**Measurement of serum ferritin levels**

A 10-ml blood sample was collected from the majority of the subjects (*N*=37) for the *in vitro* quantitative determination of serum ferritin as a representative measure of the body’s iron reserves. The sample was first centrifuged at 24,000 rpm for a period of 10 min to separate hematocrit from plasma, which was subsequently stored in 1000μl aliquots at −70 °C. Serum ferritin levels were quantified based on the electochemiluminescence immunoassay, in which a voltage applied to a sample containing tagged ferritin molecules induces chemiluminescent emissions that are measured by a photomultiplier (Elecsys 2010, Roche Diagnostics GmbH, Mannheim, Germany).

**Masking of deep grey matter nuclei**

Masks of the striatum (caudate-putamen), globus pallidus and thalamus were obtained via the FSL-FIRST Bayesian model-based subcortical segmentation algorithm (18), which was applied on optimized hybrid-contrast MP2RAGE-QSM images (19, 20). Robust co-registration between skull-stripped MP2RAGE and FLASH data was achieved using rigid-body linear transformation of the *T*1-weighted data onto N4 bias-field-corrected FLASH magnitude data (http://github.com/stnava/ANTs). Given the difficulty of segmenting brainstem and cerebellar nuclei on *T*1-weighted data due to lack of contrast, in addition to the infeasibility of performing manual segmentation of multiple nuclei in many subjects, we utilized an atlas-based registration approach to achieve accurate delineations of brainstem/cerebellar nuclei. Specifically, the diffeomorphic greedy-SyN ANTs non-linear transformation model (http://github.com/stnava/ANTs) was used to compute a nonlinear transformation warp between MP2RAGE and MNI space, which was used to map each subject’s QSM data into standard space for subsequent calculation of a population-specific average image. The standardized QSM template exhibited high contrast in brainstem/cerebellar regions and was used to carefully delineate masks of the STN, SN, RN and DN (Fig. 1). All masks were delineated by the same operator and were subsequently warped back into native QSM space. The same atlas-based registration procedure was applied to obtain subject-specific masks of the lateral ventricles, which were used for referencing the QSM data to cerebrospinal fluid (CSF; median values). All masks were thresholded at 0.5 to ensure maximal inclusion of GM tissue while limiting partial-volume effects. Following visual inspection of all the masks for quality, median susceptibility values from all ROIs were computed for further analysis.

**Quality Control**

To account for potential differences in the severity of motion-related image artifacts, we used a step-wise, multivariate outlier-detection approach implementing a robust Mahalanobis distance framework (15) to remove low-quality data based on *(a)* structural image quality indices calculated on the magnitude structural images and *(b)* susceptibility values extracted from subcortical nuclei. In general, Mahalanobis distance calculates how far each observation is to the center of a joint distribution, which can be thought of as the centroid in multivariate space. Robust distances are estimated from minimum covariance determinant estimators rather than the sample covariance. Data were regarded as outliers if the robust Mahalanobis distance was greater than the 97.5% quantile of the chi-square distribution (Supplementary Figure S1). In the first step, multivariate outliers were detected based on *(a)* the Shannon entropy focus criterion (EFC), which is an index for image ghosting and blurring (21); *(b)* the quality index 1 (QI1), which is an index for image degradation resulting from bulk motion, residual magnetization, incomplete spoiling and ghosting (22); and *(c)* the smoothness of voxels calculated as the full width at half maximum (FWHM) of the spatial distribution of image intensity values in voxel units (https://github.com/preprocessed-connectomes-project). This step was implemented on the whole sample and identified one severely affected dataset, which was marked for removal (Supplementary Figure S1A). To ensure that the remaining datasets did not contain further outliers, multivariate robust squared Mahalanobis distance outlier detection was additionally performed on vectors of median Δχ values extracted from the subcortical masks for each sample separately. This procedure identified four outlier datasets within the patient sample, which were marked for removal (Supplementary Figure S1B). Following quality control, group comparisons of magnitude image quality metrics (signal-to-noise ratio; SNR; contrast-to-noise ratio, CNR; foreground-to-background ratio, FBER; voxel smoothness; EFC; QI1) revealed no significant differences between patients and controls (Supplementary Table S1).