FULL PAPER

QSM reconstruction challenge 2.0: A realistic in silico head phantom for MRI data simulation and evaluation of susceptibility mapping procedures

José P. Marques[1] | **Jakob Meineke**[2] | **Carlos Milovic**[3,4,5] | **Berkin Bilgic**[6,7,8] | **Kwok-Shing Chan**[1] | **Renaud Hedouin**[1,9] | **Wietske van der Zwaag**[10] | **Christian Langkammer**[11] | **Ferdinand Schweser**[12,13]

[1]Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, the Netherlands

[2]Philips Research, Hamburg, Germany

[3]Department of Electrical Engineering, Pontificia Universidad Catolica de Chile, Santiago, Chile

[4]Biomedical Imaging Center, Pontificia Universidad Catolica de Chile, Santiago, Chile

[5]Department of Medical Physics and Biomedical Engineering, University College London, London, United Kingdom

[6]Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, Massachusetts, USA

[7]Department of Radiology, Harvard Medical School, Boston, Massachusetts, USA

[8]Harvard-MIT Health Sciences and Technology, MIT, Cambridge, Massachusetts, USA

[9]Centre Inria Rennes - Bretagne Atlantique, Rennes, France

[10]Spinoza Center for Neuroimaging, Amsterdam, the Netherlands

[11]Department of Neurology, Medical University of Graz, Graz, Austria

[12]Buffalo Neuroimaging Analysis Center, Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, The State University of New York, Buffalo, New York, USA

[13]Center for Biomedical Imaging, Clinical and Translational Science Institute, University at Buffalo, The State University of New York, Buffalo, New York, USA

Correspondence
José P. Marques, Donders Institute for Brain, Cognition and Behavior, Radboud University, Kapittelweg 29, 6525 EN Nijmegen, the Netherlands.
Email: j.marques@donders.ru.nl

Purpose: To create a realistic in silico head phantom for the second QSM reconstruction challenge and for future evaluations of processing algorithms for QSM.

Methods: We created a digital whole-head tissue property phantom by segmenting and postprocessing high-resolution (0.64 mm isotropic), multiparametric MRI data acquired at 7 T from a healthy volunteer. We simulated the steady-state magnetization at 7 T using a Bloch simulator and mimicked a Cartesian sampling scheme through Fourier-based processing. Computer code for generating the phantom and performing the MR simulation was designed to facilitate flexible modifications of the phantom in the future, such as the inclusion of pathologies as well as the simulation of a wide range of acquisition protocols. Specifically, the following parameters and effects were implemented: TR and TE, voxel size, background fields, and RF phase biases. Diffusion-weighted imaging phantom data are provided, allowing future investigations of tissue-microstructure effects in phase and QSM algorithms.

Results: The brain part of the phantom featured realistic morphology with spatial variations in relaxation and susceptibility values similar to the in vivo setting. We demonstrated some of the phantom's properties, including the possibility of generating phase data with nonlinear evolution over TE due to partial-volume effects or complex distributions of frequency shifts within the voxel.

Conclusion: The presented phantom and computer programs are publicly available and may serve as a ground truth in future assessments of the faithfulness of quantitative susceptibility reconstruction algorithms.

1 | INTRODUCTION

Quantitative susceptibility mapping has proven to be a valuable tool for assessing iron concentrations in the deep gray matter,[1-3] estimating vessel oxygenation and geometry,[4,5] differentiating blood and calcium products,[6,7] and studying demyelinating lesions in the white matter.[8-11] However, several recent methodical investigations have suggested that study outcomes may depend on the particular processing algorithms chosen for QSM.[12-14] Quantitative susceptibility mapping typically involves the following steps: coil combination,[12] phase unwrapping,[14] multi-echo combination,[12] background field removal,[14] and, finally, the estimation of susceptibility maps.[13,15-17] Processing artifacts and inaccuracies at any of these five processing stages can propagate into the computed susceptibility maps.

The first QSM reconstruction challenge (RC1) in 2016[18] aimed to provide initial insights on the accuracy of various proposed algorithms for estimating susceptibility from background-corrected frequency maps (ie, the last processing step of QSM). One of the key conclusions of RC1 was that the choice of the algorithm and the used parameter settings can have a substantial, nonnegligible effect on the appearance and accuracy of computed susceptibility maps. However, following completion of the challenge, it was also recognized that the particular gold-standard (reference) susceptibility maps used for evaluating the challenge submissions limited the interpretability of the challenge outcomes. The reference maps were generated from multiple acquisitions in which the subject had rotated the head toward 12 different orientations. From these data, two reference maps were created: one calculated with the susceptibility tensor imaging[19] technique and one by calculation of susceptibility through multiple orientation sampling (COSMOS).[20] Meanwhile, only one of the 12 field maps was provided to the challenge participants. The rationale of this approach was that RC1 would yield the most objective and meaningful results if algorithms were evaluated using real-world in vivo data. However, at the completion of RC1, it was observed[21] that a nonnegligible discrepancy existed between the provided frequency map and the frequency map obtained when the field perturbation was forward-simulated based on the provided reference susceptibility maps. It was speculated that a part of the discrepancies was related to unaccounted microstructure effects on in vivo brain phase images.[22] Current single-orientation QSM algorithms assume that frequency contrast is caused entirely by variations in bulk magnetic susceptibility, and all other contrast mechanisms are neglected. Consequently, the discrepancy between the provided field map and the gold-standard susceptibility reference rendered it challenging or even impossible to achieve a reconstruction from the field map that was close to the reference used. It turned out that the best-performing RC1 submissions (ie, those with the smallest error metrics) were overregularized and had a nonnatural appearance.

The goal of the second reconstruction challenge (RC2) in 2019 was to address the identified limitations of RC1 and provide more meaningful insights on the current state-of-the-art in QSM algorithms, to identify their strengths and limitations in different scenarios and inform and coordinate future methodological research efforts. During the planning phase for RC2, the challenge committee concluded that the systematic evaluation of the accuracy and robustness of QSM methods should focus on synthetic (in silico) phantoms with realistic forward simulations rather than on real-world data. The challenge was designed with two stages: stage 1 mimicked the clinical setting in which the ground truth was unknown to participants; in stage 2 the ground truth was made available, and thus allowed for systematic parameter optimizations to obtain the *best possible* quality metrics that can be obtained with each reconstruction algorithm. The results of RC2 were reported in a separate manuscript.[23]

In this paper, we present the modular framework designed to generate the realistic digital head phantoms for RC2. Methodological researchers may use the RC2 phantom in their studies to evaluate existing and future QSM algorithms and

compare their results with RC2 submission. The comparison of their metrics with those of RC2 submissions will facilitate the objective evaluation of methodological improvements and algorithm performance across labs. As more advanced physical models are incorporated into the QSM algorithms, researchers may extend the phantom according to their needs. Code and data are freely available and have been designed to facilitate adding new features to the phantom, such as calcifications and hemorrhages or microstructure effects. The software package may also be used to optimize acquisition protocols, and prepare and test complete QSM reconstruction pipelines. In combination with other software, the package will allow us to evaluate the effect of image distortions or blurring on QSM.

2 | METHODS

2.1 | Design considerations

2.1.1 | Limitations of previous evaluation strategies

In the literature, most QSM algorithms were evaluated based on their visual appearance,[5,24,25] based on the RMS error (RMSE) of reconstructions of simple digital piece-wise constant phantoms consisting of geometrical shapes[25-30] or simplistic head phantoms.[22,31] Evaluation of the susceptibility quantification accuracy and precision typically relied on phantoms made of agar or aqueous solutions with varying concentrations of contrast agents such as gadolinium[20,26,32-35] or iron oxide particles.[5,28,36-38] Such measurements have been of great importance in establishing that QSM linearly maps the magnetic susceptibility property and that measurements across different platforms can be compared. In vivo, QSM accuracy has often been evaluated[27,31] by using previously published iron concentrations in the deep gray-matter nuclei[39] as a surrogate gold standard. This approach suffers from the large variability in iron concentrations across subjects.

A major limitation of most previously used digital phantoms and liquid or gel phantoms was that they had piece-wise constant susceptibility distributions, which are particularly easy to invert for methods with total variation regularization ("inverse crime"). To address this limitation, validation has also been performed by injecting gadolinium into tissue samples[20] or using air bubbles or glass beads.[40] In the first case experiments, however, the ground truth is again not available as the agents diffuse within the tissue. Therefore, it has to be reverted to visual inspection. In vivo, as an alternative to visual inspection, maps have been compared with a COSMOS reconstruction of the same subject,[28-30,41,42] due to their reduced level of streaking artifacts, similar as in RC1. However, using a COSMOS solution as gold standard implicitly assumes that the measured phase satisfies the COSMOS field model. Specifically, COSMOS assumes that (1) susceptibility is isotropic throughout the brain; (2) the dipole model with the sphere of Lorentz approximation can be used throughout the brain[22]; and (3) microstructure-related frequency effects[43] and chemical exchange[44,45] do not exist. Because these assumptions are simplistic, COSMOS does not generate an appropriate ground-truth susceptibility map for single-orientation QSM, and therefore cannot be considered a good gold-standard method.

2.1.2 | Design considerations for RC2

This RC2 committee's decision to use a digital phantom resulted from the realization that a true gold-standard technique for in vivo QSM did not exist. Without a gold-standard technique, it was not possible to obtain a meaningful in vivo reference susceptibility map through measurements. The committee had also discussed the design of real-world test objects (phantoms) that are consistent with the QSM phase model.[33,34] Based on the committee members' experience with phantom design and a literature research of previously used phantoms, it was concluded that the inclusion of sufficiently complex morphology and fine-scale susceptibility features would be prohibitively challenging. It was unanimously concluded that it would be most reasonable to focus the committee's efforts on a digital phantom that could be adapted and extended to the community's evolving needs in the future. For real-world data, available references usually represent only approximations of the ground truth (gold standard). On the contrary, in silico phantoms provide a genuine ground truth. In silico phantoms also allow for a controlled investigation of the effect of deviations from the underlying QSM model on the reconstruction performance. In addition to the ability to model different biophysical phase contributions, digital models also allow a controlled inclusion of measurement-related phase errors. For example, field measurements close to the brain surface are affected by nuisances, such as signal dropout and the nonlinearity of the phase evolution due to the nonnegligible higher-order spatial terms inside the pixel[46] that make the measured field deviate from the actual voxel-average field. Similar limitations are present when developing background field-removal methods. Despite this known limitation, only a few methods have the possibility of explicitly accounting for field-map uncertainty,[27,32] while remaining methods address this problem by increasing the brain mask erosion.[24,26,47]

As a first step toward a future systematic evaluation of all of these experimental aspects influencing QSM reconstruction quality, the RC2 in silico phantoms enforced consistency of the provided frequency map with the physical model used by current QSM algorithms. Moreover, the RC2 phantoms were

designed to feature a realistic brain morphology and naturally varying susceptibility distribution within anatomical regions.

2.2 | Data acquisition

We acquired MRI data from a human volunteer (female, 38 years old), who gave informed consent, and the experiment was approved by the local medical ethical committees (Amsterdam University Medical Center and Radboud University Medical Center). We used a 7T scanner to obtain relaxation-rate maps and a 3T scanner to obtain DTI data and bone–air tissue interfaces. To generate the brain phantom, we acquired inherently co-registered quantitative maps of R_1,[48] R_2^*, χ, and M_0 maps using the MP2RAGEME[49] sequence on a 7T (Philips Achieva; Amsterdam, Netherlands) scanner. The main sequence parameters were $TR/TI_1/TI_2 = 6.72/0.67/3.86$ seconds. The first and second TIs were acquired with $TE_1 = 3$ ms and $TE_{1/2/3/4} = 3/11.5/20/28.5$ ms and flip angles $\alpha_1/\alpha_2 = 7°/6°$, respectively. The acquisition was performed sagittally with FOV $= 205 \times 205 \times 164$ mm^3 and matrix size $= 320 \times 320 \times 256$, resulting in an isotropic resolution of 0.64 mm and a total acquisition time of 16:30 minutes.

To generate a bone and air model, we acquired T_1-weighted (0.93 mm isotropic) data at very short TE using the PETRA[50] sequence with the following parameters at 3 T (Siemens PrismaFit, Munich, Germany): $TR_1/TR_2/TI_1/TI_2 = 3/2250/1300/900$ ms; flip angle $= 6°$, TE $= 0.07$ ms; matrix size $= 320 \times 320 \times 320$; and total acquisition time $= 5:57$ minutes.

To add microstructure effects to the phantom, we acquired DTI data using two simultaneous multislice EPI-based data sets with opposed phase-encoding directions. The main sequence parameters were TR/TE $= 3520/74$ ms, simultaneous multislice factor $= 3$, in-plane acceleration $= 2$, matrix size $= 140 \times 140 \times 93$, and FOV $= 210 \times 210 \times 139.5$ mm^3, resulting in 1.5-mm isotropic image resolution. The diffusion-weighted parameters were b $= 0/1250/2500$ s/mm^2 and 12/90/90 directions, respectively, resulting in a total acquisition time of 12:10 minutes. Diffusion data were processed using FSL software (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/); eddy_correct and top up were used to undistort the DWI data. Data were coregistered to the 7T anatomical space, and the FMRIB Diffusion Toolbox was used to extract tensor information (eg, fractional anisotropy, main eigenvector orientation).

2.3 | Tissue segmentation

Figure 1 shows a pictorial representation of the pipeline used for the segmentation. The 7T T_1 maps, derived from the MP2RAGE data set, were segmented into 28 tissue classes, including left and right splitting, using the `cbstools` atlas-based pipeline (https://www.cbs.mpg.de/institute/software/cbs-tools[51] Classes were then reclustered into 16 tissue clusters: CSF (initially split into four classes); gray matter (initially split into eight classes, left and right cortical, cerebellar, amygdala, and hippocampus); caudate; putamen; thalamus; white matter (encephalus, cerebellum, and brain stem); and large blood vessels. Deep gray-matter structures not clearly distinguishable on T_1 maps (red nucleus, substantia nigra, globus pallidus, and dentate nucleus) were manually segmented using the active contours function implemented in ITK-Snap (version 3.6)[52] on R_2^* and χ maps. A calcification present in the subject's interhemispheric fissure was identified and segmented using the M_0 map. An initial vein-and-artery mask was computed based on Frangi-filtered R_2^* maps as described in Chan et al.[53] The region outside the brain was segmented into bone, air, and tissue using a model-based segmentation approach with deformable surface meshes,[54] using the PETRA sequence as input. The resulting segmentations of nasal cavities and auditory canals were refined manually using ITK-Snap for the computation of realistic background fields.

After the combination of the individual tissue brain masks into a piece-wise constant whole-head phantom, we used R_2^* and R_1 maps to correct the label boundaries in various brain regions using customized thresholds (Supporting Information Table S1). The initial tissue segmentations were based on one single quantitative parameter (R_1 for tissues compartments, R_2^* for veins), and because smooth surfaces had been enforced in some regions, this resulted in segmentation mismatches that benefited from this second processing iteration.

2.4 | Susceptibility map

The susceptibility map was simulated by assigning tissue-typical susceptibility values taken from literature,[55,56] $\overline{\chi}_{tissue}$ (Table 1), to the various tissue segments. We modulated the susceptibility values in each region using the image intensities on R_1 and R_2^* maps according to the following equation:

$$\chi(r)_{tissue} = \overline{\chi}_{tissue} + a_{tissue}\left(R_2^*(r) - \overline{R}_{2tissue}^*\right) + b_{tissue}\left(R_1(r) - \overline{R}_{1tissue}\right), \quad (1)$$

where $\overline{R}_{2tissue}^*$ and $\overline{R}_{1tissue}$ are the mean apparent transverse and longitudinal relaxation rates of that given tissue segment class. There were three main motivations for using such an expression to compute our ground-truth susceptibility map:

- Modulation-avoided susceptibility values were constant throughout anatomical regions (piece-wise constant).
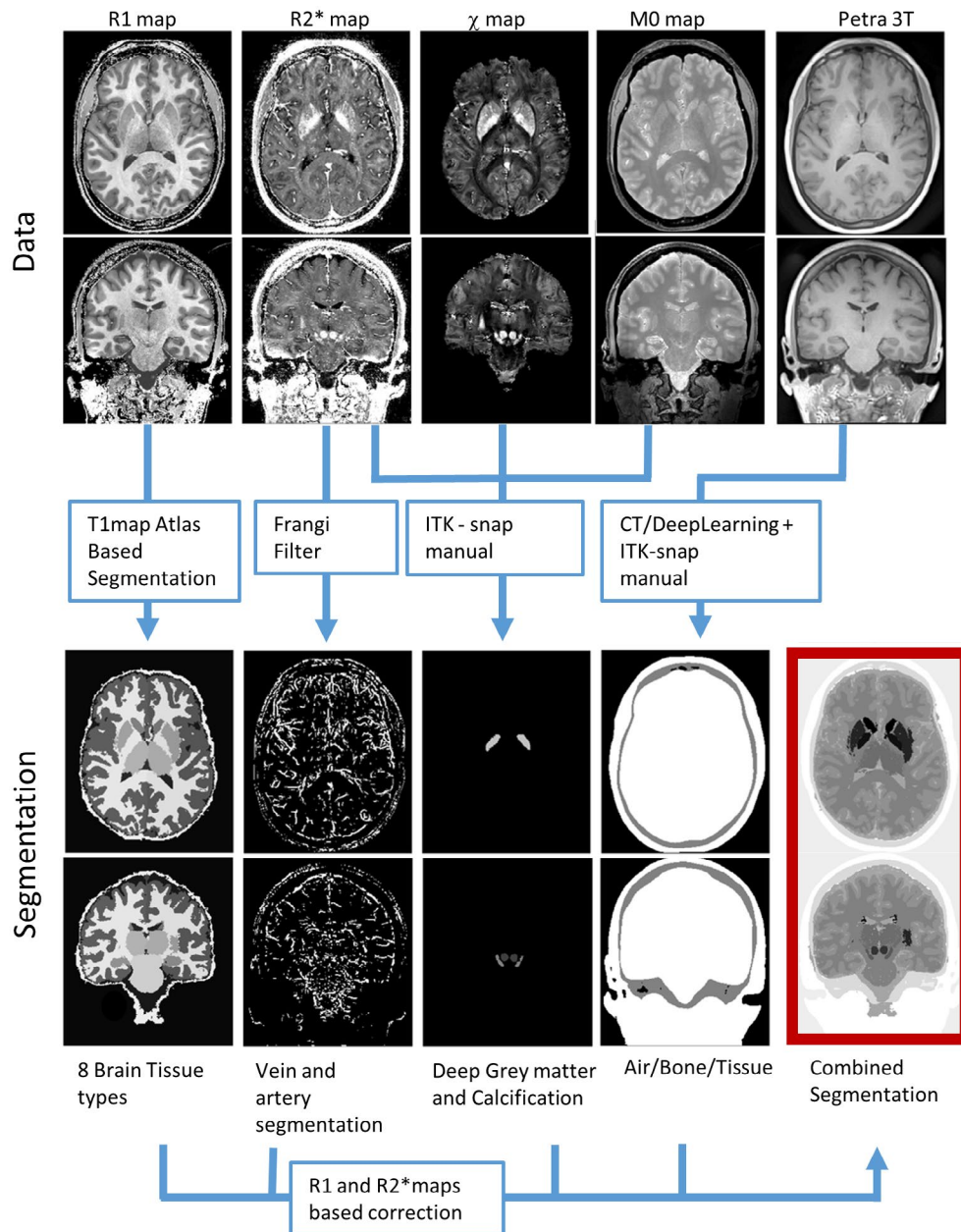
FIGURE 1 The process used to obtain the head segmentation: R_1 map (obtained from MP2RAGEME) was used to create an atlas-based segmentation using CBS tools; R_2^* were processed with a Frangi filter for vein segmentation; a semimanual approach using ITK snap was used for segmentation of the deep gray-matter nuclei based on the R_2^* and a susceptibility map computed using HEIDI; and the M_0 map was used to segment the calcification. Finally, PETRA data were used to obtain air, bone, and tissue masks using a CT-based deep-learning algorithm followed by manual ITK snap. Then, the various tissue segmentations were fine-tuned using denoised R_1 and R_2^* maps, using manually defined thresholds. The various masks were combined to generate a whole-head segmentation with 16 different tissue types

Absence of modulation would be both unrealistic and advantageous to algorithms with gradient-based regularization terms;

- From a practical perspective, R_1 and R_2^* were the only two "bias field" free maps available at high resolution that could be used to create an anatomically valid intensity modulation; and
- Both transverse and longitudinal relaxation rates, like magnetic susceptibility, are known to have a linear dependence

on the concentration of paramagnetic and diamagnetic perturbers when dealing with simple liquid solutions. The main difference to susceptibility is that relaxation rates are agnostic to the sign of the magnetic perturber; particularly in brain tissues, both R_1 and R_2^* have been shown to have a linear dependence on the concentrations of iron and myelin.[57]

Although it is reasonable to assume that the susceptibility map (assuming any other tissue properties constant) could be

TABLE 1 Parameters used to create the two magnetic-susceptibility head models released in the QSM challenge

Label name	Model 1			Model 2		
	Mean susceptibility (ppm)	a_{tissue} (ppm/Hz)	b_{tissue} (ppm/Hz)	Mean susceptibility (ppm)	a_{tissue} (ppm/Hz)	b_{tissue} (ppm/Hz)
Caudate	0.044[55,56]	−0,012	1.118	0. 044	−0.011	1.230
Globus pallidus	0.131[55]	−0.026	0.843	0.121	−0.023	0.927
Putamen	0.038[55]	−0.025	1.852	0.043	−0.022	2.038
Red nucleus	0.100[55]	−0.044	1.780	0.090	−0.040	1.958
Dentate nucleus	0.152[55]	−0.064	1.708	0.162	−0.058	1.879
Substantia nigra & subthalamic nucleus	0.111[55]	−0.075	1.491	0.121	−0.068	1.640
Thalamus	0.020[55]	−0.086	1.275	0.025	−0.078	1.402
White matter	−0.030[55]	−0.078	1.147	0.005	−0.070	1.262
Gray matter	0.020[22]	−0.095	1.402	0.020	−0.085	1.543
CSF	0.019[55]	−0.006	0.067	0.019	−0.006	0.073
Blood	0.190	−0.058	0.047	0.170	−0.052	0.052
Fat	0.019[56]	0.000	0.000	0.019	0.000	0.000
Bone	−2.100[56]	0.000	0.000	−2.100	0.000	0.000
Air	9.200[56]	0.000	0.000	9.200	0.000	0.000
Muscle	0.000	0.000	0.000	0.000	0.000	0.000
Calcification	−3.300[56]	−0.012	0.000	0.019	−0.011	0.000

Note: The values in the three columns correspond to the parameters described in Equation 1, the assumed mean magnetic susceptibility of the tissue, and the R_2^* (a_{tissue}) and R_1 (a_{tissue}) modulation terms, respectively. Model 1 values were chosen from literature and doing the fitting described in Equation 1. Model 2 mean susceptibility values were ad hoc modification of those found in literature, while the R_1 and R_2^* modulation were changed by plus and minus 10%, respectively.
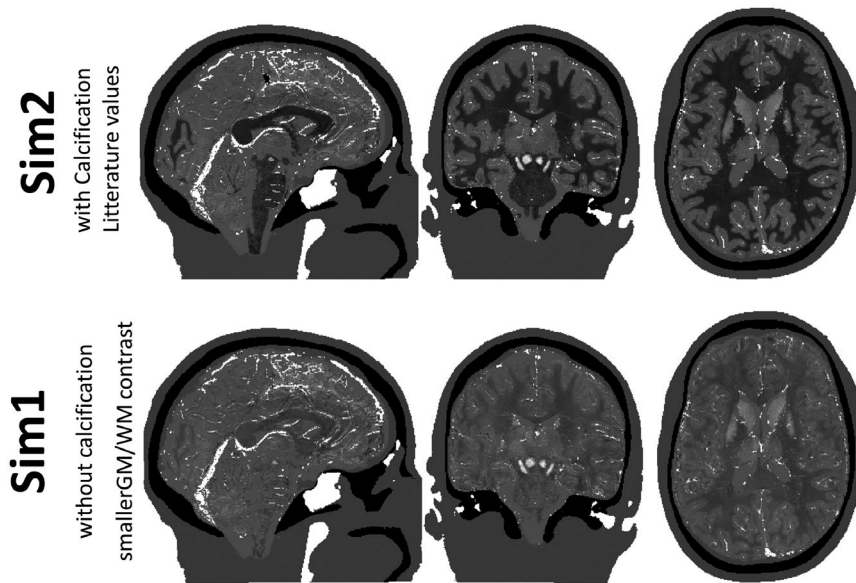

FIGURE 2 View of three slices in sagittal, coronal, and axial directions of the two digital phantoms created for the QSM challenge 2.0. Top and Bottom maps were obtained using the parameters described in Table 1 for models 1 and 2, respectively

given by a linear combination of these two maps, the main aim was to introduce a realistic texture. To obtain proportionality parameters, a_{tissue} and b_{tissue}, resulting in realistic susceptibility variations, Equation 1 was inverted for each brain tissue class using as $\chi(r)_{tissue}$, the HEIDI susceptibility map calculated from the original data. The coefficients $\overline{\chi}_{tissue}$, a_{tissue}, and b_{tissue} used for the two phantoms in the QSM challenge 2.0 are shown in Table 1 and Figure 2. Note that having different proportionality parameters for each tissue (in addition to a different mean value per tissue type) results in a susceptibility map that cannot be derived simply from the magnitude signal variations. Because the measured relaxation rates of blood (both in arteries and veins) are prone to errors (due to inflow effects on R_1 and flow effects on R_2^* maps),

the proportionality values were relatively small for the blood pool, rendering these compartments piece-wise constant. Bone, calcification, and other non-brain-tissue compartments were made piece-wise constant (the lack of a susceptibility map outside the brain prevented the derivations of a_{tissue} and b_{tissue}). Close to air–tissue boundaries, where strong field gradients are present, tissue R_2^* values are overestimated (see R_2^* maps in Figure 1 over the ear canals).[58] A low-pass-filtered version of the gradient of the acquired field map was used to differentiate regions where the R_2^* values could be trusted from those where they were unreliable. In the latter regions, we forced $a_{tissue} = 0$ (ignore R_2^* contribution when generating the susceptibility phantom). To avoid discontinuities between high (> 0.08 ppm/mm) and low (< 0.3 ppm/mm) field gradient regions, a smooth transition was created by mixing the two combinations (from full Equation 1 and $a_{tissue} = 0$, respectively). Please refer to the provided code for more details on the implementation.

To avoid unrealistically sharp edges of magnetic susceptibility at the interfaces between tissue regions, we have introduced partial-voluming in those transitions. Note that transitions between brain tissues tend not to be sharp (for example, gray-matter layers on the white-matter side are highly myelinated,[59] as are the outer parts of the thalamus), whereas between tissues and blood, CSF, air, bone and muscle, the interfaces will be sharp. The probability of a voxel being a given tissue, P_{tissue}, was computed by smoothing each binary brain tissue mask using a 3D Gaussian kernel with a FWHM of 1.2 voxels. This smoothing was not applied to veins or non-brain-tissue masks. The probability was computed as $P_{tissue}(r) = S_{tissue}(r) / \sum_{tissue=1}^{16} S_{tissue}(r)$, where S_{tissue} is either the smoothed or unsmoothed mask of a given compartment, depending on it being brain tissue or non-brain tissue. The susceptibility phantom was then given by

$$\chi(r) = \sum_{tissue=1}^{16} P_{tissue}(r) \, \chi(r)_{tissue}. \qquad (2)$$

The importance of moving from a piece-wise constant (where a_{tissue} and b_{tissue} are set to zero) to a contrast-modulated constant (where P_{tissue} is simply a binary mask) or the probabilistic formalism of the susceptibility distribution can be appreciated in Figure 3. Yellow arrows highlight the transition between gray and white matter that becomes smoother, and green arrows highlight smoothing out small segmentation errors within the thalamus. In contrast, Figure 3 shows that most vessel structures have remained sharp, with only some minor reduction in susceptibility value. The R_2^* maps tend to enlarge venous structures due to blooming artifacts. By not smoothing the blood compartment mask when computing the final susceptibility map, this effect was not further extended. However, because neighboring tissues have been smoothed into the blood compartment, the blood partial volume in blood vessels is reduced, resulting in a lower susceptibility the smaller the vessel is, mimicking a realistic scenario.
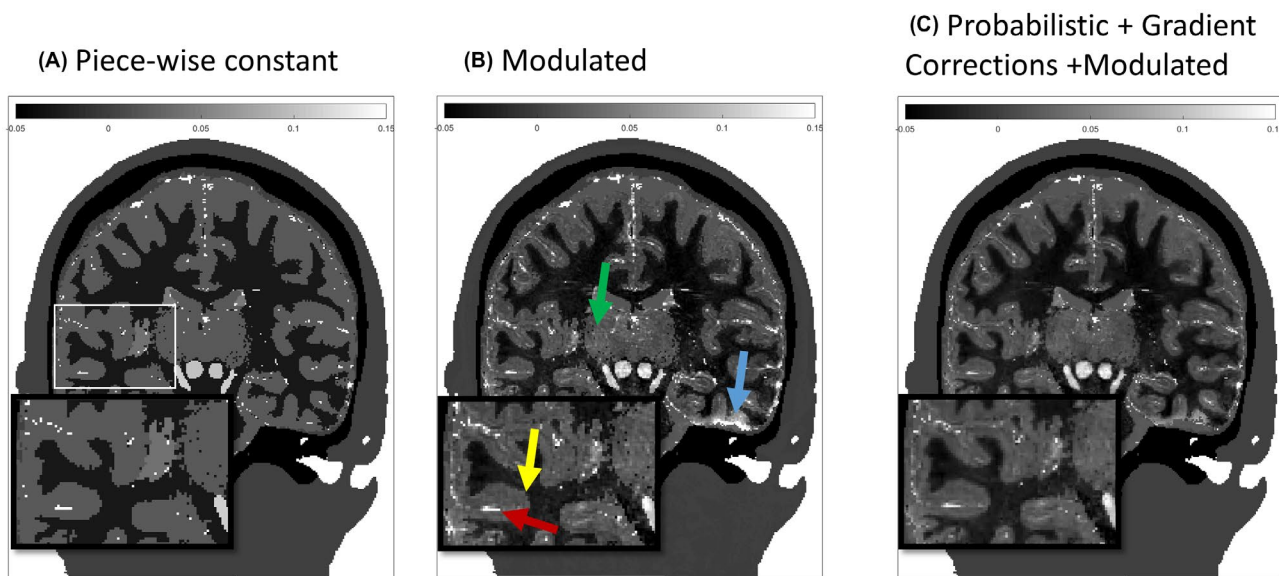

FIGURE 3 Intermediate stages of the of the creation of the in silico susceptibility head phantom: traditional piece-wise constant approach (A), modulated model as described in Equation 1 with the values presented in Table 1 (B), and finally adding the probabilistic modulation described in Equation 2 and masking regions of error bound R_2^* (C). Green and yellow arrows highlight transitions between tissue types that improved using the probabilistic approach applied to the tissue compartments. Note that the probabilistic smoothing was only applied to the brain tissues; as a result, veins retain shape in the susceptibility map only with a reduced magnetic susceptibility. Blue arrows highlight regions where the large field gradient masking approach was able to avoid abnormally large susceptibility values.

2.5 | Data simulation

Spoiled gradient-recalled-echo data can be simulated using the steady-state equation:

$$S = M_0(r) \sin(\alpha) \frac{\left(1 - e^{-TR \cdot R_1}\right)}{1 - \cos(\alpha) e^{-TR \cdot R_1}} e^{-TE \cdot R_2^* + i(\phi_0(r) + TE \cdot \Delta\omega_\chi)},$$
(3)

where ϕ_0 is an initial phase distribution originated from the transceiver phase, and $\Delta\omega_\chi$ is the frequency shift directly attributed to magnetic susceptibility. To simulate S for given TR, TE, and flip angles (α), we used the M_0, R_1, and R_2^* maps derived from the MP2RAGEME[49,60] sequence, where $\phi_0(r)$ is the TE = 0 phase (a manually selected 3D second-order polynomial was used, ensuring 2π phase variation inside the brain region), and the frequency shift, $\Delta\omega(r)$, was calculated according to

$$\Delta\omega(r) = D(r) \otimes \chi(r),$$
(4)

where $D(r)$ is the magnetic field dipole along the z-direction with Lorentzian correction. The convolution was performed in k-space using the formulation proposed in Marques and Bowtell[61] and Salomir et al.[62] To avoid aliasing artifacts associated with the discrete Fourier transform (circular convolution), which would appear as unrealistic background fields, we padded the phantom with zeros along each dimension (factor of 2) before evaluating Equation 4. Such a formulation of the signal equation explicitly neglects chemical shift (associated with spins from, for example, fat) and any chemical exchange effects on the frequency. Furthermore, because no gradient waveform was defined explicitly, image distortion effects were not simulated, and the impact of blood flow on phase data was not accounted for.

A digital phantom enables simulating the MR signal with and without background fields (fields generated by tissues and other sources located outside the brain). The latter effectively mimics "perfect" background-field correction. Using the whole-head susceptibility phantom allowed the creation of realistic background fields. To create a phantom without background fields, referred to as "local field" hereafter, all voxels outside the brain were set to zero, and the susceptibility distribution within the brain was demeaned as follows:

$$\begin{cases} \chi(r) = \chi(r) - \chi(\overset{!!}{Mask}) & \text{inside Mask} \\ \chi(r) = 0 & \text{outside Mask} \end{cases}.$$
(5)

Although not pursued for the QSM challenge purposes, field shimming was simulated by fitting the frequency map with second-order and third-order Legendre polynomials.

2.5.1 | Simulation of different acquisition protocols

For the demonstration of the acquisition protocol simulation with the code described in the Supporting Information, we chose two example protocols designed for different applications:

- (P1): Optimal for the observation of cortical gray/white-matter contrast in both the magnitude and phase data. In this case, the longest TE was chosen to be close to that of the T_2^* of cortical gray matter (33 ms).[63]; and
- (P2): Optimal for the quantification of deep gray-matter susceptibility. In this case, the longest TE should be at least that of the of the region with the highest iron concentration, which was the globus pallidus in the generated phantom (14 ms).

For the sake of simplicity, both protocols had the same echo spacing of 8 ms as the original volunteer data set. The TR of the acquisition was chosen as short as possible, assuming a readout acquisition window of 8 ms (P1: TR = 16 ms; P2: TR = 40 ms). We neglected dead times associated with phase-encoding preparation, flow compensation, rewinding, RF excitation, saturation, and crushers. The flip angle was chosen at the Ernst angle for the globus pallidus ($T_1 = 1100$ ms; $\alpha = 8$) and such that T_1-weighted contrast on magnitude images was maximized between white matter ($T_1 = 1100$ ms) and cortical gray matter ($T_1 = 1900$ ms; $\alpha = 23$) in protocols P1 and P2, respectively. This resulted in the following protocols:

- P1: $TE_1/TE_5 = 4/36$ ms; and
- P2: $TE_1/TE_2 = 4/12$ ms.

We mimicked k-space sampling by cropping the Fourier spectrum of the original 0.65-mm resolution data to an effective spatial resolution of 1 mm isotropic. We applied the same approach to down-sample the ground-truth susceptibility map. In the case of the susceptibility maps, the sharp edges between structures as well as the orders of magnitude–larger susceptibility differences between air/bone and tissue resulted in severe Gibbs ringing artifacts, which were removed using subvoxel shifts.[64] This step was repeated in all three spatial directions. Further processing consisted of spatial unwrapping of echo differences using SEGUE,[65] combination of resulting field maps using the optimum weights[12,41] and, when necessary, removal of background fields using the Laplacian boundary value method.[26]

2.5.2 | Quantitative susceptibility mapping reconstruction optimization

To demonstrate the applicability of the current framework for the QSM challenge or for QSM reconstruction optimization

purposes, we performed a simulation with only local fields (see Equation 5). The protocol used was that of the QSM challenge (TR = 50 ms; $TE_{1/2/3/4}$ = 4/12/20/28 ms; α = 15[23]). The QSM reconstructions using truncated k-space division,[66] closed form L2,[67] fast algorithm for nonlinear susceptibility inversion, FANSI[68] and iLSQR[69] as implemented in the SEPIA toolbox,[70] were performed with varying regularization parameters. The reconstructions were evaluated using the reconstructions metrics created for the challenge (Table 2). For a more detailed description, see Supporting Information Section 2.

2.5.3 | Adding microstructural effects to the obtained contrast

Microstructural effects are known to affect the observed phase. One of the driving factors of the microstructural effects is white-matter fiber orientation. The provided data and code include a simple first-order approximation of these microstructure effects, which is TE-independent. Wharton et al[22] demonstrated that the typical impact at 7 T for a protocol with TEs of 7 ms and 13 ms was given by

$$\Delta\omega = \Delta\omega_\chi - 5 * \left(\sin^2(\theta) - 2/3\right) FA_{norm} - 3, \quad (6)$$

where FA_{norm} is the fractional anisotropy divided by 0.59 (the average anisotropy observed in a human optic nerve), and θ is the angle between the white-matter fiber and the static magnetic field. Both of these quantities can be derived from the acquired diffusion data. Such a correction to the frequency shift was applied only within the segmented white-matter mask.

3 | RESULTS

3.1 | Simulations of different acquisition protocols

Figure 4 shows phantom 1 with the two different sequence parameters created by the proposed simulation toolbox (see Supporting Information Section 1 or data sharing collection for code). The top row shows an example slice of the simulated data from the P1 protocol aimed at computing QSM in cortical gray and white matter. In this case, the longer TE matched the T_2^* of that gray matter, resulting in both significant signal decay in deep gray matter and a large number of phase wraps close to tissue/air boundaries. The flip angle used (23°) was set to increase T_1 contrast in gray versus white-matter boundaries at short TEs, as can be clearly appreciated on the top-left figure. Such information can be used to inform the QSM algorithm regarding expected morphological features. The bottom row shows the images associated with the P2 protocol, aimed at measuring the susceptibility values in deep gray-matter regions. The TE range is smaller, so that the magnitude signal in the globus pallidus has not yet disappeared at the last TE, and the tissue contrast in the magnitude data is considerably weaker because of the smaller flip angle (8°). From the data shown in Figure 4, it can be expected that P1 will benefit more from a morphological informed

TABLE 2 Metrics provided with toolbox and challenge for optimization and evaluation of QSM reconstruction

Metric name	Description
nRMSE	Whole-brain RMSE after demeaning (ie, the subtraction of the mean within the mask)
rmse_detrend_Tissue	Normalized RMSE relative to ground truth (after demeaning and detrending) in gray/white-matter mask. Detrending was performed by compensating the estimated the slope of a linear fit of the reconstructed QSM voxel values against those of the ground truth in the tissue region of interest. The reconstruction was then divided by this factor to ensure that proportionality errors (measured by other metrics such as DeviationFromLinearSlope) do not affect the RMSE calculation
rmse_detrend_blood	RMSE relative to ground truth (after detrending) using a one-pixel dilated vein mask
rmse_detrend_DGM	RMSE relative to ground truth (after detrending) in a deep gray-matter mask (substantia nigra and subthalamic nucleus, red nucleus, dentate nucleus, putamen, globus pallidus, and caudate)
DeviationFromLinearSlope	Absolute difference between the slope of the average value of the six deep gray-matter regions versus the prescribed mean value and 1.0
CalcStreak	Estimation of the streaking artifact in a region of interest surrounding the calcification by means of the SD of the difference map between reconstruction and the ground truth. The region of interest was a hollow rectangular prism, with its inner boundary being two voxels away from the edge of the calcification and the outer boundary six voxels away from the inner boundary
CalcMoment	Volumetric susceptibility moment of the reconstructed calcification, compared with the ground truth (computed in the high-resolution phantom to be −49.8 ppm). This metric has been suggested to be more robust in regions of punctuated large susceptibility sources, where there is no signal in the region of the perturber[40]

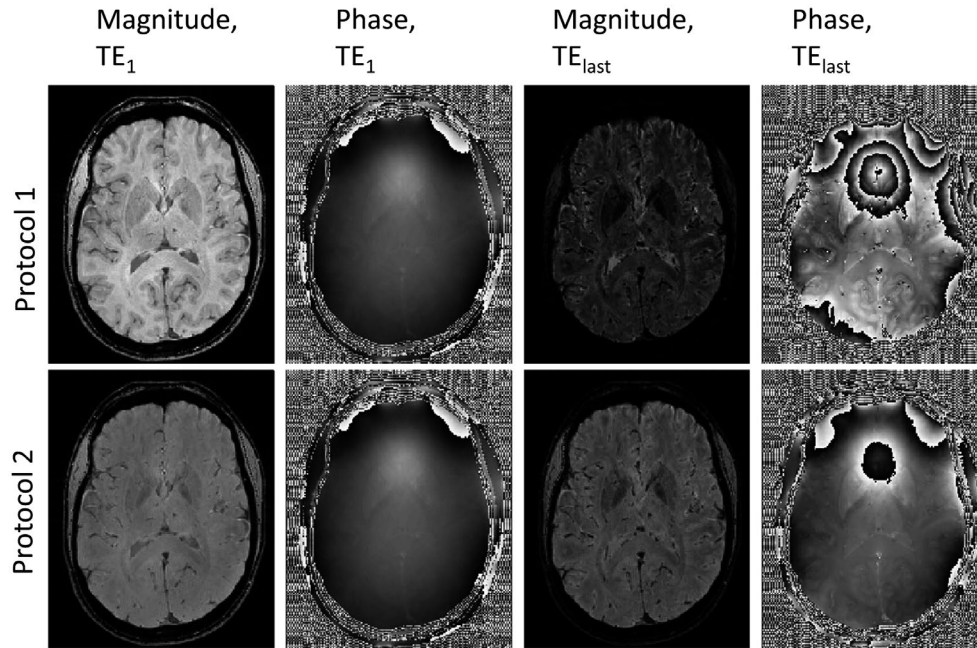*Note:* All of the RMSE metrics were multiplied by 100.

FIGURE 4 Transverse slices of the simulated data of model 1 using protocol 1 (top row) and protocol 2 (bottom row). First two columns show magnitude and phase images, respectively, at the first echo time (where the different T_1-weighting is clearly visible) and the last two columns show the magnitude and phase images associated with the last echo time of the respective protocol

reconstruction than P2, and that the use of nonlinear fit[71] for the calculation of the field map will also be particularly relevant, as noise will dominate the later echoes in P1.

Figure 5 compares the field map obtained with the P1 simulation from the whole-head phantom to a field map obtained directly from the susceptibility brain phantom using Equation 5 (ground-truth field map). When carrying out the signal simulation with the whole-head phantom, both the base resolution and the 1-mm resolution field maps are dominated by the background components arising from the air/bone/tissue interfaces, as can be seen in the transverse slice above the sphenoid sinus (first column, indicated by the black arrow). The differences with the ground-truth field map (second column) are to a large extent explained by the quadratic field used to represent the procedure of shimming. Once Laplacian boundary value was applied to the total field to obtain the tissue-specific field contributions (third column), the original resolution field map (top row) showed localized, smoothly varying differences relative to the ground truth, which have been described previously.[14] Both the base resolution and 1-mm resolution at first appear to have very similar properties, yet the normalized RMSE (nRMSE) from the down-sampled data set (bottom row) demonstrated additional deviations from the ground truth (nRMSE was 40% higher than that of the high-resolution data set). These discrepancies are caused both by incomplete background-field removal (see gray arrows) and errors around veins. In such regions, the field map measured from the gradient-echo data naturally deviates from the mean field in that pixel, because of the

reduced signal in veins (once partial volume is introduced by the reduced resolution, the field estimation is biased toward the tissue compartments).

To disentangle the effects of background field correction and MRI signal simulation on the deviations observed relative to the ground truth, we repeated the signal simulation with the local fields. In that case, the slowly varying smooth deviations disappeared and the high-resolution phantom did not demonstrate substantial deviations relative to the ground truth. The differences in the field computed at base resolution without background fields (top right panel) is at the numerical precision level, yet the nRMSE is still not negligible (nRMSE = 27) because of the errors present in the calcification region without signal and its immediate surrounding where spatial unwrapping fails.

To further investigate the sources of errors discussed previously, Figure 6 evaluates the phase evolution in three voxels: two in the surrounding of the calcification and one in the white matter. Figure 6C shows that, in the case of a homogeneous tissue region, there is a perfect match between low-resolution and high-resolution phase evolutions as well as the fitted frequency (based on the five-echo simulation) and the ground-truth frequency (computed from the susceptibility map). Figure 6B shows a region closer to the calcification; like in the case for the high-resolution data (light gray lines), in the case of the low-resolution data (dark gray) there is a larger error both with respect to fitted frequency (dashed line) and ground-truth frequency evolution. It is also clear that the phase evolution in the low-resolution data is no longer linear
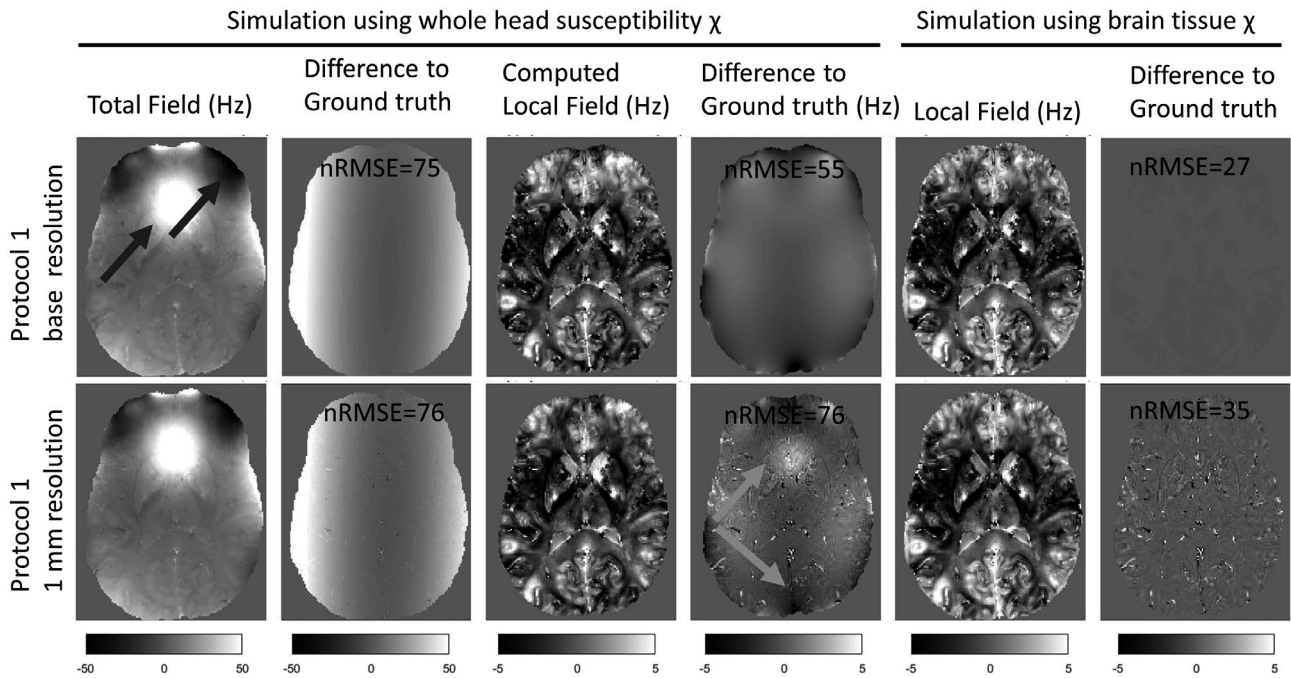
FIGURE 5 Transverse slices through derived field maps associated with protocol 1 data computed at the base resolution (top row) and after down-sampling to 1 mm (bottom row); black arrows highlight large background fields induced by air–tissue interfaces. The first and fifth columns show the field extracted from the complex signal when the whole-head and brain-only models were used, respectively, to compute the frequency shift. The third column shows the local field computed after background-field removal. The second, fourth, and sixth columns show the differences relative to the corresponding ground-truth field distributions. Gray arrows highlight the incomplete background-field removal that is exacerbated following down-sampling. The ground-truth field maps were computed using the forward dipole formulation (Equation 4) on the whole head (second column) and brain tissues alone (fourth and sixth columns) susceptibility models. It is clear that the normalized RMSE error (nRMSE), once the down-sampling is performed, is dominated by partial-volume effects in and around veins (noise pattern on the bottom of both the fourth and sixth columns) and imperfect background-field removal

(Figure 6B), as predicted from theory, due to partial-volume effects and the varying intravoxel frequency gradients.[46] In a pixel in the vicinity of the calcification, the errors are further enhanced for the high-resolution data, where unwrapping errors can introduce errors on the fitted frequency (the data are still fitted accordingly, but do not correspond to the ground-truth field). Figure 6E shows the mean squared difference map associated with the frequency fit on the low-resolution data; the errors are are predominantly found in regions of rapidly changing magnetic fields, around the calcification and close to tissue/air/bone interfaces and surrounding vessels.

3.2 | Evaluation of QSM reconstructions

Figure 7 shows the reconstructions with minimum nRMSE for the four algorithms tested. It appears that the direct methods (truncated k-space division and closed-form L2) still have some broad streaking artifacts in regions surrounding both the calcification and deep gray-matter regions, whereas in the iterative methods these were reduced. It is interesting to note that the total variation regularized nature of the FANSI clearly contributed to a better reconstruction of the

superior cerebellar vein when compared with the iLSQR (see the Supporting Information for the performance on remaining metrics and the QSM 2.0 challenge report, where these artifacts were addressed by more thoroughly optimized algorithms).[23]

3.3 | Comparison of simulations including microstructure to real data

Figure 8 shows an example of the originally acquired and the simulated magnitude and phase data, respectively. The magnitude data (first column) shows similar high-resolution features at the matched TE, despite the acquisition protocols not being identical (the simulations only support multi-echo gradient-echo acquisitions rather than MP2RAGEME, as used for data acquisition). It can be visually observed that the simulated data suffer from reduced bias field inhomogeneity; this is a result of a bias field correction applied to the computed M_0 map obtained from the MP2RAGEME. This choice was justified by two factors. First, the magnitude bias field observed after SENSE[72] reconstruction does not reflect the local SNR but a mix of the volume and surface
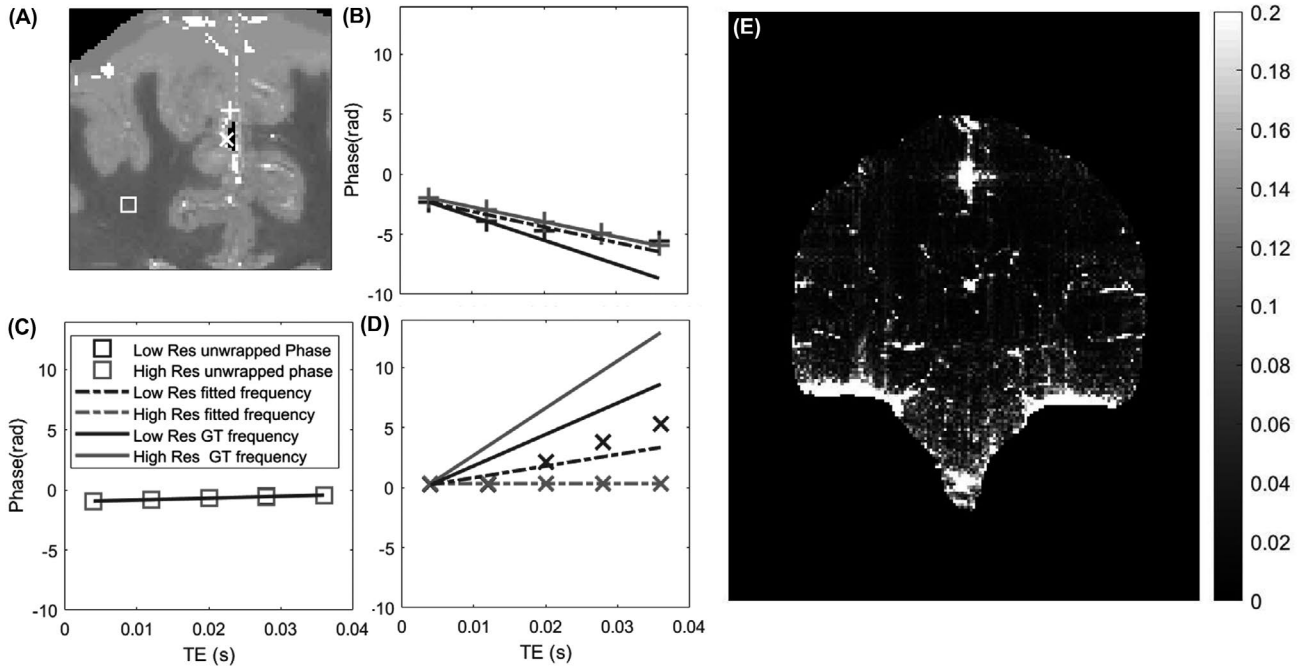
FIGURE 6 A, Coronal view of the susceptibility phantom with three locations highlighted (square, region in the middle of white matter; cross and plus, two regions close to the calcification). B-D, Plots of the unwrapped phase at the three locations (b, cross; c, square; and d, plus) as a function of TE. Unwrapped phase on the high-resolution data (light gray) and low-resolution data (dark gray) are shown using the respective markers, whereas dashed lines correspond to the fitted frequency for each point at each resolution, and the continuous line shows the ground-frequency evolutions. E, Mean squared difference map across TEs between fitted phase (dashed line in [B]-[D]) and measured phase (after unwrapping) on the coronal slice, highlighting tissue bone interfaces as well as regions surrounding the calcification. Abbreviation: GT, ground truth

coil sensitivities as well as the transmit coil inhomogeneities. Second, we wanted to separate the physical simulation from the interaction with the hardware. In the nonbackground field-phase data, the background field associated with nasal sinus and ear canals is weaker on the simulated data than on the measured data (as can be appreciated by the larger number of phase wraps on the latter). This can be attributed to two features: the imperfect bone/air segmentation or an underestimation of the susceptibility differences between tissue and bone or air. The latter is also supported by observing in the third column (after background field removal) that the field surrounding the calcification is smaller in the simulations than that what is observed in the experimental data.

When comparing gray/white-matter contrast on both the phase data (second column) and tissue frequency (third column), the simulation appears to visually approximate the acquired data better when the microstructural correction term (middle row) is added to the simulated data, as expressed in Equation 6. These data can now be used to test how different reconstruction pipelines are biased due to microstructural effects. Note that, because our susceptibility phantom is based on reported values of susceptibility rather than this particular subject susceptibility values, no quantitative evaluation of the similarity can be performed.

4 | DISCUSSION

In this paper and the accompanying shared data set and code (described in greater detail in the Supporting Information), we have presented and disseminated a realistic human brain phantom that can be used by both the QSM and the MRI communities to simulate multi-echo gradient-echo data and evaluate QSM pipelines in a controlled manner.

The data and code provided allow users to:

- Create new susceptibility phantoms with different levels of spatial modulation for each compartment; this can be performed by simply changing the values presented in Table 1 that control both the mean value and spatial modulation present within each tissue;
- Create realistic gradient-echo multi-echo data at 7 T and, to some extent, at other fields (it should be noted that, unlike susceptibility, relaxation times are field-dependent and their field dependence is tissue-dependent);
- Assess the effect of protocol changes (as well as nuisance factors such as RF phase, B_0 shimming, and noise) on the quality of the obtained QSM maps; and
- Assess the effect of changing some of the background field removal and QSM algorithm–specific options while having a ground truth to test it against.
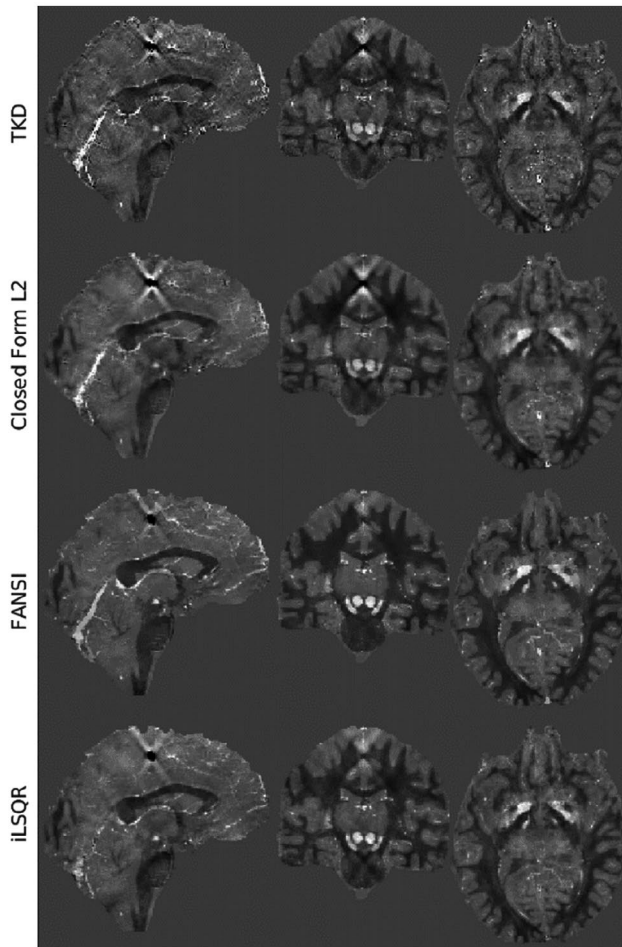
FIGURE 7 Examples of the optimum nRMSE reconstructions in three orthogonal planes obtained from the 1-mm data set based on SIM2 (see Figure 2) with peak SNR = 100 used for the QSM 2.0 challenge. The sagittal and coronal slices were chosen to cross the calcification region, to highlight remaining streaking artifacts. The four different rows correspond to the four different reconstruction pipelines tested. Abbreviations: FANSI, fast algorithm for nonlinear susceptibility inversion; iLSQR, improved least squares algorithm; TKD, truncated k-space division

This digital phantom will be important for QSM users and researchers deciding on acquisition protocols. Protocol considerations such as effect of the number of TEs and their range, as well as the degree of T_1 weighting and resolution on the ability to accurately measure QSM in a given brain region, can be quickly tested in this framework. This would allow us to extend the analysis done by Karsa et al[73] on the effects of FOV and anisotropic voxels sizes, or the analysis by Biondetti et al on the effects of Laplacian-based single echo versus multiple-echo techniques.[74]

Although the simulation framework is useful for optimizing protocols, the simulation in its current form is a static one. Consequently, flow artifacts (which build up when a large number of echoes are used), respiration-related

B_0 fluctuations,[75] and spatial distortions associated with the readout bandwidth are not considered. The latter two would be relatively straightforward to implement. Spatial distortions associated with different readouts can be obtained in a computationally efficient manner by using our provided phantom data in, for example, freely available software packages such as *JEMRIS* (http://www.jemris. org/).[76] Wave controlled aliasing in parallel imaging[77] and 3D EPI[78,79] have been used for QSM, but not much research has been done to quantify the effect of their blurring or distortions on the performance of either the background field removal or the susceptibility maps. Respiration artifacts would simply require a library of respiration fields over time, as acquired with field cameras.[80] Flow artifacts would be more complex to simulate, because the current vessel segmentation does not distinguish arteries and veins, and it would be difficult to have a local flow velocity and pulsatility estimation.

The dipole model used in QSM assumes a sphere of Lorentz approximation,[61] which does not hold true, particularly in white matter. The phantom released for RC2 purposes explicitly circumvented this limitation by ensuring perfect consistency with the QSM model in Equation 4 (ie, no microstructural effects were present). With the released phantom data set, we include diffusion data (both raw data on the 1.5-mm space and its derivatives co-registered to the phantom space after distortion correction). These data can be used to compute the frequency perturbation, as shown in Figure 8, or the hollow cylinder model can be used to explicitly introduce TE-varying perturbation, similarly to what has recently been done in the context of myelin water imaging,[22,81-84] and study the bias introduced by these effects on the reconstructed QSM maps. A critical challenge for more advanced modeling is the resolution of the diffusion acquisition, despite using state-of-the-art hardware and MR sequences. Here, we have simply interpolated our 1.5-mm DWI to the anatomical space and expect this to be sufficient to develop and validate QSM methods that account for microstructural effects in white matter.[85]

Quantitative susceptibility mapping is gaining interest in the context of neurological disorders such as multiple sclerosis, Parkinson's disease, and Alzheimer's disease, and other clinical applications such as hemorrhages or tumor imaging with iron oxide nanoparticles.[86] For the latter applications, relaxation and susceptibility values in the form of, for example, lesions in strategic locations can be added to the current phantom. Simulated data might also be relevant in the case of group studies of diseases, in which differences in the deep gray-matter nuclei were found[87,88] and the optimum QSM reconstruction parameters might improve the limits to detect those changes. Such a question can be addressed with the digital phantom by changing the parameters of Equation 1 for a
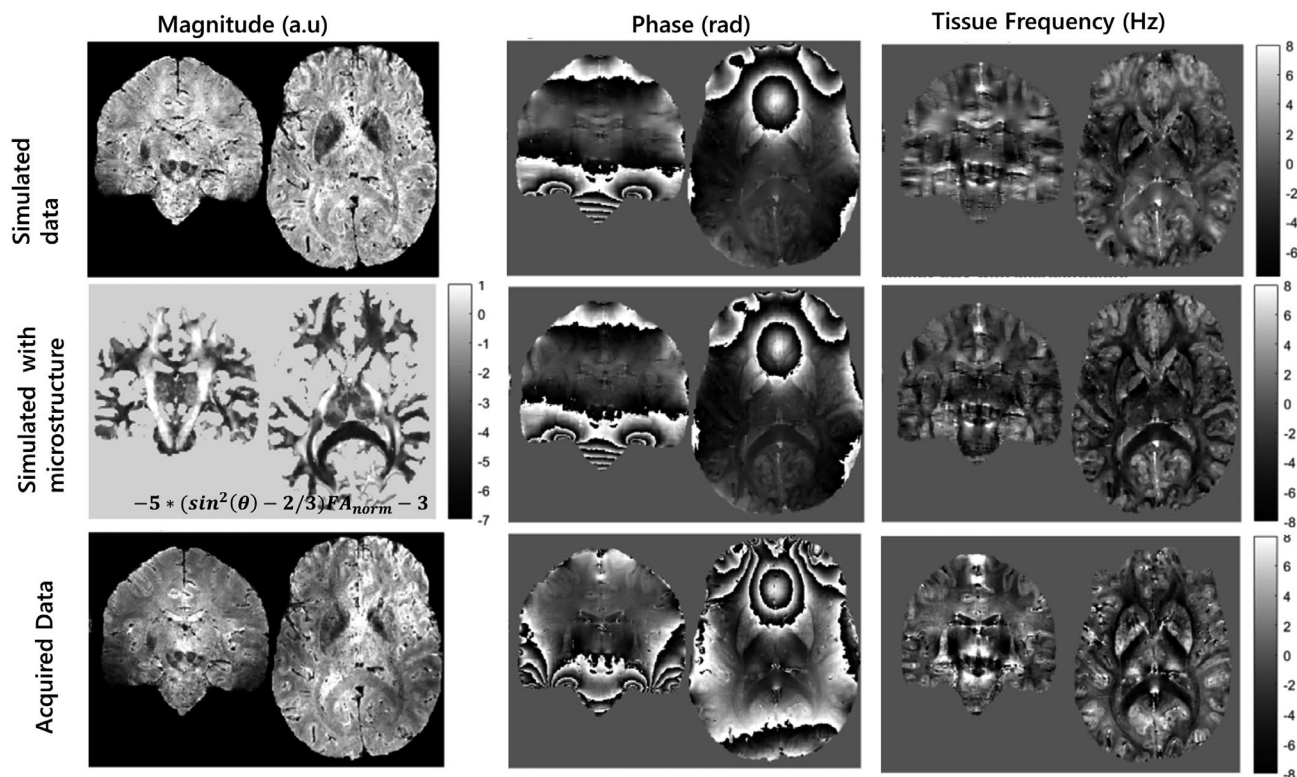
FIGURE 8 Effect of adding a microstructure correction term to the simulated field map. The first and second columns show the simulated magnitude and phase data using the whole-brain susceptibility model at TE = 20 ms, whereas the third column shows the computed tissue-frequency map (after brain masking and background-field removal). The first row shows the simulated data using protocol 1 without the addition of the microstructure effects. The second row shows the effect of adding the microstructural effects, as described in Equation 6 (first column), both on the third TE and the computed tissue-frequency map. The third row shows, for visual comparison purposes, the experimental data acquired at 7 T

given set of structures and find the QSM pipeline that better quantifies those changes.

5 | CONCLUSIONS

The presented realistic and modular phantom aims to enable researchers to optimize reconstruction as well as acquisition parameters. As such, the phantom served as a ground truth for the QSM RC2. Its modular design allows us to add microstructure effects a posteriori,[22] as well as include new nuisances such as hemorrhages or fine vessels with realistic relaxation and susceptibility properties. We foresee that this brain model will be an important tool for the evaluation of various processes associated with QSM processing and interpretation.

CONFLICT OF INTEREST
Jakob Meineke is an employee of Philips Research.

DATA AVAILABILITY STATEMENT
The code used to create the phantom, as well as to generate the various simulations and figures described in this paper can be found on the data sharing collection: https://doi.org/10.34973/m20r-jt17.

ORCID
José P. Marques https://orcid.org/0000-0001-8157-8864
Jakob Meineke https://orcid.org/0000-0001-8663-1468
Carlos Milovic https://orcid.org/0000-0002-1196-6703
Berkin Bilgic https://orcid.org/0000-0002-9080-7865
Kwok-Shing Chan https://orcid.org/0000-0001-8427-169X
Renaud Hedouin https://orcid.org/0000-0002-4341-9907
Wietske van der Zwaag https://orcid.org/0000-0003-3223-9721
Christian Langkammer https://orcid.org/0000-0002-7097-9707
Ferdinand Schweser https://orcid.org/0000-0003-0399-9211

REFERENCES
1. Langkammer C, Schweser F, Krebs N, et al. Quantitative suscep-tibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. *Neuroimage*. 2012;62:1593-1599.
2. Lotfipour AK, Wharton S, Schwarz ST, et al. High resolution mag-netic susceptibility mapping of the substantia nigra in Parkinson's disease. *J Magn Reson Imaging*. 2012;35:48-55.

3. Li W, Wu B, Batrachenko A, et al. Differential developmental trajectories of magnetic susceptibility in human brain gray and white matter over the lifespan. *Hum Brain Mapp*. 2014;35:2698-2713.

4. Fan AP, Bilgic B, Gagnon L, et al. Quantitative oxygenation venography from MRI phase. *Magn Reson Med*. 2014;72:149-159.

5. Liu S, Brisset J-C, Hu J, Haacke EM, Ge Y. Susceptibility weighted imaging and quantitative susceptibility mapping of the cerebral vasculature using ferumoxytol. *J Magn Reson Imaging*. 2018;47:621-633.

6. Deistung A, Schweser F, Wiestler B, et al. Quantitative susceptibility mapping differentiates between blood depositions and calcifications in patients with glioblastoma. *PLoS One*. 2013;8:e57924.

7. Chen W, Zhu W, Kovanlikaya IIhami, et al. Intracranial calcifications and hemorrhages: characterization with quantitative susceptibility mapping. *Radiology*. 2014;270:496-505.

8. Liu C, Li W, Johnson GA, Wu B. High-field (9.4 T) MRI of brain dysmyelination by quantitative mapping of magnetic susceptibility. *Neuroimage*. 2011;56:930-938.

9. Chen W, Gauthier SA, Gupta A, et al. Quantitative susceptibility mapping of multiple sclerosis lesions at various ages. *Radiology*. 2014;271:183-192.

10. Zhang Y, Gauthier SA, Gupta A, et al. Longitudinal change in magnetic susceptibility of new enhanced multiple sclerosis (MS) lesions measured on serial quantitative susceptibility mapping (QSM). *J Magn Reson Imaging*. 2016;44:426-432.

11. Li XU, Harrison DM, Liu H, et al. Magnetic susceptibility contrast variations in multiple sclerosis lesions. *J Magn Reson Imaging*. 2016;43:463-473.

12. Robinson SD, Bredies K, Khabipova D, Dymerska B, Marques JP, Schweser F. An illustrated comparison of processing methods for MR phase imaging and QSM: combining array coil signals and phase unwrapping. *NMR Biomed*. 2017;30:e3601.

13. Wang Y, Liu T. Quantitative susceptibility mapping (QSM): decoding MRI data for a tissue magnetic biomarker. *Magn Reson Med*. 2015;73:82-101.

14. Schweser F, Robinson SD, de Rochefort L, Li W, Bredies K. An illustrated comparison of processing methods for phase MRI and QSM: removal of background field contributions from sources outside the region of interest. *NMR Biomed*. 2017;30. https://doi.org/10.1002/nbm.3604

15. Kee Y, Liu Z, Zhou L, et al. Quantitative susceptibility mapping (QSM) algorithms: mathematical rationale and computational implementations. *IEEE Trans Biomed Eng*. 2017;64:2531-2545.

16. Jung W, Bollmann S, Lee J. Overview of quantitative susceptibility mapping using deep learning: current status, challenges and opportunities. *NMR Biomed*. 2020:e4292. https://doi.org/10.1002/nbm.4292

17. Haacke EM, Liu S, Buch S, Zheng W, Wu D, Ye Y. Quantitative susceptibility mapping: current status and future directions. *Magn Reson Imaging*. 2015;33:1-25.

18. Langkammer C, Schweser F, Shmueli K, et al. Quantitative susceptibility mapping: report from the 2016 reconstruction challenge. *Magn Reson Med*. 2018;79:1661-1673.

19. Liu C. Susceptibility tensor imaging. *Magn Reson Med*. 2010;63:1471-1477.

20. Liu T, Spincemaille P, de Rochefort L, Kressler B, Wang Y. 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*. 2009;61:196-204.

21. Milovic C, Tejos C, Acosta-Cabronero J, et al. The 2016 QSM challenge: lessons learned and considerations for a future challenge design. *Magn Reson Med*. 2020;84:1624-1637.

22. Wharton S, Bowtell R. Effects of white matter microstructure on phase and susceptibility maps. *Magn Reson Med*. 2015;73:1258-1269.

23. QSM Challenge 2.0 Organization Committee, Bilgic B, Langkammer C, et al. QSM reconstruction challenge 2.0: design and report of results. *bioRxiv*. 2020:2020.11.25.397695.

24. Li W, Wu B, Liu C. Quantitative susceptibility mapping of human brain reflects spatial variation in tissue composition. *Neuroimage*. 2011;55:1645-1656.

25. Shmueli K, de Zwart JA, van Gelderen P, Li T-Q, Dodd SJ, Duyn JH. Magnetic susceptibility mapping of brain tissue in vivo using MRI phase data. *Magn Reson Med*. 2009;62:1510-1522.

26. Zhou D, Liu T, Spincemaille P, Wang Y. Background field removal by solving the Laplacian boundary value problem. *NMR Biomed*. 2014;27:312-319.

27. Sun H, Wilman AH. Background field removal using spherical mean value filtering and Tikhonov regularization. *Magn Reson Med*. 2014;71:1151-1157.

28. Wharton S, Bowtell R. Whole-brain susceptibility mapping at high field: a comparison of multiple- and single-orientation methods. *Neuroimage*. 2010;53:515-525.

29. Schweser F, Deistung A, Sommer K, Reichenbach JR. Toward online reconstruction of quantitative susceptibility maps: superfast dipole inversion. *Magn Reson Med*. 2013;69:1582-1594.

30. Bollmann S, Rasmussen KGB, Kristensen M, et al. DeepQSM—using deep learning to solve the dipole inversion for quantitative susceptibility mapping. *Neuroimage*. 2019;195:373-383.

31. Schweser F, Sommer K, Deistung A, Reichenbach JR. Quantitative susceptibility mapping for investigating subtle susceptibility variations in the human brain. *Neuroimage*. 2012;62:2083-2100.

32. Liu T, Khalidov I, de Rochefort L, et al. A novel background field removal method for MRI using projection onto dipole fields (PDF). *NMR Biomed*. 2011;24:1129-1136.

33. Olsson E, Wirestam R, Lind E. MRI-based quantification of magnetic susceptibility in gel phantoms: assessment of measurement and calculation accuracy. *Radiol Res Pract*. https://www.hindawi.com/journals/rrp/2018/6709525/. Published July 30, 2018. Accessed December 11, 2020.

34. Deh K, Kawaji K, Bulk M, et al. Multicenter reproducibility of quantitative susceptibility mapping in a gadolinium phantom using MEDI+0 automatic zero referencing. *Magn Reson Med*. 2019;81:1229-1236.

35. Erdevig HE, Russek SE, Carnicka S, Stupic KF, Keenan KE. Accuracy of magnetic resonance based susceptibility measurements. *AIP Adv*. 2017;7:56718.

36. Mills PH, Wu Y-JL, Ho C, Ahrens ET. Sensitive and automated detection of iron-oxide-labeled cells using phase image cross-correlation analysis. *Magn Reson Imaging*. 2008;26:618-628.

37. de Rochefort L, Brown R, Prince MR, Wang Y. Quantitative MR susceptibility mapping using piece-wise constant regularized inversion of the magnetic field. *Magn Reson Med*. 2008;60:1003-1009.

38. Liu T, Spincemaille P, de Rochefort L, Wong R, Prince M, Wang Y. Unambiguous identification of superparamagnetic iron oxide particles through quantitative susceptibility mapping of the nonlinear response to magnetic fields. *Magn Reson Imaging*. 2010;28:1383-1389.

39. Hallgren B, Sourander P. The effect of age on the non-haemin iron in the human brain. *J Neurochem*. 1958;3:41-51.

40. Liu S, Neelavalli J, Cheng Y-CN, Tang J, Mark HE. Quantitative susceptibility mapping of small objects using volume constraints. *Magn Reson Med*. 2013;69:716-723.

41. Khabipova D, Wiaux Y, Gruetter R, Marques JP. A modulated closed form solution for quantitative susceptibility mapping—a thorough evaluation and comparison to iterative methods based on edge prior knowledge. *Neuroimage*. 2015;107:163-174.

42. Jung W, Yoon J, Ji S, et al. Exploring linearity of deep neural network trained QSM: QSMnet+. *Neuroimage*. 2020;211:116619.

43. Luo J, He X, Yablonskiy DA. Magnetic susceptibility induced white matter MR signal frequency shifts-experimental comparison between Lorentzian sphere and generalized Lorentzian approaches. *Magn Reson Med*. 2014;71:1251-1263.

44. Luo J, He X, d'Avignon DA, Ackerman JJH, Yablonskiy DA. Protein-induced water 1H MR frequency shifts: contributions from magnetic susceptibility and exchange effects. *J Magn Reson*. 2010;202:102-108.

45. Zhong K, Leupold J, von Elverfeldt D, Speck O. The molecular basis for gray and white matter contrast in phase imaging. *Neuroimage*. 2008;40:1561-1566.

46. Yablonskiy DA, Haacke EM. Theory of NMR signal behavior in magnetically inhomogeneous tissues: the static dephasing regime. *Magn Reson Med*. 1994;32:749-763.

47. Schweser F, Deistung A, Lehr BW, Reichenbach JR. Quantitative imaging of intrinsic magnetic tissue properties using MRI signal phase: an approach to in vivo brain iron metabolism? *Neuroimage*. 2011;54:2789-2807.

48. Marques JP, Kober T, Krueger G, van der Zwaag W, Van de Moortele P-F, Gruetter R. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *Neuroimage*. 2010;49:1271-1281.

49. Caan MWA, Bazin P-L, Marques JP, de Hollander G, Dumoulin SO, van der Zwaag W. MP2RAGEME: T1, T2*, and QSM mapping in one sequence at 7 Tesla. *Hum Brain Mapp*. 2019;40:1786-1798.

50. Grodzki DM, Jakob PM, Heismann B. Ultrashort echo time imaging using pointwise encoding time reduction with radial acquisition (PETRA). *Magn Reson Med*. 2012;67:510-518.

51. Bazin P-L, Weiss M, Dinse J, Schäfer A, Trampel R, Turner R. A computational framework for ultra-high resolution cortical segmentation at 7 Tesla. *Neuroimage*. 2014;93(Pt 2):201-209.

52. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*. 2006;31:1116-1128.

53. Chan K-S, Norris DG, Marques JP. Structure tensor informed fibre tractography at 3T. *Hum Brain Mapp*. 2018;39:4440-4451.

54. Wenzel F, Meyer C, Stehle T, et al. Rapid fully automatic segmentation of subcortical brain structures by shape-constrained surface adaptation. *Med Image Anal*. 2018;46:146-161.

55. Deistung A, Schäfer A, Schweser F, Biedermann U, Turner R, Reichenbach JR. Toward in vivo histology: a comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and R2*-imaging at ultra-high magnetic field strength. *Neuroimage*. 2013;65:299-314.

56. Buch S, Liu S, Ye Y, Cheng Y-CN, Neelavalli J, Haacke EM. Susceptibility mapping of air, bone, and calcium in the head. *Magn Reson Med*. 2015;73:2185-2194.

57. Stüber C, Morawski M, Schäfer A, et al. Myelin and iron concentration in the human brain: a quantitative study of MRI contrast. *Neuroimage*. 2014;93(Pt 1):95-106.

58. Wen J, Cross AH, Yablonskiy DA. On the role of physiological fluctuations in quantitative gradient echo MRI: implications for GEPCI, QSM, and SWI. *Magn Reson Med*. 2015;73:195-203.

59. Marques JP, Khabipova D, Gruetter R. Studying cyto and myeloarchitecture of the human cortex at ultra-high field with quantitative imaging: R1, R2* and susceptibility. *Neuroimage*. 2017;147:152-163.

60. Marques JP. *JosePMarques/MP2RAGE-related-scripts*. 2019. https://github.com/JosePMarques/MP2RAGE-related-scripts. Accessed September 30, 2018.

61. Marques JP, Bowtell R. Application of a Fourier-based method for rapid calculation of field inhomogeneity due to spatial variation of magnetic susceptibility. *Concepts Magn Reson Part B Magn Reson Eng*. 2005;25:65-78.

62. Salomir R, de Senneville BD, Moonen CT. A fast calculation method for magnetic field inhomogeneity due to an arbitrary distribution of bulk susceptibility. *Concepts Magn Reson Part B Magn Reson Eng*. 2003;19B:26-34.

63. Marques JP, Norris DG. How to choose the right MR sequence for your research question at 7T and above? *Neuroimage*. 2018;168:119-140.

64. Kellner E, Dhital B, Kiselev VG, Reisert M. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn Reson Med*. 2016;76:1574-1581.

65. Karsa A, Shmueli K. SEGUE: a speedy region-growing algorithm for unwrapping estimated phase. *IEEE Trans Med Imaging*. 2019;38:1347-1357.

66. Wharton S, Schäfer A, Bowtell R. Susceptibility mapping in the human brain using threshold-based k-space division. *Magn Reson Med*. 2010;63:1292-1304.

67. Bilgic B, Chatnuntawech I, Fan AP, et al. Fast image reconstruction with L2-regularization. *J Magn Reson Imaging*. 2014;40:181-191.

68. Milovic C, Bilgic B, Zhao B, Acosta-Cabronero J, Tejos C. Fast nonlinear susceptibility inversion with variational regularization. *Magn Reson Med*. 2018;80:814-821.

69. Li W, Wang N, Yu F, et al. A method for estimating and removing streaking artifacts in quantitative susceptibility mapping. *Neuroimage*. 2015;108:111-122.

70. Chan K-S, Marques JP. SEPIA—susceptibility mapping pipeline tool for phase images. *Neuroimage*. 2020:117611.

71. Liu T, Wisnieff C, Lou M, Chen W, Spincemaille P, Wang Y. Nonlinear formulation of the magnetic field to source relationship for robust quantitative susceptibility mapping. *Magn Reson Med*. 2013;69:467-476.

72. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med*. 1999;42:952-962.

73. Karsa A, Punwani S, Shmueli K. The effect of low resolution and coverage on the accuracy of susceptibility mapping. *Magn Reson Med*. 2019;81:1833-1848.

74. Biondetti E, Karsa A, Thomas DL, Shmueli K. Investigating the accuracy and precision of TE-dependent versus multi-echo QSM using Laplacian-based methods at 3 T. *Magn Reson Med*. 2020;84:3040-3053.

75. Van de Moortele P-F, Pfeuffer J, Glover GH, Ugurbil K, Hu X. Respiration-induced B0 fluctuations and their spatial distribution in the human brain at 7 Tesla. *Magn Reson Med*. 2002;47:888-895.

76. Stöcker T, Vahedipour K, Pflugfelder D, Shah NJ. High-performance computing MRI simulations. *Magn Reson Med*. 2010;64:186-193.

77. Bilgic B, Gagoski BA, Cauley SF, et al. Wave-CAIPI for highly accelerated 3D imaging. *Magn Reson Med*. 2015;73:2152-2162.

78. Langkammer C, Bredies K, Poser BA, et al. Fast quantitative susceptibility mapping using 3D EPI and total generalized variation. *Neuroimage*. 2015;111:622-630.

79. Wicaksono KP, Fushimi Y, Nakajima S, et al. Two-minute quantitative susceptibility mapping from three-dimensional echo-planar imaging: accuracy, reliability, and detection performance in patients with cerebral microbleeds. *Invest Radiol*. 2021;56:69-77.

80. Vannesjo SJ, Wilm BJ, Duerst Y, et al. Retrospective correction of physiological field fluctuations in high-field brain MRI using concurrent field monitoring. *Magn Reson Med*. 2015;73:1833-1843.

81. Sati P, Silva AC, van Gelderen P, et al. In vivo quantification of T2* anisotropy in white matter fibers in marmoset monkeys. *Neuroimage*. 2012;59:979-985.

82. Wharton S, Bowtell R. Fiber orientation-dependent white matter contrast in gradient echo MRI. *Proc Natl Acad Sci USA*. 2012;109:18559-18564.

83. Nam Y, Lee J, Hwang D, Kim D-H. Improved estimation of myelin water fraction using complex model fitting. *Neuroimage*. 2015;116:214-221.

84. Chan K-S, Marques JP. Multi-compartment relaxometry and diffusion informed myelin water imaging—promises and challenges of new gradient echo myelin water imaging methods. *Neuroimage*. 2020;117159.

85. Schweser F, Zivadinov R. Quantitative susceptibility mapping (QSM) with an extended physical model for MRI frequency contrast in the brain: a proof-of-concept of quantitative susceptibility and residual (QUASAR) mapping. *NMR Biomed*. 2018;31:e3999.

86. Deh K, Zaman M, Vedvyas Y, et al. Validation of MRI quantitative susceptibility mapping of superparamagnetic iron oxide nanoparticles for hyperthermia applications in live subjects. *Sci Rep*. 2020;10:1171.

87. Acosta-Cabronero J, Williams GB, Cardenas-Blanco A, Arnold RJ, Lupson V, Nestor PJ. In vivo quantitative susceptibility mapping (QSM) in Alzheimer's disease. *PLoS One*. 2013;8:e81093.

88. Damulina A, Pirpamer L, Soellradl M, et al. Cross-sectional and longitudinal assessment of brain iron level in Alzheimer disease using 3-T MRI. *Radiology*. 2020;296:619-626.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

FIGURE S1 A,B,D,E, Normalized RMS error (nRMSE) as a function of the regularization value. A,B, Numerical performance of four different QSM algorithms when the input data were the high-resolution simulated data with no additional noise or the down-sampled 1-mm data with added noise, resulting in an image peak SNR of 100. D,E, Performance of the improved least squares algorithm (iLSQR) and fast algorithm for nonlinear susceptibility inversion (FANSI) algorithms, respectively, for different types of input data (high-resolution and 1-mm data shown in solid and dashed lines, respectively); without and with added noise are shown in green and red. C, Examples of the optimum nRMSE reconstructions of the 1-mm data set with peak SNR = 100 using the four reconstruction algorithms tested in the manuscript

FIGURE S2 A-C, Plots of the various metrics as a function of the used regularization value for the iLSQR (continuous line) and FANSI (dashed line) algorithms. The different plots show nRMSE reconstructions metrics (A), calcification metric (B), and linearity of the susceptibility values (C) measured in deep gray matter with respect to ground truth

TABLE S1 Ad hoc segmentation correction based on relaxometry values. Note: These corrections were only applied in regions where the magnetic field gradient, as computed from the multi-echo data, was not expected to corrupt the R_2^* values