0.53 (0.36–0.70)
RTS (µg/g ww)	183	0.71 (0.47–0.87)	0.76 (0.57–0.89)	0.74 (0.59–0.85)	0.67 (0.44–0.84)
Ferritin (ng/mL)	878	0.88 (0.66–0.97)	0.52 (0.34–0.70)	0.67 (0.52–0.79)	0.56 (0.37–0.72)
AAS ≥ 4 mg/g dw					
MRI (s^{-1})	40.7	0.67 (0.30–0.90)	1.00 (0.90–1.00)	0.95 (0.84–0.99)	1.00 (0.51–1.00)
RTS (µg/g ww)	397	0.83 (0.44–0.97)	0.86 (0.71–0.94)	0.86 (0.72–0.93)	0.50 (0.24–0.76)
Ferritin (ng/mL)	878	1.00 (0.61–1.00)	0.42 (0.27–0.58)	0.50 (0.36–0.65)	0.22 (0.11–0.41)

Values in parentheses are 95% confidence intervals. AAS: atomic absorption spectroscopy; RTS: room temperature susceptometry; ww: wet weight; PPV: positive predictive value.

and Fig. 6. AUROCs according to the histologically confirmed steatosis grade for MRS measurements were 0.98 (0.95–1.00) for ≥S1, 0.91 (95%CI 0.81–1.00) for ≥S2, and 0.90 (95%CI 0.79–1.00) for S3, while AUROCs for CAP were 0.79 (95%CI 0.62–0.95), 0.84 (95%CI 0.72–0.97) and 0.80 (95%CI 0.65–0.94), respectively. By performing the Delong-test it was found, that only for steatosis grade S1, ROC of MRS was significantly different from CAP (p = 0.023).

The diagnostic accuracy for MRS measurements was 0.98 (95%CI 0.87–1.00), 0.88 (95%CI 0.74–0.95), and 0.78 (95%CI 0.63–0.88) for steatosis grades S1, S2, and S3, while the accuracy for CAP measurement was 0.74 (95%CI 0.59–0.85), 0.64 (95%CI 0.48–0.77), and 0.74 (95%CI 0.59–0.85). Steatosis grading by ultrasound reached accuracies of 0.74 (95%CI 0.59–0.85), 0.76 (95%CI 0.65–0.89), and 0.67 (95%CI 0.51–0.79), respectively.

The optimal cut-off values for determining a histologically proven hepatic steatosis with MRS were 4.5%, 10.7%, and 15.9% for S1, S2, and S3, respectively. For CAP measurements, cut-off values of 276, 282, and 289 dB/m were found.

The box plots given in Fig. 7 underline the positive correlation between histologically confirmed steatosis grade and MRS-based determination of the fat fraction.

4. Discussion

Results of this study indicate, that single-voxel MRS for determination of fat and iron content of the liver correlated well with quantitative iron and fat determination by AAS and histology. It performed better than RTS, ultrasound, CAP, and ferritin measurements.

Non-invasive methods to assess hepatic fat and iron are of clinical interest for follow-up studies where repeated biopsy is not indicated.

For iron overload, the preferred serum markers are ferritin and transferrin saturation, even if the limitations have been shown before [2,33]. Accordingly, this study demonstrated a weak non-significant positive correlation between serum ferritin and AAS histology results. Compared to established clinical serum ferritin reference values, we found a relatively high optimal cutoff value for iron overload. One reason could be the heterogeneity of the patient cohort including patients with chronic liver disease, where ferritin levels do not mandatorily reflect iron overload and are influenced by inflammation and liver damage as well [2].

Compared to serum ferritin, RTS had a better correlation to the biopsy-confirmed iron overload. RTS is less expensive than SQUID and does not have the limiting requirement of a high magnetic field, but it has not entered clinical routine and is rarely available [16]. However, the single-voxel MRS performed in this study correlated best with the gold standard AAS.

Previous studies have shown that the presence of fat can influence R2* from gradient echo sequences [34–36], not all of them with AAS as reference standard, using instead semiquantitative methods (Prussian blue staining). In our study, a single-voxel MRS was performed, separating the water resonance from the confounding lipid peak. However, for patients with a relatively low liver AAS-determined iron content of ≤2 mg/g dw, we observed a moderate positive correlation between MRI-determined iron content and hepatic fat fraction as well.

For the diagnosis of hepatic steatosis, the most accessible imaging technique for many patients is certainly the ultrasound examination. This study indicates that there is only a weak positive correlation between ultrasound grading and histology. This goes in line with previous studies showing that ultrasound is a good method to diagnose severe hepatic steatosis, while there is a significant overlap between normal parenchyma and mild to moderate steatosis [37].

CAP results showed a better correlation than standard ultrasound. Compared to a metaanalysis by Shi et al. [10], the diagnostic accuracy in this study was lower, probably due to the different cut-off values for the grading of the hepatic steatosis 1–3. Overall, CAP and ultrasound reached comparable results for detecting any steatosis, but MRS clearly

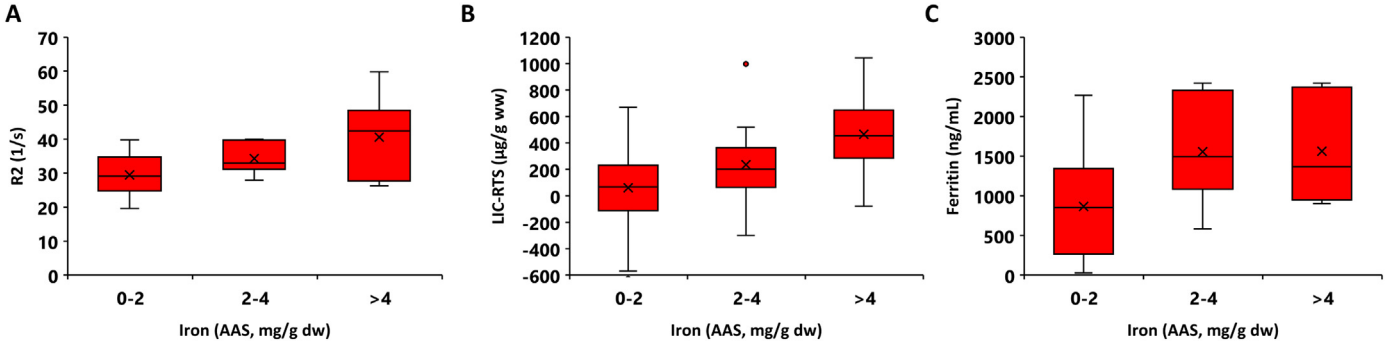

Fig. 5. Box plot analysis for iron measurements. Boxes represent the middle 50% of the data, whiskers reach to maximal 1.5 times the interquartile range above or below the box, other data points are shown as outliers. The horizontal line indicates the median, the x indicates the mean. For MRS and RTS, a linear relation to the AAS grades can be seen while no difference for moderate and severe iron overload can be seen for ferritin. AAS: atomic absorption spectroscopy; dw: dry weight; LIC-RTS: liver iron concentration measured by room temperature susceptometry.

Table 3
Diagnostic performance of fat measurements.

	Cutoff	Sensitivity	Specificity	Accuracy	PPV
≥S1					
MRI fat fraction (%)	4.5	1.00 (0.89–1.00)	0.90 (0.60–0.98)	0.98 (0.87–1.00)	0.97 (0.84–0.99)
CAP (dB/m)	276	0.86 (0.69–0.95)	0.40 (0.17–0.69)	0.74 (0.59–0.85)	0.81 (0.63–0.91)
US	1	0.76 (0.58–0.88)	0.70 (0.40–0.89)	0.74 (0.59–0.85)	0.88 (0.70–0.96)
≥S2					
MRI fat fraction (%)	10.7	1.00 (0.81–1.00)	0.79 (0.60–0.91)	0.88 (0.74–0.95)	0.76 (0.55–0.89)
CAP (dB/m)	282	0.73 (0.48–0.89)	0.58 (0.39–0.76)	0.64 (0.48–0.77)	0.52 (0.32–0.72)
US	2	0.93 (0.70–0.99)	0.71 (0.51–0.85)	0.80 (0.65–0.89)	0.67 (0.45–0.83)
S3					
MRI fat fraction (%)	15.9	1.00 (0.65–1.00)	0.73 (0.56–0.85)	0.78 (0.63–0.88)	0.44 (0.23–0.67)
CAP (dB/m)	289	0.33 (0.10–0.70)	0.82 (0.66–0.91)	0.74 (0.59–0.85)	0.25 (0.07–0.59)
US	3	1.00 (0.61–1.00)	0.61 (0.44–0.75)	0.67 (0.51–0.79)	0.32 (0.15–0.54)

Values in parentheses are 95% confidence intervals. S1/2/3: histologically confirmed hepatic steatosis; CAP: controlled attenuation parameter; US: ultrasound; PPV: positive predictive value.

correlated best with the histological grading of hepatic steatosis.

In general, MRI examinations are currently not performed routinely when chronic liver diseases, hepatic steatosis, or hepatic iron overload are suspected. Examination slots are rare, and compared to CT, ultrasound and blood samples, MRI is more expensive. Prospectively, abdominal MRI to assess fat and iron could be considered for follow-up examinations in cases of inconsistent results of the more available techniques. MRI measurements allow assessing iron content in other organs like heart and spleen in the same examination, which can be advantageous when suspecting systemic iron overload.

One of the limitations of this study is that results are only valid for the used MRI 1.5 Tesla MRI scanner. It has been shown that vendor-specific influences are possible [38]. Theoretically, one would expect different T2 values but similar results for fat fractions using 1.5 T and 3 T MRI scanners. Due to a lack of data, we cannot confirm this thesis with our experimental data. However, previous studies proved high reproducibility across field strength [39]. Due to a lack of data, we did not compare the described methods with CT-based quantification of fat and iron (e.g. dual energy CT). Furthermore, biopsy locations were not matched with the ROI used for the MRS. This may have increased the variation in the current study, meaning that the actual accuracy of MRS could be higher than reported in this study.

The used quantitative measurements for iron quantification rely on the measurement of the magnitude of the transverse relaxation. This is justified by the dominating influence of the diffusion effects in comparison with the extent of the iron depositions. Furthermore, the considered diffuse forms of parenchymal iron overload in steatosis or hemochromatosis allowed assuming a single global value of hepatic iron content. In contrary, for pathologies associated with focal iron overload or iron sparing regions e. g. dysplastic and regenerative nodules or periportal cavernomatous transformation [40], a liver iron map reflecting local iron concentrations would be advantageous. An accurate method to obtain these maps is quantitative susceptibility mapping (QSM), considering both the magnitude and the phase of the measured signal [41,42]. Compared to the HISTO sequence acquired in 15 s in one breath hold, QSM has longer acquisition times, depending on the algorithm used. Furthermore, there is no reference value and no fast and robust commercially available software for clinical routine. However, a correlation between QSM and other MRI-based iron quantification methods has been shown as well as a good correlation between QSM, $R2^*$ and Superconducting Quantum Interference Device (SQUID) [43,44]. At least in mouse models, the feasibility of studying liver iron overload with QSM and $R2^*$ at ultra-high field (7 T) has been demonstrated [45]. QSM still struggles to be widely used in the liver as the presence of fat causes additional phase shifts that have to be accounted for in the calculation of phase maps [46,47]. Very high iron concentrations cause rapid signal decay and therefore potentially very noisy phase images, and respiratory motion has to be adequately compensated for, otherwise causing severe artifacts.

In conclusion, this prospective study showed for a cohort 42 patients that single-voxel T2-corrected MRS to asses hepatic fat and iron correlated better than the bedside methods RTS, ultrasound, CAP, and serum ferritin with the histologically confirmed hepatic steatosis and iron overload.

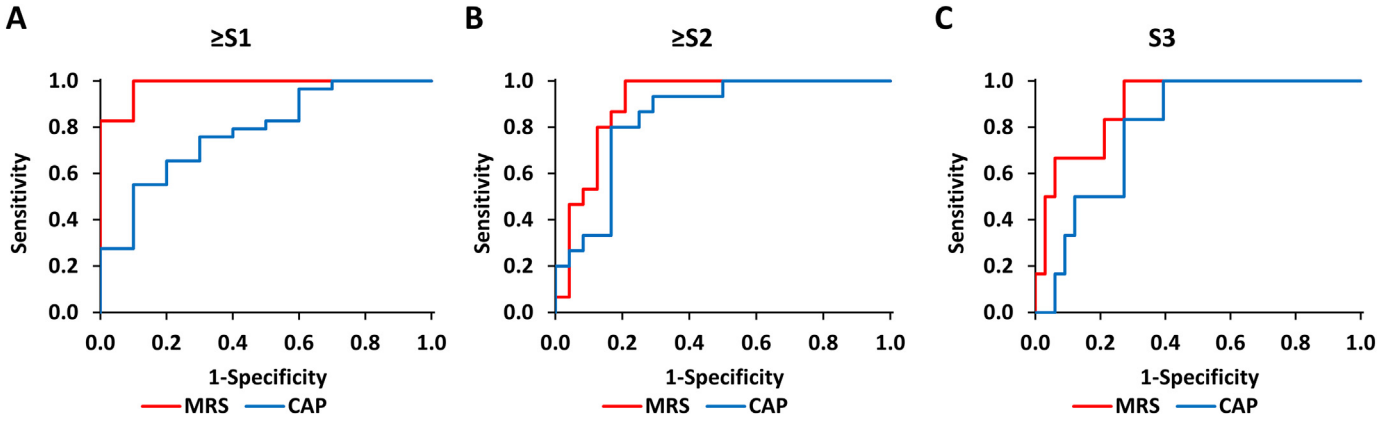

Fig. 6. Receiver operating curves (ROC) for fat measurements with MRS and CAP for histological steatosis grades ≥S1, ≥S2 and S3. MRS showed better performance than CAP for all steatosis grades: AUROCs for MRS were 0.98 (95%CI 0.95–1.00) for ≥S1, 0.91 (95%CI 0.81–1.00) for ≥S2 and 0.90 (95%CI 0.79–1.00) for S3. AUROCs for CAP were 0.79 (95%CI 0.62–0.95) for ≥S1, 0.84 (95%CI 0.72–0.97) for ≥S2 and 0.80 (95%CI 0.65–0.94) for S3. Curves were only significantly different for steatosis grade S1 ($p = 0.023$, Delong-test). S1/2/3: histologically confirmed hepatic steatosis; CAP: controlled attenuation parameter.

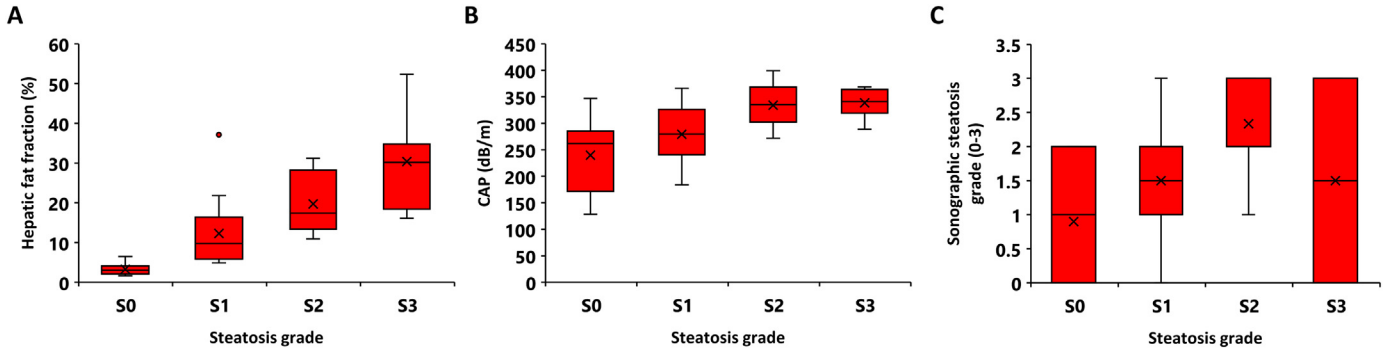

Fig. 7. Box plot analysis for fat measurements. Boxes represent the middle 50% of the data, whiskers reach to maximal 1.5 times the interquartile range above or below the box, other data points are shown as outliers. The horizontal line indicates the median, the x indicates the mean. For MRS and RTS, a linear relation to the AAS grades can be seen while no difference for moderate and severe iron overload can be seen for ferritin. MRS shows a linear increase with steatosis grade while CAP seems to saturate at S2. The steatosis grading with ultrasound shows incorrect grading for S2 and S3. Steatosis grade: histologically confirmed hepatic steatosis; CAP: controlled attenuation parameter; US: ultrasound.

Grant support

This work was supported by grants from the Deutsche Forschungsgemeinschaft (Contract Grant No. DFG ZI 1295/2-1) and the Dietmar Hopp Stiftung.

Declaration of competing interest

We declare the following potential conflict of interest: The German Cancer Research Center has a cooperation agreement with Siemens AG, Germany, which includes a specific collaboration agreement with the aim of a joint collaboration in the field of CT and MR imaging.

References

[1] Czaja AJ. Review article: iron disturbances in chronic liver diseases other than haemochromatosis - pathogenic, prognostic, and therapeutic implications. Aliment Pharmacol Ther 2019;49(6):681–701.
[2] Bacon BR, Adams PC, Kowdley KV, Powell LW, Tavill AS. American Association for the Study of Liver D. Diagnosis and management of hemochromatosis: 2011 practice guideline by the American Association for the Study of Liver Diseases. Hepatology 2011;54(1):328–43.
[3] Vag T, Kentouche K, Krumbein I, et al. Noninvasive measurement of liver iron concentration at MRI in children with acute leukemia: initial results. Pediatr Radiol 2011;41(8):980–4.
[4] Mueller S, Rausch V. The role of iron in alcohol-mediated hepatocarcinogenesis. Adv Exp Med Biol 2015;815:89–112.
[5] European Association For The Study Of The L. EASL clinical practice guidelines for HFE hemochromatosis. J Hepatol 2010;53(1):3–22.
[6] Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41(6):1313–21.
[7] Adams P, Barton JC, McLaren GD, et al. Screening for iron overload: lessons from the hemochromatosis and iron overload screening (HEIRS) study. Can J Gastroenterol 2009;23(11):769–72.
[8] Hong CW, Marsh A, Wolfson T, et al. Reader agreement and accuracy of ultrasound features for hepatic steatosis. Abdom Radiol (NY) 2019;44(1):54–64.
[9] Sasso M, Beaugrand M, de Ledinghen V, et al. Controlled attenuation parameter (CAP): a novel VCTE guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: preliminary study and validation in a cohort of patients with chronic liver disease from various causes. Ultrasound Med Biol 2010;36(11):1825–35.
[10] Shi KQ, Tang JZ, Zhu XL, et al. Controlled attenuation parameter for the detection of steatosis severity in chronic liver disease: a meta-analysis of diagnostic accuracy. J Gastroenterol Hepatol 2014;29(6):1149–58.
[11] Fung EB, Fischer R, Pakbaz Z, et al. The new SQUID biosusceptometer at Oakland: first year of experience. Neurol Clin Neurophysiol 2004;2004:5.
[12] Nielsen P, Engelhardt R, Dullmann J, Fischer R. Non-invasive liver iron quantification by SQUID-biosusceptometry and serum ferritin iron as new diagnostic parameters in hereditary hemochromatosis. Blood Cells Mol Dis 2002;29(3):451–8.
[13] Fischer R, Harmatz PR. Non-invasive assessment of tissue iron overload. Hematology Am Soc Hematol Educ Program 2009:215–21.
[14] Brittenham GM, Badman DG, National Institute of D, Digestive, Kidney Diseases W. Noninvasive measurement of iron: report of an NIDDK workshop. Blood 2003;101(1):15–9.
[15] Avrin WF, Kumar S. Noninvasive liver-iron measurements with a room-temperature susceptometer. Physiol Meas 2007;28(4):349–61.
[16] Mueller J, Raisi H, Rausch V, et al. Sensitive and non-invasive assessment of hepatocellular iron using a novel room-temperature susceptometer. J Hepatol 2017;67(3):535–42.
[17] Hajek M, Dezortova M, Wagnerova D, et al. MR spectroscopy as a tool for in vivo determination of steatosis in liver transplant recipients. Magn Reson Mater Phy 2011;24(5):297–304.
[18] Hernando D, Levin YS, Sirlin CB, Reeder SB. Quantification of liver iron with MRI: state of the art and remaining challenges. J Magn Reson Imaging 2014;40(5):1003–21.
[19] Horvat N, Monti S, Oliveira BC, Rocha CCT, Giancipoli RG, Mannelli L. State of the art in magnetic resonance imaging of hepatocellular carcinoma. Radiol Oncol 2018;52(4):353–64.
[20] Joe E, Kim SH, Lee KB, et al. Feasibility and accuracy of dual-source dual-energy CT for noninvasive determination of hepatic iron accumulation. Radiology 2012;262(1):126–35.
[21] Sharma P, Altbach M, Galons JP, Kalb B, Martin DR. Measurement of liver fat fraction and iron with MRI and MR spectroscopy techniques. Diagn Interv Radiol 2014;20(1):17–26.
[22] Buschle LR, Kurz FT, Kampf T, Triphan SMF, Schlemmer HP, Ziener CH. Diffusion-mediated dephasing in the dipole field around a single spherical magnetic object. Magn Reson Imaging 2015;33(9):1126–45.
[23] Ziener CH, Bauer WR, Jakob PM. Transverse relaxation of cells labeled with magnetic nanoparticles. Magn Reson Med 2005;54(3):702–6.
[24] Kurz FT, Kampf T, Heiland S, Bendszus M, Schlemmer HP, Ziener CH. Theoretical model of the single spin-echo relaxation time for spherical magnetic perturbers. Magn Reson Med 2014;71(5):1888–95.
[25] Ziener CH, Kurz FT, Kampf T. Free induction decay caused by a dipole field. Phys Rev E Stat Nonlin Soft Matter Phys 2015;91(3):032707.
[26] Ziener CH, Kampf T, Herold V, Jakob PM, Bauer WR, Nadler W. Frequency autocorrelation function of stochastically fluctuating fields caused by specific magnetic field inhomogeneities. J Chem Phys 2008;129(1):014507.
[27] Mulkern R, Haker S, Mamata H, et al. Lung parenchymal signal intensity in MRI: a technical review with educational aspirations regarding reversible versus irreversible transverse relaxation effects in common pulse sequences. Concepts Magn Reson Part A Bridg Educ Res 2014;43A(2):29–53.
[28] Pineda N, Sharma P, Xu Q, Hu X, Vos M, Martin DR. Measurement of hepatic lipid: high-speed T2-corrected multiecho acquisition at 1H MR spectroscopy–a rapid and accurate technique. Radiology 2009;252(2):568–76.
[29] de Ledinghen V, Hiriart JB, Vergniol J, Merrouche W, Bedossa P, Paradis V. Controlled attenuation parameter (CAP) with the XL probe of the Fibroscan((R)): a comparative study with the M probe and liver biopsy. Dig Dis Sci 2017;62(9):2569–77.
[30] Durango E, Dietrich C, Seitz HK, et al. Direct comparison of the FibroScan XL and M probes for assessment of liver fibrosis in obese and nonobese patients. Hepat Med 2013;5:43–52.
[31] Hong R, Schubert WK. Menghini needle biopsy of the liver. Am J Dis Child 1960;100:42–6.
[32] Thaler H. Experiences with the Menghini method of liver biopsy. Wien Klin Wochenschr 1958;70(34):622–5.
[33] Bonkovsky HL, Slaker DP, Bills EB, Wolf DC. Usefulness and limitations of laboratory and hepatic imaging studies in iron-storage disease. Gastroenterology 1990;99(4):1079–91.
[34] Bashir MR, Wolfson T, Gamst AC, et al. Hepatic R2* is more strongly associated with proton density fat fraction than histologic liver iron scores in patients with nonalcoholic fatty liver disease. J Magn Reson Imaging 2019;49(5):1456–66. https://doi.org/10.1002/jmri.26601.
[35] Karlsson M, Ekstedt M, Dahlstrom N, et al. Liver R2* is affected by both iron and fat: a dual biopsy-validated study of chronic liver disease. J Magn Reson Imaging 2019;50(1):325–33. https://doi.org/10.1002/jmri.26601.
[36] Mamidipalli A, Hamilton G, Manning P, et al. Cross-sectional correlation between

hepatic R2* and proton density fat fraction (PDFF) in children with hepatic steatosis. J Magn Reson Imaging 2018;47(2):418–24.
[37] Williamson RM, Perry E, Glancy S, et al. The use of ultrasound to diagnose hepatic steatosis in type 2 diabetes: intra- and interobserver variability and comparison with magnetic resonance spectroscopy. Clin Radiol 2011;66(5):434–9.
[38] Hutton C, Gyngell ML, Milanesi M, Bagur A, Brady M. Validation of a standardized MRI method for liver fat and T2* quantification. PLoS One 2018;13(9):e0204175.
[39] Artz NS, Haufe WM, Hooker CA, et al. Reproducibility of MR-based liver fat quantification across field strength: same-day comparison between 1.5T and 3T in obese subjects. J Magn Reson Imaging 2015;42(3):811–7.
[40] Idilman IS, Akata D, Ozmen MN, Karcaaltincaba M. Different forms of iron accumulation in the liver on MRI. Diagn Interv Radiol 2016;22(1):22–8.
[41] Deistung A, Schweser F, Reichenbach JR. Overview of quantitative susceptibility mapping. NMR Biomed 2017;30(4).
[42] Liu S, Buch S, Chen Y, et al. Susceptibility-weighted imaging: current status and future directions. NMR Biomed 2017;30(4). https://doi.org/10.1002/nbm.3552.
[43] Sharma SD, Fischer R, Schoennagel BP, et al. MRI-based quantitative susceptibility mapping (QSM) and R2* mapping of liver iron overload: comparison with SQUID-based biomagnetic liver susceptometry. Magn Reson Med 2017;78(1):264–70.
[44] Sharma SD, Hernando D, Horng DE, Reeder SB. Quantitative susceptibility mapping in the abdomen as an imaging biomarker of hepatic iron overload. Magn Reson Med 2015;74(3):673–83.
[45] Simchick G, Liu Z, Nagy T, Xiong M, Zhao Q. Assessment of MR-based R2* and quantitative susceptibility mapping for the quantification of liver iron concentration in a mouse model at 7T. Magn Reson Med 2018;80(5):2081–93.
[46] Hernando D, Kramer JH, Reeder SB. Multipeak fat-corrected complex R2* relaxometry: theory, optimization, and clinical validation. Magn Reson Med 2013;70(5):1319–31.
[47] Hernando D, Kuhn JP, Mensel B, et al. R2* estimation using “in-phase” echoes in the presence of fat: the effects of complex spectrum of fat. J Magn Reson Imaging 2013;37(3):717–26.