Multi-echo Quantitative Susceptibility Mapping: How to Combine Echoes for Accuracy and Precision at 3 T

Emma Biondetti[1,2], Anita Karsa[2], Francesco Grussu[3,4], Marco Battiston[3], Marios C. Yiannakas[3], David L. Thomas[5,6], and Karin Shmueli[2]

[1]Institute for Advanced Biomedical Technologies, Department of Neuroscience, Imaging and Clinical Sciences, "D'Annunzio University" of Chieti-Pescara, 66100 Chieti, Italy
[2]Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom
[3]NMR Research Unit, Queen Square MS Centre, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, University College London, London WC1N 3BG, United Kingdom
[4]Radiomics Group, Vall d'Hebron Institute of Oncology, Vall d'Hebron Barcelona Hospital Campus, 08035 Barcelona, Spain
[5]Dementia Research Centre, UCL Queen Square Institute of Neurology, University College London, London WC1E 6BT, United Kingdom
[6]Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, University College London, London WC1E 6BT, United Kingdom

Corresponding author
Emma Biondetti, PhD
Institute for Advanced Biomedical Technologies, Department of Neuroscience, Imaging and Clinical Sciences, "G. D'Annunzio University" of Chieti-Pescara, via Luigi Polacchi 11, 66100 Chieti, Italy
E-mail: emma.biondetti@unich.it

Funding information
E. Biondetti was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) (award number: 1489882). A. Karsa was supported by the EPSRC-funded UCL Centre for Doctoral Training in Medical Imaging (EP/L016478/1) and the Department of Health's National Institute for Health Research funded Biomedical Research Centre at University College London Hospitals. D. L. Thomas was supported by the UCL Leonard Wolfson Experimental Neurology Centre (PR/ylr/18575). The Queen Square MS Centre, where part of the MRI scans for this work were performed, is supported by grants from the UK MS Society and by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. F. Grussu was supported by PREdICT, a study at

the Vall d'Hebron Institute of Oncology in Barcelona funded by AstraZeneca, and currently receives funding from the postdoctoral fellowships program Beatriu de Pinós (2020 BP 00117), funded by the Secretary of Universities and Research (Government of Catalonia).

Disclosures

The authors declare no conflicts: AstraZeneca was not involved in any aspect concerning this work and has not influenced its content or the decision to submit it for publication.

Abstract

Purpose: To compare different multi-echo combination methods for MRI quantitative susceptibility mapping (QSM). Given the current lack of consensus, we aimed to elucidate how to optimally combine multi-echo gradient-recalled echo (GRE) signal phase information, either before or after applying Laplacian-base methods (LBMs) for phase unwrapping or background field removal.

Methods: Multi-echo GRE data were simulated in a numerical head phantom, and multi-echo GRE images were acquired at 3 T in ten healthy volunteers. To enable image-based estimation of GRE signal noise, five volunteers were scanned twice in the same session without repositioning. Five QSM processing pipelines were designed: one applied nonlinear phase fitting over echo times (TEs) before LBMs; two applied LBMs to the TE-dependent phase and then combined multiple TEs via either TE-weighted or signal-to-noise ratio (SNR)-weighted averaging; two calculated TE-dependent susceptibility maps via either multi-step or single-step QSM and then combined multiple TEs via magnitude-weighted averaging. Results from different pipelines were compared using visual inspection; summary statistics of susceptibility in deep gray matter, white matter, and venous regions; phase noise maps (error propagation theory); and, in the healthy volunteers, regional fixed bias analysis (Bland-Altman) and regional differences between the means (nonparametric tests).

Results: Nonlinearly fitting the multi-echo phase over TEs before applying LBMs provided the highest regional accuracy of χ and the lowest phase noise propagation compared to averaging the LBM-processed TE-dependent phase. This result was especially pertinent in high-susceptibility venous regions.

Conclusion: For multi-echo QSM, we recommend combining the signal phase by nonlinear fitting before applying LBMs.

Abstract word count: 250/250 words

Keywords: magnetic resonance imaging, quantitative susceptibility mapping, multi-echo QSM

Word count (body of the text, excluding the title page, abstract, figure captions and references): 4991/5000

Introduction

MRI quantitative susceptibility mapping (QSM) aims to determine the underlying spatial distribution of tissue magnetic susceptibility (χ) from gradient-recalled echo (GRE) phase data (ϕ):

$$\phi(\boldsymbol{r}, TE) = \gamma \Delta B_{Tot}(\boldsymbol{r}, TE) TE + \phi_0(\boldsymbol{r}) \#(1)$$

where $\boldsymbol{r}$ is a vector of image space coordinates, γ the proton gyromagnetic ratio, TE the echo time, ΔB_{Tot} the χ-induced total field perturbation along the scanner's z-axis, and ϕ_0 the TE-independent phase offset at a nominal TE=0 ms.

For QSM, the acquired phase must be spatially (single-echo data) or spatiotemporally (multi-echo data) unwrapped to resolve 2π aliasing. The unwrapped ϕ is proportional to ΔB_{Tot} (Equation 1), which is a combination of background (ΔB_{Bg}) and local field contributions (ΔB_{Loc}):

$$\Delta B_{Tot}(\boldsymbol{r}) = \Delta B_{Bg}(\boldsymbol{r}) + \Delta B_{Loc}(\boldsymbol{r}) \#(2)$$

ΔB_{Bg} are induced by the global geometry, air-tissue interfaces, and any field inhomogeneities. ΔB_{Loc} reflect the tissue χ inside the region of interest (ROI), e.g., the brain. For QSM, ΔB_{Bg} must be removed from ΔB_{Tot}. The resulting ΔB_{Loc} map is in the following relationship with χ:

$$\Delta B_{Loc}(\boldsymbol{r}) = d(\boldsymbol{r}) \star \chi(\boldsymbol{r}) \#(3)$$

where d is the magnetic dipole and $\star$ denotes a spatially-dependent convolution. Based on Equation 3, the local distribution of tissue χ, i.e., the QSM map, is calculated by solving an ill-posed ΔB_{Loc}-to-χ problem.

Recently, the QSM community critically reviewed how to best perform phase unwrapping (1) and ΔB_{Bg} removal (2). Moreover, it promoted two challenges to compare algorithms for ΔB_{Loc}-to-χ inversion, but a consensus has yet to be reached (3,4). A further open question toward QSM standardization is how and at which stage of the processing pipeline multi-echo data from different echoes should be combined. This question is relevant because phase unwrapping, ΔB_{Bg} removal, or both, are often performed using Laplacian-based methods (LBMs) (1,2).

Laplacian phase unwrapping aims to calculate the 2π-aliasing-free phase as (5):

$$\nabla^{-2}(\nabla^2 \phi) = \nabla^{-2}\left(\cos\phi_w \, \nabla^2(\sin\phi_w) - \sin\phi_w \, \nabla^2(\cos\phi_w)\right) \#(4)$$

with ∇^2 and ∇^{-2} the forward and inverse Laplace operators, and ϕ_w the aliased phase.

Laplacian ΔB_{Bg} removal relies on the harmonicity of ΔB_{Bg} inside the ROI (i.e., $\nabla^2 \Delta B_{Bg}(\boldsymbol{r}) = 0, \ \boldsymbol{r} \in ROI$) and aims to solve this equation (2):

$$\nabla^{-2}\left(\nabla^2\left(\Delta B_{Loc}(\boldsymbol{r})\right)\right) = \nabla^{-2}\left(\nabla^2\left(\Delta B_{Loc}(\boldsymbol{r}) + \Delta B_{Bg}(\boldsymbol{r})\right)\right), \qquad \boldsymbol{r} \in ROI \#(5)$$

The inverse discrete Laplace operator is not well-defined and requires regularization, which is equivalent to spatially high pass filtering the phase or local field (2). However, the known varying frequency content at different TEs (6,7) could lead to different degrees of LBM-induced high-pass filtering at different echoes, alter the linearity of Equation 1, and thus introduce inaccuracies in the estimated ΔB_{Loc} and χ maps. To investigate this issue, this study aimed to compare existing strategies for combining the multi-echo signal phase (see the Theory section for further details) when using LBMs for phase unwrapping or ΔB_{Bg} removal.

Five processing pipelines for QSM were designed incorporating LBMs for both phase unwrapping and ΔB_{Bg} removal and combining the signal from different TEs by fitting or averaging before or after applying LBMs. These pipelines were applied to both numerically simulated data and images acquired *in vivo*. Results from each pipeline were compared qualitatively by visual inspection and quantitatively via analysis of the regional χ bias and precision, and noise propagation.

Theory

Multi-echo combination

Previous studies employing multi-step reconstruction pipelines for QSM at 3 T have combined the signal from multiple echoes by either averaging (8-11) or fitting (12-15) before or after applying LBMs. For multi-echo combination in QSM, this study focuses on approaches based on weighted averaging (8-11) or complex nonlinear fitting (NLFit) (12,13) which outperform approaches based on unweighted averaging (16) or linear fitting (14).

Fitting

Multi-echo combination via nonlinear fitting (NLFit) has formulated the temporal evolution of the complex signal as a nonlinear least squares problem (12):

$$\Delta B_{Tot}^{*}(\boldsymbol{r}), \phi_0^{*}(\boldsymbol{r}) = \mathrm{argmin}_{\Delta B_{Tot}(\boldsymbol{r}), \phi_0(\boldsymbol{r})} \left\| S(\boldsymbol{r}, TE_i) - M(\boldsymbol{r}, TE_i) \exp\left(i\phi(\boldsymbol{r}, TE_i)\right) \right\|_2^2 \#(6)$$

where S denotes the acquired complex signal, M the signal magnitude, ϕ the signal phase (Equation 1), and TE_i the i-th echo time. This approach aims to mitigate noise in ΔB_{Tot}, by correctly modeling as normally distributed the noise in the real and imaginary parts of the complex signal. Unlike weighted-averaging-based approaches, NLFit enables estimating ϕ_0. Notably, the input phase to nonlinear fitting is minimally processed as it only requires to be temporally unwrapped, thus avoiding the application of LBMs before multi-echo combination.

Weighted averaging

Multi-echo combination (with n echoes) via weighted averaging has been performed using either TE-based weighting factors (8,9) or phase SNR-based weighting factors (10).

TE-based weighted averaging (TE-wAvg) (8,9) accounts for the phase at shorter TEs being affected by larger noise levels than at longer TEs. TE-wAvg calculates a combined ΔB_{Tot} as:

$$\Delta B_{Tot}(\boldsymbol{r}) = \frac{1}{\gamma}\sum_{i=1}^{n} w(\boldsymbol{r}, TE_i)\frac{\phi(\boldsymbol{r}, TE_i)}{TE_i} \#(7)$$

with weights equal to:

$$w(\boldsymbol{r}, TE_i) = \frac{TE_i}{\sum_{j=1}^{n} TE_j} \#(8)$$

TE-wAvg requires temporal unwrapping of the input multi-echo phase, resulting in a combined ΔB_{Tot}, which still contains ΔB_{Bg} contributions.

SNR-based weighted averaging (SNR-wAvg) (10) accounts for different tissue types reaching optimal SNR at different TEs and calculates a combined local field map (ΔB_{Loc}) as:

$$\Delta B_{Loc}(\boldsymbol{r}) = \frac{1}{\gamma}\sum_{i=1}^{n} w(\boldsymbol{r}, TE_i)\frac{\Delta B_{Loc}(\boldsymbol{r}, TE_i)}{TE_i} \#(9)$$

with weights equal to:

$$w(\boldsymbol{r}, TE_i) = \frac{TE_i \exp\left(-TE_i R_2^*(\boldsymbol{r})\right)}{\sum_{j=1}^{n} TE_j \exp\left(-TE_j R_2^*(\boldsymbol{r})\right)} \#(10)$$

In Equation 10, R_2^* denotes the map of voxel-wise transverse relaxation rates, which is related to the signal magnitude M by:

$$M(\boldsymbol{r}, TE) = M_0(\boldsymbol{r}) \exp(-R_2^*(\boldsymbol{r})TE) \#(11)$$

with M_0 the initial transverse magnetization. SNR-wAvg requires performing temporal and spatial unwrapping as well as background field removal on the input multi-echo phase, resulting in a background-field-free field map.

Alternatively, a distinct χ map has been calculated at each TE and multi-echo combination of χ over time has been performed via weighted averaging using magnitude-based weighting factors (Susc-wAvg) (11). Based on the inversely proportionality of the phase noise and magnitude SNR (17), this method aims to improve the SNR of the combined χ map as:

$$\chi(\boldsymbol{r}) = \sum_{i=1}^{n} w(\boldsymbol{r}, TE_i)\chi(\boldsymbol{r}, TE_i) \#(12)$$

with weights equal to:

$$w(\boldsymbol{r}, TE_i) = \frac{M^2(\boldsymbol{r}, TE_i)TE_i^2}{\sum_{j=1}^{n}\left(M^2\left(\boldsymbol{r}, TE_j\right)TE_j^2\right)} \#(13)$$

Noise propagation

Previous studies using fitting (13) or SNR-wAvg (10) have calculated expressions for the noise in the total field map ($\sigma(\Delta B_{Tot})$). For TE-wAvg or SNR-wAvg, expressions for $\sigma(\Delta B_{Tot})$ have not been calculated and were, therefore, derived here. For each multi-echo combination method, expressions for noise propagation from ΔB_{Tot} to the corresponding ΔB_{Loc} and χ images were also derived here.

Based on error propagation, the noise in ΔB_{Tot} calculated using a linear least squares fitting approach (13) is (see Supporting Information):

$$\sigma\left(\Delta B_{Tot}^{Fit}\right) = \frac{\sqrt{\sum_{i=1}^{n} \frac{1}{\sigma^2\left(\phi(TE_i)\right)}}}{\gamma\sqrt{\sum_{i=1}^{n} \frac{1}{\sigma^2\left(\phi(TE_i)\right)} \sum_{i=1}^{n} \frac{TE_i^2}{\sigma^2\left(\phi(TE_i)\right)} - \left(\sum_{i=1}^{n} \frac{TE_i}{\sigma^2\left(\phi(TE_i)\right)}\right)^2}} \#(14)$$

Equation 14 also corresponds to the a priori noise estimate found in nonlinear least squares fitting (12), thus $\sigma\left(\Delta B_{Tot}^{NLFit}\right) = \sigma\left(\Delta B_{Tot}^{Fit}\right)$.

Based on Equations 7-10, ΔB_{Tot} calculated using TE-wAvg and ΔB_{Loc} calculated using SNR-wAvg respectively have variances equal to:

$$\sigma^2\left(\Delta B_{Tot}^{TE-wAvg}\right) = \sigma^2\left(\frac{\sum_{i=1}^{n} \phi(TE_i)}{\gamma \sum_{i=1}^{n} TE_i}\right) \#(15)$$

$$\sigma^2\left(\Delta B_{Loc}^{SNR-wAvg}\right) = \sigma^2\left(\frac{1}{n\gamma} \sum_{i=1}^{n} \frac{TE_i \exp(-TE_i R_2^*)}{\sum_{j=1}^{n} TE_j \exp\left(-TE_j R_2^*\right)} \frac{\phi(TE_i)}{TE_i}\right) \#(16)$$

Assuming that noise in the single-echo phase is temporally uncorrelated and based on error propagation, the noise in ΔB_{Tot}/ΔB_{Loc} calculated using TE-wAvg/SNR-wAvg is respectively equal to:

$$\sigma\left(\Delta B_{Tot}^{TE-wAvg}\right) = \frac{\sum_{i=1}^{n} \sigma\left(\phi(TE_i)\right)}{\gamma \sum_{i=1}^{n} TE_i} \#(17)$$

$$\sigma\left(\Delta B_{Loc}^{SNR-wAvg}\right) = \frac{1}{\gamma} \sum_{i=1}^{n} \frac{\exp(-TE_i R_2^*)}{\sum_{j=1}^{n} TE_j \exp\left(-TE_j R_2^*\right)} \sigma\left(\phi(TE_i)\right) \#(18)$$

Equations 17 and 18 omit the $\boldsymbol{r}$ dependency, because they combine the multi-echo phase voxel by voxel.

Noise in the local field map

Based on error propagation and the orthogonality (18) of ΔB_{Loc} and the ΔB_{Bg} in the ROI, e.g., the brain:

$$\sigma(\Delta B_{Loc}) = \sqrt{\sigma^2(\Delta B_{Tot}) - \sigma^2\left(\Delta B_{Bg}\right)} \Leftrightarrow \sigma(\Delta B_{Loc}) \leq \sigma(\Delta B_{Tot}) \#(19)$$

where ⇔ denotes an "if and only if" relationship, and $\sigma(\Delta B_{Loc}) = \sigma(\Delta B_{Tot})$ is the worst-case scenario.

Noise in the susceptibility map

Owing to the circular convolution theorem, the deconvolution operation in Equation 3 can be performed by point-wise division in the Fourier domain. Thus, if the regularized inverse dipole kernel in k-space $\widetilde{D}^{-1}$ can be analytically derived independent of the Fourier transforms of χ or ΔB_{Loc}, $\sigma(\chi)$ can be calculated as (see Supporting Information):

$$\sigma\big(\chi(\boldsymbol{r})\big) = \frac{1}{B_0}\sqrt{FT^{-1}\left(\widetilde{D}^{-2}(\boldsymbol{k})FT\left(\sigma^2\big(\Delta B_{Loc}(\boldsymbol{r})\big)\right)\right)} \#(20)$$

where FT and FT^{-1} respectively denote the direct and inverse Fourier transforms, and $\boldsymbol{k}$ denotes k-space coordinates. Deriving an analytical expression for $\widetilde{D}^{-1}$ is possible, for example, when considering thresholded k-space division (TKD), or the Tikhonov-regularized minimal norm solution (13,19,20).

For weighted averaging of TE-dependent χ (Equations 12 and 13), based on phase error propagation over time and Equation 20, $\sigma\big(\Delta B_{Loc}(TE_i)\big)$ equals:

$$\sigma\big(\Delta B_{Loc}(TE_i)\big) = \frac{\sigma\big(\phi(TE_i)\big)}{\gamma TE_i} \#(21)$$

Based on Equations 12, 13, 20, and 21 (see Supporting Information), $\sigma(\chi^{Susc-wAvg})$ is the square root of:

$$\begin{aligned}
\sigma^2(\chi^{Susc-wAvg}) = \\
&= \frac{1}{(B_0\gamma n)^2}\sum_{i=1}^{n}\chi^2(TE_i)\left(\frac{4M^2(TE_i)TE_i^4\sigma^2\big(M(TE_i)\big)}{\sum_{j=1}^{n}\big(M^2(TE_j)TE_j^2\big)^2}\right.\\
&+ \frac{M^4(TE_i)TE_i^4}{\sum_{j=1}^{n}\big(M^2(TE_j)TE_j^2\big)^4}\sum_{j=1}^{n}4M^2(TE_j)TE_j^4\sigma^2\left(M(TE_j)\right)\\
&\left.- \frac{2M^2(TE_i)TE_i^2}{\big(\sum_{j=1}^{n}M^2(TE_j)TE_j^2\big)^3}4M^2(TE_i)TE_i^4\sigma^2\big(M(TE_i)\big)\right)\\
&+ \left(\frac{M^2(TE_i)TE_i^2}{\sum_{j=1}^{n}M^2(TE_j)TE_j^2}\right)^2\sigma^2\big(\chi(TE_i)\big)
\end{aligned} \quad (22)$$

Notably, Equation 22 requires analytically describing $\sigma(\chi)$ (Equation 20).

Methods

Where not otherwise stated, image analysis was performed on a 64-bit Windows 11 Pro operating system (Intel(R) Core(TM) i5-9400 CPU@2.90GHz processor; 16 GB RAM) using Matlab (R2021b, The MathWorks, Natick, MA). Preliminary versions of this study were presented at the 2016 and 2018 annual meetings of the International Society for Magnetic Resonance in Medicine (21,22).

In vivo data acquisition

Multi-echo 3D GRE imaging of ten healthy volunteers (average age/age range: 26/22-30 years, five females) was performed in two centers (University College London Hospital and Queen Square Multiple Sclerosis Centre, University College London) equipped with the same 3 T MRI system (Philips Achieva, Philips Healthcare, Best, NL; 32-channel head coil). Five subjects were acquired in each center. All the volunteers provided written informed consent, and the local research ethics committees approved the experimental sessions. Images were acquired using a transverse orientation, field of view=240x240x144 mm^3, voxel size=1-mm isotropic, flip angle=20□, repetition time=29 ms, five evenly spaced echoes (TE_1/TE spacing=3/5.4 ms), bandwidth=270 Hz/pixel, SENSE (23) factors=2/1.5, flyback gradients=on, no flow compensating gradients, total scan duration =04:37 min:s (24). Five subjects were scanned twice within the same session without repositioning to enable image-based calculation of magnitude and phase SNRs.

Data simulation from a numerical head phantom

To ensure the availability of ground-truth χ values against which to test the accuracy of QSM pipelines, a Zubal numerical head phantom was used (25) with the following regions of interest (ROIs): the caudate nucleus (CN), globus pallidus (GP), putamen (PU), thalamus (TH), superior sagittal sinus (SSS), grey and white matter (GM and WM), and cerebrospinal fluid (CSF). To match the acquisitions *in vivo*, the original 1.5-mm isotropic phantom was resampled to a 1-mm isotropic resolution with matrix size=384x384x192 voxels. Compared to our previous study (26), the numerical phantom was updated to achieve realistic regional means ± standard deviations (SDs) for χ, the proton density (M_0), and the transverse relaxation rate ($T_2^* = 1/R_2^*$) (see Supporting Information). Simulated multi-echo complex data were generated based on these ground-truth spatially variable χ, M_0 and T_2^* distributions, as:

$$S(\boldsymbol{r}, TE_i) = M(\boldsymbol{r}, TE_i)\exp\big(i\phi(\boldsymbol{r}, TE_i)\big) \#(23)$$

with ϕ and M respectively described by Equations 1 and 11, and TEs matched to the *in vivo* acquisitions. Random zero-mean Gaussian noise with a SD = 0.07 was added to the real

and imaginary parts of the noise-free signal independently (17,26). The random noise matrix was regenerated at each TE.

Data preprocessing

A brain mask was calculated for each subject by applying FSL BET (27,28) with robust brain center estimation (threshold=0.3) to the magnitude image at the longest TE. This choice of TE accounted for the greater amount of signal dropout near regions of high-χ gradients compared to shorter TEs.

A whole-brain mask for the Zubal phantom was calculated by applying FSL BET with robust brain center estimation (threshold=0.5) to the T_2^* map of the numerical phantom (26).

Processing pipelines for QSM

Five distinct processing pipelines (Figure 1) were applied to both the numerically simulated and the healthy volunteer data, and the time required to run each pipeline was measured using Matlab's stopwatch timer. Three of these pipelines (NLFit, TE-wAvg, and SNR-wAvg) combined the phase across TEs at different stages before performing the ΔB_{Loc}-to-χ inversion. Two other pipelines (Susc-wAvg and Susc-TGV-wAvg) first calculated a distinct χ map at each TE and then combined the χ maps. The following paragraphs describe each processing pipeline in detail.

The NLFit pipeline (26) first combined the complex GRE signal by nonlinear fitting over TEs (12) using the Cornell QSM software package's *Fit_ppm_complex* function (29). It then applied simultaneous spatial phase unwrapping and ΔB_{Bg} removal using Sophisticated Harmonic Artifact Reduction for Phase data (SHARP) (30), a direct solver of Equation 5 (30). SHARP was chosen because it has been widely used in the literature on QSM and is both robust and numerically efficient (2). Moreover, in our recent study comparing multi-echo and TE-dependent QSM, a multi-echo pipeline incorporating SHARP gave highly accurate multi-echo QSM values (26). SHARP was applied using the minimum-size 3-voxel isotropic 3D Laplacian kernel (30), a threshold for truncated singular value decomposition (tSVD) equal to 0.05, and a brain mask eroded by five voxels.

The TE-wAvg processing pipeline first applied Laplacian unwrapping to the phase at each TE using a threshold for tSVD equal to 10^{-10} (i.e., the default value in (29)). Second, it calculated ΔB_{Tot} by averaging the unwrapped phase according to Equations 7 and 8 (8,9). Then, it calculated ΔB_{Loc} by applying SHARP with the same parameters as in NLFit.

The SNR-wAvg processing pipeline first applied simultaneous phase unwrapping and ΔB_{Bg} removal to the phase at each TE using SHARP with the same parameters as in NLFit. An R_2^* map was calculated by voxel-wise fitting Equation 11 using Matlab's *nlinfit* function, where

initial values for R_2^* and M_0 were calculated by linearly fitting the log-linearized version of Equation 11. The SNR-wAvg pipeline then calculated ΔB_{Loc} by averaging the unwrapped and background-field-free phase according to Equations 9 and 10 (10).

In the NLFit, TE-wAvg, and SNR-wAvg pipelines, ΔB_{Loc}-to-χ inversion was performed using Tikhonov regularization with correction for susceptibility underestimation and using the L-curve method to determine the optimal value for the regularization parameter (13,30,31). This inversion method was chosen because it is computationally efficient and substantially reduces streaking artifacts relative to TKD (32).

The Susc-wAvg processing pipeline calculated a distinct χ map at each TE by applying simultaneous phase unwrapping and ΔB_{Bg} removal using SHARP as in NLFit and performing the ΔB_{Loc}-to-χ inversion using Tikhonov regularization as in NLFit, TE-wAvg and SNR-wAvg. This pipeline then calculated a combined χ map according to Equations 12 and 13 (11).

The Susc-TGV-wAvg processing pipeline applied one-step TGV (33) to the phase at each TE, and then calculated a combined χ map as in Susc-wAvg. The TGV method was tested because it avoids stair-casing artifacts in the resulting χ map while correctly preserving structural borders (33). Moreover, in our recent study comparing multi-echo and TE-dependent QSM, TGV provided highly accurate TE-dependent QSM images (26). TGV (v1.0.0_20210629) was run in Neurodesk (v20220302, https://neurodesk.github.io/) with the default parameter values $(\alpha_1, \alpha_0) = (0.0015, 0.005)$, which are optimal for medical imaging applications (33).

ROI segmentation in the healthy volunteer images

Regional χ values were compared within the simulated and *in vivo* data, similarly to our previous study (26). ROIs similar to those in the numerical phantom were segmented *in vivo*: the CN, GP, PU, TH, posterior corona radiata (PCR) as a WM ROI, and the straight sinus (StrS) as a venous ROI. Briefly, for each subject, the CN, GP, PU, TH and PCR were segmented based on the Eve χ atlas (34), whose GRE magnitude image was aligned to each subject's fifth-echo magnitude image using NiftyReg (35,36) (TE_{Eve}/TE_5=24/24.6 ms). The quality of ROI alignment was assessed by visual inspection. The ITK-SNAP active contour segmentation tool (37) was used to segment the StrS over several slices based on the fifth-echo magnitude image, which presented the best contrast between the StrS and the surrounding brain tissue.

Quantitative evaluation of the measured χ

In the numerical phantom simulations, each QSM pipeline's performance relative to the ground truth was visually assessed by calculating a difference image between the corresponding $\Delta B_{Loc}/\chi$ map and $\Delta B_{Loc}^{True}/\chi^{True}$. Here, ΔB_{Loc}^{True} referred to the ground-truth local field calculated using the reference scan method (18,38), and χ^{True} to the ground-truth magnetic susceptibility distribution with realistic regional means ± SDs of χ (Supporting Information Figure S1B). Means and SDs of χ were calculated for each pipeline in each ROI with $\chi^{True} \neq 0$, i.e., the CN, GP, PU, TH, SSS, and WM. The root mean squared errors (RMSEs) of both ΔB_{Loc} and χ relative to χ^{True} were also calculated throughout the brain volume (26). Notably, RMSEs for ΔB_{Loc} could not be calculated for the one-step Susc-TGV-wAvg pipeline.

In the volunteers, owing to the lack of a ground truth, representative susceptibility difference images were calculated relative to χ^{NLFit}, because the NLFit pipeline performed multi-echo combination at the earliest possible stage, and relative to $\chi^{TE-wAvg}$, because the TE-wAvg pipeline had the lowest local field RMSE in the numerical phantom simulations (see the Results). Regional means and SDs of χ were calculated for each processing pipeline and compared against χ values in subjects of a similar age from the QSM literature. RMSEs could not be calculated because of the lack of a ground truth. For visualization purposes, the pooled averages and SDs were calculated (26) after verifying that all *intra*subject SDs of χ were larger than the *inter*subject SD of χ.

Noise propagation maps

Only the healthy volunteers scanned twice were considered for this analysis. To enable image SNR calculation, in one healthy volunteer, five 20x20-voxel ROIs were drawn on a sagittal slice of the first-echo magnitude image (39), including both the GM and WM and excluding regions with artifacts induced by SENSE, motion, or flow. All five ROIs were applied across the other four volunteers by using rigid alignment transforms (NiftyReg (36)) between the first-echo magnitude images.

In each subject, ROI-based magnitude (M_{ROI}), magnitude noise ($\sigma_{ROI}(M)$) and phase noise values ($\sigma_{ROI}(\phi)$) were calculated based on the SNR difference method (40):

$$M_{ROI}(TE,\boldsymbol{r}) = \frac{1}{2} mean\big(M_1(TE,\boldsymbol{r}) + M_2(TE,\boldsymbol{r})\big), \qquad \boldsymbol{r} \in ROI \#(24)$$

$$\sigma_{ROI}\big(M(TE,\boldsymbol{r})\big) = \sqrt{\frac{R}{2}} SD\big(M_1(TE,\boldsymbol{r}) - M_2(TE,\boldsymbol{r})\big), \qquad \boldsymbol{r} \in ROI \ \#(25)$$

$$\sigma_{ROI}\big(\phi(TE,\boldsymbol{r})\big) = \sqrt{\frac{R}{2}}SD\big(\phi_1(TE,\boldsymbol{r}) - \phi_2(TE,\boldsymbol{r})\big), \qquad \boldsymbol{r} \in ROI \#(26)$$

Here, M_1/ϕ_1 and M_2/ϕ_2 respectively denote the magnitude/phase images from the first and second scan, $R = 3$ is the 2D SENSE factor calculated by multiplying the SENSE factors applied along the two phase encoding directions (41). The values calculated based on Equations 24-26 were averaged across the five ROIs, to calculate summary values of magnitude, magnitude noise, and phase noise at each TE.

For both the numerical phantom simulations and the healthy volunteers, a phase noise $\sigma(\phi)$ map was calculated at each TE as (17,42):

$$\sigma\big(\phi(TE,\boldsymbol{r})\big) = c(TE)\frac{1}{SNR(TE,\boldsymbol{r})} = c(TE)\frac{\sigma_{ROI}\big(M(TE,\boldsymbol{r})\big)}{M(TE,\boldsymbol{r})} \#(27)$$

where $c(TE)$ was a constant equal to 1 (by definition) in the numerical phantom simulations and equal to:

$$c(TE) = \frac{\sigma_{ROI}(TE)}{\frac{1}{SNR_{ROI}(M)}} = \frac{\sigma_{ROI}\big(\phi(TE)\big)\sigma_{ROI}\big(M(TE)\big)}{M(TE)} \#(28)$$

in the healthy volunteers. Notably, calculating phase noise analytically as in Equation 27 enabled the direct comparison of noise propagation between numerical simulations and data acquired *in vivo*.

For the NLFit, TE-wAvg and SNR-wAvg pipelines, the $\sigma(\Delta B_{Loc})$ map was calculated based on the multi-echo $\sigma(\phi)$ maps according to Equations 14 and 17-19. Then, the $\sigma(\chi)$ map was calculated according to Equation 20 with $B_0 = 3\ \mathrm{T}$ and the Tikhonov-regularized inverse magnetic dipole kernel with regularization parameter α (13):

$$\widetilde{D}^{-2}(\boldsymbol{k};\alpha) = \frac{D(\boldsymbol{k})}{D(\boldsymbol{k})^2 + \alpha} \#(29)$$

and correction for χ underestimation (30). For the Susc-wAvg pipeline, the $\sigma(\chi)$ map was calculated based on Equation 22. $\sigma(\chi)$ maps were not calculated for the Susc-TGV-wAvg pipeline, because TGV estimates χ iteratively (33) and an analytical expression for $\sigma(\chi)$ could not be derived.

To compare $\sigma(\chi)$ maps across pipelines, a line profile was traced in the same location of all $\sigma(\chi)$ images, and the $\sigma(\chi)$ value of each voxel along this profile was plotted. To compare the noise intensity and its variability between processing pipelines, the mean and SD of this representative line profile were also calculated.

Statistical analysis

Statistical analyses were performed based on the healthy subject data. For each ROI and each pair of pipelines, Bland-Altman analysis of the average χ was used to assess if pairs of pipelines systematically produced different results. For each ROI, statistically significant differences between pipelines were tested by considering the corresponding distributions of average χ values across subjects. To assess whether to apply parametric paired *t*-tests or nonparametric sign tests, the normal distribution of the differences between paired χ values was assessed using the Shapiro-Wilk test. All statistical tests were two-tailed, and an uncorrected *p*-value < 0.05 was considered significant.

Results

Pooling of χ measurements

For each ROI and each processing pipeline, all *intra*subject SDs of χ were larger than the *inter*subject SD. Thus, pooled means and SDs were calculated (26).

Performance of pipelines for multi-echo QSM

The Susc-wAvg and Susc-TGV-wAvg processing pipelines were the longest to run (Supporting Information Table S2), as they calculated a QSM map at each TE. The longer processing times required for the numerical phantom data were linked to the larger matrix size (384x384x192) compared to acquisitions *in vivo* (240x240x144).

In the numerical phantom, Figure 2 shows the ground-truth χ (A, G), the QSM images calculated by each processing pipeline (B-L), their difference relative to the ground truth (M-V), and the RMSEs of χ throughout the brain volume (bottom row). Analogous results for ΔB_{Loc} are shown in Supporting Information Figure S3. In the numerical phantom simulations, the χ map calculated using the NLFit pipeline had the largest RMSE (109.4%) followed, in decreasing order, by the SNR-wAvg (94.3%), TE-wAvg (93.5%), Susc-wAvg (91.5%) and Susc-TGV-wAvg pipelines (80.4%) (Figure 2, bottom row). The ΔB_{Loc} map calculated using the Susc-wAvg pipeline had the largest RMSE (average across TEs: 85.0%) followed, in decreasing order, by the NLFit (81.2%), TE-wAvg and Susc-wAvg pipelines (both 71.8%).

In the numerical phantom and for each processing pipeline, Figure 3A shows the regional means and SDs of χ. The error between the calculated and ground-truth χ appeared similar for all processing pipelines, although slightly larger SDs were always observed for the NLFit pipeline (Figures 2M-V and 3A). The SSS, which was the structure with the largest χ^{True}, showed the largest susceptibility errors for all processing pipelines (arrowheads in Figures 2G-L, R-V, and Figure 3A).

For one representative volunteer, Figures 4A-J show the susceptibility images calculated using each processing pipeline. Additionally, differences images are shown relative to the χ^{NLFit} (K-R) and the $\chi^{TE-wAvg}$ maps (S-Z). Here, susceptibility differences between processing pipelines were most prominent in the StrS (arrowheads in Figures 4F-J, O-R, W-Z). In the healthy volunteers, Figure 3B shows the pooled regional means and SDs of χ calculated by each processing pipeline. The average χ measured in the deep-GM ROIs and in the PCR had values within the ranges reported by previous studies (34,43-47): 0.01–0.13 ppm for the CN, 0.06–0.29 ppm for the GP, 0.02–0.14 ppm for the PU, -0.02–0.08 ppm for the TH, and -0.06–0.03 ppm for the PCR. In the StrS, only χ^{NLFit} had an average value close to the previously reported range for venous blood, namely, 0.17–0.58 ppm (32,46,48,49) (Figure 3B).

In the CN, GP, and venous ROIs, the regional χ measured *in vivo* had a relative accuracy across pipelines similar to the numerical phantom simulations: the NLFit and Susc-TGV-wAvg pipelines respectively provided the highest and lowest means of χ, whereas the TE-wAvg, SNR-wAvg, and Susc-wAvg pipelines provided intermediate values (Figure 3). Slightly different trends were observed in the PU, TH, and WM ROIs, where, *in vivo*, the NLFit and Susc-TGV-wAvg pipelines provided the lowest (absolute) means of χ, and the TE-wAvg, SNR-wAvg, and Susc-wAvg pipelines provided higher values (Figure 3B). In the numerical phantom simulations, the NLFit pipeline always had the largest SD of χ (Figure 3A). In contrast, *in vivo*, the NLFit pipeline had the smallest SD of χ in the CN, PU, TH, and PCR, and SDs of χ comparable to other pipelines in the GP and StrS (Figure 3B).

Phase noise propagation into the χ maps

Figures 5 and 6 show the estimated $\sigma(\chi)$ maps and profiles in the numerical phantom simulations and a representative healthy subject, respectively. All $\sigma(\chi)$ images contained some degree of streaking artifacts, especially *in vivo*, as these are a manifestation of error propagation from ΔB_{Loc} to χ caused by the dipole kernel null space (20). In the phantom, the NLFit and Susc-wAvg pipelines had the $\sigma(\chi)$ line profiles with the lowest means and SDs (Figures 5A, D, E, H, I). The NLFit and SNR-wAvg pipelines always resulted in the lowest streaking artifacts burden (Figures 5A, E, C, G, I and 6A, E, C, G, I). Streaking artifacts were more severe in the TE-wAvg and Susc-wAvg pipelines especially near high-χ venous structures.

Statistical Analysis

In the healthy volunteers, for all processing pipelines and ROIs, the Shapiro-Wilk test always rejected the hypothesis of normally distributed paired differences of χ. Therefore, pairwise

comparisons between pipelines were always evaluated using the nonparametric sign test. Significant differences between pipelines are shown in Figure 3B, whereas the between-pipeline biases are shown in Figure 7. A lower threshold equal to |0.01| ppm (|·| denotes the absolute value) was chosen because, below this level, *inter*pipeline differences cannot be disentangled from the *intra*pipeline variability (i.e., the regional SD).

A bias greater than |0.01| ppm was observed for the NLFit pipeline relative to all other pipelines in the CN, and StrS; for the Susc-TGV-wAvg pipeline relative to all other pipelines in the GP, and relative to the TE-wAvg and Susc-wAvg pipelines in the StrS (Figure 7). These results suggests that, for accurate χ quantification, some of the significant differences detected by the sign test, e.g., between the TE-wAvg, SNR-wAvg and Susc-wAvg pipelines, may be negligible (Figures 3B and 7).

Discussion

Aiming to elucidate the optimal strategy for multi-echo combination for QSM, this study compared multi-echo combination methods applied at different stages of the QSM processing pipeline before or after LBMs for phase unwrapping or background field removal. Each pipeline was applied to numerically simulated data and images from healthy volunteers.

The higher relative accuracy of the NLFit pipeline in the numerical phantom simulations and *in vivo* suggests that, for QSM, combining the temporally unwrapped multi-echo phase before applying LBMs for spatial phase unwrapping or ΔB_{Bg} removal is preferable to averaging the TE-dependent LBM-processed phase or χ. This suggestion appears to conflict with the higher RMSEs associated with the NLFit pipeline compared to other pipelines in the phantom simulations (Figure 2 and Supplementary Information Figure S3). However, the RMSE jointly reflects systematic and random errors, as it measures the bias between the estimated and true value, but also reflects the variability of the estimated χ relative to its average value (3). Thus, the RMSE must always be interpreted in combination with complementary measurements of bias and precision. Furthermore, RMSEs of ΔB_{Loc} are difficult to interpret. In contrast with RMSEs of χ, they allow to compare different pipelines without the effect of ΔB_{Loc}-to-χ inversion. However, they are voxel-based measures based on a signal that is intrinsically nonlocal, because ΔB_{Loc} variations extend beyond the anatomical region of χ shift that generated them (19). Thus, the best set of metrics for comparing images generated by a processing pipeline for QSM is still an active area of research (3,4).

There are several potential explanations as to why LBMs applied before multi-echo combination reduce the overall accuracy of QSM. Firstly, in contrast with path-based or

region-growing-based phase unwrapping, Laplacian phase unwrapping usually removes some ΔB_{Bg} components from the input phase image (30). Thus, the consistently lower accuracy of χ calculated using the TE-wAvg processing pipeline was probably driven by the incorrect assumption that the TE-dependent LBM-unwrapped phase corresponded to the true unwrapped phase. LBMs for ΔB_{Bg} removal also applied tSVD (with a larger truncation threshold) but here the high pass filtering effect was expected, as background fields are slowly varying. Finally, the similar accuracy and values of χ calculated using the TE-wAvg, SNR-wAvg and Susc-wAvg pipelines suggest a negligible difference between averaging the Laplacian unwrapped phase over TEs before (TE-wAvg pipeline) or after TE-dependent Laplacian ΔB_{Bg} removal (SNR-wAvg and Susc-wAvg pipelines).

In the numerical phantom simulations, all processing pipelines resulted in higher SDs of χ compared to the ground-truth, suggesting that noise in the multi-echo signal phase was amplified by all pipelines. This result is in line with the known noise amplification of ill-posed inverse problems. However, the estimated SDs of χ varied across pipelines. In the simulations, the Susc-TGV-wAvg pipeline had the smallest SDs of χ (Figure 3A). However, *in vivo*, the NLFit pipeline generally had the smallest SDs of χ (Figure 3B). Both the TGV reconstruction pipeline and shortcomings of the numerically simulated data could explain these discrepancies in the performance of the Susc-TGV-wAvg pipeline. The numerically simulated data were generated based on a digital phantom which, despite varying regional χ values in a realistic fashion (see Supporting Information) ultimately still appeared as a smooth piece-wise constant model (Figures 2A, G). As previously observed (3), piece-wise constant geometries allow good recovery of the underlying χ distribution using TGV-based algorithms, because the piece-wise constant constraints exactly match the underlying χ distribution. However, in regions with flow, anisotropic χ distributions, or microstructure, these numerical models are likely to depart from a realistic representation of the tissue χ. To overcome the limitations of this assumption, future studies could exploit a newly developed realistic head phantom for QSM, which does not have a piecewise constant χ distribution and incorporates microstructural effects (50).

In both numerical simulations and healthy volunteers, based on the line profiles traced on the noise maps and streaking artifact reduction, the NLFit pipeline had better noise mitigation (Figures 5 and 6). This result suggests that combining the temporally unwrapped multi-echo phase by nonlinear complex fitting, designed to account for noise in the complex signal (12), results in better noise management than combining the multi-echo phase by averaging. As previously shown (12), errors in the combined field map mainly result from both noise in the signal phase and phase unwrapping errors near high-χ regions (e.g., the

veins). Both sources of error were successfully managed by nonlinear complex fitting, as shown by the dramatic reduction of streaking artifacts in Figure 6A. In line with previous observations (51), this result also suggests that the regularization strategy employed by the ΔB_{Loc}-to χ step can mitigate artifactual streaking errors only when major sources of error in the field map have been tightly constrained. At visual inspection *in vivo*, the SNR-wAvg was the second-best pipeline for the mitigation of streaking artifacts (Figure 6). Thus, if applying nonlinear complex fitting is not possible, averaging using the SNR-weighting-based method could offer the best alternative for noise reduction.

All these indications do not necessarily apply to χ estimation in WM tissue. Indeed, due to WM's ordered microstructure, a comprehensive estimation of χ in WM requires acquiring GRE images at multiple head orientations and modelling χ as a tensor (52,53). Thus, further studies are needed to evaluate the applicability of these results to WM tissue.

In the present study, all experiments were limited to one field strength (i.e., 3 T). As tissue relaxation times (e.g., T_2^*) shorten with increasing field strength but the signal phase at a given TE increases, further work is needed to assess the relevance of these results at ultrahigh fields. Finally, it must be noted that in the numerical phantom simulations, all processing pipelines underestimated χ^{True} (Figure 3). However, in QSM, some degree of underestimation is always expected, because of the ill-posed nature of the ΔB_{Loc}-to-χ inverse problem (20).

Conclusions

The higher accuracy of regional χ values and better noise management of the NLFit pipeline suggest that, for QSM, combining the multi-echo phase by nonlinearly fitting over TEs before applying LBMs is preferable to combining the TE-dependent LBM-processed phase or χ by averaging.

Data availability statement

The Matlab code used to run the analyses in this study is available at https://github.com/emmabiondetti/multi-echo-qsm. The repository's main page lists all dependencies on code developed by other groups.

Acknowledgments

We thank all the healthy volunteers who participated in this study. We also thank Dr. Rosa Cortese and Dr. Floriana De Angelis (Queen Square MS Centre, UCL Queen Square Institute of Neurology, UCL, London, UK) for their help with the MRI scans. We are indebted to Prof. Claudia Gandini Wheeler-Kingshott (NMR Unit, Queen Square MS Centre, UCL

Queen Square Institute of Neurology, UCL, London, UK) for her support with the MRI scans performed at the Queen Square MS Centre, UCL.

References

1. 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(4).
2. 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(4).
3. Langkammer C, Schweser F, Shmueli K, Kames C, Li X, Guo L, Milovic C, Kim J, Wei H, Bredies K, Buch S, Guo Y, Liu Z, Meineke J, Rauscher A, Marques JP, Bilgic B. Quantitative susceptibility mapping: Report from the 2016 reconstruction challenge. Magn Reson Med 2018;79(3):1661-1673.
4. Q. S. M. Challenge 2.0 Organization Committee, Bilgic B, Langkammer C, Marques JP, Meineke J, Milovic C, Schweser F. QSM reconstruction challenge 2.0: Design and report of results. Magn Reson Med 2021;86(3):1241-1255.
5. Schofield MA, Zhu Y. Fast phase unwrapping algorithm for interferometric applications. Opt Lett 2003;28(14):1194-1196.
6. Cronin MJ, Wang N, Decker KS, Wei H, Zhu WZ, Liu C. Exploring the origins of echo-time-dependent quantitative susceptibility mapping (QSM) measurements in healthy tissue and cerebral microbleeds. Neuroimage 2017;149:98-113.
7. Sood S, Urriola J, Reutens D, O'Brien K, Bollmann S, Barth M, Vegh V. Echo time-dependent quantitative susceptibility mapping contains information on tissue properties. Magn Reson Med 2017;77(5):1946-1958.
8. Wei H, Dibb R, Zhou Y, Sun Y, Xu J, Wang N, Liu C. Streaking artifact reduction for quantitative susceptibility mapping of sources with large dynamic range. NMR Biomed 2015;28(10):1294-1303.
9. Li W, Wang N, Yu F, Han H, Cao W, Romero R, Tantiwongkosi B, Duong TQ, Liu C. A method for estimating and removing streaking artifacts in quantitative susceptibility mapping. Neuroimage 2015;108:111-122.
10. Wu B, Li W, Avram AV, Gho SM, Liu C. Fast and tissue-optimized mapping of magnetic susceptibility and T2* with multi-echo and multi-shot spirals. Neuroimage 2012;59(1):297-305.

11. Chen Y, Liu S, Wang Y, Kang Y, Haacke EM. STrategically Acquired Gradient Echo (STAGE) imaging, part I: Creating enhanced T1 contrast and standardized susceptibility weighted imaging and quantitative susceptibility mapping. Magn Reson Imaging 2018;46:130-139.
12. 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(2):467-476.
13. 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(2):273-281.
14. 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(4):1003-1009.
15. Karsa A, Punwani S, Shmueli K. An optimized and highly repeatable MRI acquisition and processing pipeline for quantitative susceptibility mapping in the head-and-neck region. Magn Reson Med 2020;84(6):3206-3222.
16. Qiu D, Brown RC, Sun B, Palasis S, Burns TG, Jones RA. Abnormal Iron Levels in the Brain of Pediatric Sickle Cell Disease Patients: a Study using Quantitative Susceptibility Mapping (QSM). Proceedings of the International Society for Magnetic Resonance in Medicine 2014;22:897.
17. Conturo TE, Smith GD. Signal-to-noise in phase angle reconstruction: Dynamic range extension using phase reference offsets. Magnetic Resonance in Medicine 1990;15(3):420-437.
18. Liu T, Khalidov I, de Rochefort L, Spincemaille P, Liu J, Tsiouris AJ, Wang Y. A novel background field removal method for MRI using projection onto dipole fields (PDF). NMR Biomed 2011;24(9):1129-1136.
19. Shmueli K, de Zwart JA, van Gelderen P, Li TQ, Dodd SJ, Duyn JH. Magnetic susceptibility mapping of brain tissue in vivo using MRI phase data. Magn Reson Med 2009;62(6):1510-1522.
20. Wang Y, Liu T. Quantitative susceptibility mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. Magn Reson Med 2015;73(1):82-101.
21. Biondetti E, Thomas DL, Shmueli K. Application of Laplacian-based Methods to Multi-echo Phase Data for Accurate Susceptibility Mapping. Proceedings of the International Society for Magnetic Resonance in Medicine 2016;24:1547.
22. Biondetti E, Karsa A, Thomas DL, Shmueli K. Evaluating the Precision of Multi-Echo Combination Methods for Susceptibility Mapping by Analysing the Propagation of

Single-Echo Phase Noise into Multi-Echo Field and Susceptibility Maps. Proceedings of the International Society for Magnetic Resonance in Medicine 2018;26:2191.

23. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: Sensitivity encoding for fast MRI. Magnetic Resonance in Medicine 1999;42(5):952-962.
24. Berg RC, Preibisch C, Thomas DL, Shmueli K, Biondetti E. Investigating the effect of flow compensation and quantitative susceptibility mapping method on the accuracy of venous susceptibility measurement. Neuroimage 2021;240:118399.
25. Zubal IG, Harrell CR, Smith EO, Rattner Z, Gindi G, Hoffer PB. Computerized three-dimensional segmented human anatomy. Med Phys 1994;21(2):299-302.
26. 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.
27. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. Fsl. Neuroimage 2012;62(2):782-790.
28. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp 2002;17(3):143-155.
29. Cornell MRI Research Group. http://pre.weill.cornell.edu/mri/pages/qsm.html. Updated January 15, 2020. Accessed January 21, 2022.
30. Schweser F, Deistung A, Sommer K, Reichenbach JR. Toward online reconstruction of quantitative susceptibility maps: superfast dipole inversion. Magn Reson Med 2013;69(6):1582-1594.
31. Hansen PC, O'Leary DP. The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems. SIAM Journal on Scientific Computing 1993;14(6):1487-1503.
32. Biondetti E, Rojas-Villabona A, Sokolska M, Pizzini FB, Jager HR, Thomas DL, Shmueli K. Investigating the oxygenation of brain arteriovenous malformations using quantitative susceptibility mapping. Neuroimage 2019;199:440-453.
33. Langkammer C, Bredies K, Poser BA, Barth M, Reishofer G, Fan AP, Bilgic B, Fazekas F, Mainero C, Ropele S. Fast quantitative susceptibility mapping using 3D EPI and total generalized variation. Neuroimage 2015;111:622-630.
34. Lim IA, Faria AV, Li X, Hsu JT, Airan RD, Mori S, van Zijl PC. Human brain atlas for automated region of interest selection in quantitative susceptibility mapping: application to determine iron content in deep gray matter structures. Neuroimage 2013;82:449-469.
35. Modat M, Ridgway GR, Taylor ZA, Lehmann M, Barnes J, Hawkes DJ, Fox NC, Ourselin S. Fast free-form deformation using graphics processing units. Comput Methods Programs Biomed 2010;98(3):278-284.

36. Ourselin S, Roche A, Subsol G, Pennec X, Ayache N. Reconstructing a 3D structure from serial histological sections. Image and Vision Computing 2001;19(1-2):25-31.
37. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31(3):1116-1128.
38. 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(1):196-204.
39. Lerski RA, de Wilde J, Boyce D, Ridgway J. Quality control in magnetic resonance imaging. York, UK: The Institute of Physics and Engineering in Medicine; 1998.
40. Dietrich O, Raya JG, Reeder SB, Reiser MF, Schoenberg SO. Measurement of signal-to-noise ratios in MR images: influence of multichannel coils, parallel imaging, and reconstruction filters. J Magn Reson Imaging 2007;26(2):375-385.
41. Weiger M, Pruessmann KP, Boesiger P. 2D SENSE for faster 3D MRI. MAGMA 2002;14(1):10-19.
42. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magn Reson Med 1995;34(6):910-914.
43. Li W, Wu B, Liu C. Quantitative susceptibility mapping of human brain reflects spatial variation in tissue composition. Neuroimage 2011;55(4):1645-1656.
44. Wharton S, Bowtell R. Whole-brain susceptibility mapping at high field: a comparison of multiple- and single-orientation methods. Neuroimage 2010;53(2):515-525.
45. Bilgic B, Pfefferbaum A, Rohlfing T, Sullivan EV, Adalsteinsson E. MRI estimates of brain iron concentration in normal aging using quantitative susceptibility mapping. Neuroimage 2012;59(3):2625-2635.
46. 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.
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(4):2789-2807.
48. Fan AP, Bilgic B, Gagnon L, Witzel T, Bhat H, Rosen BR, Adalsteinsson E. Quantitative oxygenation venography from MRI phase. Magn Reson Med 2014;72(1):149-159.

49. Cetin S, Bilgic B, Fan A, Holdsworth S, Unal G. Vessel Orientation Constrained Quantitative Susceptibility Mapping (QSM) Reconstruction. Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, editors. Cham: Springer International Publishing; 2016.
50. Marques JP, Meineke J, Milovic C, Bilgic B, Chan KS, Hedouin R, van der Zwaag W, Langkammer C, Schweser F. QSM reconstruction challenge 2.0: A realistic in silico head phantom for MRI data simulation and evaluation of susceptibility mapping procedures. Magn Reson Med 2021;86(1):526-542.
51. Acosta-Cabronero J, Milovic C, Mattern H, Tejos C, Speck O, Callaghan MF. A robust multi-scale approach to quantitative susceptibility mapping. Neuroimage 2018;183:7-24.
52. Li W, Liu C, Duong TQ, van Zijl PC, Li X. Susceptibility tensor imaging (STI) of the brain. NMR Biomed 2017;30(4).
53. Chen WC, Foxley S, Miller KL. Detecting microstructural properties of white matter based on compartmentalization of magnetic susceptibility. Neuroimage 2013;70:1-9.

Figure captions

Figure 1. Processing pipelines for multi-echo QSM. For each multi-echo combination method (NLFit [nonlinear phase fitting], TE-wAvg [TE-weighted phase averaging], SNR-wAvg [SNR-weighted phase averaging], Susc-wAvg and Susc-TGV-wAvg [magnitude-weighted susceptibility averaging]) the processing steps are described as separate processing streams.

Figure 2. χ maps calculated using distinct multi-echo combination methods in the numerical phantom simulations. The same transverse and sagittal slices are shown for the ground-truth susceptibility map **(A, G)**, and for the susceptibility maps calculated using NLFit **(B, H)**, TE-wAvg **(C, I)**, SNR-wAvg **(D, J)**, Susc-wAvg **(E, K)**, and Susc-TGV-wAvg (**F, L)**. The figure also shows the difference between each susceptibility map and the ground truth **(M-V)**. The bottom row shows the root mean squared errors (RMSEs) of χ for each pipeline. In all the sagittal images **(G-L, R-V)**, the yellow and blue arrowhead respectively point at the same posterior and anterior locations in the superior sagittal sinus.

Figure 3. Means and SDs of χ in the phantom and healthy volunteer ROIs. The means and SDs (error bars) of χ are shown in each ROI of the numerically simulated **(A)** and pooled healthy volunteer data **(B)** for each processing pipeline. In the numerical phantom the ground-truth χ is also shown. In the healthy volunteers, significant differences between pairs of pipelines are denoted using the symbols * (p-value < 0.05) and ** (p-value < 0.01). Abbreviations: CN, caudate nucleus; GP, globus pallidus; PCR, posterior corona radiata;

PU, putamen; StrS, straight sinus; TH, thalamus; WM, white matter; SSS, superior sagittal sinus.

Figure 4. χ maps calculated using distinct multi-echo combination methods in a representative healthy volunteer. The same transverse and sagittal slices are shown for the susceptibility maps calculated using NLFit **(A, F)** [nonlinear phase fitting], TE-wAvg **(B, G)** [TE-weighted phase averaging], SNR-wAvg **(C, H)** [SNR-weighted phase averaging], Susc-wAvg **(D, I)** and Susc-TGV-wAvg **(E, J)** [magnitude-weighted susceptibility averaging]. The figure also shows the differences between the TE-wAvg, SNR-wAvg, Susc-wAvg and Susc-TGV-wAvg maps and the NLFit map **(K-R)**, and the differences between the NLFit, SNR-wAvg, Susc-wAvg and Susc-TGV-wAvg maps and the TE-wAvg map **(S-Z)**. In all the sagittal images **(F-J, O-R, W-Z)**, the yellow arrowheads point at the same location in the StrS.

Figure 5. Susceptibility noise profiles in the numerical phantom simulations. The same sagittal slice is shown for susceptibility noise ($\sigma(\chi)$) images calculated using the NLFit, TE-wAvg, SNR-wAvg and Susc-wAvg pipelines **(A-D)**. The susceptibility noise is plotted for a line profile traced on the $\sigma(\chi)$ images **(E-H)** including the high-χ SSS. All line profiles are also shown combined in the same plot **(I)**. The mean and the standard deviation of $\sigma(\chi)$ are shown for each line profile.

Figure 6. Susceptibility noise profiles in a representative healthy volunteer. The same sagittal slice is shown for susceptibility noise ($\sigma(\chi)$) images calculated using the NLFit, TE-wAvg, SNR-wAvg and Susc-wAvg pipelines **(A-D)**. The susceptibility noise is plotted for a line profile traced on the $\sigma(\chi)$ images **(E-H)** including the high-χ StrS. All line profiles are also shown combined in the same plot **(I)**. The mean and the standard deviation of $\sigma(\chi)$ are shown for each line profile.

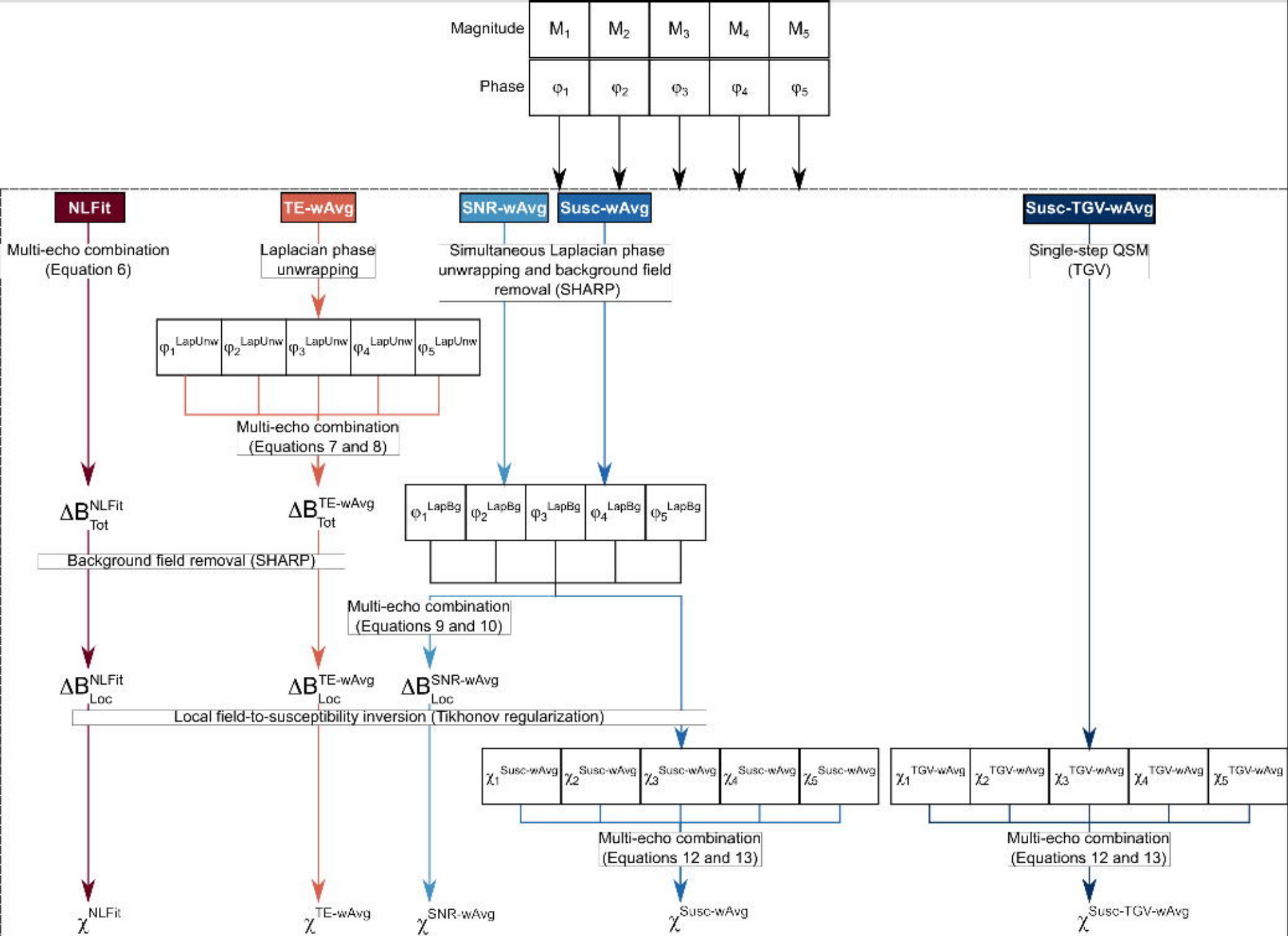
Figure 7. Bias between multi-echo pipelines for QSM in each ROI. The mean and SDs (error bars) of the bias are shown in each healthy volunteer ROI for all pairs of multi-echo processing pipelines. The grey band denotes the [-0.01 – 0.01] ppm interval. If the mean of the bias was within this interval, the difference between the corresponding pair of QSM pipelines was considered negligible.

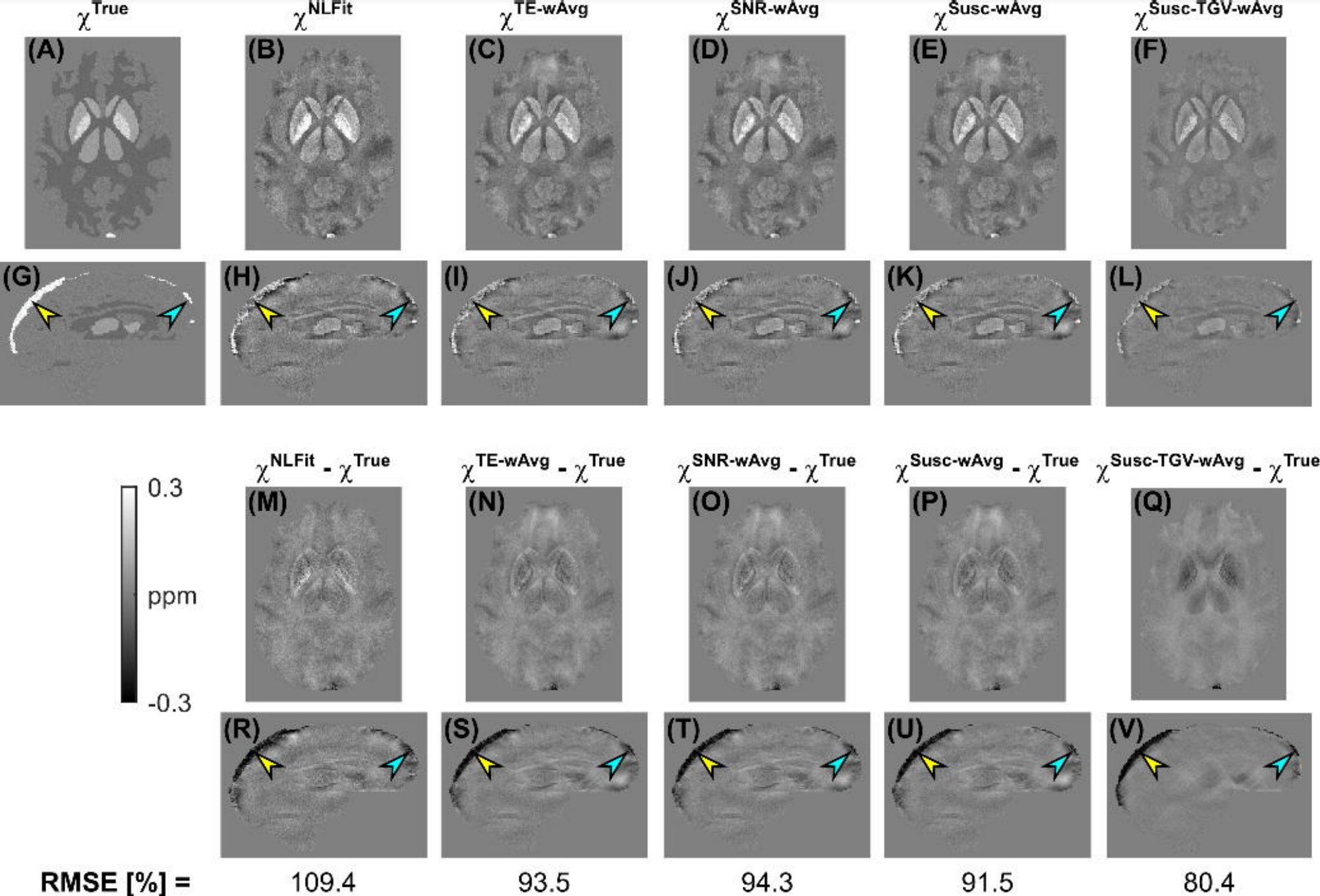
Supporting Information Figure S1. Properties of the numerical phantom. M_0 in arbitrary units (a. u.), T_2^* in ms and χ in parts per million (ppm) assigned to various ROIs in the numerical phantom are shown in **(A)**. The location of these ROIs is shown in the χ **(B)** and T_2^* maps **(C)** of the numerical phantom.

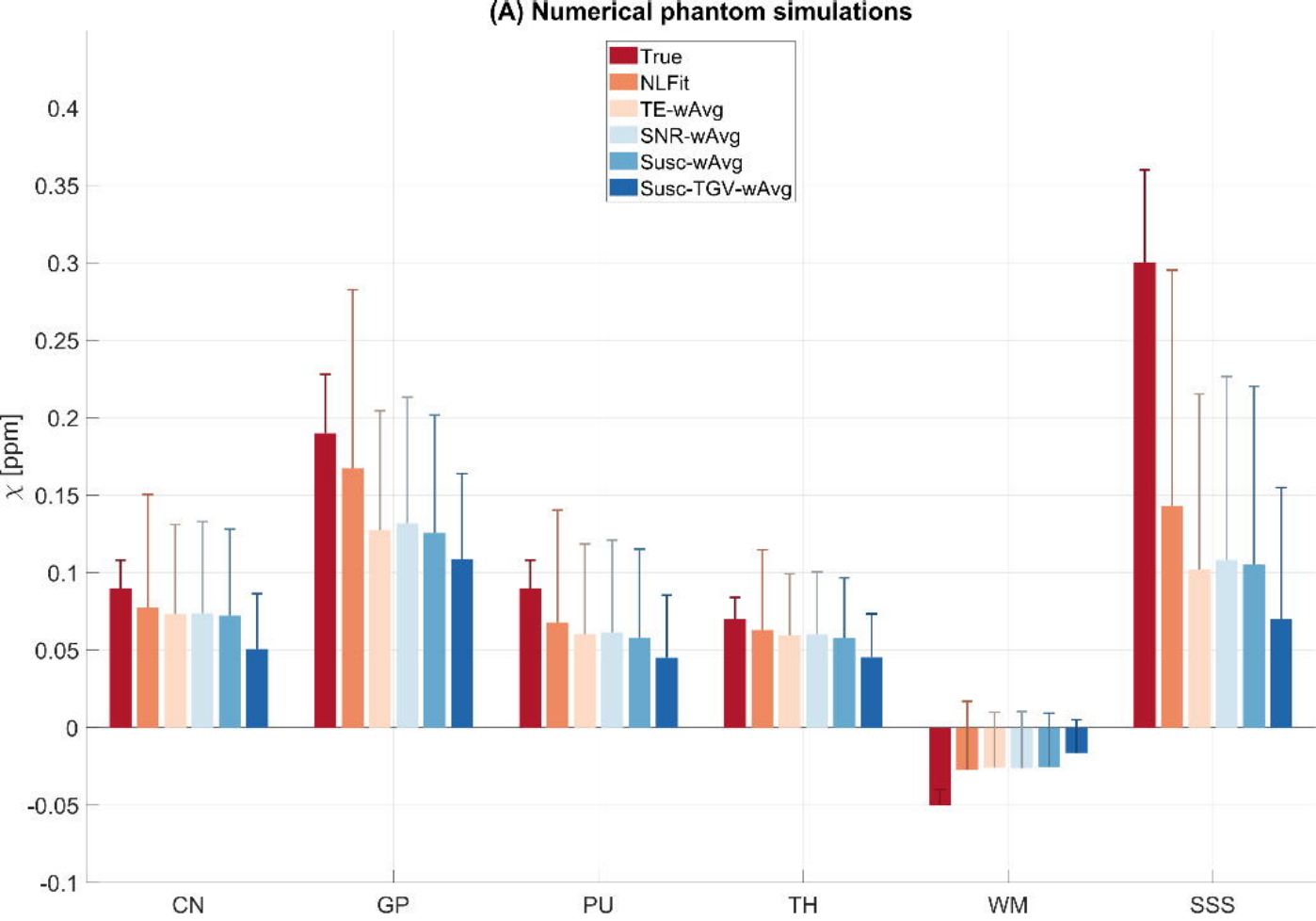
Supporting Information Table S2. Total image processing time. The table shows the time required to run each pipeline for the numerical phantom simulations and data acquired *in*

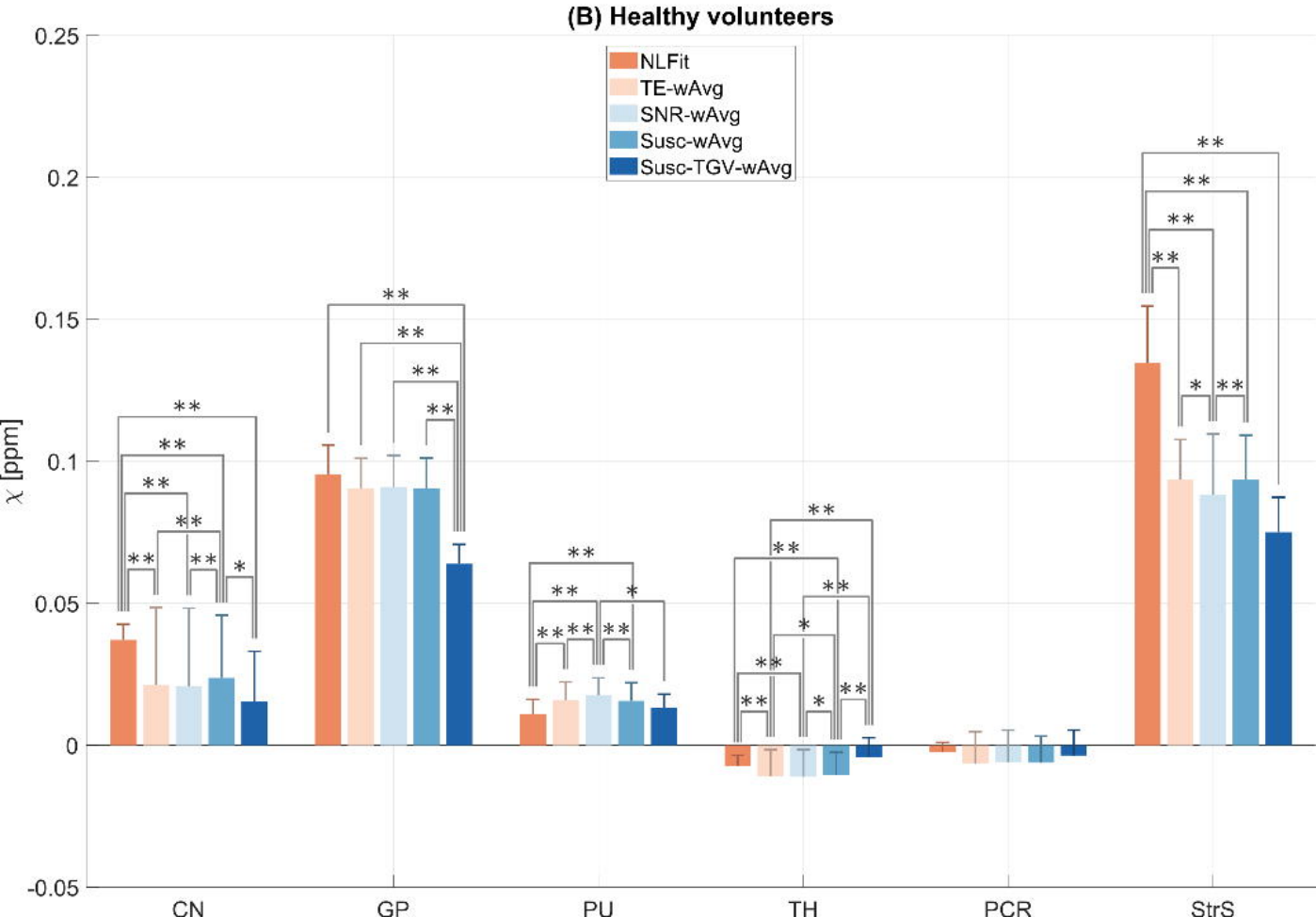
vivo. The time reported for the SNR-wAvg pipeline does not include the time required for R_2^* mapping, as optimizing this step was outside the scope of the present study. For the Susc-TGV-wAvg pipeline, approximate timings are reported because image reconstruction was performed using first Neurodesk (QSM calculation at each TE) and then Matlab (multi-echo combination).

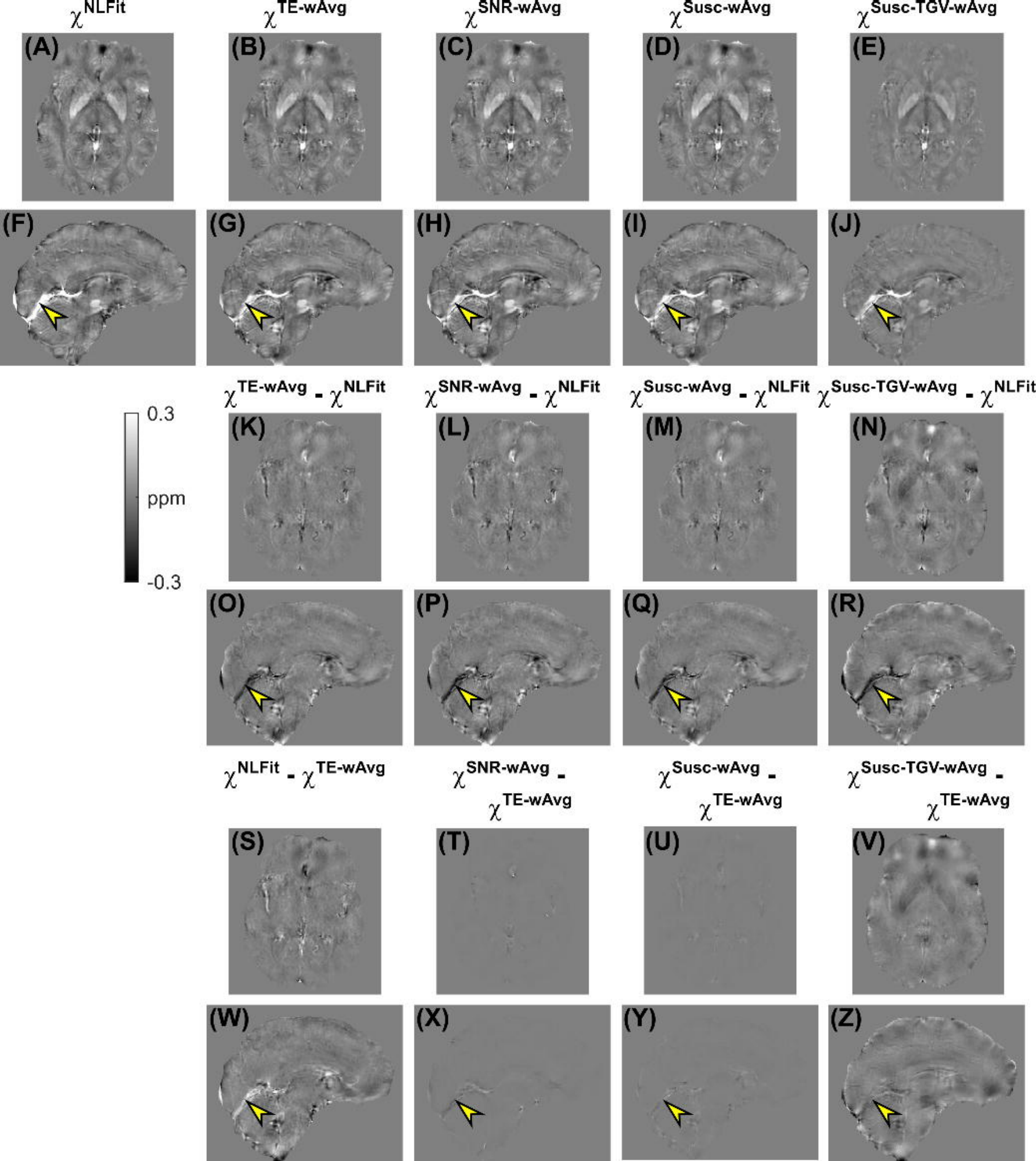
Supporting Information Figure S3. ΔB_{Loc} maps calculated using distinct multi-echo combination methods in the numerical phantom simulations. The same transverse and sagittal slices are shown for the ground-truth local field map **(A, G)**, and for the local field maps calculated using NLFit **(B, H)**, TE-wAvg **(C, I)**, SNR-wAvg **(D, J)**, and Susc-wAvg at each TE **(E, O)**. The figure also shows the difference between each local field map and the ground truth **(P-E2)**. The bottom row shows the root mean squared errors (RMSEs) of ΔB_{Loc} for each pipeline.

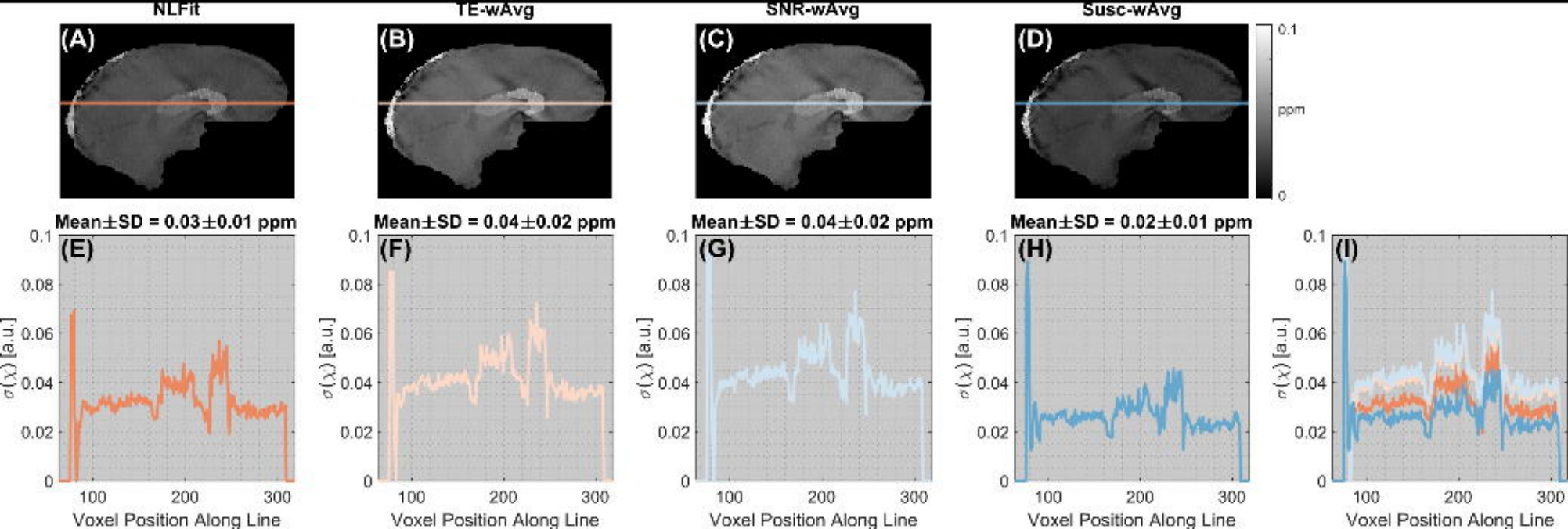

Magnitude
M1 M2 M3 M4 M5
Phase
φ1 φ2 φ3 φ4 φ5
NLFit
TE-wAvg
SNR-wAvg
Susc-wAvg
Susc-TGV-wAvg
Multi-echo combination (Equation 6)
Laplacian phase unwrapping
Simultaneous Laplacian phase unwrapping and background field removal (SHARP)
Single-step QSM (TGV)
φ1 LapUnw φ2 LapUnw φ3 LapUnw φ4 LapUnw φ5 LapUnw
Multi-echo combination (Equations 7 and 8)
ΔB Tot NLFit
ΔB Tot TE-wAvg
φ1 LapBg φ2 LapBg φ3 LapBg φ4 LapBg φ5 LapBg
Background field removal (SHARP)
Multi-echo combination (Equations 9 and 10)
ΔB Loc NLFit
ΔB Loc TE-wAvg
ΔB Loc SNR-wAvg
Local field-to-susceptibility inversion (Tikhonov regularization)
χ1 Susc-wAvg χ2 Susc-wAvg χ3 Susc-wAvg χ4 Susc-wAvg χ5 Susc-wAvg
χ1 TGV-wAvg χ2 TGV-wAvg χ3 TGV-wAvg χ4 TGV-wAvg χ5 TGV-wAvg
Multi-echo combination (Equations 12 and 13)
Multi-echo combination (Equations 12 and 13)
χ NLFit
χ TE-wAvg
χ SNR-wAvg
χ Susc-wAvg
χ Susc-TGV-wAvg

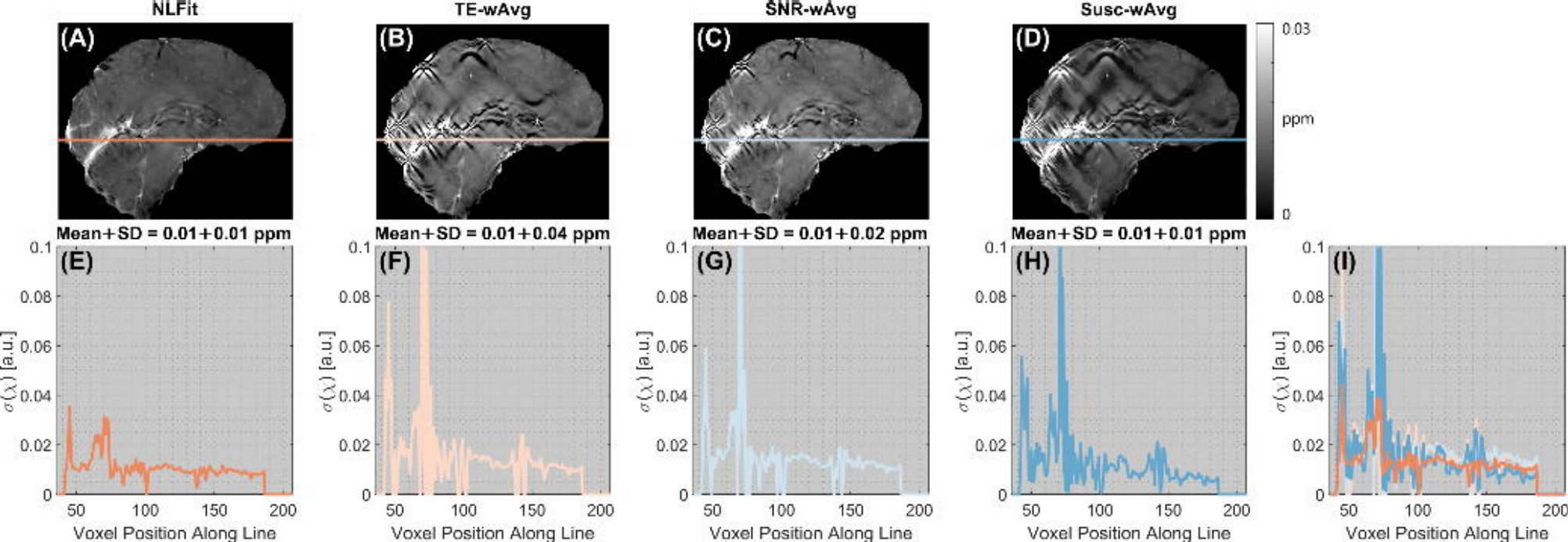

χ^True
χ^NLFit
χ^TE-wAvg
χ^SNR-wAvg
χ^Susc-wAvg
χ^Susc-TGV-wAvg
(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
(I)
(J)
(K)
(L)
0.3
ppm
-0.3
χ^NLFit - χ^True
χ^TE-wAvg - χ^True
χ^SNR-wAvg - χ^True
χ^Susc-wAvg - χ^True
χ^Susc-TGV-wAvg - χ^True
(M)
(N)
(O)
(P)
(Q)
(R)
(S)
(T)
(U)
(V)
RMSE [%] =
109.4
93.5
94.3
91.5
80.4

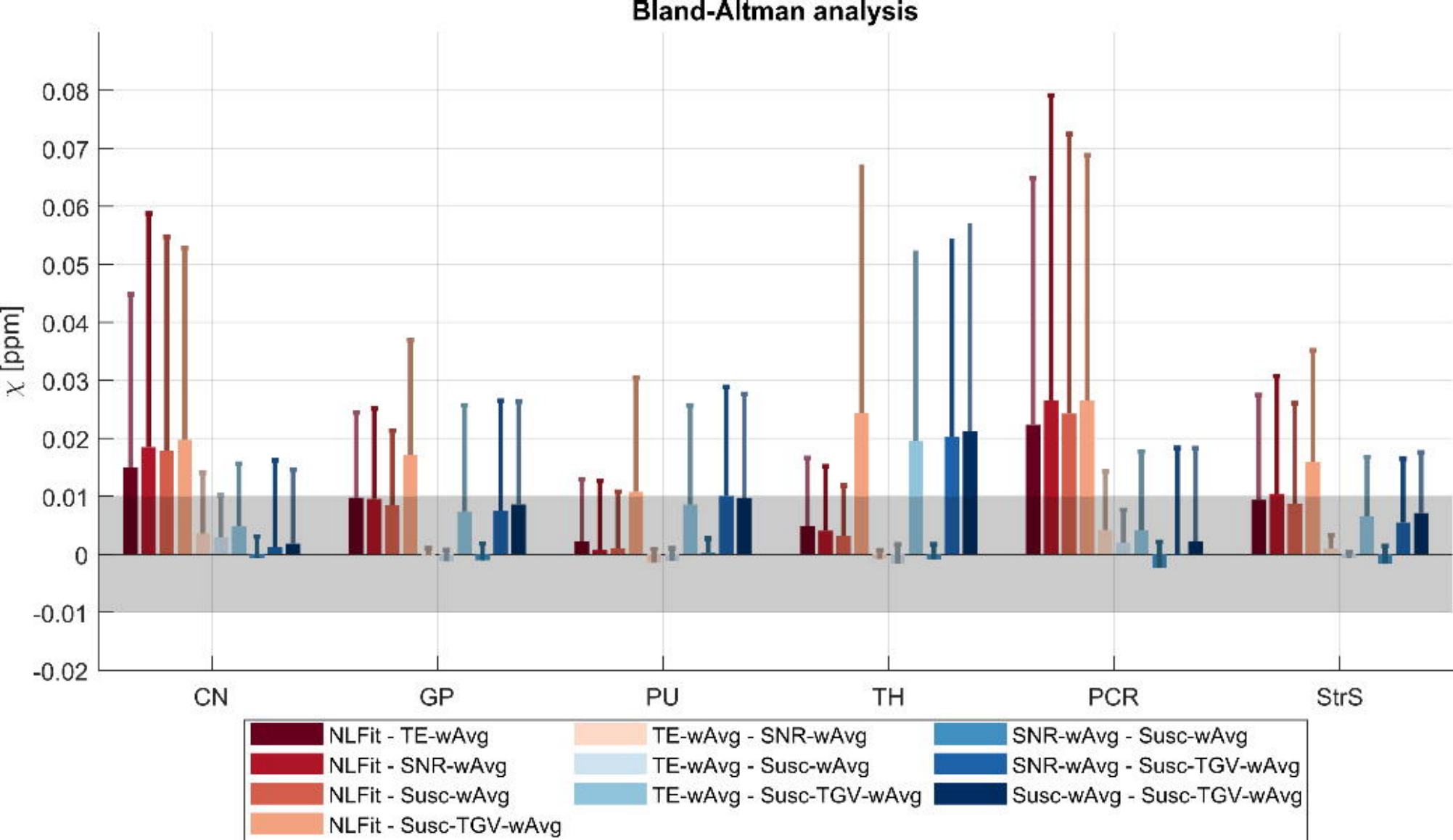

(A) Numerical phantom simulations
True
NLFit
TE-wAvg
SNR-wAvg
Susc-wAvg
Susc-TGV-wAvg
χ [ppm]
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
CN
GP
PU
TH
WM
SSS


(B) Healthy volunteers
NLFit
TE-wAvg
SNR-wAvg
Susc-wAvg
Susc-TGV-wAvg
χ [ppm]
0.25
0.2
0.15
0.1
0.05
0
-0.05
CN
GP
PU
TH
PCR
StrS


χNLFit
χTE-wAvg
χSNR-wAvg
χSusc-wAvg
χSusc-TGV-wAvg
(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
(I)
(J)
χTE-wAvg - χNLFit
χSNR-wAvg - χNLFit
χSusc-wAvg - χNLFit
χSusc-TGV-wAvg - χNLFit
0.3
ppm
-0.3
(K)
(L)
(M)
(N)
(O)
(P)
(Q)
(R)
χNLFit - χTE-wAvg
χSNR-wAvg - χTE-wAvg
χSusc-wAvg - χTE-wAvg
χSusc-TGV-wAvg - χTE-wAvg
(S)
(T)
(U)
(V)
(W)
(X)
(Y)
(Z)


NLFit
TE-wAvg
SNR-wAvg
Susc-wAvg
(A)
(B)
(C)
(D)
0.1
ppm
0
Mean±SD = 0.03±0.01 ppm
Mean±SD = 0.04±0.02 ppm
Mean±SD = 0.04±0.02 ppm
Mean±SD = 0.02±0.01 ppm
(E)
(F)
(G)
(H)
(I)
σ(χ) [a.u.]
0.1
0.08
0.06
0.04
0.02
0
100
200
300
Voxel Position Along Line


NLFit
TE-wAvg
SNR-wAvg
Susc-wAvg
(A)
(B)
(C)
(D)
0.03
ppm
0
Mean+SD = 0.01+0.01 ppm
Mean+SD = 0.01+0.04 ppm
Mean+SD = 0.01+0.02 ppm
Mean+SD = 0.01+0.01 ppm
(E)
(F)
(G)
(H)
(I)
σ(χ) [a.u.]
0.1
0.08
0.06
0.04
0.02
0
50
100
150
200
Voxel Position Along Line


Bland-Altman analysis
χ [ppm]
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
CN
GP
PU
TH
PCR
StrS
NLFit - TE-wAvg
NLFit - SNR-wAvg
NLFit - Susc-wAvg
NLFit - Susc-TGV-wAvg
TE-wAvg - SNR-wAvg
TE-wAvg - Susc-wAvg
TE-wAvg - Susc-TGV-wAvg
SNR-wAvg - Susc-wAvg
SNR-wAvg - Susc-TGV-wAvg
Susc-wAvg - Susc-TGV-wAvg