



# A dedicated phantom for exploring the interplay of fat and paramagnetic substances in quantitative susceptibility mapping

Simon Graf<sup>1,2</sup> · Josefine Trapp<sup>1</sup> · Maik Rothe<sup>1,2</sup> · Alexander Gussew<sup>1,2</sup> · Walter A. Wohlgemuth<sup>1,2</sup> · Andreas Deistung<sup>1,2</sup>

Received: 24 January 2025 / Revised: 14 April 2025 / Accepted: 25 April 2025 / Published online: 2 June 2025  
© The Author(s) 2025, corrected publication 2025

## Abstract

**Objective** Accurate quantitative tissue characterization in organs with considerable fat content, like the liver, requires thorough understanding of fat's influence on the MR signal. To continue the investigations into the use of quantitative susceptibility mapping (QSM) in abdominal regions, we present a dedicated phantom that replicates liver-like conditions in terms of effective transverse relaxation rates ( $R_2^*$ ) and proton density fat fractions.

**Materials and methods** The spherical agar phantom consists of nine smaller spheres (diameter: 3 cm) doped with a paramagnetic substance (iron nanoparticles or manganese chloride) and fat (peanut oil), embedded in a large agar sphere (diameter: 14 cm), ensuring no barriers exist between the enclosed spheres and their surrounding medium. Concentrations were selected to represent both healthy and pathologic conditions. 3T MRI measurements for relaxometry, fat–water imaging, and QSM were conducted with the head coil and for  $^1\text{H}$ -spectroscopy with the knee coil at three time points, including a scan–rescan assessment and a follow-up measurement 14 months later.

**Results** The phantoms' relaxation and magnetic properties are in similar range as reported for liver tissue. Substantial alterations in local field and susceptibility maps were observed in regions with elevated fat and iron content, where fat correction of the local field via chemical shift-encoded reconstruction effectively reduced streaking artifacts in susceptibility maps and substantially increased susceptibility values. Linear regression analysis revealed a consistent linear relationship between  $R_2^*$  and magnetic susceptibility, as well as iron concentration and magnetic susceptibility. The relaxation, fat, and susceptibility measurements remained stable across scan–rescan assessment and long-term follow-up.

**Discussion** We developed a versatile phantom to study fat–iron interactions in abdominal imaging, facilitating the optimization and comparison of susceptibility processing methods in future research.

**Keywords** Quantitative susceptibility mapping · Relaxometry · Phantom · Water–fat separation · Iron

## Introduction

The human body is composed of numerous tissue types, each with distinct characteristics and functions, and relies on multiple essential elements. Iron, for instance, is crucial for oxygen transport and DNA synthesis [1] and is simultaneously

under strict regulation, since excess iron can lead to various iron-related disorders such as iron overload [2] or neurodegenerative diseases [3, 4]. The total body iron level is typically assessed with the plasma ferritin marker; however, it is affected by conditions such as acute infections or chronic inflammation, altering the relationship between plasma ferritin and total iron levels [5, 6]. In the human liver, iron ions ( $\text{Fe}^{3+}$ ) are stored primarily in the cores of ferritin shells and hemosiderin [7, 8] and its iron content is directly linked to the total amount of iron in the body [9]. Therefore, measuring liver iron content (LIC) is commonly seen as the most accurate way to evaluate the total body iron content. LIC has historically been assessed using liver biopsy and destructive biochemical analysis [10]. However, the clinical utility of liver biopsy for LIC quantification is limited due to its

✉ Simon Graf  
simon.graf@uk-halle.de

<sup>1</sup> University Clinic and Outpatient Clinic for Radiology, University Hospital Halle (Saale), Ernst-Grube-Str. 40, 06120 Halle (Saale), Germany

<sup>2</sup> Halle MR Imaging Core Facility, Medical Faculty of Martin-Luther-University Halle-Wittenberg, Ernst-Grube-Str. 40, 06120 Halle (Saale), Germany

invasive nature and substantial sampling variability [11–13], making it unsuitable for repeated measurements required for treatment monitoring. As a result, non-invasive methods, such as magnetic resonance imaging, have largely replaced biopsy for LIC assessment. Determination of the transverse relaxation rate ( $R_2$ ) or effective transverse relaxation rate ( $R_2^*$ ) in conjunction with dedicated calibration curves are the most widely used MRI approaches for quantifying LIC [14, 15]. However, these MRI-based estimations of LIC are prone to inaccuracies arising from breathing motion, and especially the presence of abdominal adipose tissue [13, 16], as well as discrepancies in the LIC versus  $R_2/R_2^*$  calibration curves with respect to the used acquisition and processing parameters [17]. Direct measurement of tissue magnetic susceptibility can offer advantages over MRI relaxometry for LIC measurement. While  $R_2^*$  is influenced by the microscopic distribution of iron, the magnetic susceptibility remains unaffected [18, 19]. Moreover, magnetic susceptibility measurements have the capability to circumvent the necessity for calibration curves, providing a direct correlation with LIC [20, 21].

A way to derive the magnetic susceptibility distribution *in vivo* using MRI is possible by applying quantitative susceptibility mapping (QSM) [22, 23]. The method relies on sophisticated image processing of raw gradient-recalled echo phase images and has been conducted successfully across various regions of the brain to investigate iron deposits [24–26], myelination [27], microbleeds, and venous vasculature [28, 29] and to differentiate between iron and calcium deposits [30, 31]. While QSM has demonstrated its utility and reliability across a wide range of neurological and neurodegenerative applications, its use beyond the central nervous system remains limited, with only initial applications in the heart, liver, prostate, and cartilage [32]. In this context, susceptibility-based characterization of abdominal organs, particularly the liver, is of special interest as it offers an alternative method for measuring LIC. A few studies have already been applied to investigate hepatic iron content [33–35] and hepatic fibrosis [36, 37] using QSM.

In contrast to brain tissue, abdominal tissue is characterized by substantial contributions of fat. The coexistence of fat and water leads to signal intensity oscillations due to the constructive and destructive interference of these two components during multi-echo gradient-echo acquisitions. Furthermore, adipose tissue induces chemical shifts, which introduce non-susceptibility-related contributions to the gradient-echo phase data. This phase data is used in QSM to derive magnetic field maps, which are subsequently employed in the field-to-susceptibility inversion process. The presence of non-susceptibility-related phase contributions can increase the likelihood of artifacts, such as

streaking, in susceptibility maps. To mitigate these effects, abdominal QSM studies frequently employ water–fat separation techniques as a preprocessing step to correct for fat-related contributions [33–35].

In quantitative MRI, a dedicated phantom facilitates the precise evaluation of developed methods [38], which in our case considers the influence of fat and subsequent fat-correction techniques on local field and susceptibility maps to enhance the reliability of LIC quantification. Zhao et al. [39] replicated the presence of fat, iron, and fibrosis and performed comprehensive relaxometry and proton density fat fraction (PDFF) evaluations. Specifically targeting QSM, Li et al. [19] built a phantom that embedded balloons, filled with gadolinium (as paramagnetic substance), fat, collagen, and water, in a water-filled container to perform  $R_2^*$  and susceptibility mapping. Kim et al. [40] proposed a spherical gelatin phantom, in which gelatinous inclusions of iron oxide-based contrast agents as paramagnetic source were integrated, to evaluate QSM processing. The main limitation of the phantoms proposed in [19] [39] for QSM processing is their use of plastic vials as vessel for the different solutions embedded in a water-filled container. While suitable for relaxometry and the advantage of a relatively straightforward assembly, the plastic interfaces of the vials can induce Gibbs ringing and magnitude signal loss, which may impede consistent phase unwrapping (e.g., when using region-growing approaches) or amplify non-local artifacts if not properly addressed. The phantom presented in [40] demonstrates the feasibility of designing a spherical phantom containing inclusions without an additional interface to the surrounding medium; however, the  $R_2^*$  values are below  $23\text{ s}^{-1}$  at 3T, mimicking LIC only within the healthy range [41]. In addition, the authors did not include a fat source.

Inspired by the works of Kim et al. [40] and Zhao et al. [39], we developed a novel spherical phantom simulating both healthy and pathological liver tissue, specifically targeting abdominal QSM processing, which mimics both healthy and pathological liver tissue. We removed plastic interfaces in the phantom and embedded nine different solutions of agar base doped with fat and iron nanoparticles within an enclosed agar sphere. The phantom was comprehensively characterized using relaxometry ( $R_1$ ,  $R_2$ ,  $R_2^*$ ), water–fat imaging,  $^1\text{H}$ -MR spectroscopy, and QSM (single- and multi-orientation), including a 14-month follow-up measurement and a scan–rescan assessment. This work focuses on the design, construction, and reproducibility characterization (short-term reproducibility and long-term stability) of the phantom, to assess its feasibility for relaxometry and QSM processing. Additionally, we explored the use of manganese chloride as a paramagnetic substance as a potential cost-effective alternative to iron nanoparticles.

## Theory

The complex-valued MR signal obtained with gradient echo measurements in body parts containing water and lipid components is well established. Here, we briefly introduce the relationship between gradient-echo phase images, magnetic susceptibility, and the chemical shift in adipose tissue.

The complex-valued MR signal within the voxel obtained with gradient echo measurements in body parts containing water and lipid components is given by [42]:

$$S(\rho_W, \rho_F, f_B, R_2^*) = \left( \rho_W + \rho_F \sum_{p=1}^P \alpha_p e^{i2\pi f_{F,p} t} \right) e^{i(\phi_0 + 2\pi f_B t)} e^{-R_2^* t}, \quad (1)$$

where  $\rho_W$  and  $\rho_F$  are the amplitudes of the water and lipid signal, respectively,  $\phi_0$  is the initial phase introduced by the radiofrequency response of the sample,  $f_B$  is the frequency shift due to field inhomogeneities, e.g., due to susceptibility effects or imperfect shimming,  $R_2^*$  is the effective transverse relaxation rate of the voxel,  $p$  is the number of lipid peaks,  $f_{F,p}$  is the lipid peak frequencies relative to the water peak,  $\alpha_p$  is the relative amplitude of the lipid peaks ( $\sum_{p=1}^P \alpha_p = 1$ ), and  $t$  is the echo time. The lipid peak frequencies  $f_{F,p}$  and the corresponding amplitude  $\alpha_p$  may be determined using  $^1\text{H}$ -MR spectroscopy to reduce the number of unknowns when solving for  $\rho_W$ ,  $\rho_F$ ,  $f_B$ , and  $R_2^*$ . In case only the magnitude is available, the phase-dependent term in Eq. 1 can be discarded while accounting for the lipid-induced contribution based on the previously determined amplitudes and frequencies of the lipid peaks. When interested in magnetic susceptibility, magnitude and phase data have to be considered in Eq. 1.

The gradient echo phase  $\phi(\vec{r}, t)$  is given by [22, 43]:

$$\phi(\vec{r}, t) = \phi_0(\vec{r}) + 2\pi\Delta f(\vec{r})t \approx \phi_0(\vec{r}) + \gamma \left( \left( \frac{2\pi}{\gamma} f_B(\vec{r}) \right) (1 - \sigma(\vec{r})) - B_0 \sigma(\vec{r}) \right) t \quad (2)$$

with  $\phi_0(\vec{r})$  being the initial phase [44] (see Eq. 1),  $\Delta f(\vec{r})$  the measured frequency variation,  $\gamma$  the gyromagnetic ratio,  $f_B(\vec{r}) = \frac{\gamma}{2\pi} (B_{inh}(\vec{r}) + B_\chi(\vec{r}))$  (see Eq. 1),  $\sigma(\vec{r})$  the chemical shift [45], and  $\vec{r}$  the spatial location vector. With respect to Eq. 1,  $\sigma$  denotes the spectral peak area-weighted average chemical shift that can be defined as  $\sigma = \bar{f}_F$  in the voxel. Chemical shift encoded reconstruction incorporating fat correction [46–48] effectively eliminates the spatially varying field contributions arising from the chemical shift,  $\sigma(\vec{r})$ . The magnetic susceptibility  $\chi(\vec{r})$  can be extracted from the phase data via deconvolution with the unit dipole response  $d_z(\vec{r})$ , in case of appropriate shimming, and if magnetic field contributions arising from magnetic sources outside a defined volume of interest (VOI, e.g., the area of the brain

if interest in brain imaging)  $B_{\chi,\text{ext}}(\vec{r})$  are removed properly, e.g., with sophisticated harmonic artifact removal for phase data (SHARP) [49], only the sources within the VOI  $B_{\chi,\text{int}}(\vec{r})$  contribute:

$$\tilde{B}(\vec{r}) \approx B_0 (\chi(\vec{r}) \otimes d_z(\vec{r})). \quad (3)$$

Since the phase is influenced by the chemical shift due to lipid sources (Eqs. 1 and 2), accurate fat correction is essential to determine the susceptibility-induced magnetic field map [32]. A detailed derivation of Eqs. 2 and 3 is provided in the supplementary.

To this day, the COSMOS approach (calculation of susceptibility through multiple orientation sampling) is the gold standard method for QSM measurements [50]. While its feasibility in a clinical setting is already limited for brain MRI, it is barely possible for abdominal measurements. For phantom measurements, however, it provides a viable opportunity to measure and compute ground truth susceptibility data.

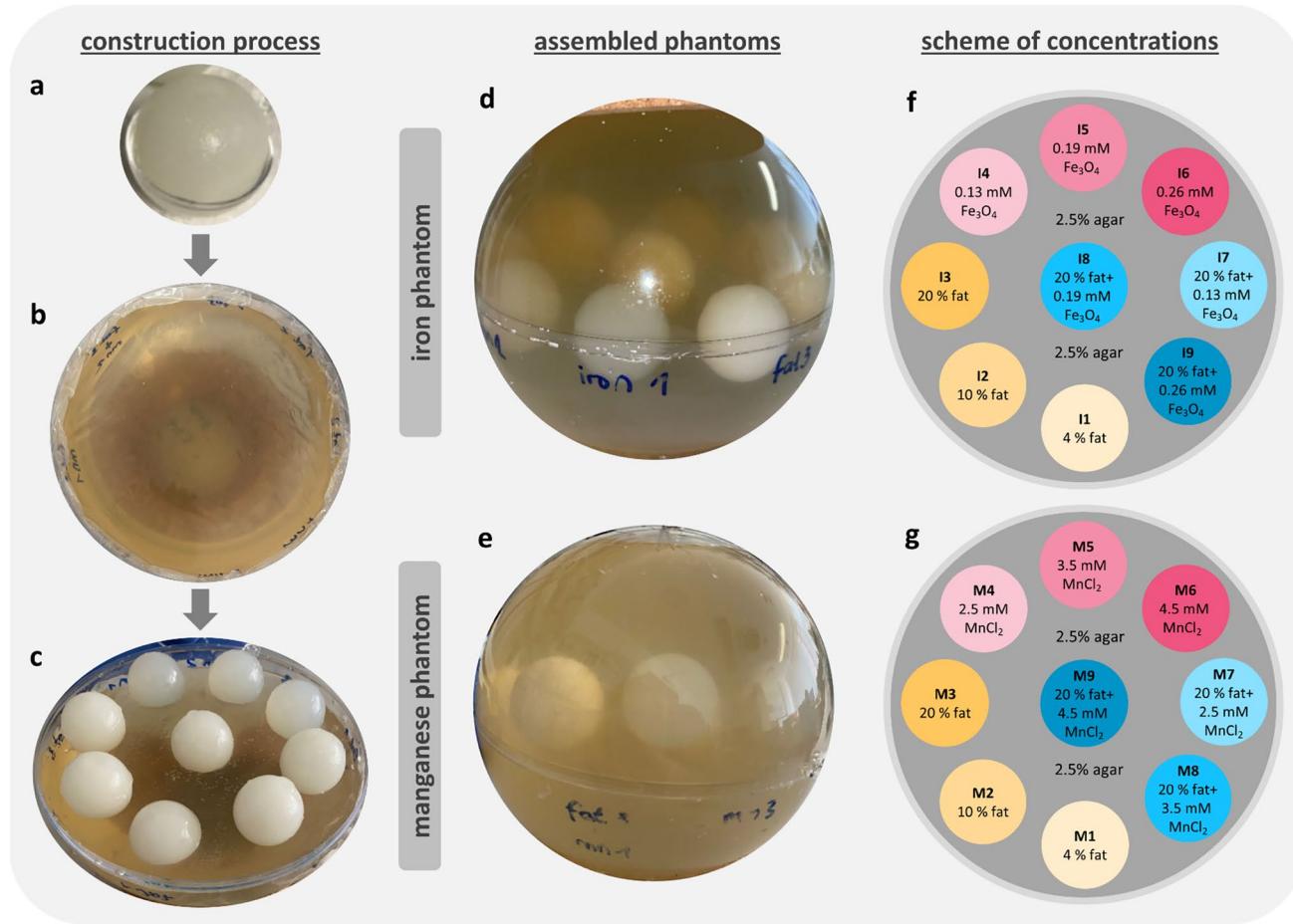
## Material and methods

To investigate the impact of fat on the susceptibility map, a comprehensive multi-purpose phantom was constructed. Typical proton density fat fractions (PDFF) values range from under 5% for healthy tissue to over 30% in patients with steatosis grade 3 [51–53]. For hepatic iron overload, healthy  $R_2^*$  values are below  $126 \text{ s}^{-1}$ , increasing to approximately  $240 \text{ s}^{-1}$ ,  $1160 \text{ s}^{-1}$ , or even higher for mild, moderate, and severe overload, respectively [41, 54]. The phantom design was based on the following prerequisites: (i) inclusion of fat and iron sources simulating liver tissue, (ii)  $R_2^*$  rates and PDFF within healthy to mildly pathologic ranges (PDFF up

to 20%,  $R_2^*$  up to  $250 \text{ s}^{-1}$ ), (iii) spherical design allowing high reproducibility, and (iv) absence of a structural barrier between the different compositions.

## Spherical agar phantom

The generated spherical phantoms are illustrated in Fig. 1 and consist of nine small agar spheres doped with varying amounts of paramagnetic substances and fat. The base solution used for phantom construction was a 2.5 wt. % water–agar solution (agar–agar powder, Carl Roth GmbH + Co. KG) for the large sphere (diameter of 14 cm), and 2 wt. % water–agar solution for the small spheres



**Fig. 1** Phantom construction process and schematic representation of the phantoms. The small spheres are prepared in small polystyrene shells (a). The lower hemisphere is coated with cling foil during the cooling process (b). After removing the small spheres from their shells, they are placed on the lower phantom hemisphere (c), the outer shell is closed and filled with agar to the rim for the iron (d) and

manganese phantoms (e). Both spherical phantoms differ only in the used paramagnetic substance, iron nanoparticles, (f) and manganese chloride (g). The spheres are numbered from 1 to 9, with each number corresponding to an increasing concentration of fat, paramagnetic substance, or both

(diameter of 3 cm). The phantom size of 14 cm approximates the average liver diameter (male: 14 cm, female: 14.5 cm) [55]. Sodium azide, 0.1 wt. % (Sigma-Aldrich, Merck KGaA), was used to preserve the phantom and prevent contamination (e.g., bacteria or mold). A coating agent (Imprägnol, Brauns-Heitmann GmbH & Co. KG) was used to prevent diffusion of water between the different phantom areas, maintaining the adjusted concentrations. To increase the melting point of the small spheres, prevent melting and maintain the different volume integrities, 1% formaldehyde ( $\text{CH}_2\text{O}$ ) (Carl Roth GmbH + Co. KG) was added to the small spheres. The contrast agent Gadovist (1 mmol  $\text{mL}^{-1}$ , Bayer Vital GmbH) was used to modulate the longitudinal relaxation rate in the large and small spheres. Iron nanoparticles ( $5\text{mg mL}^{-1}$  iron oxide (II, III), magnetic nanoparticle solution, Sigma-Aldrich, Merck KGaA) or  $\text{MnCl}_2$  (manganese(II) chloride monohydrate, Carl Roth

GmbH + Co. KG) were used as paramagnetic substances to elevate  $R_2^*$ . The spherical phantom with iron nanoparticles is herein referred to as the iron phantom, while the one with  $\text{MnCl}_2$  is referred to as the manganese phantom. As fat source, peanut oil (Kunella Feinkost GmbH, Cottbus, Germany) was chosen, since the proton NMR spectrum is similar to that of human adipose tissue [56]. The adjusted concentrations for the different phantom parts are given in Table 1.

## Phantom assembly

For the construction of the phantom, nine vials (S1–S9) were used to prepare the individual solutions. Different amounts of paramagnetic substances (iron nanoparticles or  $\text{MnCl}_2$ ) were added by pipetting to the vials (S4–S9). A manganese

**Table 1** Composition of the spherical iron and manganese phantom

Volumes	Both phantoms			Iron phantom			Manganese phantom		
	Gadovist [mM]	NaN <sub>3</sub> [wt%]	CH <sub>2</sub> O [vol%]	Peanut oil [vol %]			Peanut oil [vol %]	Agar [wt %]	Fe <sub>3</sub> O <sub>4</sub> [mM]
				0	10	20			
Small sphere	Lipid only	S1 0.25	0.1	1	4	2	4	2	
		S2 0.25			10	10	10		
		S3 0.25			20	20	20		
	Agent only	S4 0.25			0	0	0	2.5	
		S5 0.25				0.13	0.13	3.5	
		S6 0.25				0.19	0.19	4.5	
		S7 0.25				0.26	0.26	4.5	
		S8 0.25				0.26	0.26	4.5	
		S9 0.25				0	0	0	
Big sphere			0.1	0	0	2.5	0	2.5	0

solution was prepared with water, whereas the iron nanoparticles were already delivered in a solution. Formaldehyde was added to all vials. The water–agar base solution was prepared with distilled water, Gadovist, and sodium azide. The components were heated to approximately 90 °C with constant stirring. The agar powder was added to the solution and stirred until fully dissolved. Peanut oil was added to a beaker and stirred until emulsification, indicated by the solution turning white. Four different base solutions (2wt. % agar with 0, 4, 10, and 20% fat) were necessary for the small spheres of the phantom. The base solution was added to the respective vials and stirred for 15 s on a vortex to ensure proper mixing. After that, the solutions were poured into the small spheres. Once cooled, the small spheres were removed from the polystyrene shells, sprayed with the coating agent, and left to dry for approximately 24 h (Fig. 1a). On the next day, the large sphere was constructed in two steps: first, the lower hemisphere (Fig. 1b, c) and then the final phantom assembly (Fig. 1d, e). A 2.5 wt. % base solution was prepared and half of the shell of the lower hemisphere was filled. A cling foil was used to cover the drying agar solution to prevent film formation (Fig. 1b). Preliminary studies showed the formation of artifacts at the interhemispheric interface. Thus, the lower shell was only filled half, to move the interhemispheric interface under the small spheres. While the base solution was left to cool for approximately 30 min in the lower hemisphere shell, the 2.5 wt. % base solution for the final phantom assembly was prepared. To minimize the formation of an interface, the base solution of the lower hemisphere should not be fully gelled. The cling film was removed and the small spheres were placed onto the agar solution (Fig. 1c). The spherical shell was closed with the upper lid and the sphere filled to the rim with base solution through a hole in the shell of the upper hemisphere (Fig. 1d, e). The phantom was left to cool at least for 24 h before conducting the MR measurements. The phantoms were stored in a container to minimize air contact and the hole in the shell was closed. The construction process is identical for the two versions of the built phantom and only differs in the use of the paramagnetic substance.

## Data acquisition

Measurements were conducted on a 3T Siemens Magnetom Vida Scanner (Siemens Healthcare GmbH, Erlangen, Germany) with a 64-channel phased-array head coil. A 2D turbo spin echo inversion recovery (IR-TSE) sequence was used to collect data for  $R_1$  quantification and 2D multi-echo spin echo (ME-SE) imaging was employed to determine  $R_2$ . A 3D multi-echo gradient-echo sequence (GRE) based on the volumetric interpolated breath-hold examination (VIBE) with six echoes was applied for water–fat separation,  $R_2^*$  mapping,

and QSM. Further data for QSM relying on COSMOS were acquired using the GRE-VIBE sequence in four additional different phantom positions with respect to the main magnetic field ( $\vec{B}_0$ , tilts by approximately  $\pm 60^\circ$  in the  $\vec{y}$ - $\vec{z}$  plane and  $\vec{x}$ - $\vec{z}$ , if  $\vec{z}$  is the direction of the magnetic field). MR spectroscopic measurements were conducted with the stimulated echo acquisition mode sequence (STEAM) using a transmit/receive knee coil with 18 receive channels to determine the exact lipid spectrum in the phantom to parameterize  $f_{F,p}$  and  $\alpha_p$  in Eq. 1. The detailed sequence parameters are listed in Table 2. To assess repeatability (scan–rescan) and long-term stability, the phantom was additionally measured twice with an identical scanning protocol (except MR spectroscopy) 14 months after the initial phantom construction. For the scan–rescan assessment, the phantom was completely removed from the scanner and positioned a second time.

## Data processing

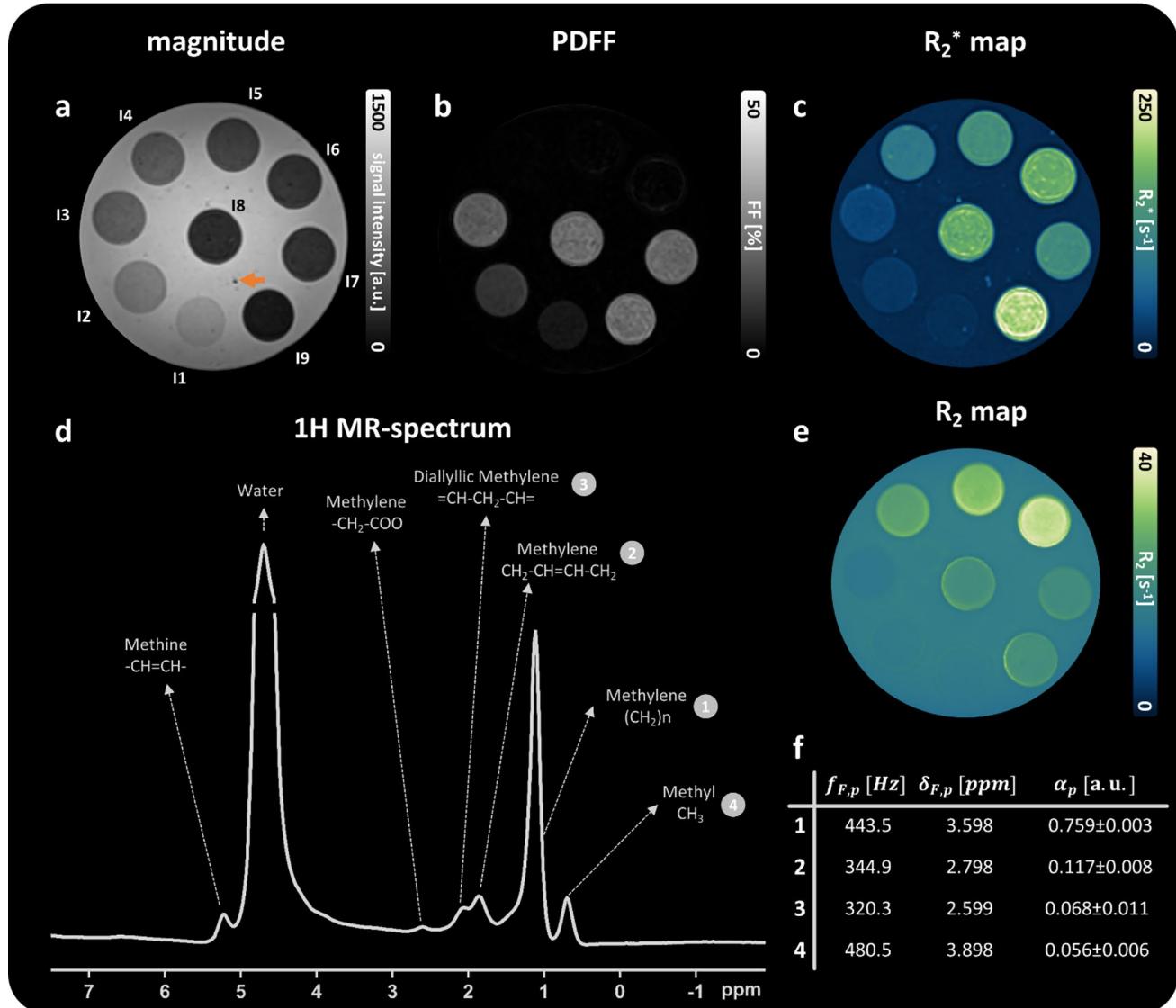
$R_1$  was determined from the IR-TSE data sets with varying inversion times using ordinary least squares fitting [57] in a VOI-based manner. The monoexponential signal decay of the multi-echo spin echo data was fitted voxel-wise in an ordinary least squares manner to compute  $R_2$ . Based on  $^1\text{H}$ -MR spectroscopy measurements, the intensities and resonance frequency of particular lipid compartments defining the lipid spectrum in Eq. 1 were obtained (Fig. 2f). All spectra were preprocessed and quantified using the jMRUI package [58, 59] (<http://www.jmrui.eu>). The relative amplitudes of the lipid spectrum  $\alpha_p$  were calculated by dividing the area under the respective peak by the sum of the area under all lipid peaks and additionally corrected for fat compartment-specific  $T_2$  attenuations based on a monoexponential fit of

signal decays in the STEAM TE series. The lipid peak frequencies  $f_{F,p}$  were determined by multiplying the chemical shift difference of the lipid peaks and the water peak to the resonance frequency of the MR scanner (123.195 MHz). The parameters were computed for spectra acquired in the iron-free spheres with the two highest fat concentrations (inclusion I2 and I3 in Fig. 1f) and subsequently averaged. The complex-valued GRE-VIBE signal decay, acquired at intervals of 1.71 ms, beginning at 1.17 ms, was processed using water–fat separation to derive fat-corrected field maps for QSM [46], while  $R_2^*$  and PDFF were determined by fitting Eq. 1 to the magnitude signal decay only using a nonlinear least squares model, incorporating the phantom-specific lipid spectrum while disregarding  $\phi_0$  and  $f_B$  (see also section Theory, [42]).

To correct for fat-induced errors in the phase-derived magnetic field map, Eq. 1 was efficiently solved using the iterative graph-cut algorithm proposed by Hernando et al. [46], and applied to the complex-valued GRE-VIBE data. This technique jointly estimates the field map and water–fat images by discretizing the field map and applying an iterative binary graph-cut approach to optimize water–fat separation while avoiding local minima. Here, spatial smoothness constraints enhance robustness against field inhomogeneities while preserving local accuracy. The resulting fat-corrected field maps were then used for further susceptibility processing. For comparison, fat-uncorrected, wrapped multi-echo phase data were unwrapped [60], divided by  $TE_i * \gamma$ , and combined across echo times to generate fat-uncorrected field maps. Both field maps were further processed with sophisticated harmonic artifact removal for phase data [49] (SHARP) using ten different spherical kernels (1–10 voxels, regularized at a high-pass cutoff frequency of  $0.02 \text{ mm}^{-1}$ ) to compute the local magnetic field perturbation. The

**Table 2** Sequence parameters for the phantom measurements

Parameters	$R_1$	$R_2$	$R_2^*$ and PDFF	Spectroscopy
Sequence	IR-TSE	ME-SE	GRE-VIBE	STEAM
TR [ms]	10,000	3000	11.5	3000
TI [ms]	50, 150, 300, 600, 900, 1200, 1500	–	–	–
No. of echoes	1	14	6	7
TE [ms]	57	9.4	1.17	20, 30, 40, 50, 60, 100, 200
$\Delta TE$ [ms]		9.4	1.71	–
TM [ms]	–	–	–	10
Flip angle [°]	128	180	9	90
Voxel size [mm <sup>3</sup> ]	$1.0 \times 1.0 \times 2.0$	$0.8 \times 0.8 \times 4.0$	$0.8 \times 0.8 \times 0.8$	$14 \times 14 \times 14$
Field of view [mm <sup>3</sup> ]	$192 \times 192 \times 88$	$250 \times 250 \times 60$	$190 \times 194 \times 128$	–
Bandwidth [Hz px <sup>-1</sup> ]	137	781	740	acq. BW.: 10 000 Hz
Acquisition time [min:s]	4:52 per TI	12:05	2:20	1:12 per spectrum
No. of spectra	–	–	–	6 per VOI



**Fig. 2** Comprehensive characterization of the iron phantom. One representative slice of the GRE-VIBE magnitude image at  $TE=8.01$  ms (a), the fat fraction map (b), the  $R_2^*$  map (c), and  $R_2$  map (e) are presented. The proton MR spectrum (d, 6 averages) for inclusion I3 and the resulting fat characterizing parameters (lipid peak frequencies in Hz [ $f_{F,p}$ ] and ppm [ $\delta_{F,p}$ ]), relative amplitudes [ $\alpha_p$ ]) are shown in (d) and (f), respectively. The numbers adjacent to the individual lipid peaks in d refer to the numbers in f. The orange arrow in a indicates an air inclusion. The small spherical inclusions are exemplarily labeled in (a)

cies in Hz [ $f_{F,p}$ ] and ppm [ $\delta_{F,p}$ ], relative amplitudes [ $\alpha_p$ ]) are shown in (d) and (f), respectively. The numbers adjacent to the individual lipid peaks in d refer to the numbers in f. The orange arrow in a indicates an air inclusion. The small spherical inclusions are exemplarily labeled in (a)

field-to-susceptibility inversion of the local field maps was conducted using homogeneity-enabled incremental dipole inversion (HEIDI) [61] for single-orientation inversion. COSMOS measurements were conducted to derive the magnetic susceptibility from multiple phantom orientations. The local field maps (corrected and uncorrected for fat contributions) were also computed for the four other GRE-VIBE acquisitions with varying phantom positions to  $\vec{B}_0$  and then linearly transformed (6 degrees of freedom) using advanced normalization tools (ANTs [62], <https://stnava.github.io/>

ANTs/) to the data set of the first measurement. Simultaneous inversion of the aligned local fields (5 data sets) to magnetic susceptibility was performed to derive the COSMOS maps. We referenced all susceptibility maps to the average susceptibility of 2.5% agar base solution (see Fig. 1f, g) and stated susceptibility values in parts-per-billion (ppb).

Volumes of interests (VOIs) were manually drawn in the nine spherical inclusions and their embedding medium for both the iron and manganese phantom using ITK-SNAP [63] (<http://www.itksnap.org/pmwiki/pmwiki.php>). Means

and standard deviations were computed for the different VOIs. A one-way repeated measures ANOVA was performed separately for  $R_1$ ,  $R_2$ ,  $R_2^*$ , PDFF, and susceptibility values to assess statistical significance in repeatability and long-term stability. The analysis included the values of the nine different spherical inclusions measured across three time points.

## Results

### Phantom characterization

Parametric maps of the iron phantom and their VOI-based analysis are shown in Fig. 2 and Table 3, respectively. The high-resolution GRE magnitude image (Fig. 2a) shows a homogenous background in the large agar sphere with just a few dark speckles (orange arrow) air inclusions. The measured PDFFs agree well with the set fat concentrations during phantom construction (Fig. 2b, Table 3) as also supported by the slope of 0.954 (coefficient of determination [ $R_2$ ]=0.988) of the regression line (Fig. S1). The gradation of PDFFs ranges from  $4.4\% \pm 0.6\%$  in I1 to  $17.7\% \pm 1.4\%$  in I3. Regardless of the iron concentration, the fat fraction remains approximately constant at 20% for I7–I9. The  $R_2^*$  values increase with higher iron and fat concentration ranging from  $24.2 \pm 2.1 \text{ s}^{-1}$  for the lowest fat concentration (I1) to  $211.3 \pm 18.8 \text{ s}^{-1}$  for the highest fat and iron particle concentration (I9). Iron-only spheres exhibit lower  $R_2^*$  as those mixed with 20% peanut oil (Fig. 2c).  $R_2$  increases with higher iron concentration (Fig. 2e) from  $25.2 \pm 0.7 \text{ s}^{-1}$  (I4) to  $33.9 \pm 1.2 \text{ s}^{-1}$  (I6).  $R_2$  remained approximately constant at  $15 \text{ s}^{-1}$  in the fat-only spheres (I1–I3), regardless of the respective fat concentration, and at approximately  $21 \text{ s}^{-1}$  for I7–I9.  $R_1$  varies between  $2.4 \pm 0.1$  and  $2.9 \pm 0.2 \text{ s}^{-1}$  and decreases with increasing concentrations of fat. The proton MR spectrum acquired in sphere I3 (Fig. 2d) shows six peaks originating from lipids and one at 4.8 ppm originating from water protons. The four labeled peaks were

quantifiable reliably and considered for the multi-peak fat model (Fig. 2f). The close proximity to the water peak and a low amplitude limited the evaluability of the methine ( $-\text{CH}=\text{CH}-$ ) and methylene ( $-\text{CH}_2-\text{COO}$ ) peaks, respectively. The manganese phantom is comprehensively characterized in the supplementary material (Fig. S2, Table S1).

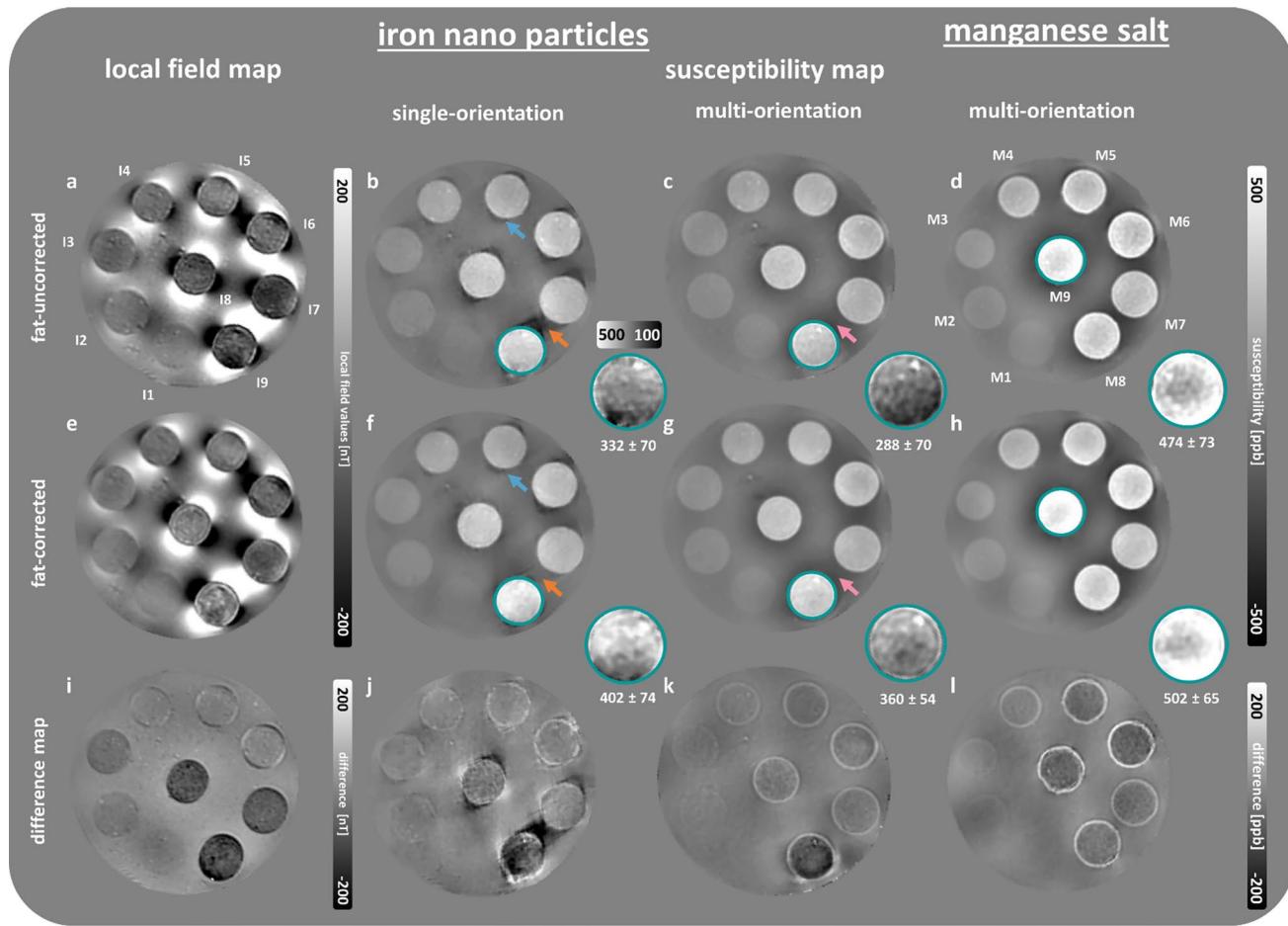
### Local field and susceptibility map analysis

The local field and susceptibility map of the iron phantom are shown in Fig. 3. As expected, the use of fat correction substantially alters the local field values, especially in spheres containing a mixture of iron nanoparticles and fat (I7–I9), which is further supported by the difference map (Fig. 3i). The susceptibility maps are homogeneous in the small spheres' surrounding agar, but exhibit particularly pronounced extraspherical contributions for spheres containing increasing concentrations of iron nanoparticles. The susceptibility values increase in fat- and iron-only spheres with higher fat and iron particle concentrations. Similarly, the susceptibilities increase with higher iron particle concentrations in the spherical inclusions with 20% fat. Substantial dipole inversion artifacts (Fig. 3, orange arrows) are visible around spheres I7–I9 without fat correction that attenuate with fat correction. As shown in the close-ups of I9 (Fig. 3, turquoise circle), fat correction leads to higher susceptibilities. The average local field and susceptibility values and respective standard deviations from VOI-based analysis are presented in Table S2. We further investigated the suitability of  $\text{MnCl}_2$  as a paramagnetic substance for phantom construction (Fig. 3d, h). The principal findings observed for the phantom containing iron nanoparticles could be replicated with the manganese phantom as well and are presented in more detail in the supplementary materials (Fig. S3, Table S3).

Linear fitting of the COSMOS susceptibilities obtained from fat-corrected field maps as a function of iron concentration (Fig. 4) yielded slopes that were very similar when

**Table 3** Means and standard deviations of relaxation rates ( $R_1$ ,  $R_2$ ,  $R_2^*$ ) and PDFF measured in volumes of interests of the individual spherical inclusions of the iron phantom. The known iron nanoparticle and fat concentrations are shown as comparison. The label of the sample refers to the ones in Fig. 1f, with BG additionally referring to a VOI in the large agar sphere

Sample	$R_1 [\text{s}^{-1}]$	$R_2 [\text{s}^{-1}]$	$R_2^* [\text{s}^{-1}]$	PDFF in %	$\text{Fe}_3\text{O}_4 [\text{mM}]$	Fat [vol %]
I1	$3.0 \pm 0.2$	$15.5 \pm 0.4$	$24.2 \pm 2.0$	$4.4 \pm 0.6$	0	4
I2	$2.7 \pm 0.1$	$15.6 \pm 0.2$	$30.5 \pm 2.0$	$9.9 \pm 0.7$	0	10
I3	$2.4 \pm 0.1$	$15.0 \pm 0.3$	$40.6 \pm 3.3$	$17.7 \pm 1.4$	0	20
I4	$2.5 \pm 0.2$	$25.2 \pm 0.7$	$94.2 \pm 7.2$	$0.7 \pm 0.6$	0.13	0
I5	$2.4 \pm 0.3$	$29.7 \pm 1.0$	$132.7 \pm 6.7$	$1.0 \pm 0.8$	0.19	0
I6	$2.4 \pm 0.3$	$33.9 \pm 1.2$	$169.3 \pm 8.7$	$1.0 \pm 0.9$	0.26	0
I7	$2.7 \pm 0.3$	$20.6 \pm 0.6$	$127.0 \pm 11.5$	$19.6 \pm 5.7$	0.13	20
I8	$2.9 \pm 0.3$	$21.0 \pm 0.6$	$168.9 \pm 13.2$	$21.3 \pm 2.2$	0.19	20
I9	$2.8 \pm 0.3$	$21.3 \pm 0.6$	$211.3 \pm 18.9$	$21.2 \pm 2.2$	0.26	20
BG	$2.8 \pm 0.1$	$16.3 \pm 0.2$	$20.8 \pm 2.1$	$0.52 \pm 0.1$	0	0



**Fig. 3** Local field (**a**, **e**) and susceptibility maps (**b**, **c**, **f**, **g**) of a representative slice of the iron phantom and exemplary susceptibility maps of the manganese phantom (**d**, **h**). Fat-uncorrected maps are presented in (**a**–**d**), followed by those experiencing fat correction (**e**–**h**), and the respective difference maps (**i**–**l**) (fat-corrected map subtracted from fat-uncorrected map). Single-orientation field-to-susceptibility inversion was conducted using HEIDI, whereas COSMOS was used for the multi-orientation approach. Enlarged views of the spherical inclusion I9, circled in turquoise and scaled between 100 and 500

focusing on pure iron (slope: 1088 ppb/mM) and focusing on iron–fat mixtures (slope: 1082 ppb/mM). The intercept, however, increased substantially from 9.15 ppb (without fat) to 86.08 ppb when fat was present.

## Susceptibility and $R_2^*$ analysis

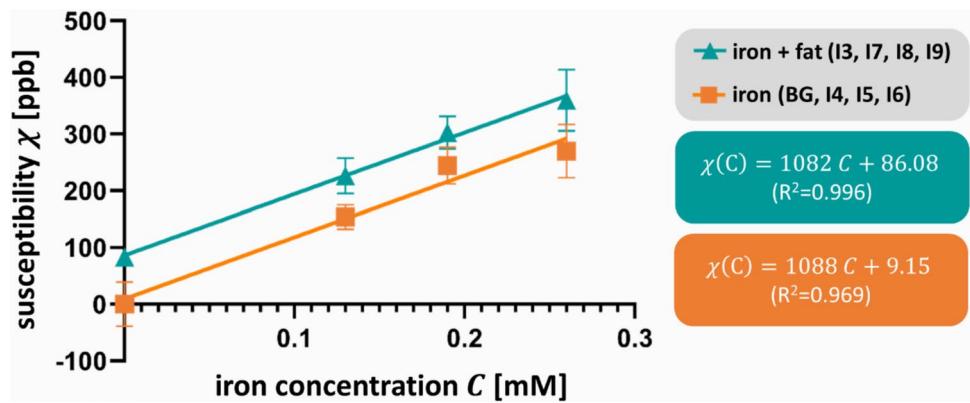
The relationship between magnetic susceptibility and  $R_2^*$  within inclusions of iron, iron and fat, as well as manganese is illustrated in Fig. 5. Overall, the data is more scattered along the susceptibility axis than along the  $R_2^*$  axis. Both,  $R_2^*$  and susceptibility, are elevated in spheres containing iron and fat mixtures (Fig. 5a) compared to the ones

ppb, highlight the changes in susceptibility due to fat correction. The mean value and standard deviation of the susceptibilities within the turquoise region are located under the close-up in ppb. All susceptibility maps are referenced to the background agar susceptibility. Orange arrows indicate a reduction of streaking (extraspherical contributions) after fat correction. In the COSMOS maps (pink arrows), these extraspherical contributions are marginal and barely change due to fat correction. The labeling of the small spherical inclusions is provided in (**a**, **d**)

with iron only (Fig. 5b). Here, the slopes of the fitting line were very similar, with values of  $0.548 \text{ ppb}^{-1} \cdot \text{s}^{-1}$  versus  $0.540 \text{ ppb}^{-1} \cdot \text{s}^{-1}$ , respectively. The values in the inclusions with manganese only (Fig. 5c) are more scattered than the ones observed in the iron-only inclusions. Interestingly, the susceptibilities in M6 are substantially higher than those in I6, while  $R_2^*$  in M6 is slightly reduced compared to I6.

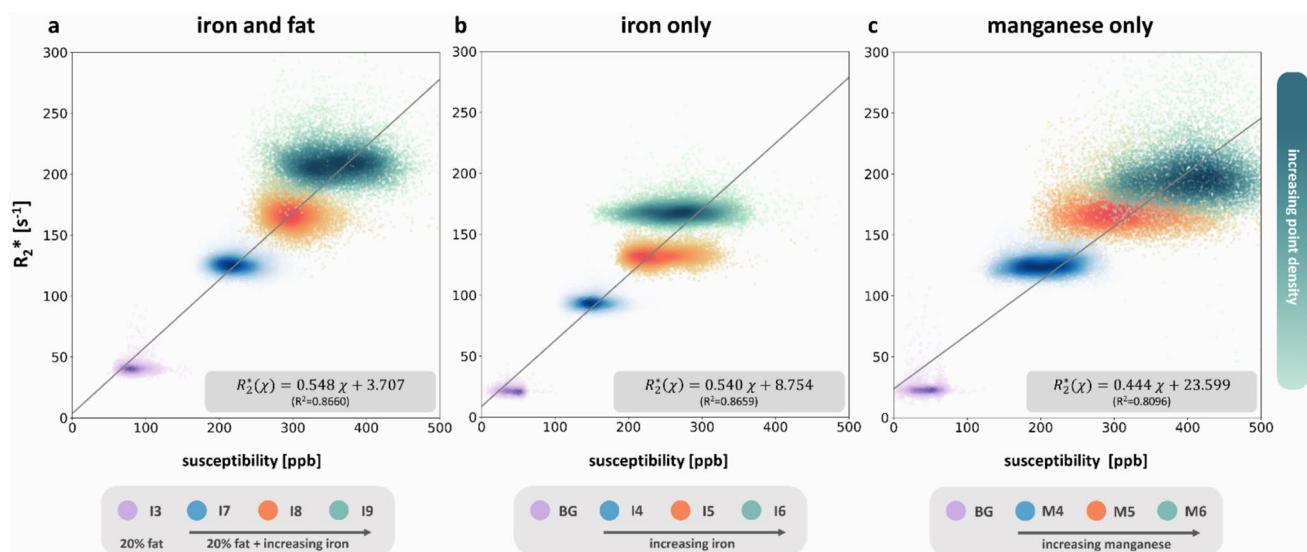
## Repeatability and long-term stability

Figure 6 summarizes the measured relaxation rates ( $R_1$ ,  $R_2$ ,  $R_2^*$ ), PDFFs, and magnetic susceptibilities from the initial measurement and two follow-up sessions conducted



**Fig. 4** Susceptibility  $\chi$  versus iron concentration  $C$  for iron-laden spheres with 20% fat (turquoise) and without fat (orange). BG, background, denotes the volume of interest placed in the large agar sphere. The average susceptibility values and the standard deviations

(error bars) of the spherical inclusions are plotted for the COSMOS approach. A linear least squares fit was applied to the data, producing the presented slope, intercept, and coefficient of determination ( $R^2$ )



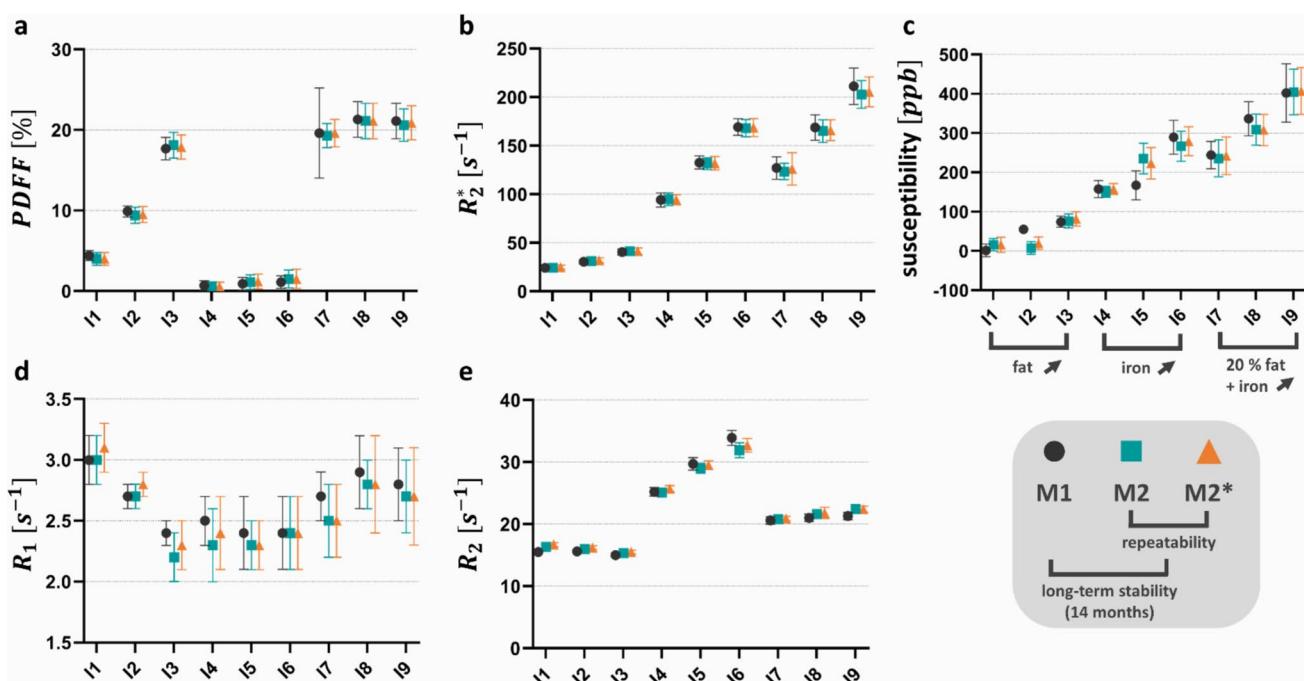
**Fig. 5** Susceptibility versus  $R_2^*$  in the presence of fat and iron (a), iron only (b), and manganese only (c). The COSMOS susceptibilities are plotted against the  $R_2^*$  values for four distinct VOIs. Purple indicates the fat-only VOI (I3) or background VOIs (BG), while blue, orange, and turquoise represent VOIs with increasing concentrations

of the paramagnetic substances. The density of data points within each VOI is visualized through color fading. A linear least squares fit was applied to the data, producing the presented slope, intercept, and coefficient of determination ( $R^2$ ) in the bottom right corner of each plot

on the same day after 14 months. The quantitative parameters exhibit similar variations across both time intervals. One-way ANOVA with repeated measurements indicated no statistically significant differences in the average  $R_2$ ,  $R_2^*$ , PDFF and susceptibilities across the individual measurements ( $p < 0.05$ ). While ANOVA indicated statistically significant  $R_1$  variations across measurements, these lie within the respective standard deviations.

## Discussion

This work proposed a dedicated phantom design, using agar spheres doped with varying concentrations of paramagnetic solution and fat surrounded in agar, to achieve an interface-free phantom, simulating liver tissue under healthy and pathologic conditions. The phantom was comprehensively characterized regarding its relaxation properties and its fat composition. Additionally, the phantom-specific fat spectrum was obtained by MR spectroscopy



**Fig. 6** Repeatability and long-term stability assessment. Average PDFF (a),  $R_2^*$  (b), susceptibility (c),  $R_1$  (d), and  $R_2$  values (e) for each spherical inclusion are plotted across different measurement sessions (M1, M2, M2\*). Error bars indicate the standard deviation. The repeatability measurements (scan–rescan) M2 (turquoise) and M2\*

(orange) were conducted 14 months after the initial measurement M1 (dark gray). The susceptibilities were derived from a single-orientation susceptibility map reconstructed using HEIDI-based field-to-source inversion

for subsequent fat correction. We employed the commonly used [33, 35] iterative graph-cut approach from Hernando et al. [46] to correct the SHARP-processed local field maps for the influence of fat.

The analysis of susceptibility maps demonstrated the influence of fat correction, particularly in regions near inclusions with iron–fat mixtures, where artifacts, notably high-frequency streaking, were markedly reduced on fat-corrected maps of the single-orientation approach (Fig. 3, orange arrows). Because of the intrinsic oversampling of the field-to-susceptibility inversion problem in COSMOS, extra-spherical artifacts were less pronounced than on the single-orientation susceptibility maps even in particular if the local field is not fat corrected. We generally noticed increased susceptibility values in iron- and manganese-laden spheres after fat correction for both the single- and multi-orientation approach.

As expected, we observed a linear relationship between susceptibility and iron concentration, with slopes representing the molar susceptibility at 3T that were similar in both fat-free and fat-containing inclusions. The slope of 1088 ppb/mM in fat-free inclusions (Fig. 4) was substantially higher compared to that in other phantom studies. For instance, Dietrich et al. [64] reported a slope of  $88 \pm 3$  ppb/mM for ferric chloride ( $\text{FeCl}_3$ ) solutions, and Gustavo Cuna et al. [65] found slopes of  $161 \pm 32$  ppb/mM

for free iron and  $40 \pm 2$  ppb/mM for clustered iron (iron absorbed in sodium polyacrylate), suggesting that differences in dopants, such as our use of super-paramagnetic iron oxide particles with core sizes of 4–6 nm as measured by transmission electron microscopy, may account for the observed deviations. Our slope is very close to that of monocrystalline iron oxide nanoparticles with the same core size (MION-46L) for which a molar susceptibility of 1382.3 ppb/mM has been reported [66, 67]. To our best knowledge, there is no study comparing liver susceptibility with biopsy-based iron measurements. However, in some studies, in vivo liver susceptibility was correlated to LIC measured via the biopsy-calibrated FerriScan technique yielding slopes of approximately 100–200 ppb\* $\text{g dw}/\text{mg}$  (slopes were estimated from the diagrams). After converting dry weight to wet weight using the conversion factor 0.2439 (1/4.1) [35, 68], and dividing it by the molar mass of iron (55.85 g/mol), these slopes were 436 ppb/mM and 873 ppb/mM. Despite the substantially smaller particle size we used to that in the liver (0.1–3.2  $\mu\text{m}$ ) [69], we observed a relationship between susceptibility and iron in a similar range to that determined *in vivo*.

In contrast to the slope in Fig. 4, our linear regression between susceptibility and iron concentration revealed markedly different intercepts for fat-free inclusions ( $i_{ff} = 9.2$  ppb) and fat-containing inclusions ( $i_{fc} = 86.1$  ppb). This difference

is expected, since fat has a slightly higher magnetic susceptibility (susceptibility of protons in fat:  $-8.44$  ppm) than water ( $\chi_{\text{H}_2\text{O}} = -9.05$  ppm) [70]. Clinically, however, such differences are critical. Using our regression equations (Fig. 4), a susceptibility of 200 ppb translates to 0.175 mM iron in fat-free tissue, but only 0.105 mM iron in tissue with 20% fat—a reduction of approximately 40%. Consequently, fat concentration must be accounted for by establishing a calibration curve between magnetic susceptibility and PDFF. Our data show a strong linear relationship between susceptibility and measured PDFF ( $\chi_{\text{fat}}(\text{PDFF}) = 4.83 \frac{\text{ppb}}{\%} \cdot \text{PDFF\%} - 6.49 \text{ ppb}$ ,  $R^2 = 0.91$ , determined in the inclusions: I1, I2, I3, the surrounding medium). Incorporating this correction into the regression model without fat content ( $c_{\text{Fe}}(\chi, \text{PDFF}) = (\chi - i_{\text{ff}} - \chi_{\text{fat}}(\text{PDFF})) / 1088 \text{ ppb/mM}$ ), a 200 ppb shift translates to an estimated iron concentration of 0.093 mM at 20% fat and 0.026 mM at 35% fat. As susceptibility-based iron quantification becomes less accurate with increasing fat content, its reliability may be compromised in patients with fatty liver, especially those with severe steatosis. Similar issues have been reported for  $R_2^*$ -based liver iron quantification, where correcting  $R_2^*$  values for PDFF effects has been proposed [71]. Overall, further research is required to explore the relationship between magnetic susceptibility and PDFF and develop a fat-compensated approach for iron quantification.

Linear regression analyses confirm that both susceptibility and  $R_2^*$  are linear indicators of iron content, suggesting that higher iron concentrations are often associated with magnetic field inhomogeneities at the voxel level. The slopes of the linear relationship between  $R_2^*$  and COSMOS susceptibility observed in this study (iron only: 0.54, iron and fat: 0.548) were notably higher than those reported in the liver *in vivo* using anisotropic GRE-VIBE imaging (Sharma et al. 2015 [35]: 0.357 s/ppb, Sharma et al. 2017 [34]: 0.294 s/ppb, Li et al. [19]: 0.333 s/ppb), likely due to differences between phantom and *in vivo* conditions, such as the homogeneous iron storage in the phantom versus ferritin cores in the liver, the smaller iron particle sizes (4–6 nm) compared to the ones in the liver (0.1–3.2  $\mu\text{m}$ ) [69], or different acquisition and post-processing protocols. In comparison, slopes between  $R_2^*$  and susceptibility in cerebral gray matter ranged from 0.117 to 0.157  $\text{ppb}^{-1} \text{s}^{-1}$  (adjusted from 7 to 3T) [72, 73], with smaller slopes potentially attributed to reduced intravoxel spin dephasing due to smaller voxel sizes. It is important to note that the fundamental mechanisms underlying susceptibility changes and their effects on MRI signals do not imply a direct relationship between magnetic susceptibility and  $R_2^*$ . While pure susceptibility differences induce a frequency shift of the Larmor frequency without affecting  $R_2^*$ , intravoxel susceptibility variations increase intravoxel dephasing, hence elevating  $R_2^*$  [73, 74].

Thus, susceptibility-based estimation of iron concentration is insensitive to the microscopic spatial distribution of iron, potentially providing a distribution independent measure of iron concentration.

The phantoms we constructed met the specified criteria and showed their suitability through a thorough characterization process. Visual appraisal revealed that the assembled phantom is made of homogeneously distributed solutions within the larger and smaller spheres that contain negligible amounts of air inclusions. The presence of air bubbles must be minimized, as they introduce susceptibility differences of about 10 ppm [20] at the air–liquid interfaces, leading to local magnetic field distortions. These distortions can cause phase wrapping artifacts, signal dephasing in gradient echo sequences, and inaccuracies in QSM. Additionally, air inclusions may interfere with water–fat separation algorithms and bias  $R_2^*$  measurements. The repeatability and long-term stability assessments demonstrated consistent quantitative parameters across sessions and a 14-month follow-up, suggesting that the phantom maintains its stability over at least this time period. The stability is likely facilitated by the phantom’s storage in an airtight container at room temperature, which prevents desiccation and preserves the integrity of its constituents. These results support the phantom’s potential as a reliable reference for evaluating new MRI approaches in longitudinal MRI studies, though further investigations over longer time frames could provide additional insights into its durability.

The quantitative analysis yielded values within the range of liver tissue, and the susceptibility maps further confirmed the feasibility of the spherical phantom design. To ultimately bridge the gap between the phantom signal and the *in vivo* liver, several adaptations are conceivable to overcome the limitations of the current phantom design. Firstly, while the phantoms replicate paramagnetic and fat-based signal contributions, the current concentrations achieve maximal  $R_2^*$  values, which are indicative of mild pathological iron overload [41]. MRI fat fractions and the grading of steatosis, which measures the percentage of cells with intracellular fat vacuoles, do not directly correlate [75]. Nonetheless, Tang et al. [51] established thresholds for steatosis grading, allowing the phantom to simulate steatosis grades 0, 1 ( $\text{PDFF} < 6.4\%$  to  $< 17.4\%$ ), and 2 ( $\text{PDFF} < 22.1\%$ ). When increasing fat concentrations in future phantoms, strategies to prevent the formation of inhomogeneous agar solutions for fat concentrations over 20% have to be considered. Here, incorporating both water-soluble and oil-soluble surfactants during phantom construction could stabilize the water–agar–fat emulsion more effectively [76]. Other Potential confounders—including Impragnol and 1% formaldehyde—were carefully managed to maintain the phantom’s structural integrity. Due to the application of Impragnol only at the surface of the small spherical inclusion, any impact on quantitative MR

parameters measured within a rather large VOI is expected to be negligible. As formaldehyde was uniformly added to all solutions of the inclusions, its influence, which at the applied concentration, is substantially lower than that of the other compounds (iron, fat,  $MnCl_2$ ) [77], remains consistent across all inclusions and does not affect the analysis of their varying concentrations. Secondly, our proposed phantom does not specifically account for liver fibrosis, which alters  $T_1$  values. This limitation could be addressed by incorporating nickel chloride, as demonstrated by Zhao et al. [39]. Given that nickel chloride introduces a distinct paramagnetic susceptibility contribution (molar susceptibility of  $\chi_{NiCl_2}$  at  $20\text{ }^\circ\text{C} = 4436 * 10^{-6}$  ml/mol) [78], its impact on the phantom's overall susceptibility has to be accounted for by adjusting the concentration of the paramagnetic solution during construction. Thirdly, on a microscopic level, the composition of the phantom and liver tissue differs. In liver tissue, iron ions are stored in ferritin shells within hepatocytes [7, 8], and fatty acids that accumulate as intrahepatocytic lipid vesicles, i.e., lipid droplets in the cytoplasm of hepatocytes, with a range of droplet sizes [79, 80]. In our phantom, however, iron nanoparticles are mixed directly into the agar–fat solution, likely causing iron particles to segregate into the agar–water phase while the hydrophobic lipids from peanut oil form globules. Although the use of surfactants might alleviate this effect at higher fat concentrations (> 20%), we assume a macroscopically homogeneous distribution of lipids and iron particles. Nevertheless, the question remains whether the droplet size distribution in the phantom matches that observed in hepatocytes. In vivo, droplet sizes vary with fat volume fraction, where smaller droplets dominate at lower fat fractions, while coalescence at higher fractions yields larger droplets. In contrast, the phantom's mixing process may yield a relatively uniform droplet size regardless of the fat concentration. To assess and compare the droplet sizes within the phantom with those reported in liver tissue, microscopy (light or electron microscopy) or laser diffraction techniques [81] would be straightforward approaches. Recently, a study for estimating lipid droplet sizes in adipose tissue using MRI was proposed [82], representing an MR technique to potentially estimate our phantom's droplet sizes.

We demonstrated that manganese chloride can also be used as paramagnetic substance to replicate the  $R_2^*$  and susceptibility effects caused by iron accumulation. Despite the lower cost of manganese chloride compared to iron nanoparticles, it exhibits different relaxivities [39, 83], which influences the quantity of the chemical required to achieve the desired relaxation rates. Moreover, Zhao et al. [39] showed that manganese chloride solely alters  $R_2^*$  of water, whereas iron microspheres affect the  $R_2^*$  of both water and fat. Other chemicals mimicking effects of LIC on MRI are also conceivable. For instance, the combination of

paramagnetic salt with iron microspheres has been shown to provide more realistic signal behaviors, closer to those of in vivo measurements, as shown by relaxometry studies [39, 84, 85]. Therefore, it would be worthwhile to explore such a configuration in spherical inclusions of the proposed phantom in future studies.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10334-025-01261-3>.

**Acknowledgements** This work was supported by the European Regional Development Fund (ERDF—IP\* 1b, ZS/2021/06/158189). MRI data was collected using an MR scanner for human research founded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft, INST 271/406-1 FUGG). We extend our gratitude to the authors of the Fat–Water Toolbox (<http://ismrm.org/workshops/FatWater12/data.htm>). We also like to thank Anne Slawig (University clinic and polyclinic for Radiology, University Hospital Halle (Saale), Germany) for the valuable discussions and insights.

**Author contributions** Conceptualization: Graf, Trapp, Deistung; Methodology: Graf, Trapp, Rothe, Gussew, Deistung. Formal analysis and investigation: Graf, Trapp, Rothe, Gussew, Deistung. Writing—original draft preparation: Graf, Deistung. Writing—review and editing: Graf, Rothe, Gussew, Wohlgemuth, Deistung. Funding acquisition: Wohlgemuth, Deistung. Resources: Wohlgemuth, Deistung. Supervision: Wohlgemuth, Deistung.

**Funding** Open Access funding enabled and organized by Projekt DEAL. Deutsche Forschungsgemeinschaft, INST 271/406-1 FUGG, European Regional Development Fund, ERDF—IP\* 1b, ZS/2021/06/158189.

**Data availability** The data supporting this study's findings are available from the corresponding author upon reasonable request.

## Declarations

**Conflict of interest** The authors declare that they no conflict of interest.

**Ethical standards** This article does not contain any studies with human participants or animals performed by any of the authors.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Lieu PT, Heiskala M, Peterson PA et al (2001) The roles of iron in health and disease. *Mol Aspects Med* 22:1–87. [https://doi.org/10.1016/S0098-2997\(00\)00006-6](https://doi.org/10.1016/S0098-2997(00)00006-6)
2. Hsu CC, Senussi NH, Fertrin KY et al (2022) Iron overload disorders. *Hepatol Commun* 6:1842–1854. <https://doi.org/10.1002/hep4.2012>
3. Hadzhieva M, Kirches E, Mawrin C (2014) Review: iron metabolism and the role of iron in neurodegenerative disorders. *Neuropathol Appl Neurobiol* 40:240–257. <https://doi.org/10.1111/nan.12096>
4. Ward RJ, Zucca FA, Duyn JH et al (2014) The role of iron in brain ageing and neurodegenerative disorders. *Lancet Neurol* 13:1045–1060. [https://doi.org/10.1016/S1474-4422\(14\)70117-6](https://doi.org/10.1016/S1474-4422(14)70117-6)
5. Brittenham GM, Cohen AR, McLaren CE et al (1993) Hepatic iron stores and plasma ferritin concentration in patients with sickle cell anemia and thalassemia major. *Am J Hematol* 42:81–85. <https://doi.org/10.1002/ajh.2830420116>
6. Nielsen P, Günther U, Dürken M et al (2000) Serum ferritin iron in iron overload and liver damage: correlation to body iron stores and diagnostic relevance. *J Lab Clin Med* 135:413–418. <https://doi.org/10.1067/mlc.2000.106456>
7. Bonkovsky HL (1991) Iron and the liver. *Am J Med Sci* 301:32–43. <https://doi.org/10.1097/00000441-199101000-00006>
8. Anderson ER, Shah YM (2013) Iron homeostasis in the liver. *Compr Physiol* 3:315–330. <https://doi.org/10.1002/cphy.c120016>
9. Angelucci E, Brittenham GM, McLaren CE et al (2000) Hepatic iron concentration and total body iron stores in thalassemia major. *N Engl J Med* 343:327–331. <https://doi.org/10.1056/NEJM200008033430503>
10. Brittenham GM, Badman DG (2003) Noninvasive measurement of iron: report of an NIDDK workshop. *Blood* 101:15–19. <https://doi.org/10.1182/blood-2002-06-1723>
11. Villeneuve JP, Bilodeau M, Lepage R et al (1996) Variability in hepatic iron concentration measurement from needle-biopsy specimens. *J Hepatol* 25:172–177. [https://doi.org/10.1016/S0168-8278\(96\)80070-5](https://doi.org/10.1016/S0168-8278(96)80070-5)
12. Emond MJ, Bronner MP, Carlson TH et al (1999) Quantitative study of the variability of hepatic iron concentrations. *Clin Chem* 45:340–346
13. Sirlin CB, Reeder SB (2010) Magnetic resonance imaging quantification of liver iron. *Magn Reson Imaging Clin N Am* 18(359–81):ix. <https://doi.org/10.1016/j.mric.2010.08.014>
14. St Pierre TG, Clark PR, Chua-Anusorn W (2004) Single spin-echo proton transverse relaxometry of iron-loaded liver. *NMR Biomed* 17:446–458. <https://doi.org/10.1002/nbm.905>
15. Reeder SB, Yokoo T, França M et al (2023) Quantification of liver iron overload with MRI: review and guidelines from the ESGAR and SAR. *Radiology* 307:e221856. <https://doi.org/10.1148/radiol.221856>
16. Hernando D, Cook RJ, Qazi N et al (2021) Complex confounder-corrected  $R_2^*$  mapping for liver iron quantification with MRI. *Eur Radiol* 31:264–275. <https://doi.org/10.1007/s00330-020-07123-x>
17. Henninger B, Alustiza J, Garbowski M et al (2020) Practical guide to quantification of hepatic iron with MRI. *Eur Radiol* 30:383–393. <https://doi.org/10.1007/s00330-019-06380-9>
18. Jensen JH, Tang H, Tosti CL et al (2010) Separate MRI quantification of dispersed (ferritin-like) and aggregated (hemosiderin-like) storage iron. *Magn Reson Med* 63:1201–1209. <https://doi.org/10.1002/mrm.22273>
19. Li J, Lin H, Liu T et al (2018) Quantitative susceptibility mapping (QSM) minimizes interference from cellular pathology in  $R_2^*$  estimation of liver iron concentration. *J Magn Reson Imaging* 48:1069–1079. <https://doi.org/10.1002/jmri.26019>
20. Schenck JF (1996) The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. *Med Phys* 23:815–850. <https://doi.org/10.1111/1.597854>
21. Liu S, Wang C, Zhang X et al (2018) Quantification of liver iron concentration using the apparent susceptibility of hepatic vessels. *Quant Imaging Med Surg* 8:123–134. <https://doi.org/10.21037/qims.2018.03.02>
22. Schweser F, Deistung A, Reichenbach JR (2016) Foundations of MRI phase imaging and processing for quantitative susceptibility mapping (QSM). *Z Med Phys* 26:6–34. <https://doi.org/10.1016/j.zemedi.2015.10.002>
23. Deistung A, Schweser F, Reichenbach JR (2017) Overview of quantitative susceptibility mapping. *NMR Biomed* 30. <https://doi.org/10.1002/nbm.3569>
24. Langkammer C, Schweser F, Krebs N et al (2012) Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. *Neuroimage* 62:1593–1599. <https://doi.org/10.1016/j.neuroimage.2012.05.049>
25. Deistung A, Jäschke D, Draganova R et al (2022) Quantitative susceptibility mapping reveals alterations of dentate nuclei in common types of degenerative cerebellar ataxias. *Brain Commun* 4:fcab306. <https://doi.org/10.1093/braincomms/fcab306>
26. Li KR, Avecillas-Chasin J, Nguyen TD et al (2022) Quantitative evaluation of brain iron accumulation in different stages of Parkinson's disease. *J Neuroimaging* 32:363–371. <https://doi.org/10.1111/jon.12957>
27. Li W, Wu B, Liu C (2011) Quantitative susceptibility mapping of human brain reflects spatial variation in tissue composition. *Neuroimage* 55:1645–1656. <https://doi.org/10.1016/j.neuroimage.2010.11.088>
28. Fan AP, Bilgic B, Gagnon L et al (2014) Quantitative oxygenation venography from MRI phase. *Magn Reson Med* 72:149–159. <https://doi.org/10.1002/mrm.24918>
29. Ward PGD, Ferris NJ, Raniga P et al (2018) Combining images and anatomical knowledge to improve automated vein segmentation in MRI. *Neuroimage* 165:294–305. <https://doi.org/10.1016/j.neuroimage.2017.10.049>
30. Schweser F, Deistung A, Lehr BW et al (2010) Differentiation between diamagnetic and paramagnetic cerebral lesions based on magnetic susceptibility mapping. *Med Phys* 37:5165–5178. <https://doi.org/10.1111/1.3481505>
31. Deistung A, Schweser F, Wiestler B et al (2013) Quantitative susceptibility mapping differentiates between blood depositions and calcifications in patients with glioblastoma. *PLoS ONE* 8:e57924. <https://doi.org/10.1371/journal.pone.0057924>
32. Dimov AV, Li J, Nguyen TD et al (2023) QSM throughout the body. *J Magn Reson Imaging*. <https://doi.org/10.1002/jmri.28624>
33. Zhao R, Velikina J, Reeder SB et al (2023) Validation of liver quantitative susceptibility mapping across imaging parameters at 1.5 T and 3.0 T using SQUID susceptometry as reference. *Magn Reson Med* 89:1418–1428. <https://doi.org/10.1002/mrm.29529>
34. Sharma SD, Fischer R, Schoennagel BP et al (2017) MRI-based quantitative susceptibility mapping (QSM) and  $R_2^*$  mapping of liver iron overload: comparison with SQUID-based biomagnetic liver susceptometry. *Magn Reson Med* 78:264–270. <https://doi.org/10.1002/mrm.26358>
35. Sharma SD, Hernando D, Horng DE et al (2015) Quantitative susceptibility mapping in the abdomen as an imaging biomarker of hepatic iron overload. *Magn Reson Med* 74:673–683. <https://doi.org/10.1002/mrm.25448>
36. Jafari R, Hectors SJ, Koehne de González AK et al (2021) Integrated quantitative susceptibility and  $R_2^*$  mapping for evaluation of liver fibrosis: an ex vivo feasibility study. *NMR Biomed* 34:e4412. <https://doi.org/10.1002/nbm.4412>

37. Qu Z, Yang S, Xing F et al (2021) Magnetic resonance quantitative susceptibility mapping in the evaluation of hepatic fibrosis in chronic liver disease: a feasibility study. *Quant Imaging Med Surg* 11:1170–1183. <https://doi.org/10.21037/qims-20-720>

38. Weingärtner S, Desmond KL, Obuchowski NA et al (2022) Development, validation, qualification, and dissemination of quantitative MR methods: overview and recommendations by the ISMRM quantitative MR study group. *Magn Reson Med* 87:1184–1206. <https://doi.org/10.1002/mrm.29084>

39. Zhao R, Hamilton G, Brittain JH et al (2021) Design and evaluation of quantitative MRI phantoms to mimic the simultaneous presence of fat, iron, and fibrosis in the liver. *Magn Reson Med* 85:734–747. <https://doi.org/10.1002/mrm.28452>

40. Kim J-H, Kim J-H, Lee S-H et al (2019) Fabrication of a spherical inclusion phantom for validation of magnetic resonance-based magnetic susceptibility imaging. *PLoS ONE* 14:e0220639. <https://doi.org/10.1371/journal.pone.0220639>

41. Labranche R, Gilbert G, Cerny M et al (2018) Liver iron quantification with MR imaging: a primer for radiologists. *Radiographics* 38:392–412. <https://doi.org/10.1148/radio.2018170079>

42. Hernando D, Kramer JH, Reeder SB (2013) Multipeak fat-corrected complex  $R_2^*$  relaxometry: theory, optimization, and clinical validation. *Magn Reson Med* 70:1319–1331. <https://doi.org/10.1002/mrm.24593>

43. Wharton S, Schäfer A, Bowtell R (2010) Susceptibility mapping in the human brain using threshold-based k-space division. *Magn Reson Med* 63:1292–1304. <https://doi.org/10.1002/mrm.22334>

44. Kim M, Gillen J, Landman BA et al (2009) Water saturation shift referencing (WASSR) for chemical exchange saturation transfer (CEST) experiments. *Magn Reson Med* 61:1441–1450. <https://doi.org/10.1002/mrm.21873>

45. Ma J (2008) Dixon techniques for water and fat imaging. *J Magn Reson Imaging* 28:543–558. <https://doi.org/10.1002/jmri.21492>

46. Hernando D, Kellman P, Haldar JP et al (2010) Robust water/fat separation in the presence of large field inhomogeneities using a graph cut algorithm. *Magn Reson Med* 63:79–90. <https://doi.org/10.1002/mrm.22177>

47. Hernando D, Hines CDG, Yu H et al (2012) Addressing phase errors in fat–water imaging using a mixed magnitude/complex fitting method. *Magn Reson Med* 67:638–644. <https://doi.org/10.1002/mrm.23044>

48. Boehm C, Diefenbach MN, Makowski MR et al (2021) Improved body quantitative susceptibility mapping by using a variable-layer single-min-cut graph-cut for field-mapping. *Magn Reson Med* 85:1697–1712. <https://doi.org/10.1002/mrm.28515>

49. Schweser F, Deistung A, Lehr BW et al (2011) Quantitative imaging of intrinsic magnetic tissue properties using MRI signal phase: an approach to in vivo brain iron metabolism? *Neuroimage* 54:2789–2807. <https://doi.org/10.1016/j.neuroimage.2010.10.070>

50. Liu T, Spincemaille P, de Rochebort L et al (2009) Calculation of susceptibility through multiple orientation sampling (COSMOS): a method for conditioning the inverse problem from measured magnetic field map to susceptibility source image in MRI. *Magn Reson Med* 61:196–204. <https://doi.org/10.1002/mrm.21828>

51. Tang A, Desai A, Hamilton G et al (2015) Accuracy of MR imaging-estimated proton density fat fraction for classification of dichotomized histologic steatosis grades in nonalcoholic fatty liver disease. *Radiology* 274:416–425. <https://doi.org/10.1148/radiol.14140754>

52. Middleton MS, Heba ER, Hooker CA et al (2017) Agreement between magnetic resonance imaging proton density fat fraction measurements and pathologist-assigned steatosis grades of liver biopsies from adults with nonalcoholic steatohepatitis. *Gastroenterology* 153:753–761. <https://doi.org/10.1053/j.gastro.2017.06.005>

53. Qu Y, Li M, Hamilton G et al (2019) Diagnostic accuracy of hepatic proton density fat fraction measured by magnetic resonance imaging for the evaluation of liver steatosis with histology as reference standard: a meta-analysis. *Eur Radiol* 29:5180–5189. <https://doi.org/10.1007/s00330-019-06071-5>

54. Verlhac S, Morel M, Bernaudin F et al (2015) Liver iron overload assessment by MRI  $R_2^*$  relaxometry in highly transfused pediatric patients: an agreement and reproducibility study. *Diagn Interv Imaging* 96:259–264. <https://doi.org/10.1016/j.diii.2014.11.021>

55. Kratzer W, Fritz V, Mason RA et al (2003) Factors affecting liver size: a sonographic survey of 2080 subjects. *J Ultrasound Med* 22:1155–1161. <https://doi.org/10.7863/jum.2003.22.11.1155>

56. Yu H, Shimakawa A, McKenzie CA et al (2008) Multiecho water–fat separation and simultaneous  $R_2^*$  estimation with multifrequency fat spectrum modeling. *Magn Reson Med* 60:1122–1134. <https://doi.org/10.1002/mrm.21737>

57. Barral JK, Gudmundson E, Stikov N et al (2010) A robust methodology for in vivo T1 mapping. *Magn Reson Med* 64:1057–1067. <https://doi.org/10.1002/mrm.22497>

58. Stefan D, Di Cesare F, Andraescu A et al (2009) Quantitation of magnetic resonance spectroscopy signals: the jMRUI software package. *Meas Sci Technol* 20:104035. <https://doi.org/10.1088/0957-0233/20/10/104035>

59. Vanhamme L, van den Boogaart A, van Huffel S (1997) Improved method for accurate and efficient quantification of MRS data with use of prior knowledge. *J Magn Reson* 129:35–43. <https://doi.org/10.1006/jmre.1997.1244>

60. Abdul-Rahman HS, Gdeisat MA, Burton DR et al (2007) Fast and robust three-dimensional best path phase unwrapping algorithm. *Appl Opt* 46:6623–6635. <https://doi.org/10.1364/ao.46.006623>

61. Schweser F, Sommer K, Deistung A et al (2012) Quantitative susceptibility mapping for investigating subtle susceptibility variations in the human brain. *Neuroimage* 62:2083–2100. <https://doi.org/10.1016/j.neuroimage.2012.05.067>

62. Tustison NJ, Cook PA, Holbrook AJ et al (2021) The ANTsX ecosystem for quantitative biological and medical imaging. *Sci Rep* 11:9068. <https://doi.org/10.1038/s41598-021-87564-6>

63. Yushkevich PA, Piven J, Hazlett HC et al (2006) User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* 31:1116–1128. <https://doi.org/10.1016/j.neuroimage.2006.01.015>

64. Dietrich O, Levin J, Ahmadi S-A et al (2017) MR imaging differentiation of  $Fe^{2+}$  and  $Fe^{3+}$  based on relaxation and magnetic susceptibility properties. *Neuroradiology* 59:403–409. <https://doi.org/10.1007/s00234-017-1813-3>

65. Gustavo Cuña E, Schulz H, Tuzzi E et al (2023) Simulated and experimental phantom data for multi-center quality assurance of quantitative susceptibility maps at 3 T, 7 T and 9.4 T. *Phys Med* 110:102590. <https://doi.org/10.1016/j.ejmp.2023.102590>

66. Deistung A, Reichenbach JR (2011) Effects of Contrast Agents in Susceptibility Weighted Imaging. In: Haacke EM, Reichenbach JR (eds) Susceptibility weighted imaging in MRI: basic concepts and clinical applications. Wiley-Blackwell, Hoboken, N.J, pp 487–515

67. Shen T, Weissleder R, Papisov M et al (1993) Monocrystalline iron oxide nanocompounds (MION): physicochemical properties. *Magn Reson Med* 29:599–604. <https://doi.org/10.1002/mrm.1910290504>

68. Zuyderhoud FM, Hengeveld P, van Gool J et al (1978) A method for measurement of liver iron fractions in needle biopsy specimens and some results in acute liver disease. *Clin Chim Acta* 86:313–321. [https://doi.org/10.1016/0009-8981\(78\)90387-x](https://doi.org/10.1016/0009-8981(78)90387-x)

69. Ghugre NR, Wood JC (2011) Relaxivity-iron calibration in hepatic iron overload: probing underlying biophysical mechanisms using a Monte Carlo model. *Magn Reson Med* 65:837–847. <https://doi.org/10.1002/mrm.22657>

70. Szczepaniak LS, Dobbins RL, Stein DT et al (2002) Bulk magnetic susceptibility effects on the assessment of intra- and extramyocellular lipids *in vivo*. *Magn Reson Med* 47:607–610. <https://doi.org/10.1002/mrm.10086>
71. Ma M, Cheng J, Li X et al (2025) Prediction of MRI  $R_2^*$  relaxometry in the presence of hepatic steatosis by Monte Carlo simulations. *NMR Biomed* 38:e5274. <https://doi.org/10.1002/nbm.5274>
72. Wang R, Xie G, Zhai M et al (2017) Stability of  $R_2^*$  and quantitative susceptibility mapping of the brain tissue in a large scale multi-center study. *Sci Rep* 7:45261. <https://doi.org/10.1038/srep45261>
73. Deistung A, Schäfer A, Schweser F et al (2013) Toward *in vivo* histology: a comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and  $R_2^*$ -imaging at ultra-high magnetic field strength. *Neuroimage* 65:299–314. <https://doi.org/10.1016/j.neuroimage.2012.09.055>
74. Colgan TJ, Knobloch G, Reeder SB et al (2020) Sensitivity of quantitative relaxometry and susceptibility mapping to microscopic iron distribution. *Magn Reson Med* 83:673–680. <https://doi.org/10.1002/mrm.27946>
75. Reeder SB, Sirlin C (2010) Quantification of liver fat with magnetic resonance imaging. *Magn Reson Imaging Clin N Am* 18:337–357. <https://doi.org/10.1016/j.mric.2010.08.013>
76. Bush EC, Gifford A, Coolbaugh CL et al (2018) Fat-water phantoms for magnetic resonance imaging validation: a flexible and scalable protocol. *J Vis Exp*. <https://doi.org/10.3791/57704>
77. Birk C, Soellradl M, Toeglhofer AM et al (2018) Effects of concentration and vendor specific composition of formalin on post-mortem MRI of the human brain. *Magn Reson Med* 79:1111–1115. <https://doi.org/10.1002/mrm.26699>
78. Nettleton HR, Sugden S (1939) The magnetic susceptibility of nickel chloride. *Proc R Soc Lond A* 173:313–323. <https://doi.org/10.1098/rspa.1939.0147>
79. Gluchowski NL, Becuwe M, Walther TC et al (2017) Lipid droplets and liver disease: from basic biology to clinical implications. *Nat Rev Gastroenterol Hepatol* 14:343–355. <https://doi.org/10.1038/nrgastro.2017.32>
80. Mashek DG (2021) Hepatic lipid droplets: a balancing act between energy storage and metabolic dysfunction in NAFLD. *Mol Metab* 50:101115. <https://doi.org/10.1016/j.molmet.2020.101115>
81. Gmach O, Bertsch A, Bilke-Krause C et al (2019) Impact of oil type and pH value on oil-in-water emulsions stabilized by egg yolk granules. *Colloids Surf, A* 581:123788. <https://doi.org/10.1016/j.colsurfa.2019.123788>
82. Weidlich D, Honecker J, Boehm C et al (2021) Lipid droplet-size mapping in human adipose tissue using a clinical 3T system. *Magn Reson Med* 86:1256–1270. <https://doi.org/10.1002/mrm.28755>
83. Ebrahimpour A, Riahi Alam N, Abdolmaleki P et al (2021) Magnetic metal-organic framework based on zinc and 5-aminolevulinic acid: MR imaging and brain tumor therapy. *J Inorg Organomet Polym* 31:1208–1216. <https://doi.org/10.1007/s10904-020-01782-5>
84. Horng DE, Hernando D, Hines CDG et al (2013) Comparison of  $R_2^*$  correction methods for accurate fat quantification in fatty liver. *J Magn Reson Imaging* 37:414–422. <https://doi.org/10.1002/jmri.23835>
85. Horng DE, Hernando D, Reeder SB (2017) Quantification of liver fat in the presence of iron overload. *J Magn Reson Imaging* 45:428–439. <https://doi.org/10.1002/jmri.25382>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.