QSM Reconstruction Challenge 2.0–Part 1: 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], Wietske van der Zwaag[9], Christian Langkammer[10] , Ferdinand Schweser[11,12]

1 - Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, 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, UK;
6 - Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA
7 - Department of Radiology, Harvard Medical School, Boston, MA, USA
8 - Harvard-MIT Health Sciences and Technology, MIT, Cambridge, MA, USA
9 - Spinoza Centre for Neuroimaging, Amsterdam, Netherlands.
10 - Department of Neurology, Medical University of Graz, Graz, Austria
11 - Buffalo Neuroimaging Analysis Center, Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, The State University of New York, Buffalo, NY, USA
12 - Center for Biomedical Imaging, Clinical and Translational Science Institute, University at Buffalo, The State University of New York, Buffalo, NY, USA

Running head:

Correspondence address:

José P. Marques
Address:
Kapittelweg 29
6525 EN Nijmegen,
The Netherlands
-Tel/Fax number: +31 024-3668491
Email:j.marques@donders.ru.nl

Abstract

Purpose: To create a realistic *in-silico* head phantom for the second QSM Reconstruction Challenge and for future evaluations of processing algorithms for Quantitative Susceptibility Mapping (QSM).

Methods: We created a whole-head tissue property model by segmenting and post-processing high-resolution, multi-parametric MRI data acquired from a healthy volunteer. We simulated the steady-state magnetization using a Bloch simulator and mimicked a Cartesian sampling scheme through Fourier-based post-processing. We demonstrated some of the phantom's properties, including the possibility of generating phase data that do not evolve linearly with echo time due to partial volume effects or complex distributions of frequency shifts within the voxel. Computer code for generating the phantom and performing the MR simulation was designed to facilitate flexible modifications of the model, such as the inclusion of pathologies, as well as the simulation of a wide range of acquisition protocols.

Results: The brain-part of the phantom features realistic morphology combined with realistic spatial variations in relaxation and susceptibility values. Simulation code allows adjusting the following parameters and effects: repetition time and echo time, voxel size, background fields, and RF phase biases. Additionally, diffusion weighted imaging data of the phantom is provided allowing future investigations of tissue microstructure effects in phase and QSM algorithms.

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 MRI reconstruction algorithms.

1 Introduction

Quantitative Susceptibility Mapping (QSM) has proven to be a valuable tool for assessing iron concentrations in the deep gray matter (DGM)(1), estimating vessel oxygenation (2), differentiating blood and calcium products (3), and studying demyelinating lesions in the white matter (WM). However, several recent methodical studies have suggested that study outcomes may be affected by the particular processing algorithms chosen for QSM (4–6). QSM typically involves the following steps in the order of application: coil combination (4); phase unwrapping (6); multi-echo combination (4); background field removal (6); and finally the estimation of susceptibility maps (5,7–9). Processing artifacts and inaccuracies at any of these 5 processing stages propagate into the computed susceptibility maps.

The first QSM Reconstruction Challenge (RC1) in 2016 (10) aimed to provide insights on the accuracy of the last processing step: the estimation of susceptibility maps from background-corrected frequency maps. One of the key conclusions of the challenge was that the choice of the algorithm and the used parameter settings can have a significant effect on the appearance and accuracy of computed susceptibility maps. However, the first challenge also raised questions about the gold-standard (reference) susceptibility map used for evaluating the challenge submissions. The reference susceptibility maps for RC1 were generated from multiple acquisitions in which the subject had rotated the head towards 12 different orientations. Specifically, two reference maps were used, one calculated with susceptibility tensor imaging (STI) (11) technique and one with the Calculation Of Susceptibility through Multiple Orientation Sampling (COSMOS) (12) technique. Meanwhile, only one field-map obtained from those twelve orientation acquisitions was provided to challenge participants. After completion of the challenge, it was observed (13) that a non-negligible discrepancy existed between the provided frequency map and the frequency map obtained when forward-simulating the field perturbations using the provided reference susceptibility maps. Part of the discrepancies could be explained by the presence of unaccounted microstructure effects on in vivo brain phase images (14). Since current single-orientation QSM algorithms assume that frequency contrast is caused entirely by variations in bulk magnetic susceptibility, their solution did not match the multiple orientation methods, which resulted in the best performing challenge submissions being over-regularized and having a non-natural appearance.

During the planning phase for the second reconstruction challenge (RC2) in 2019, we concluded that synthesized phantoms and realistic forward simulations are needed to evaluate the accuracy and robustness of QSM methods. Using *in silico* models as a basis for benchmarking instead of real data allows comparing the reconstructed maps to the ground truth. The goal of the RC2 was to provide insights on the current state of the art in QSM algorithms and identify their strengths and limitations in different scenarios. The challenge used the numerical phantoms presented in this paper and consisted of two stages: Stage 1 mimicked the clinical setting in which the ground-truth was unknown; in Stage 2 the ground-truth was available and thus allowed for a systematic parameter optimisations for the *best possible* quality metrics that can be obtained with a specific reconstruction algorithm. The challenge results are reported in a separate manuscript

This *in silico* model was developed to address several specific demands from the community developing QSM methods. Most QSM algorithms were either demonstrated based on their visual appearance, based on the RMSE using simple digital piece-wise constant phantoms (either simple shapes such cylinders or head models), by comparison to phantom data using varying quantities of contrast agents (15), or comparing the obtained *in vivo* reconstructions to that obtained by a ground-truth COSMOS reconstruction (16,17). These approaches have severe limitations, for example: digital phantoms and liquid or gel phantoms share property of having piece-wise constant susceptibility distributions which tend to favor methods based on total variation. When using a COSMOS solution as ground-truth, despite their reduced level of streaking artifacts, one is assuming that the dipole model with the sphere of Lorentz approximation used in the QSM formulation is valid through-out the brain (14). Moreover, digital models allow a controlled investigation of the effect of deviations from the underlying QSM model on the reconstruction performance, such as the inclusion of microstructure-related frequency effects. Similar limitations are present when developing background field removal methods. Most methods do not consider the degree of accuracy of the field measurements close to the edges that are affected by various factors such as signal drop-out and the non-linearity of the phase evolution due to the non-negligible higher order spatial terms inside the pixel (18). All the above aspects can be addressed using the a realistic digital phantom.

This paper presents a framework for generating realistic digital head phantoms used for RC2. The phantom features a realistic brain morphology and naturally varying susceptibility distribution within anatomical regions. As a first step toward systematically evaluating all experimental aspects that influence QSM reconstruction quality, the presented *in-silico* phantom enables enforcing consistency of the provided frequency map with the physical model underlying current QSM algorithms. Code and data are made freely available and the authors encourage researchers to use the phantom to evaluate their processing algorithms in the future. The modular design of the code facilitates adding new features to the phantom, such as calcifications and hemorrhages provided that relaxometry maps exist. Furthermore the software package may be used to optimize acquisition protocols, prepare and test QSM reconstruction pipelines, and to study the impact of effects that deviate from the QSM standard formulation (such as tissue microstructure and signal voids); or to simulate the effect of the extended readouts.

2 Methods

2.1 Data acquisition

We acquired MRI data from a human volunteer (F, 38 years) who gave informed consent, and the experiment was approved by the local Medical Ethical Committees (Amsterdam University Medical Centre and Radboud University Medical Centre). We used a 7T scanner to obtain relaxation rate maps and a 3T scanner to obtain DTI data and bone-air tissue interfaces. For generating the brain model, we acquired inherently co-registered quantitative maps of R_1 (19), R_2^*, χ and M_0 maps using the MP2RAGEME (20) sequence on a 7T (Philips Achieva) scanner. The main sequence parameters were $TR/TI_1/TI_2$=6.72/0.67/3.86s. The first and second inversion time were acquired with TE_1=3ms and $TE_{1/2/3/4}$=3/11.5/20/28.5ms and flip angles α_1/α_2 = 7/6 °, respectively. The acquisition was performed sagittally with FOV=205x205x164 mm^3 and matrix size = 320x320x256, resulting in an isotropic resolution of 0.64mm. A SENSE acceleration factor of 1.7 in both phase encoding directions resulted in a total acquisition time of 16min 30secs.

For generating a bone and air model, we acquired T_1w (0.93 mm isotropic) data at very short echo time using the PETRA (21) sequence with the following parameters at 3T (Siemens PrismaFIT): $TR_1/TR_2/TI_1/TI_2$=3/2250/1300/900 ms; flip angle=6°, TE=0.07 ms; matrix size = 320x320x320; total acquisition time 5min 57secs.

To enable additional microstructure effects to the model, we acquired DTI data using two simultaneous multi-slice EPI-based datasets with opposed phase encoding directions. The main sequence parameters were TR/TE=3520/74 ms, SMS factor 3, in-plane acceleration 2, matrix size = 140x140x93 and FOV=210x210x139.5 mm, resulting in 1.5mm isotropic image resolution. The diffusion-weighted parameters were b=0/1250/2500 s/mm² and 12/90/90 directions, respectively, resulting in a total acquisition time of 12 min 10 secs.

2.2 Tissue segmentation

Figure 1 shows a pictorial representation of the pipeline used for the segmentation. The 7T T_1 maps, derived from the MP2RAGE dataset, 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(22)). Classes were then re-clustered into 16 tissue clusters: CSF (initially split into 4 classes); grey matter (initially split into 8 classes, left and right cortical, cerebellar, amygdala and hippocampus); caudate; putamen; thalamus; white matter (encephalus, cerebellum and brain stem); and large blood vessels. Deep grey 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 ITKSnap (version 3.6) (23) on R_2* and χ maps. Subject's interhemispheric calcification 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 (24). The region outside the brain was segmented into bone, air and tissue using a model-based segmentation approach with deformable surface meshes (25) using the PETRA sequence as input. The resulting segmentations of nasal cavities and auditory canals were refined manually using ITKSnap for the computation of realistic background fields (see below).

After the combination of the individual tissue brain masks into a piece-wise constant whole-head model, we used R_2* and R_1 maps to correct the label boundaries in various brain regions

using customized thresholds (see supplementary Table 1). Note that the previous tissue segmentations were based on one single quantitative parameter (either R1 or R2*), and because smooth surfaces were enforced in some regions this resulted, segmentation overestimations that benefited from this second iteration.

2.3 Susceptibility map

The susceptibility map was simulated by assigning tissue-typical susceptibility values taken from literature (26,27), $\overline{\chi}_{tissue}$ (see Table 1). To avoid that susceptibility values were constant throughout anatomical regions (piece-wise constant), which would be unrealistic and could be advantageous to algorithms with gradient-based regularization terms, we modulated the susceptibility values in each region using the image intensities on R_1 and R_2* maps according to the following equation:

$$\chi(r)_{tissue} = \bar{\chi}_{tissue} + a_{tissue}(R_2^*(r) - \bar{R}_{2tissue}^*) + b_{tissue}(R_1\ (r) - \bar{R}_{1tissue}) \qquad \text{[Eq. 1]}$$

where $\bar{R}_{2tissue}^*$ and $\bar{R}_{1tissue}$ are the mean apparent transverse and longitudinal magnetization rate of that given segmented class. To obtain realistic proportionality parameters a_{tissue} and b_{tissue} Eq.1 was inverted for each brain tissue class using in that case a susceptibility map $\chi(r)$ computed from the original data using HEIDI.The coefficients $\bar{\chi}_{tissue}$, a_{tissue} and b_{tissue} used for the two models in the QSM challenge 2.0 can be seen in Table 1 and Figure 2. Because the measured relaxation rates of blood (both in arteries and veins) are prone to errors (due to the inflow for the R_1 and flow for R_2*), the proportionality values were relatively small for the blood pool, rendering these compartments piece-wise constant. Similarly, bone, calcification and other non-brain tissue compartments were made piece-wise smooth due to the lack of a susceptibility map outside the brain that could guide the derivations on the terms a_{tissue} and b_{tissue}. Furthermore, R_2* maps are typically overestimated close to air-tissue boundaries due to signal dropout (as can be seen in the R_2* maps in Fig.1 over the ear canals) (28). A low-pass filtered version of the gradient of the acquired field maps was used to differentiate regions where the R_2* values could be trusted from those where they had to be considered unreliable. In the latter regions, we forced $a_{tissue} = 0$ (did not account for the R_2* contribution when generating the susceptibility model. To avoid discontinuities in the transition between high field gradient and low field gradient, a smooth transition was created by mixing the two models (when Eq.1 was

used and when a_{tissue} was forced to 0) in regions with field gradients between 0.08ppm/mm and 0.3ppm/mm (highlighted by blue arrows in Fig.3 3). Please refer to provided code for more details of implementation.

To avoid unrealistically sharp edges at the interfaces between tissue regions and introduce some degree of partial volume in those regions, 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 full-width at half maximum (FWHM) of 1.2 voxels. This smoothing was not applied to veins to ensure that veins remain sharp, but that their susceptibility will vary depending on its size. The susceptibility model was then given by:

$$\chi(r) = \sum_{tissue=1}^{16} P_{tissue}(r)\, \chi(r)_{tissue} / \sum_{tissue=1}^{16} P_{tissue}(r) \qquad \text{[Eq2]}$$

The importance of moving from a piece-wise constant (where a a_{tissue} and b_{tissue} are set to zero), to a contrast modulated (where P_{tissue} is simply a binary mask) or the probabilistic formalism of the susceptibility distribution can be appreciated in Figure 3. Yellow arrows highlighting the transition between grey and white matter that becomes smoother, and green arrows smoothing out small segmentation errors within the thalamus. On the other hand, Figure 3 shows that most vessel structures have remained sharp with only some minor reduction in susceptibility value.

2.4 Data Simulation

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

$$S = M_0(r) sin(\alpha) \frac{(1-e^{-TR.R_1})}{1-cos(\alpha)e^{-TR.R_1}} e^{-TE.R_2^* + i(\phi_0(r) + TE.\Delta\omega_\chi)} \qquad \text{[Eq. 3]}$$

Here, 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 repetition times (TR), echo times (TE), and flip angles (α), we used the M_0, R_1 and R_2* maps derived from the MP2RAGEME (20,29) sequence, $\phi_0(r)$ is the TE=0 phase (a second order polynomial ensuring 2π phase variation inside the brain region was used), and the frequency shift, $\Delta\omega(r)$, was calculated according to

$$\Delta\omega(r) = D(r) \otimes \chi(r), \qquad \text{[Eq.4]}$$

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 (30,31). To avoid aliasing artifacts associated with the Discrete Fourier Transform (circular convolution), which would appear as unrealistic background fields, we padded the model with zeros along each dimension (factor of 2) before evaluating Eq. 4.

Having a digital phantom allowed the creation of signal models with and without background fields (fields generated by tissues outside the brain). Using the whole-head susceptibility model allowed the creation of realistic background fields. To create a model without background fields, named here after "local field", all voxels outside the brain were set to zero and the susceptibility distribution within the brain was demeaned:

$$\begin{cases} \chi(r) = \chi(r) - \overline{\chi(Mask)} & insideMask \\ \chi(r) = 0 & outsideMask \end{cases} \qquad \text{[Eq.5]}$$

Also relevant, but not pursued for the QSM challenge purposes, field shimming is simulated by fitting the frequency map with second or third-order Legendre polynomials.

2.4.1 Examples of acquisition protocol optimization

For the demonstration of the acquisition protocol simulation with the code in appendix 1, we chose two example protocols optimized for different applications:

(P1) A protocol optimal for the quantification of deep grey matter susceptibility. In this case, the longest TE should be at least that of the T_2^* of the region with the highest iron concentration which is the globus pallidum (14ms).

(P2) A protocol optimal for the observation of cortical grey/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 GM (33ms) (32).

For the sake of simplicity, both protocols had the same echo spacing of 8ms as the original volunteer dataset. The TR of the acquisition was chosen as short as possible assuming a readout acquisition window of 8ms (P1: TR = 16ms; P2: TR = 40ms). 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 GP (T_1=1100ms; α=8) and such that T1-weighted contrast on magnitude images was maximized between WM

(T_1=1100ms) and cortical GM (T_1=1900ms; α=23) in protocols P1 and P2, respectively. This resulted in the following protocols:

Protocol 1: TE_1/TE_2= 4/12ms ;

Protocol 2: TE_1/TE_5= 4/36ms ;

We mimicked k-space sampling by cropping the Fourier spectrum of the original 0.65mm resolution data to an effective spatial resolution of 1mm 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 artefacts, which were removed using sub-voxel shifts (33). This step was repeated in all three spatial directions.

2.4.1 QSM reconstruction optimization

To demonstrate the applicability of the current framework for QSM challenge or for QSM reconstruction optimization, we performed a simulation with only local fields (see Eq. 5). The protocol used was that of the QSM challenge (TR = 50ms; $TE_{1/2/3/4}$= 4/12/20/28ms ; α=15). Multi-echo combination was performed using optimum weights (16) assuming that the phase at TE=0 was unknown and spatial unwrapping was performed across successive echoes using SEGUE (34). QSM reconstruction using TKD (35), closed form L2 (36), FANSI (37) or iLSQR (38) with each of the QSM reconstructions being run 20 times with varying regularization parameters:

- TKD: the k-space threshold was varied from 0.02 to 0.4 in linear steps;

- Closed form L2: the regularization parameter was varied in 20 logarithmic steps 0.01 to 100;

- FANSI TV: the regularization parameter was varied from 10^{-6} to 0.1 in logarithmic steps, remaining parameters were μ_1=0.005, μ_2=1 and 50 iterations;

- iLSQR: the regularization term was varied from 10^{-3} to 1 in logarithmic steps, had a tolerance of 0.01 and was allowed 50 iterations.

The reconstructions were evaluated in respect to the ground-truth using various metrics tested for the purpose of the QSM challenge:

- nRMSE[1]: Whole-brain root-mean-squared error (after demeaning) relative to the root-mean-squared of the ground-truth times 100;

- rmse_detrend_Tissue: nRMSE relative to ground-truth (after detrending) in grey and white matter mask. By detrending it is meant the slope was between the QSM reconstruction and the ground truth QSM in the ROI was taken and the original reconstruction was then divided by this factor to ensure proportionality errors (measured elsewhere) would not be affecting the obtained results;

- 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 & sub-thalamic nucleus, red nucleus, dentate nucleus, putamen, globus pallidus and caudate);

- DeviationFromLinearSlope: Absolute difference between the slope of the average value of the 6 deep gray matter regions vs the prescribed mean value and 1.0.

- CalcStreak: Estimation of the impact of the streaking artifact in a region of interest surrounding the calcification by means of the standard deviation of the difference map between reconstruction and the ground-truth. The region of interest where this was measured was a hollow rectangular prism, with its inner boundary being 2 voxels away from the edge of the calcification and the outer boundary 6 voxels away from the inner boundary.

- CalcMoment: volumetric susceptibility moment of the reconstructed calcification, compared to the ground-truth (computed in the high resolution model to be -49.8 ppm).

The process was repeated for the "challenge protocol" with different simulated datasets:

- 1mm and 0.64mm isotropic resolution
- With infinite SNR (no noise added) or peak SNR of 100

[1] Note that all the RMSEs were multiplied by 100

2.4.2 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. Diffusion data was processed using fsl software (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/), eddy_correct and top up were used to undistort the DWI data. Subsequently the data was coregistered to the 7T anatomical space and FDT was used to extract tensor information (FA, fractional anisotropy, main eigenvector orientation).

Recent literature has been able to explain microstructure effects in GRE data through a three-compartment white matter model with intra-axonal, extra-axonal and myelin water compartments (14,39–42).

The provided data and code include a simple first-order approximation of these microstructure effects which is echo time independent. Wharton et al (14) demonstrated that the typical impact at 7T for a protocol with echo times of 7 and 13 ms was given by:

$$\Delta\omega = \Delta\omega_{\chi} - 5 * (sin^2(\theta) - 2/3)FA_{norm} \quad - 3 \qquad \text{[Eq.7]}$$

where FA_{norm} is the fractional anisotropy (FA) 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 is only applied within the segmented white matter mask.

3 Results

3.1 Acquisition Protocol Optimisation

Figure 4 shows Model 1 with the two different sequence parameters created by the proposed simulation toolbox (See Appendix 1 for code). The top row shows an example slice of the simulated data from the P2 protocol aimed at computing QSM in cortical grey and white matter. In this case, the longer echo time matched the T_2* of those grey 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 degrees) was set to increase T_1 contrast in grey vs white

matter boundaries at short echo times 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 grey matter regions. The echo time range is smaller, so that the magnitude signal in the globus pallidus has not yet disappeared at the last echo time and the tissue contrast in the magnitude data is considerably weaker because of the smaller flip angle (8 degrees). From the data shown in Figure 4 it can be expected that P1 will benefit more from a morphological informed reconstruction than P2, and that the use of non-linear fit (43) for the calculation of the field map will also be particularly relevant as noise will dominate the later echoes in P1.

In Figure 5 we compared the field map obtained with the P1 simulation from the whole head model to a field map obtained directly from the susceptibility brain model via Eq. 5 (ground truth field map). When carrying out the signal simulation with the whole-head model, both the base resolution and the 1mm 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, as pointed by the black arrow). The differences to 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 LBV 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 (44). Both the base resolution and 1mm resolution look at first sight to have very similar properties, yet the nRMSE from the down-sampled dataset (bottom row) demonstrated additional deviations from the ground truth (nRMSE was 40% higher than that of the high-resolution dataset). This discrepancies are both due to incomplete background field removal (see grey arrows) and due to errors around veins. In such regions, the field map measured from the GRE 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 towards 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 model 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 exists and its immediate surrounding where spatial unwrapping fails.

To further investigate the sources of errors discussed in the previous paragraph, 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 in the case of an 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 5 echo simulation) and the ground truth frequency (computed from the susceptibility map). Figure 6b shows that in a region closer to the calcification that is also the case for the high-resolution data (light grey lines), while in the case of the low-resolution data (dark grey) there is a larger error both in respected to fitted frequency (dashed line) and ground truth frequency evolution. Also, it is clear that the phase evolution in Fig. 6b in the low resolution data is no longer linear, as is predicted from theory due to partial volume effects and the varying intravoxel frequency gradients (18). In a pixel immediately close to the vicinity of the calcification the errors are further enhanced now also for the high resolution data where unwrapping errors can introduce errors on the fitted frequency (the data is still fitted accordingly, but it does not correspond to the ground truth field). Figure 6c shows the mean squared difference map associated with the frequency fit on the low resolution data, it is clearly visible that these errors 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 7a shows the nRMSE values associated with various reconstruction methodologies as a function of the regularization value for the high-resolution model in the absence of noise. Figure 7b shows the more realistic case where the images have been down-sampled to 1mm isotropic and noise was added to the data with a peak SNR of 100. It can be seen from a shift of the individual curves to the right that the more realistic model required stronger regularization than the perfect high-resolution model, which can be attributed to the phase noise and inconsistencies

associated with the down-sampling process. Surprisingly, the FANSI method, as opposed to the other tested methods achieved a lower nRMSE in the down-sampled model than on the original model (36 vs 43). Figure 7c shows the reconstructions with minimum RMSE of the plots in Fig. 7b for the different methods tested. It can be seen that the direct methods (TKD and closed-form L2) still have some broad striking artifacts in regions surrounding both the calcification and deep gray matter regions while in the iterative methods these were reduced. It is interesting to note that the total variation (TV) regularized nature of FANSI clearly contributed to a better reconstruction of the superior cerebellar vein when compared to iLSQR.

Figures 7d and 7e show the performance of iLSQR and FANSI respectively for the different models and incremental addition of phase inconsistencies. The addition of a small background field term does not have an impact on the optimum regularization value of any of the methods, but it is interesting to observe that iLSQR was better able to cope with this term than FANSI while the latter was less penalized by the presence of noise on original data (smaller increase in nRMSE).

Finally, more specific metrics were evaluated. These included RMSE in different tissues (Fig. 8a), calcification moment calculation and linearity susceptibility estimations in white matter. It is interesting to observe that, for each method, the optimum regularization values for the different regions remains mainly unchanged, with the minimum of each curve corresponding to either the same regularization factor or minimal deviations when compared to the broadness of the curves. Interestingly, when comparing the dependence of the calculation of the calcification moment (measured as -49.8 ppm on the high resolution model) as a function of the regularization parameter, it can be observed that the two evaluated methods have different optimum regularizations when compared to their optimum RMSE settings (iLSQR and FANSI needed a lower and a higher regularization factors, respectively). Figure 8c, shows the expected behavior where the greatest precision of the accuracy of deep gray matter susceptibility estimation (as measured by the slope between reconstructed average QSM value per deep gray matter structure and ground-truth) was obtained using the smallest amount of regularization (45,46).

3.3 Comparison of simulations including microstructure to real data

Figure 9 shows an example of the originally acquired and the simulated magnitude and phase data, respectively. Here it can be clearly seen that the grey-white matter contrast on both the phase data (second column) and tissue frequency (third column) is better approximated when the microstructural correction term is added to the data as expressed in Equation 7.

4 Discussion

In this paper and the accompanying shared dataset and code (described in greater detail in the Appendix section), 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 GRE 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 of each compartment – this can be simply performed by changing the values presented on Table 1 which control both the mean value and spatial modulation present within each tissue;

- create realistic gradient echo multi-echo data at 7T and, to some extent, also at other fields (it should be noted unlike susceptibility, relaxation times are field dependent and their field dependence is tissue dependent).

- assess the impact 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,

- assess the impact of changing some of the background field removal and QSM algorithm specific options while having a ground-truth to test it against.

Such a digital phantom will also be important for QSM users and researchers deciding on protocols and collecting patient data. Protocol considerations such as impact of the number of echo times and its 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 for example to extend the analysis done by Karsa et al (47) on the effects of field of

view and anisotropic voxels sizes, or the analysis by Biondetti et al on the effects of Laplacian based single echo vs multiple echoes techniques (48).

While the simulation framework is useful for optimizing protocols, the simulation in its current form is a static one, and consequently flow artifacts (which tend to build up particularly when large number of echoes are used), respiration related B_0 fluctuations (49) and spatial distortions associated with the readout bandwidth are not considered. The latter two would be relatively straightforward to implement. For example, respiration artifacts would require a library of respiration fields over time as acquired with field cameras, that would be resampled with the sequence TR the field would be added to the simulated down-sampled complex data, Fourier Transformed back to k-space in order to obtain the corrupted k-space line at a given time point. 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 (30), which does not hold true, particularly in white matter. The phantom released for the RC2 purposes, explicitly circumvented this limitation by ensuring perfect consistency with the QSM model in Eq. 4, i.e. no microstructural effects were present. With the released phantom dataset, we include diffusion data (both raw data on the 1.5mm space and its derivatives co-registered to the phantom space after distortion correction). These data can be used to either compute the frequency perturbation as shown in Figure 9, or the Hollow Cylinder Model can be used to explicitly introduce the echo-time varying perturbation similarly to what has been recently done in the context of myelin water imaging (42). A critical challenge for more advanced modelling is the resolution of the diffusion acquisition, despite using state-of-the-art hardware and MR sequences. At 3T resolutions below 1.5mm isotropic are not commonly obtained *in vivo* (as a reference, a one hour acquisition is used in the human connectome project to a achieve a 1.25mm isotropic resolution(50)). One avenue to obtain higher resolution DWI data are novel 3D multi-slab acquisition with gradient or rf encoding (51,52) that have significant SNR improvements in respect to conventional 2D-SMS as used in this work. Here we have simply interpolated our 1.5mm 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 (53).

QSM is gaining interest in the context of neurological disorders such as multiple sclerosis, PD and AD and other clinical applications such as hemorrhages or tumor imaging with iron oxide nanoparticles (54). For the latter applications, relaxation and susceptibility values in the form of, e.g., lesions in strategic locations can be easily added to the current phantom. Simulated data might be also relevant in the case of group studies of diseases, where differences in the deep grey matter nuclei were found (55,56) and the optimum QSM reconstruction parameters might improve the limits to detect those changes. Such question can be readily addressed with the digital phantom by changing the parameters of Eq. 1 for a given set of structures and test the QSM pipeline that can better quantify those changes.

5 Conclusion:

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 adding microstructure effects *a posteriori* (14) as well as the inclusion of 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.

Data Availability Statement

The code and datasets that supports the findings of this study are openly available at https://data.donders.ru.nl/

6 Appendix

On the data sharing collection: https://data.donders.ru.nl/login/reviewer-113366422/QRFk431i299BX-8bcY6ta6nQPP-MqSzG0DTIhkJrqBs
the code used to create the phantom described in this paper, the various simulations and figures can be found.

Here we outline some of the main functionalities provided:

- creation of a realistic magnetic susceptibility phantom;
- simulation of a GRE data with a specific protocol;
- evaluation of a QSM reconstruction using provided phantom;
- adding microstructural information;

6.1 Creation of a realistic magnetic susceptibility phantom

Example code can be found on `MacroCreateSusceptibilityPhantom.m`.

The main function responsible by the creation of a new susceptibility phantom is, `CreateOwnRealisticPhantom`, which takes as an input a matlab structure with the various fields:

- R1map_file – location of R1 map in s-1 in nifti format;
- R2starmap_file – location of R1 map in s-1 in nifti format;
- M0map_file – location of M0 map in arbitrary units and in nifti format;
- Segmentation_file – Location of segmentation file in nifti format (each segmented tissue is associated with an integer);
- ChiModulation_file - matlab file containing the various parameters relating to Equation 1 and shown on Table 1 to create a susceptbility map – these has to be in agreement with the labeling used in the Segmentation nifty;
- rawField_file – precomputed field map in nifti format;
- highGradMask_file – nifti mask, defining regions where R2* values can and can't be trusted;
- OutputChiModel_file –address where to store the output susceptibility model in ppm units;

Furthermore, code in `MacroCreateSusceptibilityPhantom.m.` , creates Figures 2 and 3 of the current manuscript.

6.2 Creation of simulated GRE data

Example code can be found in `MacroCreateSimulationdata.m`, with the main function being `CreateSimulationData` that takes as an input three Matlab structures. These Structures contain the information regarding the model parameters, the sequence parameters of the gradient echo sequence and the simulation parameters respectively.

The model parameters fields are:

- Chimap_file – location of magnetic susceptibility map in ppm units and nifti format;;
- R1map_file – location of R1 map in s-1 in nifti format;
- R2starmap_file – location of R2* map in s-1 in nifti format;;
- M0map_file – location of M0 map in arbitrary units in nifti format;
- Segmentation_file – location of segmentation file in nifti format;
- BrainMask_file = 'data/masks/BrainMask.nii.gz';

The sequence parameters fields describe the GRE contrast and are TR (repetition time, secs), TE (echo time), FlipAngle (flip angle, degrees). Finaly, the simulation parameters are :

- Res - resolution of output in mm ;
- B0 and B0_dir are the main static magnetic field strength (in Tesla) and its orientation in respect to the head model (if along z [0 0 1]);
- PhaseOffset - multiplier term of a quadratic phase over the brain aimed at simulating the RF transmit phase (0 ensures no phase offset, 1 ensures a π phase difference inside the brain mask ;
- Shimm - boolean variable to account B0 shimming with second order spherical harmonics (0 if no B0 shimms is applied, 1 - 2nd order shimms are computed inside the brain mask and applied only when doing the simulations with the whole head model)
- Output_dir - Directory where the simulation results will be written

Running the function `CreateSimulationData` creates two directories inside Output_dir, SimulatedHR (where the native resolution simulation data is stored) and Simulated_1p0mm (where the downsampled data simulated data). For each protocol two simulations are performed: one with the full head susceptibility model, and for the brain masked susceptibility model [Eq. 5].

Simulation results are stored as 3D or 4D (depending on the number of echo times) Data[Extension]_magn.nii.gz and Data[Extension]_phase.nii.gz where the Extension field is either empty or is 'BrainExtracted'. Additionally the downsampled directory contains also the new brain mask (Brain.nii.gz), the downsampled Segmented nifti (FinalSegment.nii.gz) and the downsampled susceptibility maps (Chi_crop.nii.gz).
Additionally, `MacroCreateSimulationdata` allows the reproduction of figure 4.

6.3 QSM pipeline evaluation

The script `MacroCreateQSMpipelineAndCompareHighResLowRes.m` allows the creation of figures 5 and 6 of the current manuscript. Furthermore it shows how to create a data structure in order to perform a QSM reconstruction using SEPIA (sepia-documentation.readthedocs.io/) and evaluate the output using the various metrics used in RC2 and introduced in the Methods section of this manuscript. By running the function `EvaluateRecon_ChallengeFinalMetricsRelease` that takes as input reconstruction file name and a structure with 4 fields:

- Segment - Segmentation file at a matched resolution (output of `CreateSimulationData`);
- chi_crop - Ground truth susceptibility map file obtained by k-space cropping (output of `CreateSimulationData`);
- maskEroded - brain mask where evaluation is to be performed;
- CalcMoment – calculated moment of the calcification performed on the high resolution susceptibility map;

The output of the function contains the metrics described in methods section;

6.4 Comparison of simulations including real data to real data

The script `MacroAddingMicrostructure` contains the implementation of the code needed to generate Figure 9 using the already pre-processed DWI. The addition of the microstructure data is as described in Eq. 7.

References

1. Langkammer C, Schweser F, Krebs N, et al. Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. NeuroImage 2012;62:1593–1599 doi: 10.1016/j.neuroimage.2012.05.049.

2. Fan AP, Bilgic B, Gagnon L, et al. Quantitative oxygenation venography from MRI phase. Magn. Reson. Med. 2014;72:149–159 doi: 10.1002/mrm.24918.

3. 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 doi: 10.1371/journal.pone.0057924.

4. 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 doi: 10.1002/nbm.3601.

5. Wang Y, Liu T. Quantitative susceptibility mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. Magn. Reson. Med. 2015;73:82–101 doi: 10.1002/mrm.25358.

6. 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. 2016:n/a-n/a doi: 10.1002/nbm.3604.

7. 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 doi: 10.1109/TBME.2017.2749298.

8. Jung W, Bollmann S, Lee J. Overview of quantitative susceptibility mapping using deep learning: Current status, challenges and opportunities. NMR Biomed. n/a:e4292 doi: 10.1002/nbm.4292.

9. 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 doi: 10.1016/j.mri.2014.09.004.

10. Langkammer C, Schweser F, Shmueli K, et al. Quantitative susceptibility mapping: Report from the 2016 reconstruction challenge. Magn. Reson. Med. 2017 doi: 10.1002/mrm.26830.

11. Liu C. Susceptibility tensor imaging. Magn. Reson. Med. 2010;63:1471–1477 doi: 10.1002/mrm.22482.

12. 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 doi: 10.1002/mrm.21828.

13. 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 doi: 10.1002/mrm.28185.

14. Wharton S, Bowtell R. Effects of white matter microstructure on phase and susceptibility maps. Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med. Soc. Magn. Reson. Med. 2014 doi: 10.1002/mrm.25189.

15. Kressler B, de Rochefort L, Liu T, Spincemaille P, Jiang Q, Wang Y. Nonlinear regularization for per voxel estimation of magnetic susceptibility distributions from MRI field maps. IEEE Trans. Med. Imaging 2010;29:273–281 doi: 10.1109/TMI.2009.2023787.

16. 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 doi: 10.1016/j.neuroimage.2014.11.038.

17. Wharton S, Bowtell R. Whole-brain susceptibility mapping at high field: a comparison of multiple- and single-orientation methods. NeuroImage 2010;53:515–525 doi: 10.1016/j.neuroimage.2010.06.070.

18. Yablonskiy DA, Haacke EM. Theory of NMR signal behavior in magnetically inhomogeneous tissues: the static dephasing regime. Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med. Soc. Magn. Reson. Med. 1994;32:749–763.

19. 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 doi: 10.1016/j.neuroimage.2009.10.002.

20. Caan MWA, Bazin P-L, Marques JP, Hollander G de, Dumoulin SO, Zwaag W van der. MP2RAGEME: T1, T2*, and QSM mapping in one sequence at 7 tesla. Hum. Brain Mapp. 2019;40:1786–1798 doi: 10.1002/hbm.24490.

21. 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 doi: 10.1002/mrm.23017.

22. 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, Part 2:201–209 doi: 10.1016/j.neuroimage.2013.03.077.

23. 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 doi: 10.1016/j.neuroimage.2006.01.015.

24. Chan K-S, Norris DG, Marques JP. Structure tensor informed fibre tractography at 3T. Hum. Brain Mapp. 2018;39:4440–4451 doi: 10.1002/hbm.24283.

25. 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 doi: 10.1016/j.media.2018.03.001.

26. 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 doi: 10.1016/j.neuroimage.2012.09.055.

27. 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 doi: 10.1002/mrm.25350.

28. 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 doi: 10.1002/mrm.25114.

29. Marques JP. JosePMarques/MP2RAGE-related-scripts.; 2019.

30. 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 doi: 10.1002/cmr.b.20034.

31. 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 doi: 10.1002/cmr.b.10083.

32. Marques JP, Norris DG. How to choose the right MR sequence for your research question at 7T and above? NeuroImage 2018;168:119–140 doi: 10.1016/j.neuroimage.2017.04.044.

33. Kellner E, Dhital B, Kiselev VG, Reisert M. Gibbs-ringing artifact removal based on local subvoxel-shifts. Magn. Reson. Med. 2016;76:1574–1581 doi: 10.1002/mrm.26054.

34. Karsa A, Shmueli K. SEGUE: A Speedy rEgion-Growing Algorithm for Unwrapping Estimated Phase. IEEE Trans. Med. Imaging 2019;38:1347–1357 doi: 10.1109/TMI.2018.2884093.

35. Wharton S, Schäfer A, Bowtell R. Susceptibility mapping in the human brain using threshold-based k-space division. Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med. Soc. Magn. Reson. Med. 2010;63:1292–1304 doi: 10.1002/mrm.22334.

36. Bilgic B, Chatnuntawech I, Fan AP, et al. Fast image reconstruction with L2-regularization. J. Magn. Reson. Imaging JMRI 2014;40:181–191 doi: 10.1002/jmri.24365.

37. 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 doi: 10.1002/mrm.27073.

38. 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 doi: 10.1016/j.neuroimage.2014.12.043.

39. 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 doi: 10.1016/j.neuroimage.2011.08.064.

40. Wharton S, Bowtell R. Fiber orientation-dependent white matter contrast in gradient echo MRI. Proc. Natl. Acad. Sci. U. S. A. 2012;109:18559–18564 doi: 10.1073/pnas.1211075109.

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

42. 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 doi: 10.1016/j.neuroimage.2020.117159.

43. 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 doi: 10.1002/mrm.24272.

44. 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 doi: 10.1002/nbm.3604.

45. Schweser F, Deistung A, Sommer K, Reichenbach JR. Toward online reconstruction of quantitative susceptibility maps: superfast dipole inversion. Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med. Soc. Magn. Reson. Med. 2013;69:1582–1594 doi: 10.1002/mrm.24405.

46. Milovic C, Prieto C, Bilgic B, et al. Comparison of parameter optimization methods for quantitative susceptibility mapping. Magn. Reson. Med. n/a doi: 10.1002/mrm.28435.

47. 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 doi: 10.1002/mrm.27542.

48. 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. n/a doi: 10.1002/mrm.28331.

49. 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 doi: 10.1002/mrm.10145.

50. Vu AT, Auerbach E, Lenglet C, et al. High resolution whole brain diffusion imaging at 7T for the Human Connectome Project. Neuroimage 2015;122:318–31 doi: 10.1016/j.neuroimage.2015.08.004.

51. Wu W, Poser BA, Douaud G, et al. High-resolution diffusion MRI at 7T using a three-dimensional multi-slab acquisition. NeuroImage 2016;143:1–14 doi: 10.1016/j.neuroimage.2016.08.054.

52. Haldar JP, Liu Y, Liao C, Fan Q, Setsompop K. Fast submillimeter diffusion MRI using gSlider-SMS and SNR-enhancing joint reconstruction. Magn. Reson. Med. 2020;84:762–776 doi: 10.1002/mrm.28172.

53. 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 doi: 10.1002/nbm.3999.

54. 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 doi: 10.1038/s41598-020-58219-9.

55. 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 doi: 10.1371/journal.pone.0081093.

56. 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 doi: 10.1148/radiol.2020192541.

Tables

Table 1

	Model 1			Model 2		
Label Name	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[+*T†]	-0,012	1,118	0,044	-0,011	1,230
Globus Pallidus	0,131[+]	-0,026	0,843	0,121	-0,023	0,927
Putamen	0,038[+]	-0,025	1,852	0,043	-0,022	2,038
Red Nucleus	0,100[+]	-0,044	1,780	0,090	-0,040	1,958
Dentate Nucleus	0,152[+]	-0,064	1,708	0,162	-0,058	1,879
Substantia Nigra & Sub-Thalamic Nucleus	0,111[+]	-0,075	1,491	0,121	-0,068	1,640
Thalamus	0,020[+]	-0,086	1,275	0,025	-0,078	1,402
White Matter	-0,030[+]	-0,078	1,147	0,005	-0,070	1,262
Grey Matter	0,020[*]	-0,095	1,402	0,020	-0,085	1,543
Cerebro Spinal Fluid	0,019[+]	-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[†]	0,000	0,000	0,019	0,000	0,000
Bone	-2,100[†]	0,000	0,000	-2,100	0,000	0,000
Air	9,200[†]	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[†]	-0,012	0,000	0,019	-0,011	0,000

Parameters used to create the two magnetic susceptibility head models released in the QSM challenge. 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 ([+] (26), [*] (14) [†] (27)) and doing the fitting described in Eq. 1. Model 2 mean susceptibility values were adhoc modification of those found in literature, while the R_1 and R_2* modulation were changed by plus and minus 10% respectively.

Figures

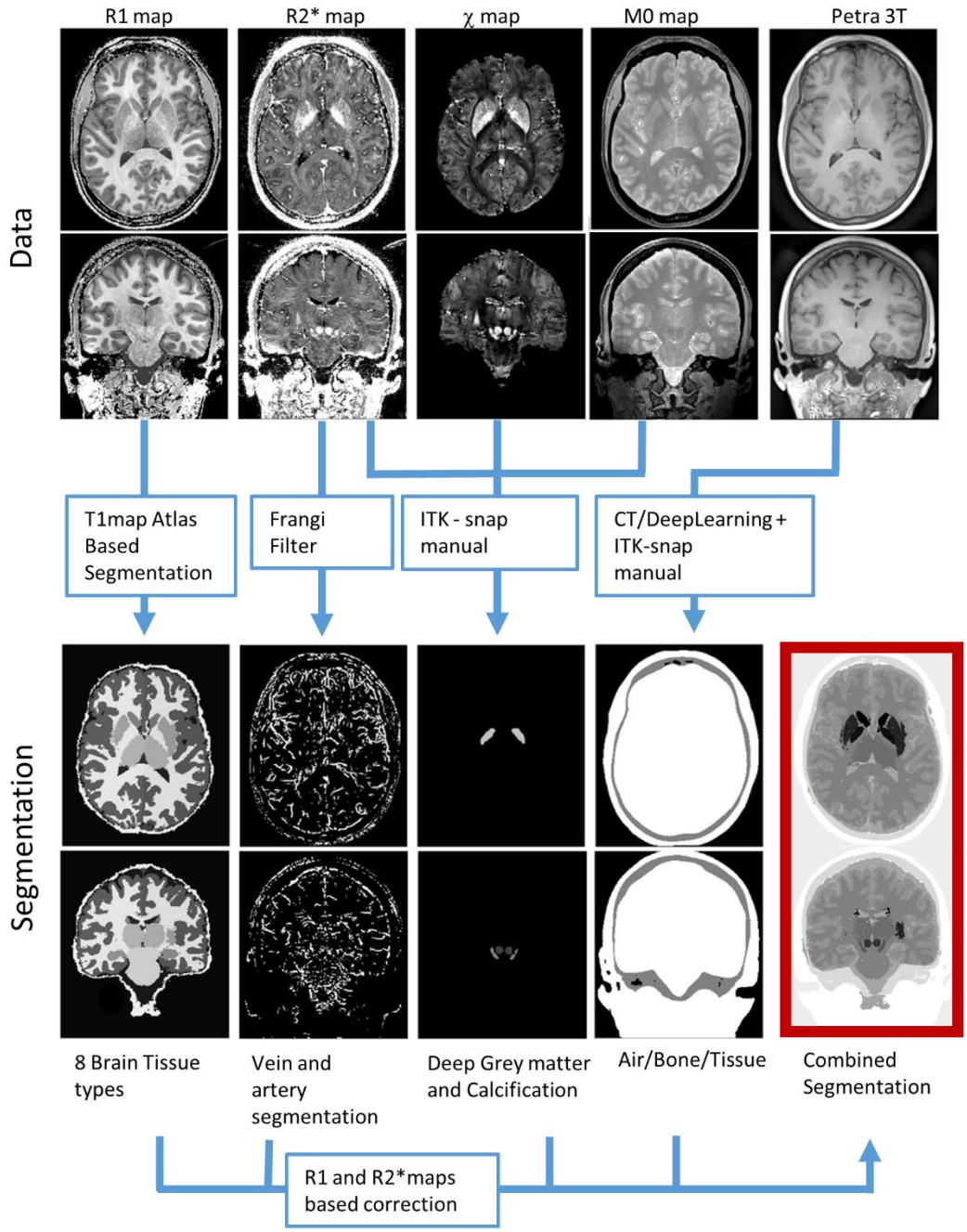

Figure 1 Diagram representing the process used to obtain the head segmentation: R_1 map (obtained from the MP2RAGEME) was used to create an atlas-based segmentation using CBS tools; R_2* were processed with a Frangi filter for vein segmentation; A semi-manual approach using ITK snap was used for segmentation of the deep gray matter nuclei based on the R_2* and a susceptibility map computed using HEIDI, and the M0 map were used to segment the calcification. Finally, PETRA data was 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.

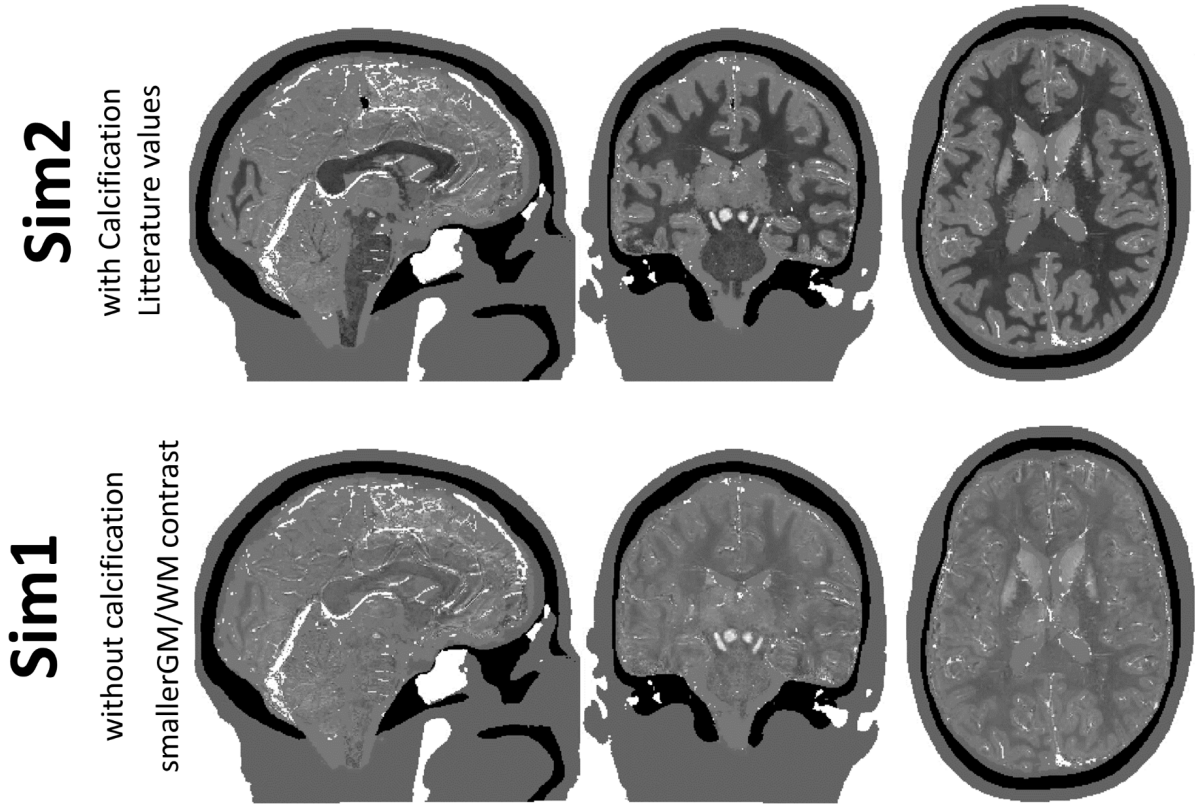

Figure 2 View of three slices in sagittal, coronal and axial directions of the two digital phantoms created for the QSM challenge 2.0. Top and Bottom maps were obtained using the parameters described on Table 1 for Model 1 and 2 respectively.

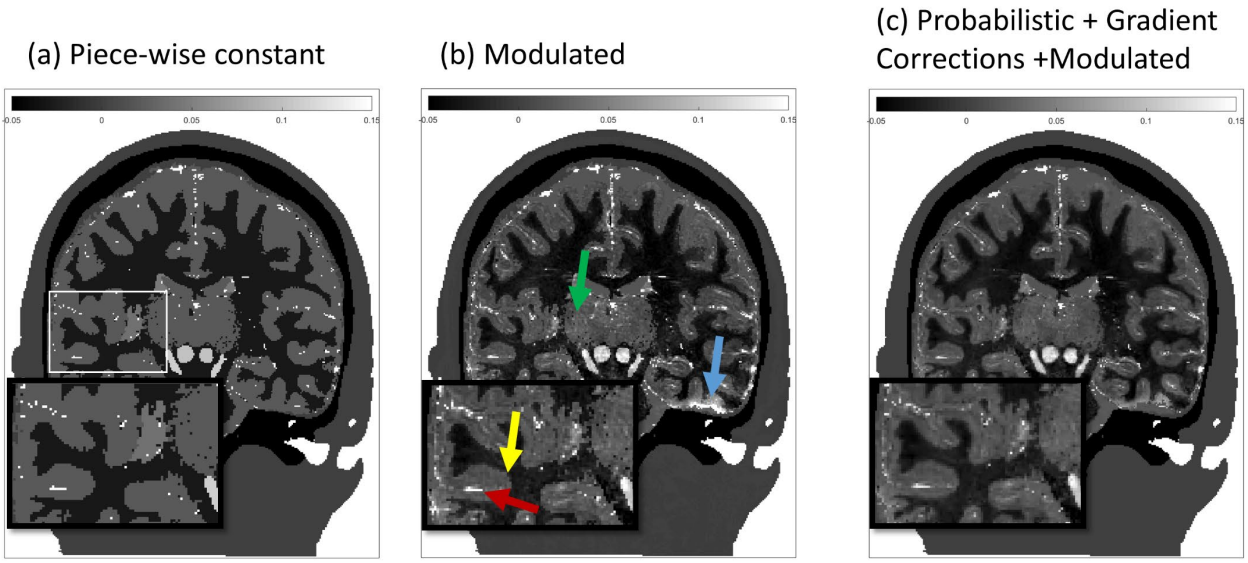

Figure 3 Intermediate stages of the in silico susceptibility head phantom creation: (a) traditional piece-wise constant approach, (b) modulated model as described in Eq. 1 with the values presented in Table 1 and (c) finally adding the probabilistic modulation described in Eq. 2 and masking regions of error bound R2*. Green and yellow arrows highlight transitions between tissue types that get improved via the probabilistic approach applied to the tissue compartments. Note that the probabilistic smoothing was only applied to the brain tissues and CSF, as a result, veins remain shape in the susceptibility map only with a reduced magnetic susceptibility. Blue arrow highlights regions where the large filed gradient masking approach was able to avoid abnormally large susceptibility values.

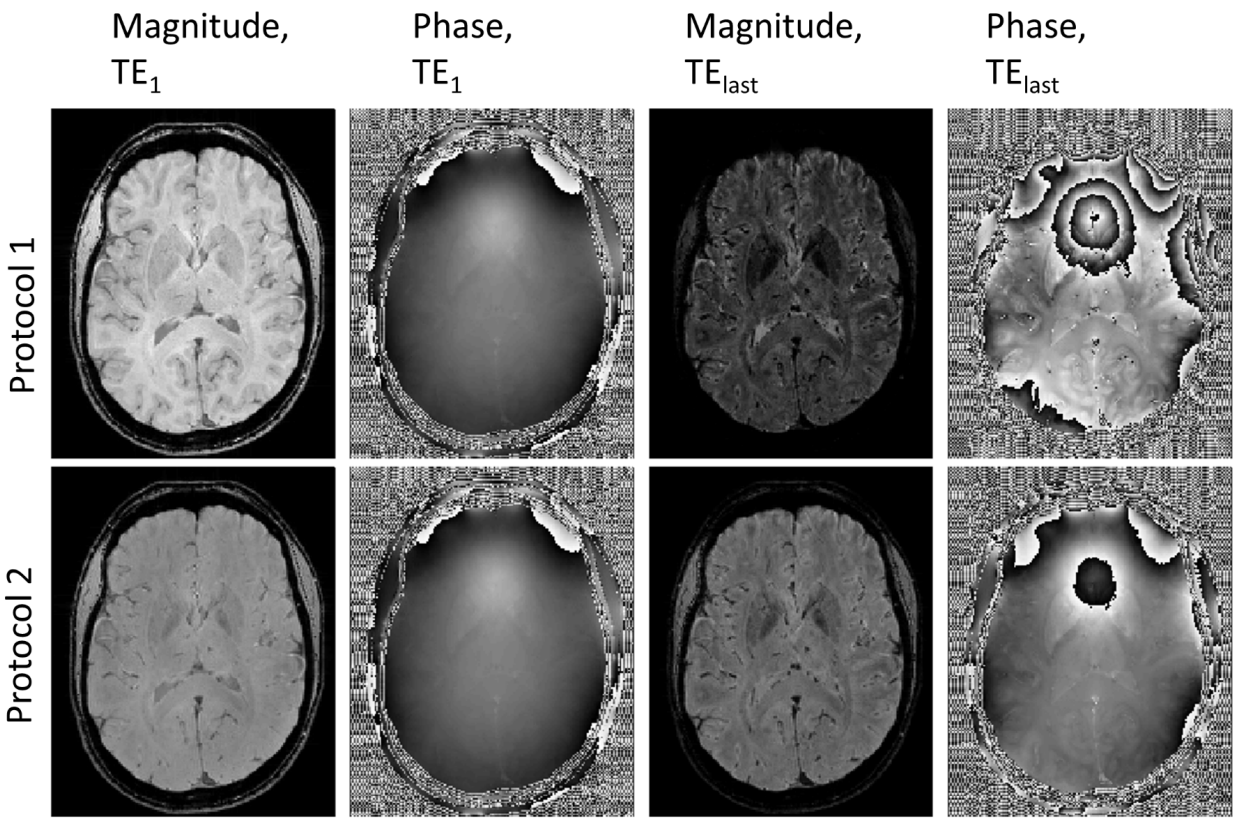

Figure 4 Transverse slices of the simulated data of Model 1 using Protocol 1 (top row) and Protocol 2 (bottoms row). First two columns show magnitude and phase images, respectively, at the first echo time (where the different T_1-weighting is clearly visible) and the last two columns show the magnitude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

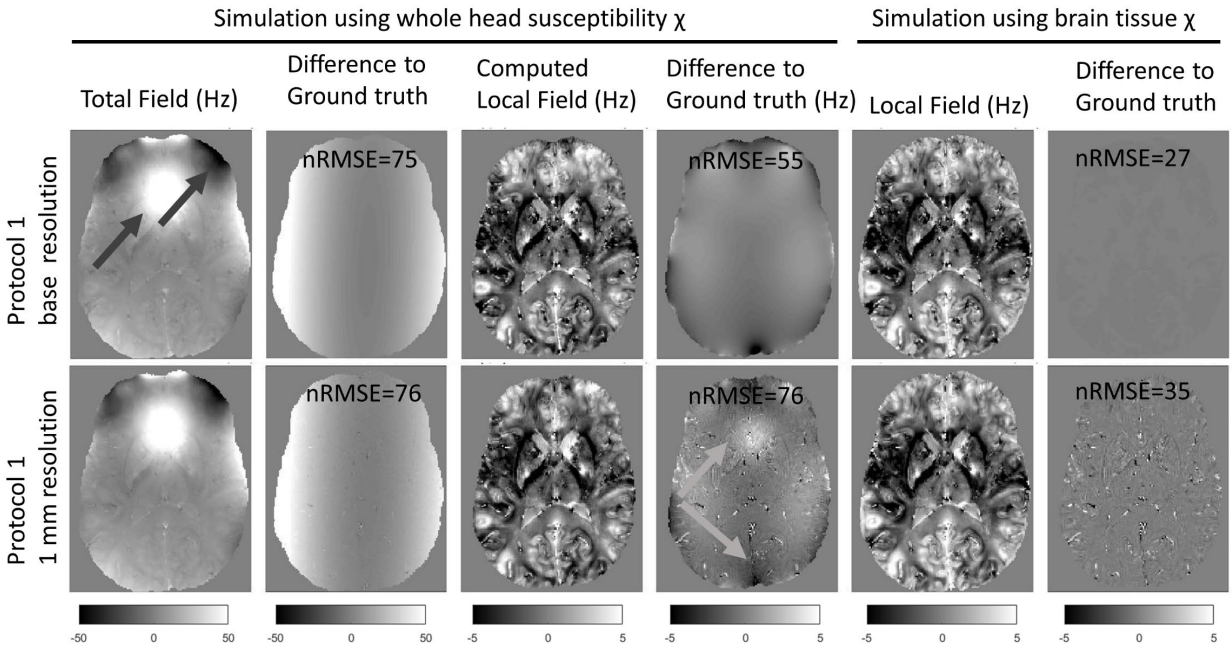

Figure 5 Transverse slices through derived field-maps associated with Protocol 1 data computed at the base resolution (top row) and after down-sampling to 1mm (bottom row), black arrows highlight large background fields induced by air-tissue interfaces. The first and fifth columns shows the field extracted from the complex signal when the whole-head and the 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. Grey arrows highlight the incomplete background field removal that is exacerbated upon downsampling. The ground-truth field maps were computed using the forward dipole formulation (Eq. 4) on the whole head (second column) and brain tissues alone (fourth and sixth columns) susceptibility models. It is clear that the nRMSE, once the down-sampling is performed, is dominated by partial volume effects in and around veins (noise pattern on both bottom 4th and 6th column) and imperfect background field removal.

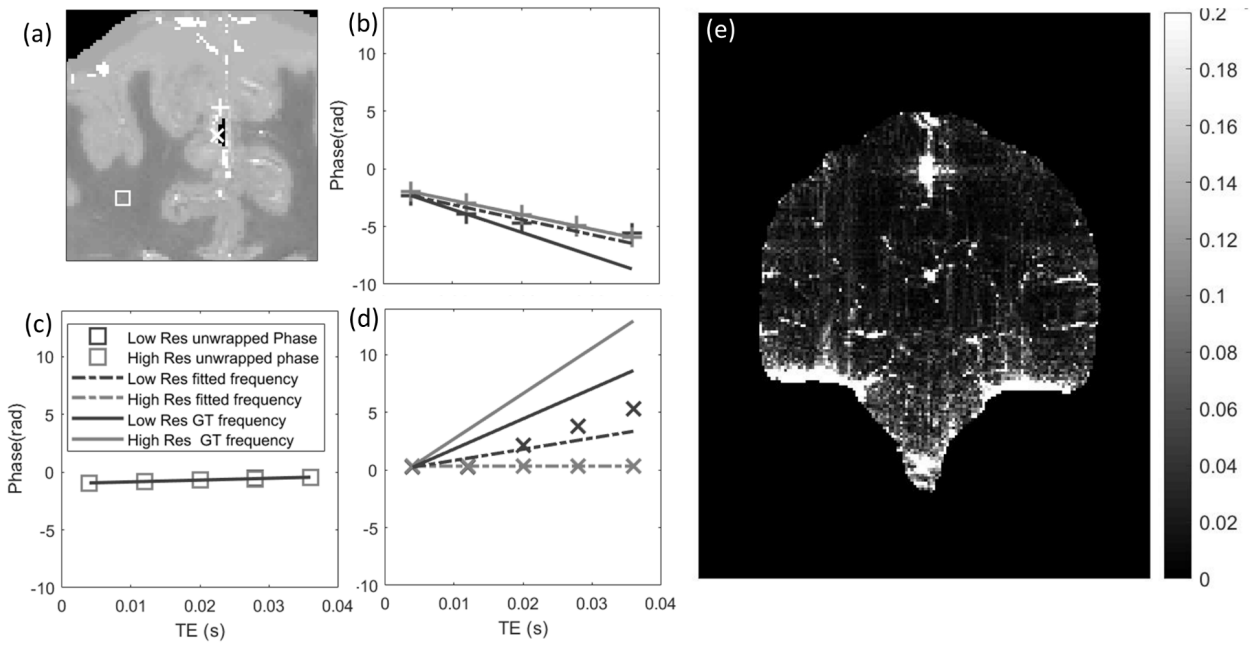

Figure 6 (a) Coronal view of the susceptibility phantom with three locations highlighted (square: region in the middle of white matter; cross and plus: two regions close to the calcification). (b,c,d) Plots of the unwrapped phase at the three locations (b – cross, c –square and d - plus) as a function of echo time. Unwrapped phase on the high-resolution data (light grey) and low-resolution data (dark grey) are shown using the respective markers, while dashed lines corresponds to the fitted frequency for each point at each resolution and the continuous line shows the ground frequency evolutions. (e) Mean squared difference map between fitted phase (dashed line in panel b-d) and measured phase (after unwrapping) on the coronal slice, highlighting tissue bone interfaces as well as regions surrounding the calcification.

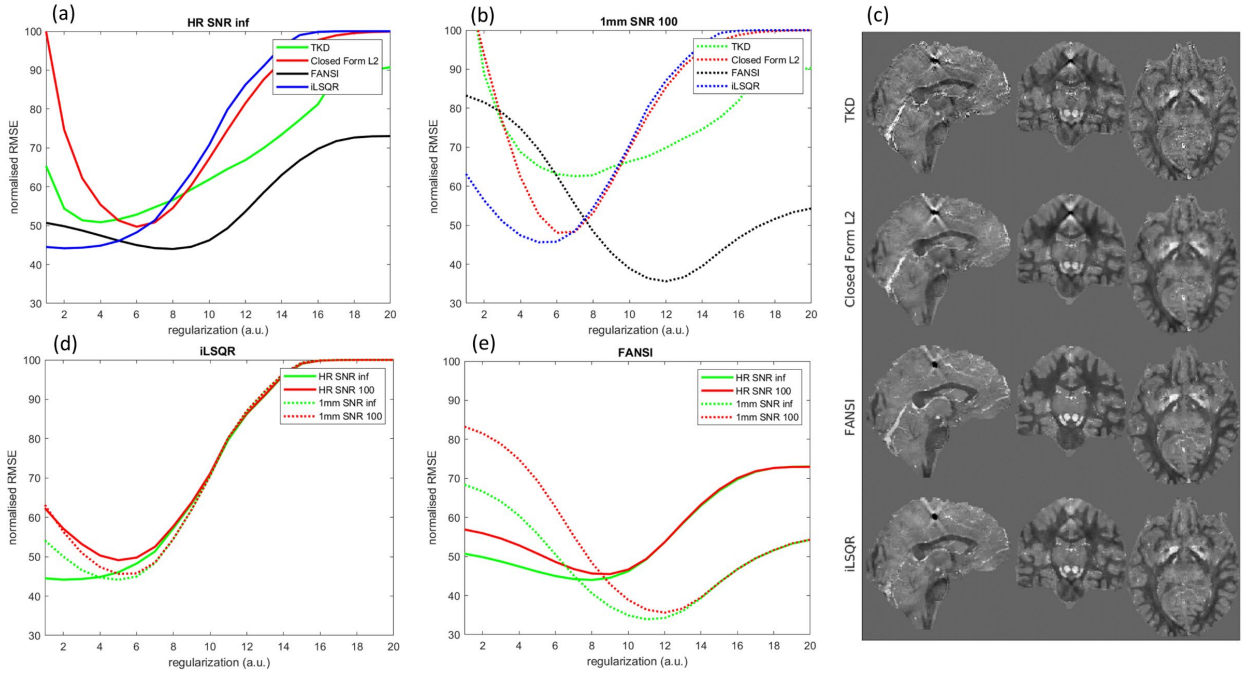

Figure 7 Shows various plots of the normalized RMSE as a functions of the regularization value (a,b,d,e). Plots (a) and (b) show the numerical performance of 4 different QSM algorithms when the input data was the high resolution simulated data with no additional noise or the down-sampled 1mm data with added noise resulting in an image peak SNR of 100. Plots (d) and (e) show the performance of iLSQR and FANSI respectively for different types of input data: high resolution and 1mm data shown in solid and dashed lines respectively; without and with added noise shown in green and red. Panel (c) shows examples of the optimum nRMSE reconstructions of the 1mm dataset with peak SNR 100 using the 4 reconstruction algorithms tested in the manuscript.

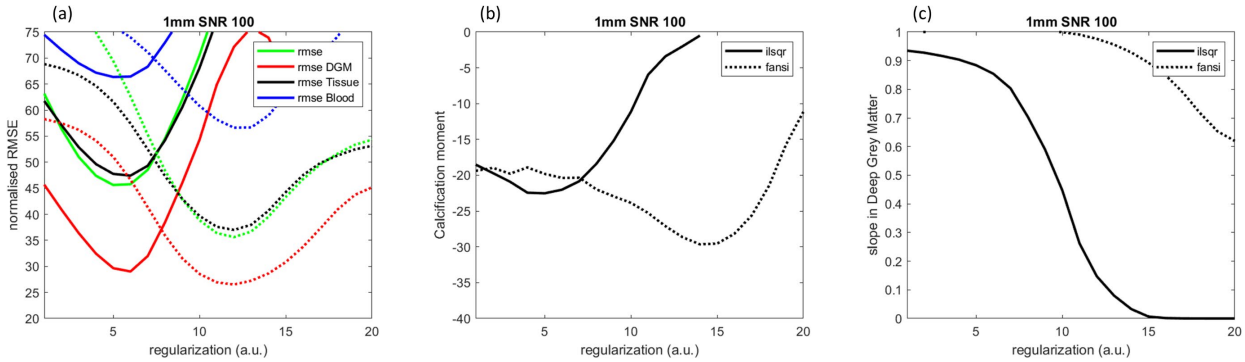

Figure 8 (a-c) Plots of the various metrics as a function of the used regularization value for the iLSQR (continuous line) and FANSI (dashed line) algorithms. The different plots show: (a)

normalized RMSE reconstructions metrics, (b) calcification metric and (c) linearity of the susceptibility values measured in deep grey matter in respect to ground truth.

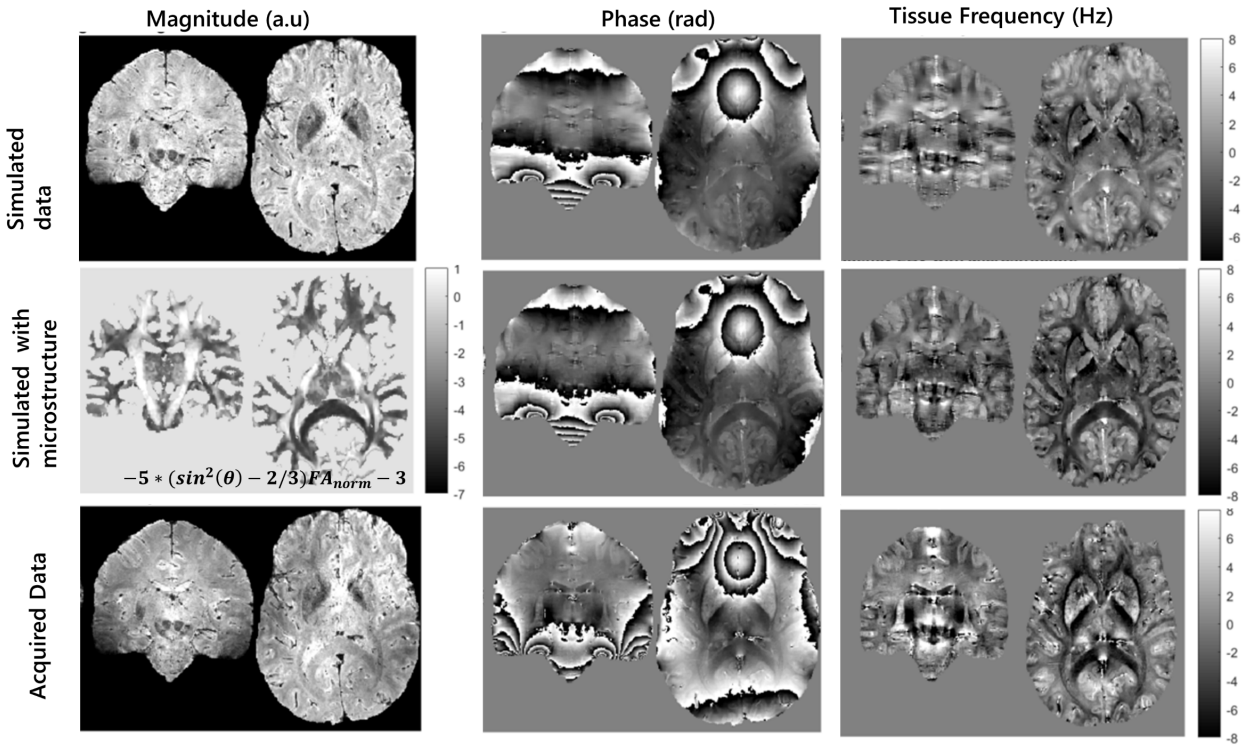

Figure 9 Shows the impact of adding a microstructure correction term to the simulated field map. First and second column show the simulated magnitude and phase data using whole brain susceptibility model at TE=20ms, while the third column shows the computed tissue frequency map (after brain masking and background field removal). First row shows the simulated data using Protocol 1 without the addition of the microstructure effects. The second row shows the impact of adding the microstructural effects as described in equation 7 (first column) both on the third echo time and the computed tissue frequency map. The third row shows, for visual comparison purposes, the experimental data acquired at 7T.