Title

Comparison of quantitative susceptibility mapping methods for iron-sensitive susceptibility imaging at 7T: An evaluation in healthy subjects and patients with Huntingtons disease.

Authors

Yao, Jingwen
Morrison, Melanie
Jakary, Angela
et al.

Publication Date

2023

DOI

10.1016/j.neuroimage.2022.119788

HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2024 November 26.

Published in final edited form as:

Neuroimage. 2023 January ; 265: 119788. doi:10.1016/j.neuroimage.2022.119788.

Comparison of Quantitative Susceptibility Mapping Methods for Imaging Brain Iron at 7T

Jingwen Yao[1], **Melanie A. Morrison**[1], **Angela Jakary**[1], **Sivakami Avadiappan**[1], **Yicheng Chen**[1,2,*], **Johanna Luitjens**[1,3], **Julia Glueck**[4], **Theresa Driscoll**[4], **Michael D. Geschwind**[4], **Alexandra B. Nelson**[4], **Javier E. Villanueva-Meyer**[1], **Christopher P. Hess**[1,4], **Janine M. Lupo**[1,2]

[1]Department of Radiology and Biomedical Imaging, UCSF, San Francisco, CA, USA

[2]UCSF/UC Berkeley Graduate Program in Bioengineering, San Francisco & Berkeley, CA, USA

[3]Department of Radiology, University Hospital, LMU Munich, Munich, Germany.

[4]Department of Neurology, University of California San Francisco, San Francisco, CA, USA

Abstract

Quantitative susceptibility mapping (QSM) is a promising tool for investigating iron dysregulation in neurodegenerative diseases, including Huntington's disease (HD). A diverse range of methods have been proposed to generate accurate and robust QSM images. In this study, we evaluated the performance of different dipole inversion algorithms for iron-sensitive susceptibility imaging at 7T on healthy subjects of large age range and patients with HD. We compared an iterative least-squares-based method (iLSQR), iterative methods that use regularization, single-step approaches, and deep learning-based techniques. Their performance was evaluated by comparing: (1) deviations from a multiple-orientation QSM reference; (2) visual appearance of QSM maps and the presence of artifacts; (3) susceptibility in subcortical brain regions with age; (4) regional brain susceptibility with published postmortem brain iron quantification; and (5) susceptibility in HD-affected basal ganglia regions between HD subjects and healthy controls. We found that single-step QSM methods with either total variation or total generalized variation constraints (SSTV/SSTGV) and the single-step deep learning method iQSM generally provided the best performance in terms of correlation with iron deposition and were better at differentiating between healthy controls and premanifest HD individuals, while deep learning QSM methods trained with multiple-orientation susceptibility data created QSM maps that were most similar to the multiple orientation reference and with the best visual scores.

Keywords

Magnetic resonance imaging (MRI); Quantitative susceptibility mapping (QSM); Huntington's disease (HD); Dipole inversion; Iron imaging; Magnetic susceptibility

Corresponding author Janine M Lupo, PhD, Professor in residence, Department of Radiology & Biomedical Imaging, University of California, San Francisco, Byers Hall, Rm 303D, MC 2532, 1700 4th Street, San Francisco, CA 94158-2330, Fax: 415-514-1028, janine.lupo@ucsf.edu.

*Currently at Meta Platforms, Inc., Mountain View, CA, USA

1. Introduction

Quantitative susceptibility mapping (QSM) is an MRI technique that non-invasively measures susceptibility in vivo. The diamagnetic and paramagnetic materials in the tissue induce local magnetic field perturbation, resulting in alternations in phase information collected by gradient recalled echo (GRE) acquisitions. Tissue susceptibility is sensitive to a variety of physiological and pathological factors, including iron concentration[1–3], myelin content[4,5], deoxyhemoglobin[6,7], lesion calcification[8,9], and myelin orientation[10,11]. QSM has been increasingly used in the evaluation of iron dysregulation in neurodegenerative diseases, especially in deep gray matter, revealing, for example, increases in tissue susceptibility indicative of increased iron deposition in the amygdala and caudate nucleus (CN) in Alzheimer's disease[12,13], the substantia nigra (SN) and red nucleus (RN) in Parkinson's disease[14], and the CN, putamen (PU), and globus pallidus (GP) in Huntington's disease (HD)[15–17]. The increased susceptibility within the basal ganglia in HD subjects has also been shown to significantly correlate with other disease characteristics, including regional atrophy[15] and the composite scales of age and genetic burden of disease[15,17].

Despite the great promise of QSM as a biomarker of iron content in the diagnosis and monitoring of progression of neurological diseases, the lack of a standardized processing method has limited its routine application in clinical settings. QSM processing typically involves coil combination of complex MR signal, phase unwrapping, background field removal, and dipole field-to-susceptibility inversion[18]. The dipole inversion process involves a mathematically ill-posed problem of deconvoluting the tissue magnetic field perturbation with the point-dipole response function. Due to the zero values of the point-dipole response in the frequency domain on two cone surfaces at the magic angle ($\theta \approx 54.7°$) with respect to the main magnetic field, the inversion problem will substantially amplify the noise close to the zero-cone surfaces, leading to streaking artifacts in the reconstructed susceptibility distribution map[18–20].

Numerous methods have been proposed to solve this problem, including acquiring multiple measurements with different orientations[10,21] and calculating susceptibility through multiple orientation sampling (COSMOS) method[21], which provides an analytical solution by filling in the missing data but is impractical clinically due to the excessive acquisition time required and associated discomfort related to rotating the anatomy of interest while inside the scanner. A variety of methods have since been offered to reconstruct the susceptibility distribution using one single MRI acquisition, the majority of which fall into the category of iterative methods incorporating regularization[18]. For example, the morphological-enabled dipole inversion (MEDI)[22] method imposes the edge information from the magnitude images as the anatomical prior using L1 regularization. QSM methods in this category have been developed with different data consistency terms (L1-norm[23]/L2-norm[24], linear[24]/nonlinear[25] consistency) and regularization terms (Tikhonov[26], total variation[19], total generalized variation[25], sparse susceptibility gradient[24], etc.). Other methods have proposed separating k-space into different subregions to estimate frequency coefficients in only ill-conditioned subregions, such as the iLSQR[20] method. More recently, new methods that integrate the background field removal step and field-to-susceptibility inversion have emerged, with the benefit of reducing the error propagation between the two processing

steps[27,28]. With the recent development of machine learning and deep learning techniques and their application to image contrast generation, the most recent approaches for the dipole inversion step either rely completely on deep learning[29–35] or use a hybrid strategy with physical models[36]. These models include networks for dipole inversion (QSMGAN[29], QSMnet[30,35], xQSM[31], DeepQSM[34]), combined background field removal and dipole inversion (autoQSM[33]), and an end-to-end network that takes raw phase images as input (iQSM[32]). In addition to the significantly shorter processing time, deep learning based QSM methods do not require parameter tweaking compared to the other QSM reconstruction techniques.

The QSM community has been increasingly active in the development of new reconstruction algorithms, with 98 submissions to the most recent 2019 QSM challenge[37] compared to 27 submitted in 2016 [38]. As QSM has become a widely adopted research tool in neurodegenerative and neurovascular diseases, determining an optimal processing pipeline is necessary for clinical translation. However, the task of finding the best algorithms is not trivial. One of the most outstanding hurdles is the lack of ground truth. Contrast agent-based phantoms are generally incapable of recapitulating the various biomaterials and microstructures that contribute substantially to the distribution of tissue susceptibility *in vivo*. Studies using post-mortem *ex vivo* tissue have enabled validation with histological ground truth, however, the fixation process alters MR properties. Although multiple orientation acquisition allows for calculation of an analytical solution of the inversion problem, the resulting susceptibility map is nonetheless affected by inaccuracies at earlier processing steps and altered susceptibility in regions with susceptibility anisotropy.

The 2016 QSM challenge utilized the susceptibility tensor[10] χ_{33} term as ground truth and evaluated 4 imaging metrics of each reconstruction method, including root mean squared error, high-frequency error norm, structural similarity index metric, and error in anatomical regions of interest. To overcome the lack of a reliable ground truth when using *in vivo* acquisitions, the 2019 QSM challenge adopted a synthesized dataset from *in silico* head phantoms[39] and evaluated the submitted algorithms using both visual assessment and quantitative metrics measuring the consistency of QSM maps with ground truth. These comparison studies, however, are limited to reflect clinical utility and relevance in their assessments and only evaluated QSM methods on either one single healthy volunteer or synthetic MRI data. More recently, Chen et al. evaluated the effects of background field removal and dipole inversion algorithms on cerebral microbleed quantification at both 3T and 7T, suggesting that the selection of the most appropriate QSM processing procedure may depend on the application of interest and scanner field strength[40].

In this study, we evaluate the performance of different QSM reconstruction algorithms in a more diverse population that includes healthy volunteers from a wide range of ages and HD subjects. To assess their potential for providing iron-sensitive brain images in routine clinical practice, especially for detecting abnormal iron deposition within the basal ganglia at 7T, we evaluated QSM methods from three different perspectives: (1) their ability to generate susceptibility values in subcortical regions with known correlations between iron and age in normal subjects confirmed by postmortem studies[41,42]; (2) their ability to distinguish between subjects with HD and healthy controls based on susceptibility values obtained

from basal ganglia regions CN, PU, and GP; and (3) their visual appearance and presence of artifacts. In a subset of subjects, we also compared QSM algorithms to a reference reconstruction using COSMOS[21].

2. Methods

2.1. Subjects

Eight healthy volunteers (mean age of 28) were scanned for the generation of QSM maps using COSMOS reconstruction. We also prospectively recruited 33 healthy volunteers (mean age of 44), 14 individuals with premanifest HD (mean age of 38, mean CAG repeat length of 42.1), and 17 patients with early manifest HD (mean age of 46, mean CAG repeat length of 42.9). The disease stages of the HD individuals were determined clinically based on the total motor score, total functional capacity, and diagnostic confidence level from the Unified Huntington Disease Rating Scale (UHDRS). The detailed criteria are included in the Supplementary Material. More detailed subject demographics are included in Figure 1. The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of University of California, San Francisco (IRB#17–23143, first approved February 2018). Informed consent was obtained from all subjects involved in the study.

2.2. MRI Acquisition

All subjects were scanned on a 7T MRI scanner (GE Healthcare, Milwaukee, WI, USA) with a 2-channel transmit and 32-channel phased-array receive coil (Nova Medical). Phase images used for QSM reconstruction were acquired using a 3D multi-echo GRE sequence with 4 echoes (TE = 6/9.5/13/16.5 ms), TR = 50 ms, flip angle = 20°, bandwidth = 50 kHz, voxel size = 0.8×0.8×1.0 mm, FOV = 240×240×148 mm, and Auto-calibrating Reconstruction for Cartesian imaging (ARC)[43] with reduction factor (R) = 3 and 16 auto-calibrating lines. For the 8 healthy volunteers, this sequence was acquired at 0.8 mm isotropic resolution with otherwise the same parameters, repeated three times with different head orientations: (1) normal position, (2) an average of 15° forward/backward head tilt, and (3) an average of 7° left/right head rotation. The data from all three acquisitions were combined to generate QSM maps using COSMOS reconstruction. The raw multi-echo GRE data were acquired using a modified susceptibility weighted angiography (SWAN) sequence that included ARC parallel imaging and was reconstructed offline, as described in 2.3. All SWAN images were acquired with pure axial slab positioning.

2.3. QSM Processing

The QSM processing pipeline is demonstrated in Figure 1. We reconstructed the magnitude and raw phase images for each echo from raw complex-value k-space data using an in-house developed program in MATLAB 2015b (MathWorks, Natick, MA), which included the ARC reconstruction[43] and coil image combination using the MCPC-3D-S method[44]. Compared to conventional adaptive coil combination, MCPC-3D-S utilizes phase information from multiple echoes to yield phase images that reflect only ΔB_0-related phase and removes any coil-related phase contributions[45]. Brain extraction was performed on magnitude images from each echo separately using a deep-learning-based HD-BET algorithm[46]. The final

brain mask was generated by calculating the intersection of all four masks. The mask was applied before the background field removal and dipole inversion steps of QSM processing. A 3D Laplacian phase unwrapping method[47] was applied to the phase images, followed by background field removal using the Sophisticated Harmonic Artifact Reduction for Phase data with varying spherical kernels (VSHARP) algorithm[47]. The resulting local field maps were averaged across all echoes for subsequent processing.

2.4. Dipole Inversion Algorithms

QSM maps were generated using ten different algorithms (parameters listed in Table 1), including iLSQR[20], iterative optimization methods with regularization (STARQSM[19], FANSI[25], HD-QSM[48], MEDI[22]), single-step methods (QSIP[27], SSTV/SSTGV[28]), and deep learning methods (QSMGAN[29], QSMnet+[30,35], xQSM[31], iQSM[32]). Single-step methods were performed after the phase-unwrapping step. iQSM was applied directly on the raw MRI phase data. All other methods were performed on the local field maps. We additionally performed COSMOS reconstruction[21] of QSM maps for the 8 subjects who received scans at 3 different head orientations. All phase unwrapping, background field removal, and dipole inversion algorithms, except for the deep learning algorithms and COSMOS, were performed using the MATLAB-based, QSM pipeline platform SEPIA.[49] In the following sections and in Table 1, the dipole inversion algorithms compared in this study are described in more detail.

2.4.1. COSMOS Reconstruction—For the 8 subjects who received 3 scans at different head orientations, the 3 magnitude images were co-registered with FSL[50] (FMRIB Software Library) *flirt*. The resulting transformation matrices were applied to the 3 brain masks, with the intersection of the masks taken as the final brain mask used in COSMOS and other dipole inversion algorithm calculations. The transformation matrices were also applied to the tissue phase images from each orientation after phase unwrapping and background field removal with VSHARP. The COSMOS QSM map was calculated from the tissue phase images using a weighted linear least squares method[21].

2.4.2. iLSQR—iLSQR[20] uses a sparse linear equation and least-squares (LSQR)-algorithm-based method, combined with estimations of the susceptibility boundaries and the susceptibility artifact from ill-conditioned k-space regions. The iLSQR algorithm was performed in SEPIA with the integrated STISuite Toolbox using the default error tolerance of 0.01.

2.4.3. Iterative Optimization Methods with Regularization

STARQSM[19]: Streaking Artifact Reduction for QSM algorithm performs a two-level QSM reconstruction algorithm by using different regularization parameters to reconstruct both large and small susceptibility values. We performed STARQSM in SEPIA with the integrated STISuite Toolbox using the default regularization parameters of 0.1 and 10^{-5}.

FANSI[25]: Fast Nonlinear Susceptibility Inversion method solves an optimization problem for nonlinear dipole inversion with a novel variable-splitting scheme. FANSI was performed

in SEPIA with the integrated FANSI Toolbox using the optimized parameter $\lambda = 1 \times 10^{-4}$ and the default assumptions of $\mu_1 = 100\,\lambda$ and $\mu_2 = 1$.

HDQSM[48]: Hybrid Data fidelity term approach for QSM consists of two stages, including an L1-norm optimization, the result of which is used as the initialization of a second L2-norm optimization. We integrated the HDQSM code into the SEPIA platform using the optimized parameter $\lambda_2 = 6.3 \times 10^{-5}$ and the default assumptions of $\lambda_1 = \sqrt{\lambda_2}$, $\mu_1 = \sqrt{\mu_2}$, and $\mu_2 = 100\lambda_2$.

MEDI[22]: Morphology-Enabled Dipole Inversion algorithm promotes sparsity of the edges in the susceptibility distribution that do not have a corresponding edge in the magnitude image. MEDI was performed in SEPIA with the integrated MEDI Toolbox using the optimized regularization parameter $\lambda = 1780$.

2.4.4. Single-Step Methods

QSIP[27]: Quantitative Susceptibility Mapping by Inversion of a Perturbation Field Model utilizes a dipole kernel in the spatial domain, and simultaneously estimates the susceptibility sources inside and outside the brain, using a tissue/air susceptibility atlas as an initial condition. We integrated the QSIP algorithm into the SEPIA platform, using the default parameters of $\lambda_1 = 1 \times 10^{-8}$, $\lambda_2 = 1 \times 10^{-6}$, and $\lambda_3 = 1 \times 10^{-10}$.

SSTV/SSTGV[28]: Single-Step Total Variation (TV) or Total Generalized Variation (TGV) employs multiple spherical mean value (SMV) kernels of varying radii together with TV and TGV penalties to jointly perform the background removal and dipole inversion in a single step. We integrated the SSTV/SSTGV algorithm into the SEPIA platform using the optimized parameters $\lambda = 0.004$ for SSTV, and $\lambda_0 = 0.0079$ and $\lambda_1 = 0.004$ for SSTGV.

2.4.5. Deep Learning Methods

QSMGAN[29]: QSMGAN uses a 3D Generative Adversarial Network (GAN) architecture with an increased receptive field. The model was used to infer QSMGAN QSM maps on a Linux workstation with 4 Nvidia V100 32G GPUs. Tissue phase volumes were first resampled to 0.8 mm isotropic resolution to match the input of the deep learning model.

QSMnet+[35]: QSNnet+ is a deep neural network with a 3D U-net architecture trained with data augmentation to improve the generalizability to a wider range of susceptibility. Inference of the model was performed on the same Linux workstation as QSMGAN. Tissue phase volumes were resampled to 1.0 mm isotropic resolution before inference.

xQSM[31]: xQSM was designed based on a U-net backbone, with additional noise regularization and modified octave convolutional layers. It was trained with synthetic and in vivo datasets. In this study, we adopted the xQSM network trained with in vivo datasets because it achieved the best results in the original study. Similarly, the tissue phase volumes were resampled to 1.0 mm isotropic resolution as input to the model.

iQSM[32]: This method utilized a Laplacian-of-Trigonometric-functions (LoT) enhanced deep neural network for quantitative field and susceptibility mapping, directly from raw MRI phase data. The iQSM architecture is composed of a tailored LoT layer and a 3D residual U-net. Again, raw MRI phase data were resampled to 1.0 mm isotropic resolution and used as input to the iQSM model.

2.5. Parameter Optimization

For open-source QSM algorithms that include regularization terms, we performed parameter tuning using a frequency equalization method[51]. The method is based on the assumption that optimized regularization can be achieved when high-frequency coefficients of QSM maps have similar local mean values close to and away from the magic angle. Figure 2(A) depicts example QSM maps calculated using HDQSM with different regularization parameters and their frequency-domain representations. Under-regularized reconstruction results showed streaking artifact in the image domain and amplification of the coefficients close to the magic angle in the frequency domain, whereas over-regularized maps showed attenuation of the coefficients in the ill-conditioned frequency space. The frequency normalization method defines ROIs with different proximities to the magic angle, M2 and M3 (Figure 2) to select the optimal regularization weight that minimizes the normalized squared difference in mask amplitudes between M2 and M3, using the equation: $\xi_{23} = \left(\frac{A_2 - A_3}{A_2 + A_3}\right)^2$, where A_i is the average of the coefficients within the mask M_i. For each method, we reconstructed the QSM maps using a range of different regularization parameters and selected the optimal regularization weight based on the frequency equalization plot (Figure 2(B)).

2.6. Image Post-Processing and Analysis

All reconstructed QSM maps were normalized to the median susceptibility within a lateral ventricle ROI before subsequent analysis[52]. The lateral ventricle was chosen as the reference region because cerebrospinal fluid QSM does not vary with age or disease and has small inter-subject variability[52]. We performed linear and nonlinear registration of the iLSQR QSM map from each subject to a QSM atlas[53] using FSL[50] and inversely warped a predefined brain region segmentation to the subject space (Figure 1, Supplementary Figure 1), followed by the extraction of the median susceptibility value within each brain ROI. In the 8 subjects with COSMOS QSM, normalized root mean square error (NRMSE) and susceptibility-optimized similarity index metric (XSIM)[37] were calculated between each QSM map and the COSMOS QSM reference.

2.7. Visual Evaluation

Visual assessment of QSM map quality was performed by two neuroradiologists with 6 and 2 years of experience. Rating criteria were adopted from the QSM reconstruction challenge 2.0[37], which assessed the level of 3 types of artifacts present on QSM maps (streaking, unnaturalness, and noise) using scores from 0 (no artifact) to 3 (worst artifact). The detailed rating criteria is provided in the Supplementary Material (Supplementary Table 1). Ratings of visual quality were performed on 120 3D QSM volumes visualized in all 3 planes from 10 randomly selected healthy volunteers, with ages ranging from 29 to 61 years old. For each subject, QSM maps derived from the 12 different reconstruction methods were

included. The 120 QSM maps were randomized in order, and the raters were blinded to the reconstruction method associated with each map. For each QSM map, we took the average of the visual scores from the two raters as the final score.

2.8. Statistical Analysis

To overcome our lack of ground truth for the quantification of iron content *in vivo*, we selected three metrics to indirectly evaluate the relationship between QSM methods and iron content: (1) Pearson's correlation was performed to correlate susceptibility in basal ganglia regions and age in healthy volunteers, based on the widely accepted assumption that a linear correlation exists between iron content in the basal ganglia and age within the age range evaluated[3,54,55]; (2) Pearson's correlation was performed between mean susceptibility across healthy volunteers and previously published unfixed postmortem iron quantification[42] for 7 different brain regions: GP, RN, SN, PU, CN, dentate nucleus (DN), and thalamus (TH), based on the assumption that the dominant contribution of susceptibility differences across gray matter regions is due to iron; and (3) Kruskal-Wallis and Dunn's post hoc tests were performed to compare the susceptibility values in basal ganglia regions CN, PU, and GP between HD subjects and healthy controls, based on the well-validated observation that premanifest and early HD subjects exhibit increased iron deposition in the basal ganglia compared to age-matched healthy controls[15–17,56].

For the comparison between HD subjects and healthy controls, we computed age-corrected susceptibility values of basal ganglia regions CN, PU, and GP using the slope of the susceptibility-age linear fitting results from healthy volunteers only, based on the assumption that the effects of HD and age on brain region susceptibility are independent. The age-corrected susceptibility values were subsequently compared between HD subjects and healthy controls using Kruskal-Wallis and Dunn's post hoc tests. The effect sizes for the comparison between healthy controls and HD cohorts were estimated using the Hedges' g calculation[57]. All brain region susceptibility values were calculated as the mean of the two ROIs from both hemispheres. We adopted a Bonferroni correction to correct for multiple comparisons across the 12 QSM methods evaluated. The corrected significance level ($\alpha_{corrected}$) was 0.0042. P-values were considered significant if < 0.0042 unless otherwise specified.

3. Results

3.1. Comparison to COSMOS Reference

COSMOS QSM and iLSQR QSM maps from a healthy volunteer are shown in Figure 3(A). The two maps were qualitatively similar, with higher susceptibility values observed in basal ganglia regions and lower values in white matter tracts. The COSMOS QSM map exhibited relatively less noise due to the three imaging acquisitions. The means and standard deviations of NRMSE and XSIM between QSM maps using different reconstruction methods and COSMOS QSM reference are plotted in Figure 3(B). Deep learning methods QSMGAN and QSMnet+, along with STARQSM, achieved the lowest NRMSE (QSMGAN: 0.51±0.03; QSMnet+: 0.64±0.03; STARQSM: 0.61±0.05) and highest XSIM (QSMGAN: 0.63±0.04; QSMnet+: 0.52±0.04; STARQSM: 0.57±0.06), whereas the

single-step deep learning iQSM method had higher NRMSE (1.08±0.07) and lower XSIM values (0.32±0.03).

3.2. Visual Evaluation

Figure 4 demonstrates the consistent contrast observed across all 12 QSM methods of one single subject and all subjects averaged (bottom row). SSTV and SSTGV methods showed slight signal variations near the cortex of the left occipital lobe and the right temporal lobe. Visual assessment of the level of streaking, unnaturalness, and noise is plotted in Figure 5(A–C). The deep learning method QSMnet+ demonstrated the least amount of streaking (0.10±0.21), followed by SSTV (0.45±0.37) and STARQSM (0.65±0.47), whereas xQSM (1.85±0.63) and FANSI (1.70±0.71) had a relatively higher average level of streaking. The levels of unnaturalness and noise showed the opposite trend for some of the methods, with iLSQR showing lower-than-average levels of unnaturalness and higher-than-average levels of noise. On the other hand, SSTGV, SSTV, and MEDI demonstrated lower-than-average levels of noise and higher-than-average levels of unnaturalness. STARQSM and the two deep learning methods trained with COSMOS (QSMGAN and QSMnet+), showed lower-than-average levels of artifact in all three categories, whereas FANSI demonstrated higher-than-average levels of artifact. The mean scores across all 10 healthy volunteers are plotted in Figure 5(D). When averaged across all three types of artifacts, QSMnet+ had the lowest overall artifact score (0.43), followed by QSMGAN (0.73) and STARQSM (0.63). When considering the highest score of artifacts, the same three methods performed the best (QSMnet+ 0.80, STARQSM 0.80, QSMGAN 0.80). Two pairs of QSM maps with scores of 0 and 3 on streaking and level of unnaturalness, respectively, from rater 1 are illustrated in Figure 5(E).

3.3. Correlation between Susceptibility and Age

Figure 6(A) shows correlation plots between putamen susceptibility and age using different QSM reconstruction methods. A significant and moderate positive correlation was observed for all 10 methods, with correlation coefficients ranging from 0.54 to 0.67. The strongest correlations between regional brain susceptibility and age were observed in basal ganglia ROIs (Figure 6(B)) for all QSM methods. When combining the subcortical nuclei (CN, PU, DN, SN, and RN) with known iron-age correlation, all methods showed similar Pearson's correlation coefficients, ranging from 0.58 to 0.69, with the highest coefficients observed using QSMnet+ (r=0.69, p<0.001) and iLSQR (r=0.67, p<0.001) (Figure 6(C)). Correlation plots in brain regions CN and SN are included in the Supplementary Material (Supplementary Figures 2, 3).

3.4. Correlation between Susceptibility and Iron Quantification

When evaluating the correlation between the average susceptibility in healthy volunteers and postmortem iron quantification in different brain regions, all methods showed significant and strong correlation with coefficients higher than 0.9, ranging from 0.93 to 0.97. Among the categories, the single-step methods (QSIP, SSTGV, SSTV) generally had the strongest correlations (Figure 7(A–C)), with r > 0.96. Out of the other methods, iLSQR and STARQSM provided a stronger correlation between susceptibility and iron. Within the deep

learning reconstruction methods, the single-step iQSM model had the highest correlation coefficient.

3.5. Comparison between HD Subjects and Healthy Volunteers

Kruskal-Wallis tests revealed significant differences in CN susceptibility among healthy controls, premanifest HD, and early manifest HD, using all QSM reconstruction methods (Figure 8(A)). Similar results were observed in PU and GP (Supplementary Figure 4, 5). When taking the mean susceptibility of the three HD-related brain regions (CN, PU, and GP), single-step methods SSTV and SSTGV, along with single-step deep learning method iQSM, showed the highest test statistics when comparing between healthy controls and premanifest HD (SSTGV: z=3.59; SSTV: z=3.41; iQSM: z=3.80), whereas all methods provided significant test statistics and large effect size when comparing between healthy subjects and early manifest HD (Figure 8(B)).

4. Discussion

This study compared different QSM dipole inversion algorithms for their similarity to a COSMOS reference, visual quality, correlation with age and iron quantification in healthy subjects, and ability to detect abnormal brain iron deposition in individuals with premanifest/early manifest HD. A summary of the six metrics compared across the various QSM method types is displayed in Table 2 and Supplementary Table 2.

The 2016 and 2019 QSM challenges provided a comprehensive quantitative comparison of a wide range of QSM reconstruction methods by comparing against either a reference image obtained from a multiple orientation acquisition of one healthy (30-year-old) volunteer (2016) or an *in silico* head phantom (2019). Although the simulated dataset overcomes the lack of a ground truth when using an *in vivo* dataset, it is a less realistic scenario where the background field and the influence of the acquisition parameters and imperfections may not be fully captured. Age-related physiological variations and pathological effects on QSM contrast were also not reflected in these current challenges. This study builds upon these contributions by providing a complementary assessment of QSM methods that focuses on their performance in detecting susceptibility changes related to iron content in normal aging and in HD subjects, which could be more closely associated with their clinical utility as imaging biomarkers for many neurological diseases.

The results demonstrated that single-step methods SSTV/SSTGV, and deep learning based iQSM, performed among the best in terms of correlating with previously reported postmortem iron concentration of different brain regions in normal subjects, and differentiating between healthy control and premanifest HD individuals. These relatively independent metrics were evaluated to overcome our lack of ground truth of iron content *in vivo*. The first metric was selected based on the assumptions that the dominant contribution of susceptibility differences across brain regions is from iron. The latter relies on the increase of basal ganglia iron in premanifest and early HD subjects observed in multiple existing imaging studies[56,58]. Single-step methods also provided a moderate positive correlation between susceptibility in the basal ganglia and age, consistent with previous literature[3,54,55]. We believe that by performing these three relatively independent analyses,

we were able to confirm the advantage of SSTV/SSTGV and iQSM over other methods in detecting physiological and pathological alterations in basal ganglia iron content at 7T. The observation that all three methods involve the integration of multiple QSM processing steps may potentially indicate the superiority of single-step QSM reconstruction methods over multi-step pipelines.

Our measurements of regional brain susceptibility is in good agreement with previously reported QSM measurements performed at the same field strength, with GP, RN, and SN showing the highest susceptibility values, followed by PU and CN, with WM and TH having the lowest susceptibility[40,59,60]. In accordance with previous studies, we found significant correlations between basal ganglia susceptibility values and age. The correlation coefficients obtained in our study (0.4–0.7) are comparable with findings from Liu et al[61], showing the highest age correlation in PU (0.7, current study $r = 0.5–0.7$), followed by RN (0.6, current study $r = 0.5–0.6$) and CN (0.6, current study 0.5–0.6). Controversial findings have been reported in the GP region. Whereas our study found a weak positive correlation with age (0.3–0.5), some studies showed no correlation with age[55,61], while others demonstrated a positive correlation using both QSM and R_2* imaging metrics[62,63]. This could be in part explained by the substantial inter-subject variability of GP non-heme iron concentration compared to age-related variability, as found in post-mortem histochemistry studies[42]. Compared to similar studies that used the postmortem iron concentrations from Hallgren et al.[42], the correlation between susceptibility and iron in different brain regions found in our study (0.93–0.98) was consistent with previously reported values (0.91–1.00)[27,61]. Lastly, we found significantly higher susceptibility values in CN, PU, and GP in premanifest and early manifest HD subjects, which has been consistently reported by other studies[15–17,58].

As a single-step method, SSTV/SSTGV has the advantage of reducing the reconstruction error compared to multi-step methods that adopt the same background filtering and dipole inversion algorithm[28]. Similarly, the deep learning method iQSM benefits from the integration of both phase unwrapping and background field removal steps[32]. Despite their better performance in characterizing iron, these single-step methods were among the least similar to the COSMOS reference, likely due to the integration of the preprocessing steps involved in the COSMOS reconstruction pipeline. We observed that in our particular dataset, SSTV/SSTGV sometimes failed to completely remove the background field variation at the edge of the brain, as observed in Figure 5(E), which may limit its application in assessing iron content in cortical gray matter without fine-tuning the SMV kernel parameters. Although QSIP, the other single-step method evaluated, provided susceptibility values that were strongly correlated with postmortem iron quantification in healthy subjects, it was less optimal in terms of differentiating between premanifest HD and healthy subjects. One possible reason for this may be because we were not able to optimize QSIP parameters for the current acquisition scheme due to the high number of regularization weights required. The reliance of QSIP on a predefined tissue-air susceptibility atlas may also affect its performance to more generalized applications.

Deep learning approaches QSMGAN and QSMnet+ outperformed the other methods when compared to a COSMOS reference, likely because both networks were trained using COSMOS QSM images as ground truth. However, this similarity to COSMOS did not

translate to better performance in other metrics except for visual assessment, which may be due to the limited age range of the healthy subjects used in training these models. A more diverse training dataset including both healthy subjects of a wider age range and pathological conditions could potentially further improve the performance of deep learning QSM reconstruction methods. Additionally, the hybrid methods combining deep learning and physics-based data consistency may also provide a more generalizable reconstruction.

Our results also confirmed the limitation of using COSMOS as ground truth for algorithm comparison, as discussed in previous studies[37,39]. Discrepancy exists between COSMOS reconstructed susceptibility maps and single-orientation susceptibility, especially in regions with anisotropic contributions of tissue susceptibility, making it a less accurate "ground truth" when evaluating single-orientation QSM methods. Like other dipole inversion algorithms, COSMOS reconstruction is also affected by the error generated from the phase unwrapping and background field removal steps, which causes additional differences when comparing between multi-step and single-step QSM methods. Using a digital phantom in the form of a recently developed simulated brain phantom[39] may be a more appropriate reference when comparing these two categories of reconstruction methods and provide complimentary information to this study. It could also serve as the ground truth to investigate potential over- or under-estimation of susceptibility using the different methods.

With the emergence of new single-step QSM reconstruction methods comes another rising question of how to interpret differences between dipole inversion algorithms in single-step algorithms compared to multi-step pipeline where the potential error induced by preprocessing steps may mask the true performance of dipole inversion techniques evaluated. Laplacian unwrapping and VSHARP background field removal were performed prior to various dipole inversion algorithms in this study for simplicity since they are the most widely used methods. Laplacian unwrapping has been shown to be more robust than other unwrapping methods[45], while VSHARP outperformed other background field removal methods in terms of vein contrast and white matter homogeneity[40]. Although we did not observe a significant presence of remnant field on the local field maps in any subjects (Supplementary Figure 6), future studies that evaluate different combinations of phase unwrapping, background field removal, and dipole inversion algorithms are still warranted to thoroughly decouple the effect of preprocessing steps and dipole inversion algorithms.

There are a few other limitations of this study. Although we performed parameter tuning of most dipole inversion methods used in this study according to Milovic et al.[45], the lack of availability to tune regularization parameters for STARQSM in the STISuite Toolbox prohibited the optimization of the regularization weights for this method. As we also utilized a relatively objective metric of frequency equalization on only one healthy volunteer, it is possible that including more subjects and other parameter tuning methods like L-curve[51] and comparing resulting imaging metrics to ground truth QSM[37] could potentially reduce bias and further improve a method's performance. Although we did not have a "ground truth" measure of iron distribution throughout the brain, or assess potential non-iron contributions to tissue susceptibility values, we did evaluate the sensitivity to iron using three relatively independent metrics. Nevertheless, further validation with histological information from either preclinical or postmortem human tissues would be

beneficial in confirming our results. Although it was not feasible to compare every QSM method presented in the literature, we selected the most widely utilized, readily available methods that were representative of the 4 main methodological groupings: iterative least-squares-based, iterative methods that use regularization, single-step approaches, and deep learning-based techniques. Our dataset and framework for evaluating clinical QSM data can easily incorporate evaluations of new QSM reconstruction methods, especially new single-step approaches, hybrid methods, and end-to-end deep learning models which incorporate multiple steps of the QSM processing pipeline, when they become available.

5. Conclusion

In conclusion, this study evaluated the performance of different dipole inversion algorithms for iron-sensitive susceptibility imaging in the brain at 7T, including iLSQR, iterative methods with regularization (STARQSM, FANSI, HDQSM, MEDI), single-step approaches (QSIP, SSTV, SSTGV), and deep learning methodologies (QSMGAN, QSMnet+, xQSM, iQSM). We found that the single-step iterative methods SSTV/SSTGV and single-step deep learning method iQSM provided overall good performance in terms of correlating with postmortem iron in normal subjects and differentiating between healthy control and premanifest Huntington's disease individuals, despite being less similar to the typical gold standard method, COSMOS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the study participants for their contribution to the research. This study is funded by NIH Grant R01 NS099564.

Abbreviations:

MRI	magnetic resonance imaging
QSM	quantitative susceptibility mapping
HD	Huntington's disease
GRE	gradient recalled echo
CN	caudate nucleus
SN	substantia nigra
RN	red nucleus
PU	putamen
GP	globus pallidus
DN	dentate nucleus

WM	white matter
TH	thalamus
MEDI	morphological-enabled dipole inversion
UHDRS	Unified Huntington Disease Rating Scale
COSMOS	calculation of susceptibility through multiple orientation sampling
VSHARP	Sophisticated Harmonic Artifact Reduction for Phase data with varying spherical kernel
NRMSE	normalized root mean square error
iLSQR	a sparse linear equation and least-squares (LSQR)-algorithm-based method
STARQSM	streaking artifact reduction for QSM
FANSI	fast nonlinear susceptibility inversion
HDQSM	hybrid data fidelity QSM
QSIP	quantitative susceptibility mapping by inversion of a perturbation field model
SSTV	single-step total variation
SSTGV	single-step total generalized variation
SMV	spherical mean value
GAN	generative adversarial network
ARC	Auto-calibrating Reconstruction for Cartesian imaging
SWAN	susceptibility weighted angiography
LoT	Laplacian-of-Trigonometric
XSIM	susceptibility-optimized similarity index metric

References

1. 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. doi:10.1016/j.neuroimage.2011.08.077 [PubMed: 21925274]
2. 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(3):1593–1599. doi:10.1016/j.neuroimage.2012.05.049 [PubMed: 22634862]
3. Acosta-Cabronero J, Betts MJ, Cardenas-Blanco A, Yang S, Nestor PJ. In Vivo MRI Mapping of Brain Iron Deposition across the Adult Lifespan. J Neurosci. 2016;36(2):364–374. doi:10.1523/JNEUROSCI.1907-15.2016 [PubMed: 26758829]

4. Granziera C, Wuerfel J, Barkhof F, et al. Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. Brain. 2021;144(5):1296–1311. doi:10.1093/brain/awab029 [PubMed: 33970206]
5. Hametner S, Endmayr V, Deistung A, et al. The influence of brain iron and myelin on magnetic susceptibility and effective transverse relaxation - A biochemical and histological validation study. NeuroImage. 2018;179:117–133. doi:10.1016/j.neuroimage.2018.06.007 [PubMed: 29890327]
6. Haacke EM, Xu Y, Cheng YCN, Reichenbach JR. Susceptibility weighted imaging (SWI). Magn Reson Med. 2004;52(3):612–618. doi:10.1002/mrm.20198 [PubMed: 15334582]
7. Cho ZH, Ro YM, Lim TH. NMR venography using the susceptibility effect produced by deoxyhemoglobin. Magn Reson Med. 1992;28(1):25–38. doi:10.1002/mrm.1910280104 [PubMed: 1435219]
8. Schweser F, Deistung A, Lehr BW, Reichenbach JR. Differentiation between diamagnetic and paramagnetic cerebral lesions based on magnetic susceptibility mapping. Med Phys. 2010;37(10):5165–5178. doi:10.1118/1.3481505 [PubMed: 21089750]
9. Chen W, Zhu W, Kovanlikaya I, et al. Intracranial calcifications and hemorrhages: characterization with quantitative susceptibility mapping. Radiology. 2014;270(2):496–505. doi:10.1148/radiol.13122640 [PubMed: 24126366]
10. Liu C. Susceptibility tensor imaging. Magn Reson Med. 2010;63(6):1471–1477. doi:10.1002/mrm.22482 [PubMed: 20512849]
11. Denk C, Torres EH, MacKay A, Rauscher A. The influence of white matter fibre orientation on MR signal phase and decay. NMR Biomed. 2011;24(3):246–252. doi:10.1002/nbm.1581 [PubMed: 21404336]
12. 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(11):e81093. doi:10.1371/journal.pone.0081093
13. Kim HG, Park S, Rhee HY, et al. Quantitative susceptibility mapping to evaluate the early stage of Alzheimer's disease. NeuroImage Clin. 2017;16:429–438. doi:10.1016/j.nicl.2017.08.019 [PubMed: 28879084]
14. He N, Ling H, Ding B, et al. Region-specific disturbed iron distribution in early idiopathic Parkinson's disease measured by quantitative susceptibility mapping. Hum Brain Mapp. 2015;36(11):4407–4420. doi:10.1002/hbm.22928 [PubMed: 26249218]
15. Chen L, Hua J, Ross CA, Cai S, van Zijl PCM, Li X. Altered brain iron content and deposition rate in Huntington's disease as indicated by quantitative susceptibility MRI. J Neurosci Res. 2019;97(4):467–479. doi:10.1002/jnr.24358 [PubMed: 30489648]
16. Domínguez JFD, Ng ACL, Poudel G, et al. Iron accumulation in the basal ganglia in Huntington's disease: cross-sectional data from the IMAGE-HD study. J Neurol Neurosurg Psychiatry. 2016;87(5):545–549. doi:10.1136/jnnp-2014-310183 [PubMed: 25952334]
17. van Bergen JMG, Hua J, Unschuld PG, et al. Quantitative Susceptibility Mapping Suggests Altered Brain Iron in Premanifest Huntington Disease. AJNR Am J Neuroradiol. 2016;37(5):789–796. doi:10.3174/ajnr.A4617 [PubMed: 26680466]
18. Deistung A, Schweser F, Reichenbach JR. Overview of quantitative susceptibility mapping. NMR Biomed. 2017;30(4). doi:10.1002/nbm.3569
19. Wei H, Dibb R, Zhou Y, et al. Streaking artifact reduction for quantitative susceptibility mapping of sources with large dynamic range. NMR Biomed. 2015;28(10):1294–1303. doi:10.1002/nbm.3383 [PubMed: 26313885]
20. 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 [PubMed: 25536496]
21. 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. doi:10.1002/mrm.21828 [PubMed: 19097205]

22. Liu T, Liu J, de Rochefort L, et al. Morphology enabled dipole inversion (MEDI) from a single-angle acquisition: comparison with COSMOS in human brain imaging. Magn Reson Med. 2011;66(3):777–783. doi:10.1002/mrm.22816 [PubMed: 21465541]
23. Milovic C, Lambert M, Langkammer C, Bredies K, Irarrazaval P, Tejos C. Streaking artifact suppression of quantitative susceptibility mapping reconstructions via L1-norm data fidelity optimization (L1-QSM). Magn Reson Med. 2022;87(1):457–473. doi:10.1002/mrm.28957 [PubMed: 34350634]
24. Bilgic B, Fan AP, Polimeni JR, et al. Fast Quantitative Susceptibility Mapping with L1-Regularization and Automatic Parameter Selection. Magn Reson Med Off J Soc Magn Reson Med Soc Magn Reson Med. 2014;72(5):1444–1459. doi:10.1002/mrm.25029
25. Milovic C, Bilgic B, Zhao B, Acosta-Cabronero J, Tejos C. Fast nonlinear susceptibility inversion with variational regularization. Magn Reson Med. 2018;80(2):814–821. doi:10.1002/mrm.27073 [PubMed: 29322560]
26. de Rochefort L, Liu T, Kressler B, et al. Quantitative susceptibility map reconstruction from MR phase data using bayesian regularization: Validation and application to brain imaging. Magn Reson Med. 2010;63(1):194–206. doi:10.1002/mrm.22187 [PubMed: 19953507]
27. Poynton CB, Jenkinson M, Adalsteinsson E, Sullivan EV, Pfefferbaum A, Wells W. Quantitative susceptibility mapping by inversion of a perturbation field model: correlation with brain iron in normal aging. IEEE Trans Med Imaging. 2015;34(1):339–353. doi:10.1109/TMI.2014.2358552 [PubMed: 25248179]
28. Chatnuntawech I, McDaniel P, Cauley SF, et al. Single-step quantitative susceptibility mapping with variational penalties. NMR Biomed. 2017;30(4). doi:10.1002/nbm.3570
29. Chen Y, Jakary A, Avadiappan S, Hess CP, Lupo JM. QSMGAN: Improved Quantitative Susceptibility Mapping using 3D Generative Adversarial Networks with increased receptive field. NeuroImage. 2020;207:116389. doi:10.1016/j.neuroimage.2019.116389
30. Yoon J, Gong E, Chatnuntawech I, et al. Quantitative susceptibility mapping using deep neural network: QSMnet. NeuroImage. 2018;179:199–206. doi:10.1016/j.neuroimage.2018.06.030 [PubMed: 29894829]
31. Gao Y, Zhu X, Moffat BA, et al. xQSM: quantitative susceptibility mapping with octave convolutional and noise-regularized neural networks. NMR Biomed. 2021;34(3):e4461. doi:10.1002/nbm.4461 [PubMed: 33368705]
32. Gao Y, Xiong Z, Fazlollahi A, et al. Instant tissue field and magnetic susceptibility mapping from MRI raw phase using Laplacian enhanced deep neural networks. NeuroImage. 2022;259:119410. doi:10.1016/j.neuroimage.2022.119410
33. Wei H, Cao S, Zhang Y, et al. Learning-based single-step quantitative susceptibility mapping reconstruction without brain extraction. NeuroImage. 2019;202:116064. doi:10.1016/j.neuroimage.2019.116064
34. Bollmann S, Rasmussen KGB, Kristensen M, et al. DeepQSM - using deep learning to solve the dipole inversion for quantitative susceptibility mapping. NeuroImage. 2019;195:373–383. doi:10.1016/j.neuroimage.2019.03.060 [PubMed: 30935908]
35. Jung W, Yoon J, Ji S, et al. Exploring linearity of deep neural network trained QSM: QSMnet. NeuroImage. 2020;211:116619. doi:10.1016/j.neuroimage.2020.116619
36. Zhang J, Liu Z, Zhang S, et al. Fidelity imposed network edit (FINE) for solving ill-posed image reconstruction. NeuroImage. 2020;211:116579. doi:10.1016/j.neuroimage.2020.116579
37. 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. doi:10.1002/mrm.28754 [PubMed: 33783037]
38. 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(3):1624–1637. doi:10.1002/mrm.28185 [PubMed: 32086836]
39. Marques JP, Meineke J, Milovic C, et al. 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. doi:10.1002/mrm.28716 [PubMed: 33638241]

40. Chen Y, Genc O, Poynton CB, Banerjee S, Hess CP, Lupo JM. Comparison of quantitative susceptibility mapping methods on evaluating radiation-induced cerebral microbleeds and basal ganglia at 3T and 7T. NMR Biomed. n/a(n/a):e4666. doi:10.1002/nbm.4666
41. Ramos P, Santos A, Pinto NR, Mendes R, Magalhães T, Almeida A. Iron levels in the human brain: A post-mortem study of anatomical region differences and age-related changes. J Trace Elem Med Biol. 2014;28(1):13–17. doi:10.1016/j.jtemb.2013.08.001 [PubMed: 24075790]
42. Hallgren B, Sourander P. The effect of age on the non-haemin iron in the human brain. J Neurochem. 1958;3(1):41–51. doi:10.1111/j.1471-4159.1958.tb12607.x [PubMed: 13611557]
43. Beatty PJ, Brau AC, Chang S, et al. A Method for Autocalibrating 2-D Accelerated Volumetric Parallel Imaging with Clinically Practical Reconstruction Times. In: Proc. Intl. Soc. Mag. Reson. Med; 2007:1.
44. Eckstein K, Dymerska B, Bachrata B, et al. Computationally Efficient Combination of Multi-channel Phase Data From Multi-echo Acquisitions (ASPIRE). Magn Reson Med. 2018;79(6):2996–3006. doi:10.1002/mrm.26963 [PubMed: 29034511]
45. 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):e3601. doi:10.1002/nbm.3601 [PubMed: 27619999]
46. Isensee F, Schell M, Pflueger I, et al. Automated brain extraction of multisequence MRI using artificial neural networks. Hum Brain Mapp. 2019;40(17):4952–4964. doi:10.1002/hbm.24750 [PubMed: 31403237]
47. Li W, Wu B, Liu C. Quantitative susceptibility mapping of human brain reflects spatial variation in tissue composition. NeuroImage. 2011;55(4):1645–1656. doi:10.1016/j.neuroimage.2010.11.088 [PubMed: 21224002]
48. Lambert M, Milovic C, Tejos C. Hybrid Data fidelity term approach for Quantitative Susceptibility Mapping. In: Proc. Intl. Soc. Mag. Reson. Med. ISMRM; 2020. https://cds.ismrm.org/protected/20MPresentations/abstracts/3205.html
49. Chan KS, Marques JP. SEPIA—Susceptibility mapping pipeline tool for phase images. NeuroImage. 2021;227:117611. doi:10.1016/j.neuroimage.2020.117611
50. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. NeuroImage. 2012;62(2):782–790. doi:10.1016/j.neuroimage.2011.09.015 [PubMed: 21979382]
51. Milovic C, Prieto C, Bilgic B, et al. Comparison of parameter optimization methods for quantitative susceptibility mapping. Magn Reson Med. 2021;85(1):480–494. doi:10.1002/mrm.28435 [PubMed: 32738103]
52. Straub S, Schneider TM, Emmerich J, et al. Suitable reference tissues for quantitative susceptibility mapping of the brain. Magn Reson Med. 2017;78(1):204–214. doi:10.1002/mrm.26369 [PubMed: 27529579]
53. Zhang Y, Wei H, Cronin MJ, He N, Yan F, Liu C. Longitudinal atlas for normative human brain development and aging over the lifespan using quantitative susceptibility mapping. NeuroImage. 2018;171:176–189. doi:10.1016/j.neuroimage.2018.01.008 [PubMed: 29325780]
54. Burgetova R, Dusek P, Burgetova A, et al. Age-related magnetic susceptibility changes in deep grey matter and cerebral cortex of normal young and middle-aged adults depicted by whole brain analysis. Quant Imaging Med Surg. 2021;11(9):3906919–3903919. doi:10.21037/qims-21-87
55. Li Y, Sethi SK, Zhang C, et al. Iron Content in Deep Gray Matter as a Function of Age Using Quantitative Susceptibility Mapping: A Multicenter Study. Front Neurosci. 2021;14. Accessed June 13, 2022. https://www.frontiersin.org/article/10.3389/fnins.2020.607705
56. Bartzokis G, Cummings J, Perlman S, Hance DB, Mintz J. Increased Basal Ganglia Iron Levels in Huntington Disease. Arch Neurol. 1999;56(5):569–574. doi:10.1001/archneur.56.5.569 [PubMed: 10328252]
57. Hedges LV. Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. J Educ Stat. 1981;6(2):107–128. doi:10.2307/1164588
58. Ravanfar P, Loi SM, Syeda WT, et al. Systematic Review: Quantitative Susceptibility Mapping (QSM) of Brain Iron Profile in Neurodegenerative Diseases. Front Neurosci. 2021;15:618435. doi:10.3389/fnins.2021.618435

59. Schneider TM, Ma J, Wagner P, et al. Multiparametric MRI for Characterization of the Basal Ganglia and the Midbrain. Front Neurosci. 2021;15. Accessed June 24, 2022. https://www.frontiersin.org/article/10.3389/fnins.2021.661504

60. 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 [PubMed: 23036448]

61. Liu M, Liu S, Ghassaban K, et al. Assessing global and regional iron content in deep gray matter as a function of age using susceptibility mapping. J Magn Reson Imaging. 2016;44(1):59–71. doi:10.1002/jmri.25130 [PubMed: 26695834]

62. Wang D, Li YY, Luo JH, Li YH. Age-related iron deposition in the basal ganglia of controls and Alzheimer disease patients quantified using susceptibility weighted imaging. Arch Gerontol Geriatr. 2014;59(2):439–449. doi:10.1016/j.archger.2014.04.002 [PubMed: 24820446]

63. Haacke EM, Miao Y, Liu M, et al. Correlation of putative iron content as represented by changes in R2* and phase with age in deep gray matter of healthy adults. J Magn Reson Imaging. 2010;32(3):561–576. doi:10.1002/jmri.22293 [PubMed: 20815053]

Highlight

- Single-step QSM susceptibility values were the most correlated with iron.
- Single-step QSM better distinguished premanifest HD subjects from healthy controls.
- COSMOS-trained QSM values were most similar to COSMOS but less correlated with iron

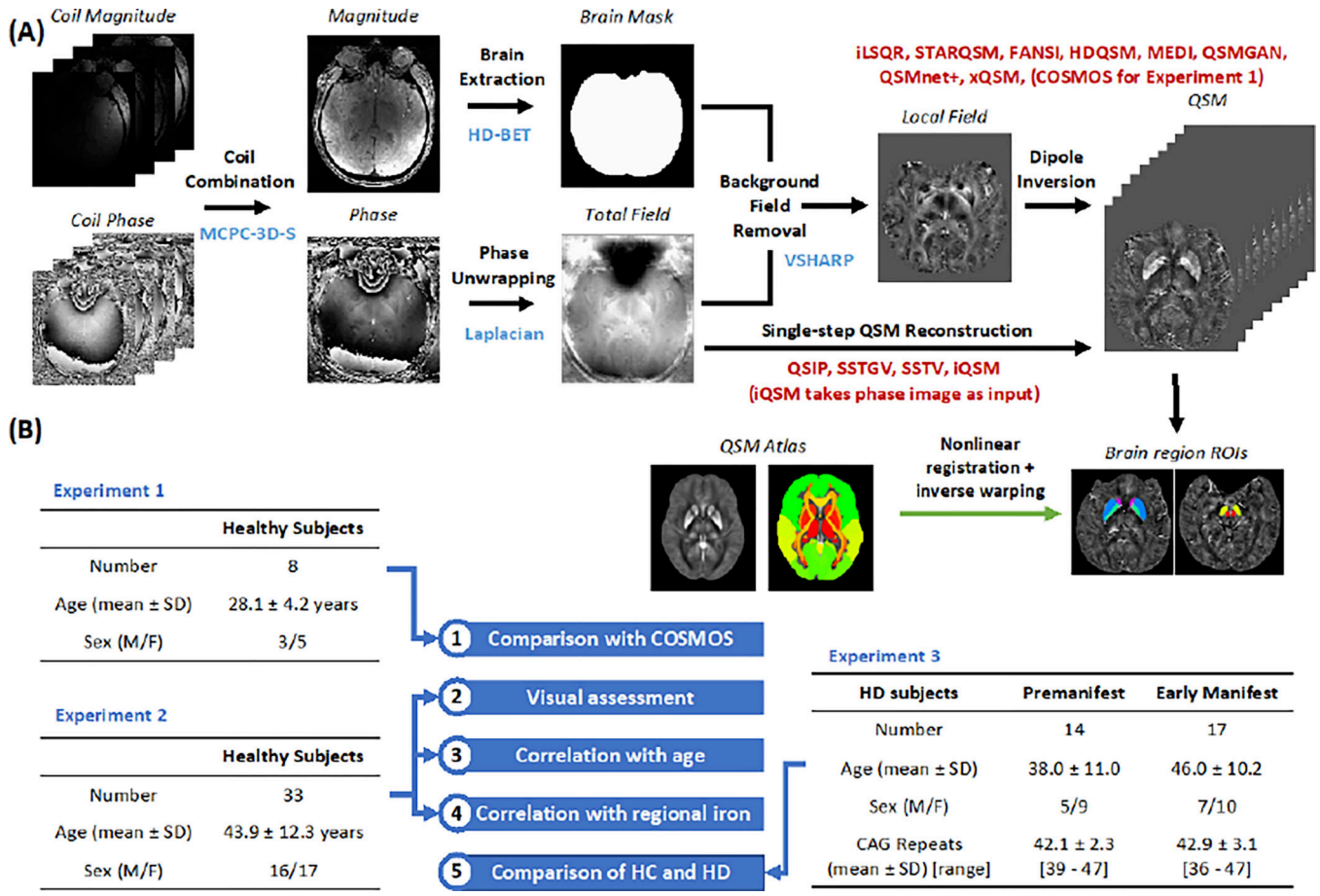

Figure 1.

QSM processing pipeline and experimental design. (A) QSM processing pipeline for all subjects. Brain region ROIs were generated for subjects in experiments 2 and 3 by registration of a QSM atlas with predefined whole brain segmentation. (B) Demographics of healthy volunteers and Huntington's disease subjects who participated in the three experiments. HC: healthy control; HD: Huntington's disease; SD: standard deviation.

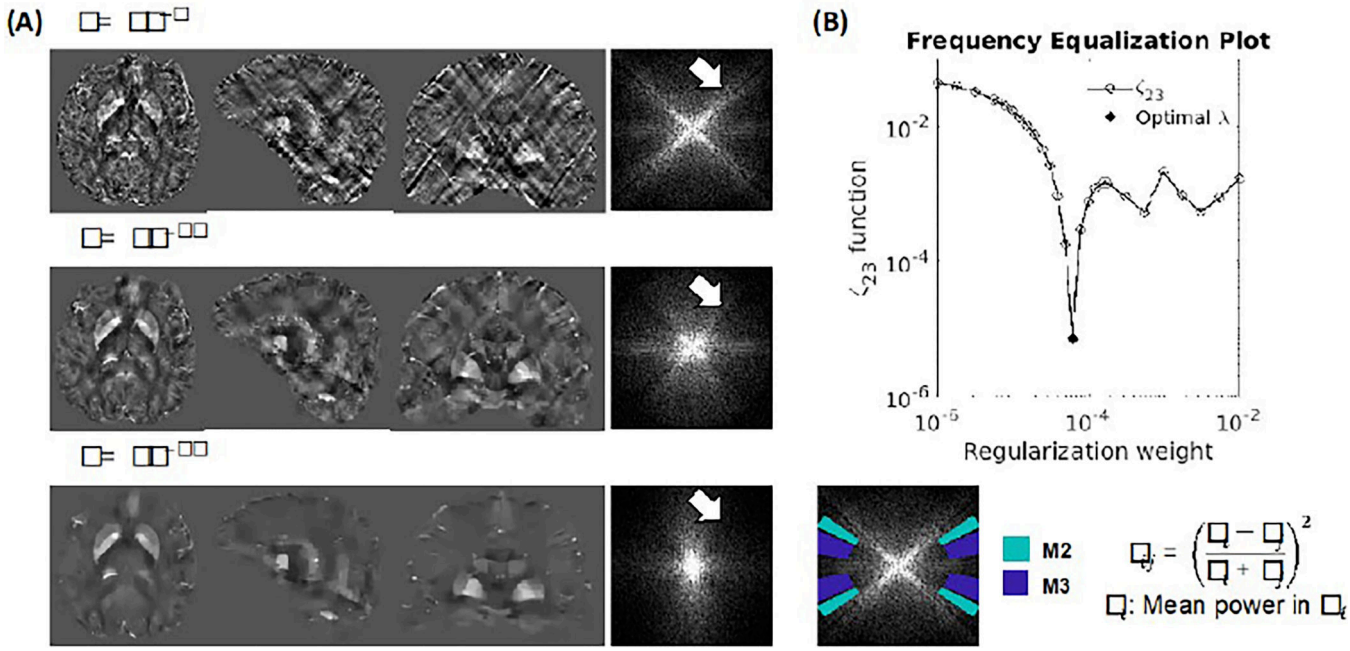

Figure 2.

Optimization of QSM reconstruction methods with regularization weights determined by frequency equalization. (A) QSM maps reconstructed using HDQSM with different regularization weights and the frequency representation of the QSM maps. In the frequency domain, the region near the magic angle is labeled with the white arrow. The amplification and attenuation of frequency coefficients in this region can be observed in cases of under-regularization and over-regularization, respectively. (B) Frequency equalization plot and the equation of the equalization metric. The solid dot represents the regularization weight corresponding to the lowest equalization metric, indicating the optimal regularization. In this case, the optimal regularization parameter is $10^{-4.2}$.

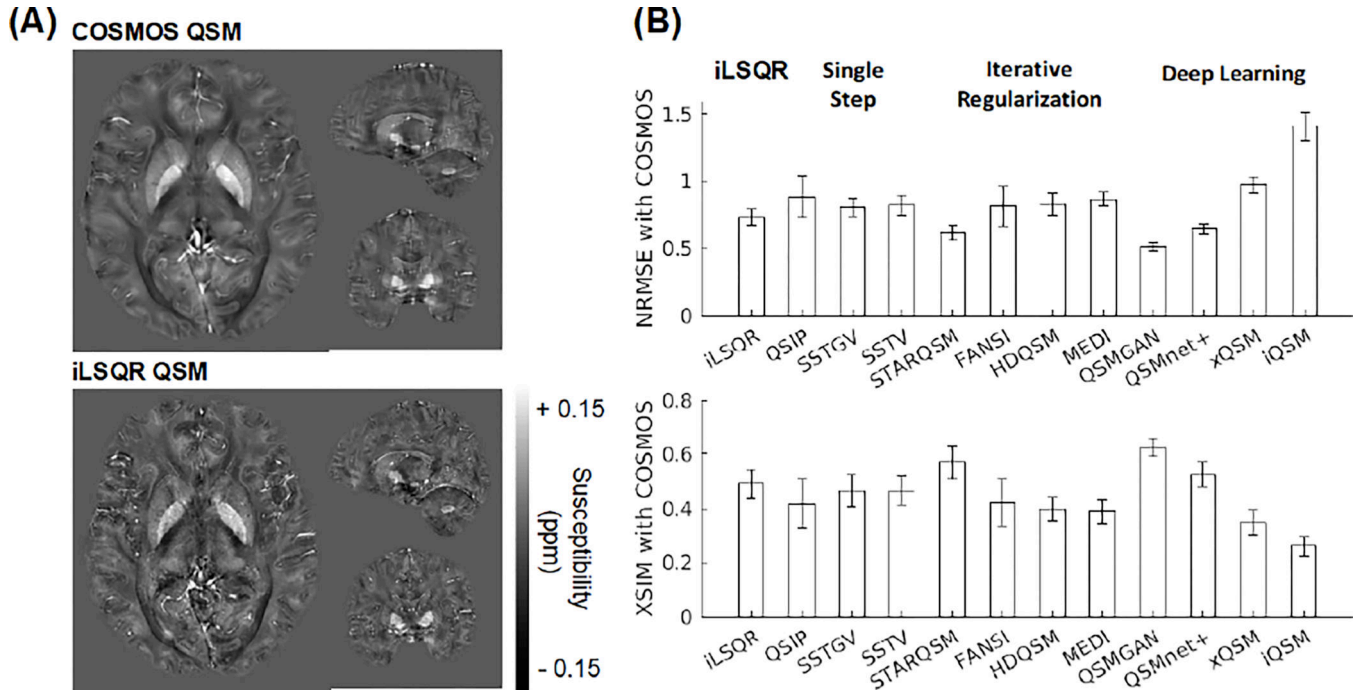

Figure 3.

Example QSM images and comparison with COSMOS reference. (A) Example COSMOS and iLSQR QSM images. (B) NRMSE and XSIM plots across all QSM methods. COSMOS: calculation of susceptibility through multiple orientation sampling; NRMSE: normalized root mean squared error; XSIM: Susceptibility-Optimized Similarity Index Metric.

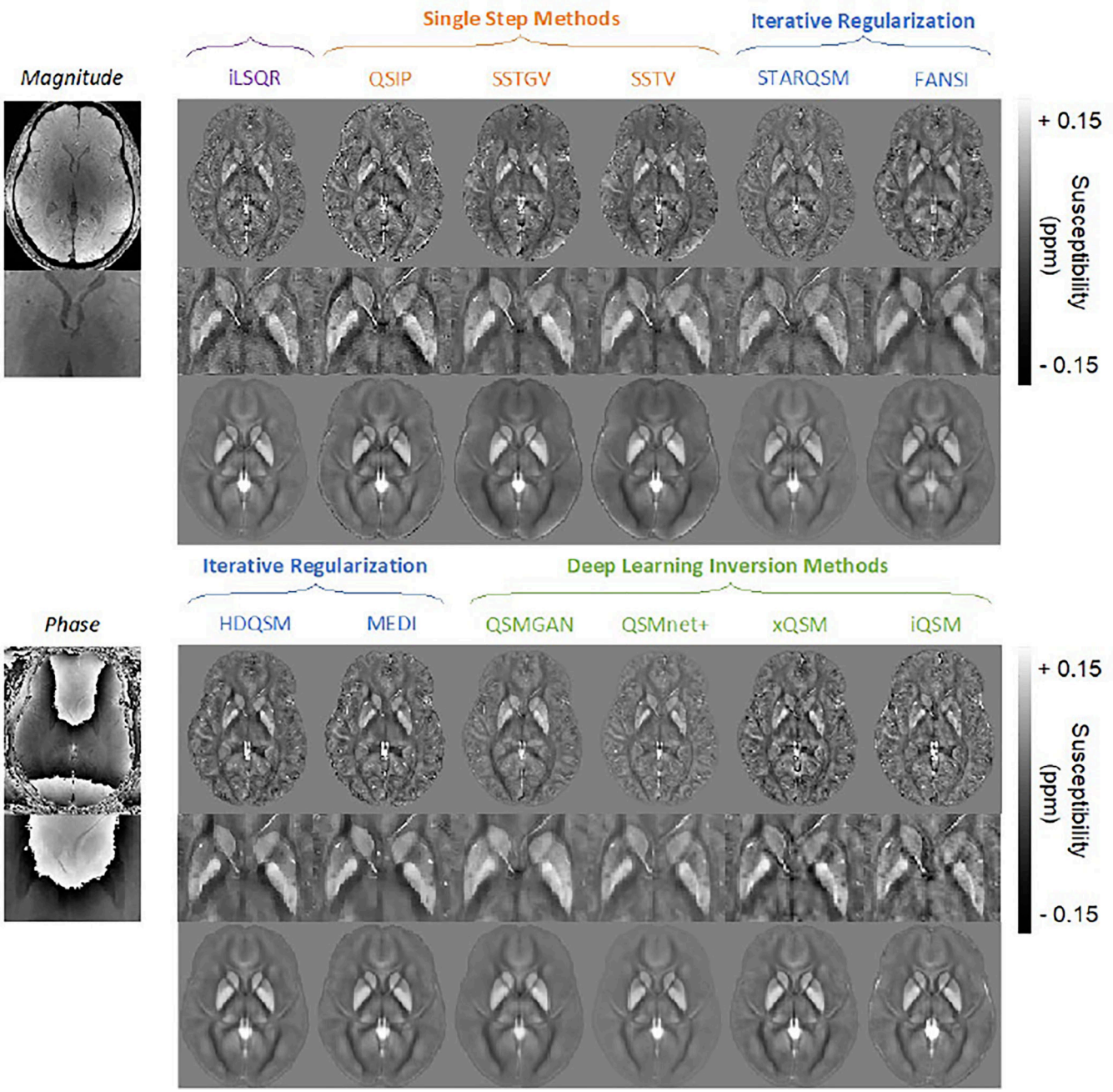

Figure 4.

Example QSM images and the zoomed-in depiction of the basal ganglia from one healthy volunteer. The bottom rows demonstrate the averaged QSM images across all healthy volunteers after registration to the QSM atlas.

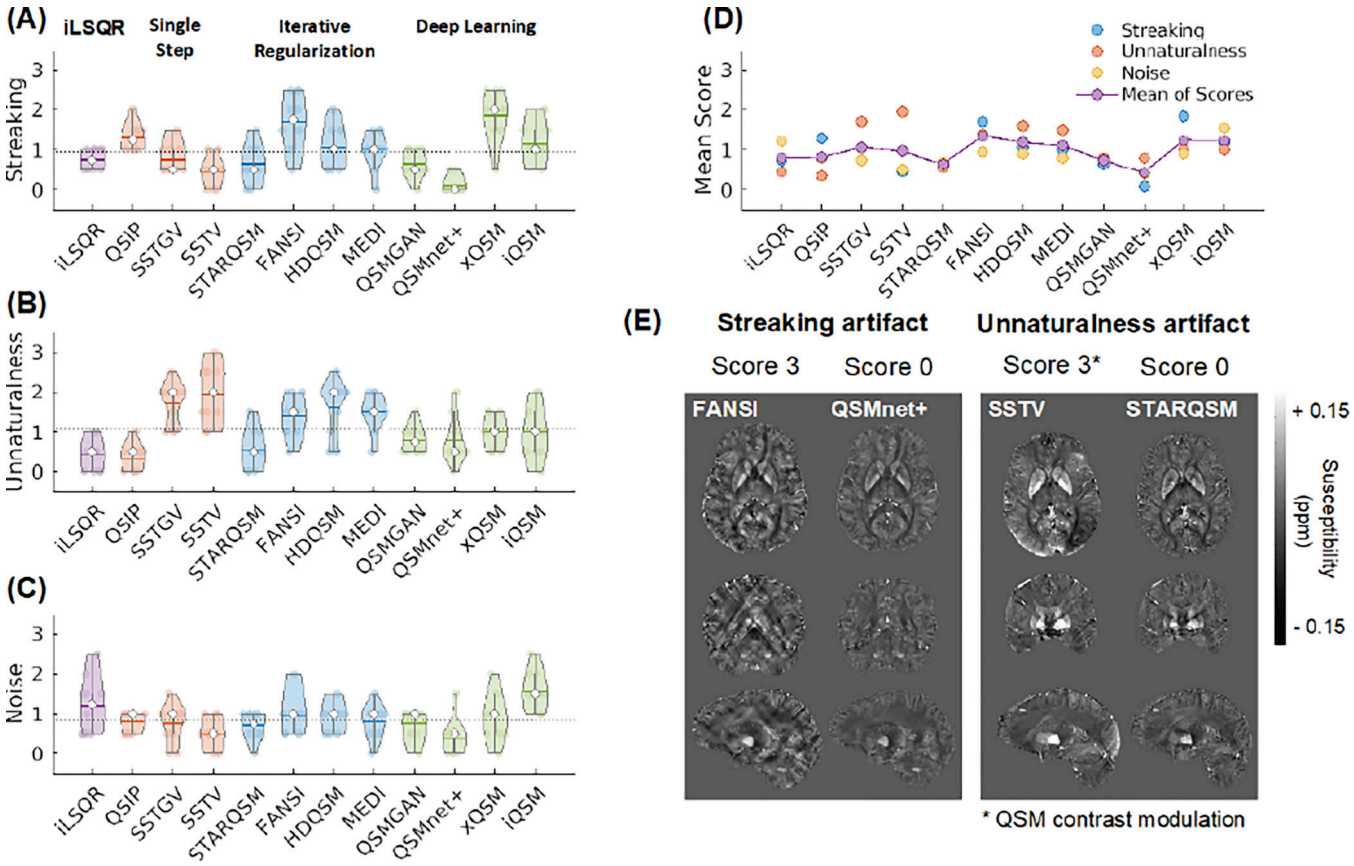

Figure 5.

Visual assessment of QSM images. (A-C) Scores of streaking, unnaturalness, and noise assessed in 10 randomly selected healthy volunteer QSM maps, with 0 representing the lowest level of artifact and 3 representing the highest level. The dashed lines represent the mean level of artifact in all 120 evaluations. (D) Mean visual assessment scores across the 10 healthy volunteer QSM images. (E) Example pairs of QSM images with score 3 and 0 (from rater 1) for streaking (left) and unnaturalness (right).

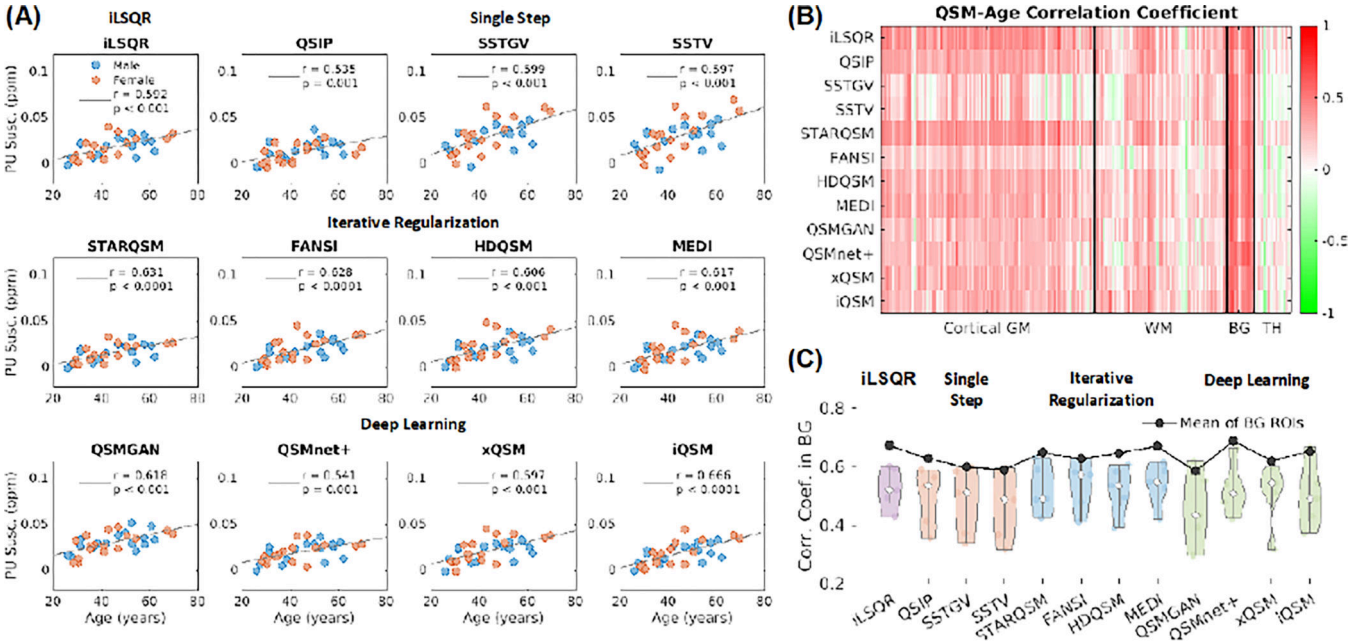

Figure 6.

Correlation between QSM values with age. (A) Correlation plots between putamen (PU) susceptibility and age. (B) Correlation coefficients in all brain ROIs. (C) Correlation coefficients in subcortical nuclei with known iron-age correlation. GM: gray matter; WM: white matter; BG: basal ganglia; TH: thalamus.

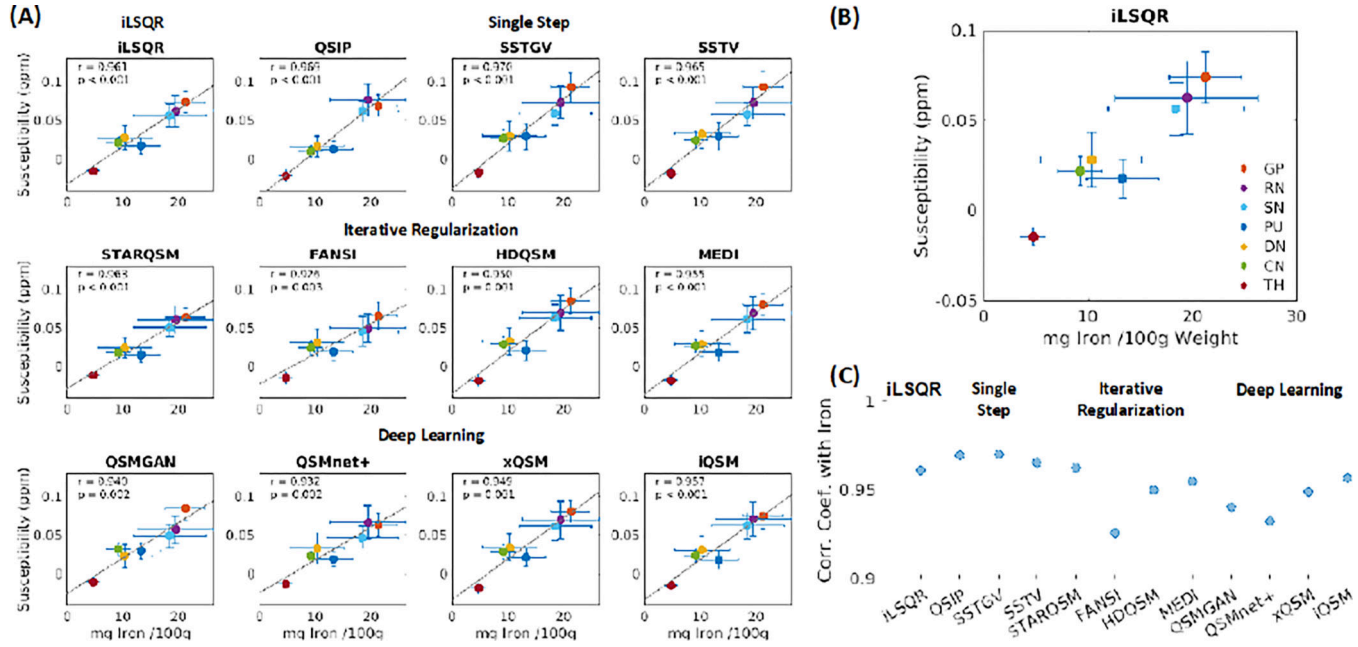

Figure 7.

Correlation between QSM values and postmortem iron quantification. (A) Regional brain susceptibility versus iron concentration. (B) Brain ROI labels on iLSQR plot. (C) Summary of the coefficients in cohort-averaged susceptibility-iron correlation. TH: thalamus; GP: globus pallidus; RN: red nucleus; SN: substantia nigra; PU: putamen; DN: dentate nucleus; CN: caudate nucleus.

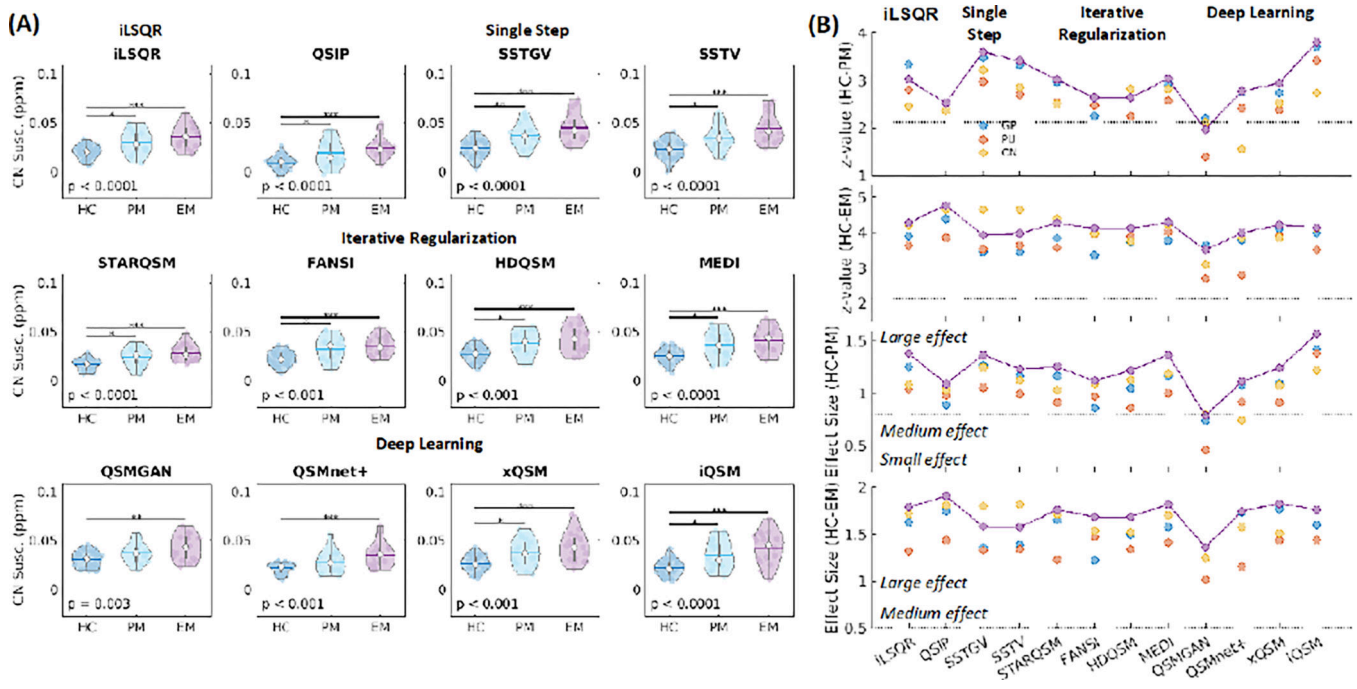

Figure 8.

Comparison of QSM values across healthy volunteers, premanifest HD subjects, and early manifest HD patients. (A) Kruskal-Wallis comparisons of CN susceptibility using different QSM methods across the three groups. The Kruskal-Wallis comparison p-values were shown in each subfigure. (B) Test statistics and estimated effect sizes between HC and PM, and between HC and EM. The dashed lines represent the critical value after Dunn's correction of multiple comparisons on z-value plots and effect size ranges on the effect size plots. HD: Huntington's disease; CN: caudate nucleus; HC: healthy control; PM: premanifest HD; EM: early manifest HD; GP: globus pallidus; PU: putamen; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Table 1.

QSM reconstruction algorithms. χ: tissue susceptibility; W: weighting; F/F^H: Fourier transform and inverse Fourier transform; D: dipole kernel; Φ: tissue field; λ: regularization weight; Ω: regularization term; LoT: Laplacian-of-Trigonometric.

QSM Method	Reconstruction	Algorithm			Parameters
Multiple orientations	**COSMOS**	$argmin_\chi \frac{1}{2}\left\|W\left(F^H D_{multi} F\chi - \Phi\right)\right\|_2^2$	Calculation of susceptibility through multiple orientation sampling and weighted least squares.		N/A
Ill-conditioned k-space estimation	**iLSQR**	Initial estimation using a sparse linear equation and least-squares (LSQR)-algorithm-based method; estimation of the susceptibility boundaries using a fast quantitative susceptibility mapping method; estimation of the susceptibility artifact from ill-conditioned k-space regions only using an iterative approach.			LSQR error tolerance = 0.01
Iterative optimization using regularization			*Data consistency term:*	*Regularization term:*	
	STARQSM		Linear L2-norm	Total variation	$\lambda_1 = 0.1$ $\lambda_2 = 1\times10^{-5}$
	FANSI	Linear form: $argmin_\chi \frac{1}{2}\left\|W\left(F^H DF\chi - \Phi\right)\right\|_p + \lambda\Omega\left(\chi\right)$ Nonlinear form: $argmin_\chi \frac{1}{2}\left\|W\left(e^{iF^H DF\chi} - e^{i\Phi}\right)\right\|_p + \lambda\Omega\left(\chi\right)$	Nonlinear L2-norm	Total variation	$\lambda = 1\times10^{-4}$ $\mu_1 = 0.1$ $\mu_2 = 1$
	HDQSM		Linear L1+L2-norm	Total variation	$\lambda_1 = 0.0079$ $\lambda_2 = 6.3\times10^{-5}$ $\mu_1 = 0.079$ $\mu_2 = 6.3\times10^{-3}$
	MEDI		Linear L2-norm	L1-norm of morphologically weighted gradients	$\lambda = 1780$
Single-step methods	**QSIP**	Minimizing (1) the disagreement between Laplacian of the estimated and true field maps; (2) the disagreement between estimated and bias field corrected filed map; and (3) the non-zero external susceptibility sources from outside the brain.			$\lambda_1 = 1\times10^{-8}$ $\lambda_2 = 1\times10^{-6}$ $\lambda_3 = 1\times10^{10}$
	SSTV/SSTGV	Combining the VSHARP background field removal with regularized dipole inversion using total variation or total generalized variation constraints.			SSTV $\lambda = 0.004$ SSTGV $\lambda_0 = 0.0079$, $\lambda_1 = 0.004$
Deep learning methods	**QSMGAN**	3D Generative Adversarial Networks. Trained on 7T COSMOS data.			N/A
	QSMnet+	3D U-net; data augmentation for a wider range of susceptibility. Trained on 3T COSMOS data.			N/A
	xQSM	3D U-net with octave convolutional and noise-regularized. Trained on single-orientation QSM data.			N/A

QSM Method	Reconstruction	Algorithm	Parameters
	iQSM	LoT enhanced 3D residual U-net. Trained on 3T single-orientation QSM data. * Single-step method.	N/A

Table 2.

Summary of the performance of different QSM methods. Six characteristics were included: Overall visual quality, XSIM compared to COSMOS QSM, correlation of basal ganglia susceptibility with age, correlation of susceptibility and iron quantification in different brain regions, and the test statistics of susceptibility comparison between HC and PM/EM. QSM methods were grouped according to their categories. HC: healthy control; PM: premanifest HD; EM: early manifest HD; XSIM: susceptibility-optimized similarity index metric.

		Visual Quality[a]	Similarity to COSMOS[b]	Correlation with Age[c]	Correlation with Iron[d]	Differentiate PM/HC[e]	Differentiate EM/HC[e]
Single-step methods	iLSQR	+	+	+	++	+	++
	QSIP	+	+	+	++	+	++
	SSTGV	+	+	+	++	++	++
	SSTV	+	+	+	++	++	++
Iterative optimization using regularization	STARQSM	++	++	+	++	+	++
	FANSI	−	+	+	+	+	++
	HDQSM	+	+	+	+	+	++
	MEDI	+	+	+	++	+	++
Deep learning methods	QSMGAN	++	++	+	+	+	++
	QSMnet+	++	++	+	+	+	++
	xQSM	−	+	+	+	+	++
	iQSM (single-step DL)	−	+	+	++	++	++

[a] Mean rating of visual quality: – > 1.2; + 0.8–1.2; ++ < 0.8.

[b] XSIM with COSMOS: – < 0.3; + 0.3–0.5; ++ > 0.5.

[c] Correlation coefficient with age: – < 0.5; + 0.5–0.7; ++ > 0.7.

[d] Correlation coefficient with iron: – < 0.9; + 0.9–0.95; ++ > 0.95.

[e] Test statistics of differentiation of HD subjects from HC: – < 1.96; + 1.96–3.29; ++ > 3.29. Note: 1.96 and 3.29 correspond to the z-score two-tailed cutoff at 0.05 and 0.001, respectively.