Supplementary Materials S1 – S5

S1: Efficiency of robust combination and the Fermi function

Robust combination and two-repeat multi-parameter mapping protocol

This section elaborates on the sensitivity of the robust combination to outliers and its dependence on the k-parameter of the Fermi function. Moreover, the heuristic selection criteria that was employed to determine the optimal k-parameter for this study is explained. Both are illustrated for the MTsat contrast depicted in Figure S1.1.
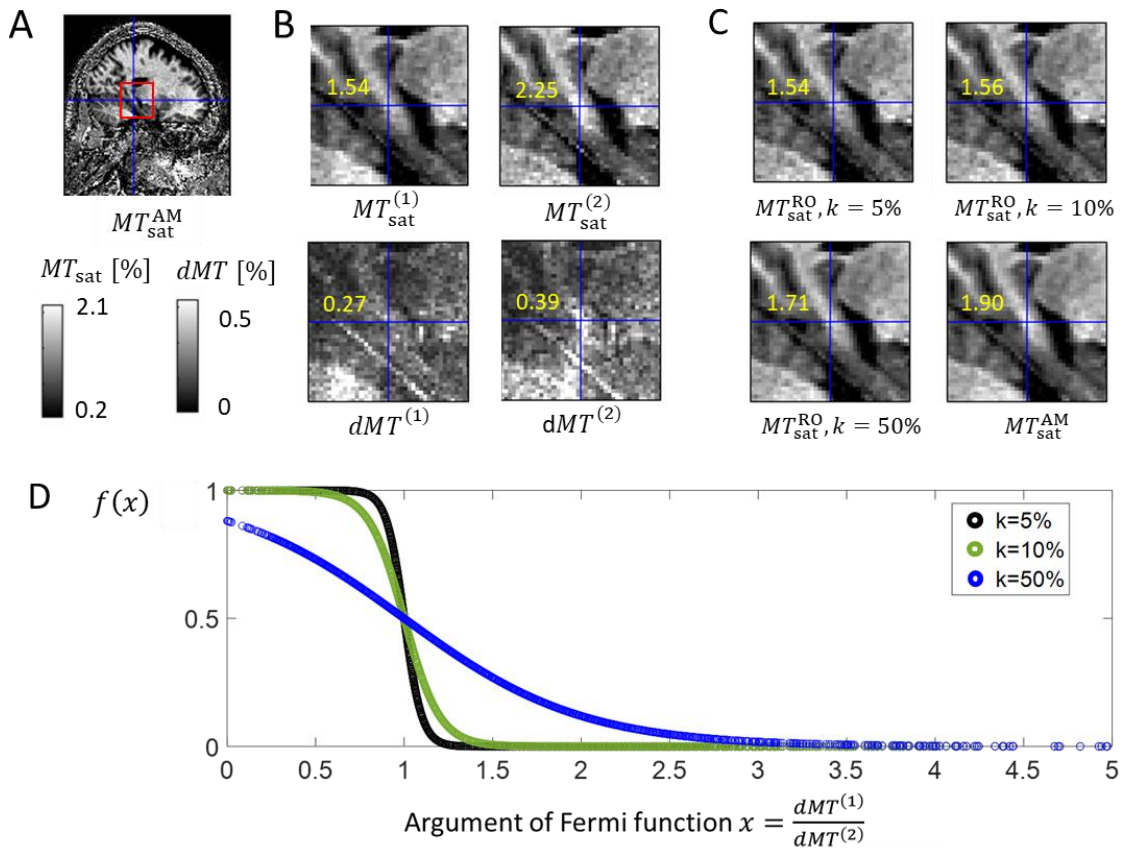

Figure S1.1: Efficiency of robust combination and how it is modulated by the Fermi-function k-parameter, exemplified for MTsat. A: Arithmetic mean of the MTsat maps from the two repeat multi-parameter mapping protocol P1. Region in red box is magnified in subfigures **B** and **C**. **B (top-left to bottom-right):** MTsat map from each run ($MT_{\text{sat}}^{(1)}, MT_{\text{sat}}^{(2)}$) and the corresponding error maps ($dMT^{(1)}, dMT^{(2)}$). The yellow numbers are the respective grayscale values at the crosshair position depicted for each map: in the second run, the erroneously high MTsat value ($MT_{\text{sat}}^{(2)} = 2.25$ compared to $MT_{\text{sat}}^{(1)} = 1.54$) is accompanied by a higher corresponding error value ($dMT^{(2)} = 0.39$ compared to $dMT^{(1)} = 0.27$). **C (top-left to bottom-right):** Robust combination using three different k-parameters (5%, 10%, and 50%) as well as arithmetic mean combination. Again, the grayscale value at the crosshair position is shown in yellow. The effectiveness of robust combination was reduced when higher k-parameters were used. **D:** Depicts the Fermi-function used for robust combination (see Eq. (7)) as a function of the ratio of errors from both repeats and for the different k-parameters used in **C**. The higher the k-parameter, the smoother the transition from one to zero in the Fermi function. The curves in **D** are exemplified for the ensemble of errors of the MTsat parameters across an entire slice (z=85, each circle represents the error values in one voxel of the slice).

Figure S1.1 illustrates, for the MTsat map, how the k-parameter in the Fermi function can tune the efficiency of the robust combination in reducing a regionally-localized artefact (Fig. S1.1A, red box) in a two-repeat multi-parameter mapping dataset. The artefact appeared in the second repeat in the form of an unusually high MTsat value ($MT_{\text{sat}}^{(2)} = 2.25$ as compared to $MT_{\text{sat}}^{(1)} = 1.54$) that was accompanied by a higher error (Fig. S1.1B, crosshair in second column). The biased MTsat value was still present in the arithmetic mean combination ($MT_{\text{sat}}^{\text{AM}} = 1.90$) but could be substantially reduced with robust combination ($MT_{\text{sat}}^{\text{RO}} = 1.54$ for $k = 5\%$). The level of artefact reduction in the robust combination (assessed by how close the MTsat value was to the repeat without artefact) depended on the k-parameter of the Fermi function: at the lowest k-parameters (Fig. S1.1C, top-left) the robust combination was the same as the unbiased MTsat value of the first run whereas at the highest k-parameter (Fig. S1.1C, bottom-left) the robust combined MTsat value was still biased, although to a smaller amount than the arithmetic mean combined MTsat. Another important observation was that the higher accuracy of the robust combination at lower k-parameter was accompanied by a higher noise level in the image: the robustly combined MTsat image with the lowest k-parameter was noisier than its counterpart with the highest k-parameter.

The reason why the increase in accuracy comes at the price of an increased noise level can be seen in Figure S1.1D. This subfigure shows that for a low k-parameter (black curve) the shape of the Fermi function becomes almost like a Heaviside function, i.e. it is 1 for the case that the error is of the first repeat is smaller than the error of the second repeat and it is 0 of the opposite. For this scenario, the robust combination is selecting only the MTsat value of one of the two repeats if there is even only a small difference in the error between the two repeats. For a large k-parameter (blue curve), the Fermi function shows an almost linear dependence on the ratio of errors of the two repeats. In this scenario, both repeats strongly contribute to the robust combination also for voxels where the errors are different in the two runs. We conclude that outliers are better reduced in robust combination with lower k-parameter but random noise is better reduced at higher k-parameter. In this study, we used an intermediate k-parameter (k=10%) to profit from both reduced sensitivity to outliers and noise.

Choice of weighting function for two-repeat protocol

Robust combination (e.g., (Mohammadi et al., 2012)) is a special case of weighted combinations that aims to reduce the influence of outliers. While in statistics weighted combinations often directly use the fitting residuals as weights (e.g., $\frac{1}{\sigma^2}$ in weighted least squares fitting (Koay et al., 2006)), robust combination methods often use nonlinear functions of the error to determine the weights such as e.g., Gaussian (Zwiers, 2010) or Lorentzian (Mohammadi et al., 2012) functions. This enables a more efficient down-weighting of outliers. However, they require determining the median of the error in the image as a global reference value for normalising the weights. This can be problematic if there is a spatially varying trend in the error maps across the brain, as found in the error maps of this study (Fig. S2.1): The error in the skull is apparently higher than in the brain, but there are local variations in the brain as well (e.g. the error in the cerebellum is higher than in the basal areas). To circumvent this problem, we used the ratio of the errors of the two repeats. As such, it is sensitive to variations of the local error across repeats without requiring a global reference value. For the two-repeat combination the Fermi function is an intuitive choice because, through the k-parameter (Fig. S1.1D), it can be easily varied between two extremes: from a binary weighting function (i.e. a

Heaviside function that selects one of the two repeats on a voxelwise basis) to an arithmetic mean combination (i.e. a constant ½-weight function that weights both repeats equally).

Generalising the robust combination to more than two repeats

The proposed robust combination can be generalised to an arbitrary number of repeats, N, by rewriting Eq. (7) as follows:

(S1) $R_1^{RO} = \frac{1}{W} \sum_{j=1}^{N} w_{R1}^j \, R_1^{(j)}$,

where $W = \sum_{j=1}^{N} w_{R1}^j$; $w_{R1}^j = \frac{1}{mw_{R1}} f\left(\frac{dR_1^{(j)}}{mdR_1}\right)$; $f(x) = \frac{1}{\exp(100\,(x-1)/k)+1}$ is the Fermi function; mdR_1 is the median error across all repeats: $mdR_1 = median\left(\left\{dR_1^{(1)}, \dots, dR_1^{(N)}\right\}\right)$; $R_1^{(j)}$, $dR_1^{(j)}$ are the longitudinal relaxation rates and their respective errors from the *j-th* repeat; and $mw_{R1} = \max_{\vec{r}}(f(\vec{r}))$ is defined as the maximum weight across voxels $\vec{r}$. The same idea can also be used to robustly combine the other two MPM parameters (PD and MT_{sat}).

S2: How to use the hMRI toolbox to generate uncertainty maps, error maps, mSNR maps and robust combinations

To generate uncertainty maps, error maps and mSNR maps, they need to be enabled. This can be done via the configuration toolbox module (see Fig. S2.1). An example of a customized defaults file which computes error maps and does robust combination can be found in the hMRI toolbox subfolder "examples" named as "hmri_local_defaults_errormaps.m".
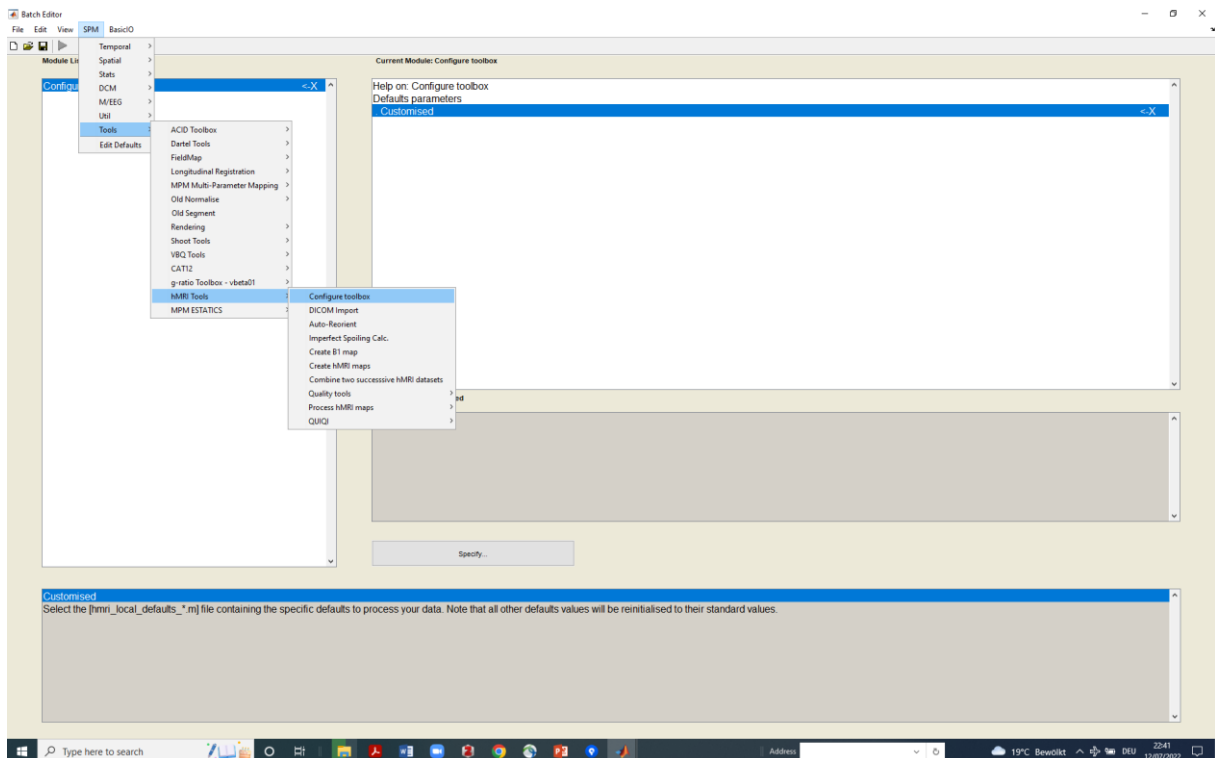

Figure S2.1: Configuration toolbox module. To generate uncertainty maps, error maps, mSNR maps, the customized defaults file can be loaded into the configuration toolbox module. The customized defaults file is named "hmri_local_defaults_errormaps.m" and can be found in the examples folder.

Robust combination of MPM parameters can be generated using a two-repeat MPM measurement and estimating the MPM parameters after enabling the generation of error maps (see Fig. S2.1). To this end, the robust combination module is used (Fig. S2.2). The three MPM parameters (MTsat, PD, R1) of the first and second repeat are loaded into the first and second entry of the robust combination module. Then, the error maps of the first and second repeat are loaded into the third and fourth entry of the robust combination module (in the same order as the MPM parameters). The fifth entry is optional and allows one to define a reference image to which both repeats are resliced when combining them. The last entry is also optional and allows to load in a brain mask.
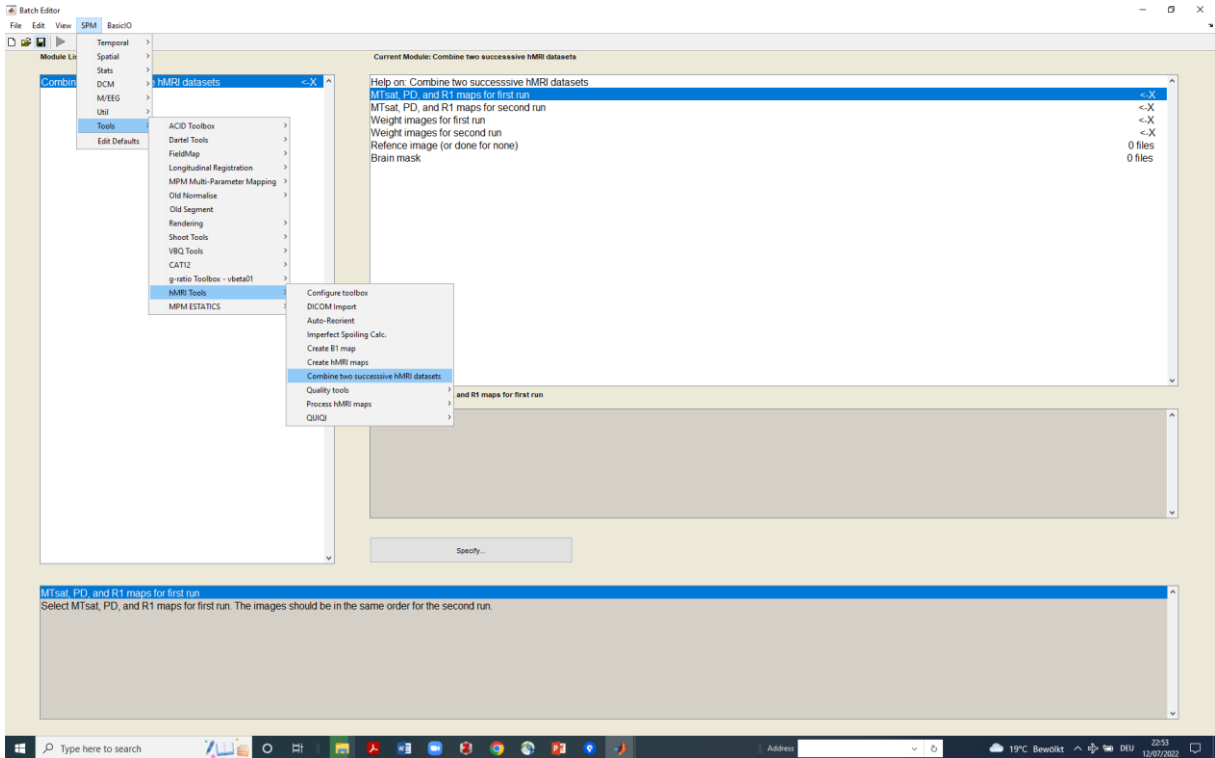
Figure S2.2: Robust combination module. The three MPM parameters (MTsat, PD, R1) of the first and second repeat are loaded into the first and second entry. Then, the error maps of the first and second repeat are loaded into the third and fourth entry (in the same order as the MPM parameters). The fifth entry is optional and allows one to define a reference image to which both repeats are resliced when combining them. The last entry is also optional and allows to load in a brain mask.

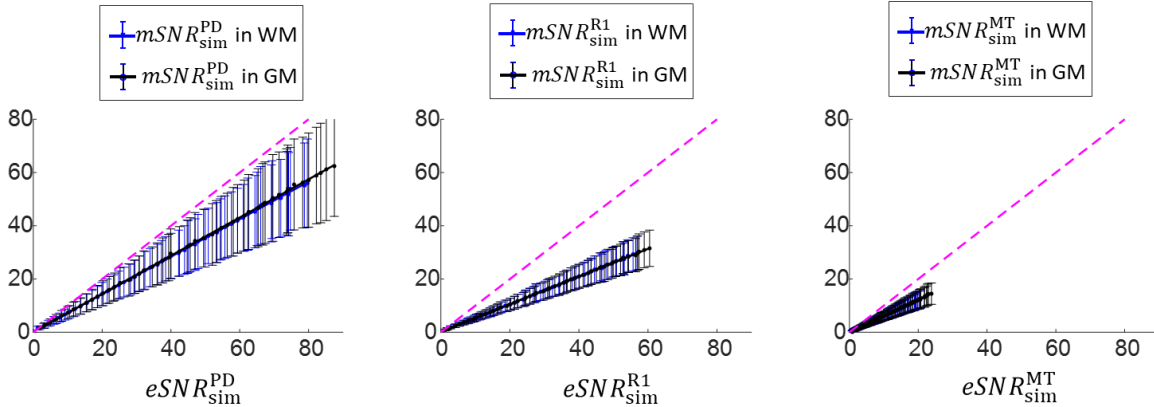# S3: Comparing estimated mSNR against experimental SNR of MPM parameters using simulations

Figure S3.1: Relating mSNR to the experimentally assessed SNR (eSNR) in simulations. Depicted is SNR_{sim}^m as a function of $eSNR_{sim}^m$ (with $m \in \{PD, R1, MT\}$) using the simulations described in the Materials and Methods section, Analysis II. $eSNR_{sim}^m$ is defined as the ratio between the ground truth parameter divided by the standard deviation of the parameter estimated from noisy data (Eq. (S3.1)). A heuristic linear relation is fitted between SNR_{sim}^m and $eSNR_{sim}^m$ data (see Table S3). In blue, the data in white matter is depicted and in black the data in grey matter. The dashed line in magenta is the unity line.

In this section, we use simulations to demonstrate how the proposed mSNR measure is related to the experimentally determined SNR ($eSNR$) of each MPM parameter. To calculate the eSNR, we divided the ground-truth of a given MPM parameter by the standard deviation of the parameters, estimated from noise data using the simulations in Analysis II:

(S3.1) $eSNR_{sim}^m = m^{GT}/std([m(1), \dots, m(N)])$,

with $m(k)$ being the m parameter estimated from the simulated MPM dataset with the k-th noise realisation, and m^{GT} being the ground truth parameter (see Table 2), where $m \in \{PD, R_1, MT_{sat}\}$ and N is the number of noise realisations, e.g., $eSNR_{sim}^{R1} = \frac{R_1^{GT}}{std([R_1(1), \dots, R_1(N)])}$.

The results are shown in Figure S3.1. We found that mSNR was linearly related to eSNR for all three MPM parameters. The apparent linear relation was assessed by:

(S3.2) $mSNR_{sim}^m = A_{sim}^m \, eSNR_{sim}^m + B_{sim}^m$.

The slope A_{sim}^m was always smaller than one and varied between 0.5 and 0.7 (Table S3).

	PD - WM	PD - GM	R1 - WM	R1 - GM	MT - WM	MT - GM
Offset B_{sim}^m	0.11	0.24	0.19	0.12	-0.04	-0.02
Slope A_{sim}^m	0.72	0.70	0.52	0.52	0.61	0.60
Table S3: Relation between model-based and experimental SNR of MPM parameters for simulated data. The coefficients of the heuristic linear models (Eq. (S3.2)) that relate $eSNR_{sim}^m$ to $mSNR_{sim}^m$, with A^m being the slope and B^m being the intercept ($m \in \{PD, R1, MT\}$). The subscript "sim" refers to the simulated data.						

S4: Spatial sensitivity of mSNR maps to coil configuration and head orientation

In this section, we investigate the effect of different head coil configurations and coil sensitivity-profiles on the mSNR maps. To this end, we used the protocol P2 to acquire MPMs of a fixed human post-mortem brain, embedded into phosphate-buffered saline (PBS) in a spherical container. Using a post-mortem brain enabled reorientation of the head by 180° and avoided confounding artefacts, e.g., subject motion. To assess the dependence of mSNR on the sensitivity profile, we acquired the mSNR maps at two different head positions using two different head coils (32 channel coil [32ch] and 64 channel coil [64ch]) and registered the second head position to the first for one of the two head coils (32ch).

Figure S4.1 shows the effect of two different head coils on the mSNR maps. The mSNR maps varied locally when different head coils were used. The regions highlighted by the black arrow in Figure S4.1B differed strongly in mSNR when the 64ch coils was used (top row) as compared to the 32ch coil (bottom row). The mSNR was, however, almost independent of the orientation of the head (Fig. S4.1B: first column vs. second column).
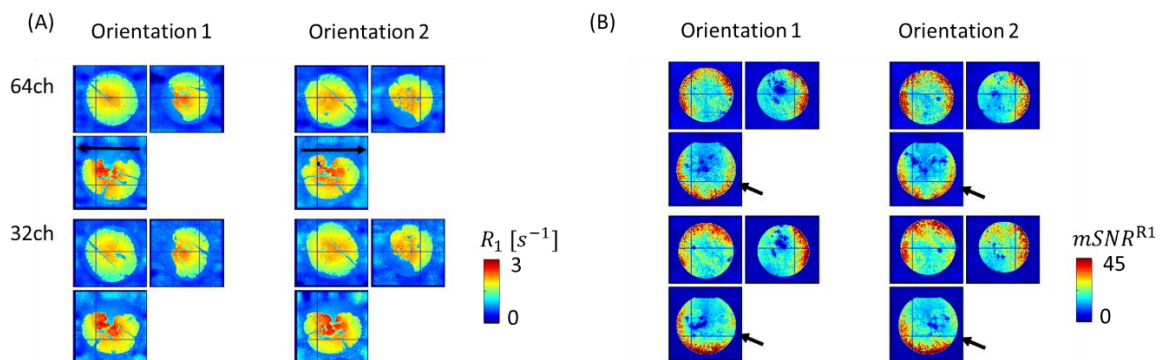

Figure S4.1: Effect of head coil configuration on mSNR – illustration for R1 parameter. Depicted is the R_1 map (A) and the associated $mSNR^{R1}$ map (B) for two different head coils (top row: 64 channel head coil and bottom row: 32 channel head-neck coil) and two different head orientations (first and second column in A and B, see arrow pointing along posterior-anterior direction in top row in A). The black arrows in (B) highlight a region where the mSNR strongly varies between the two head coils but is almost independent of the head orientations. Note that the dark regions in the mSNR maps are caused by air bubbles in the brain: areas with air bubbles have a high error and thus a reduced mSNR. For the presented analysis they can be ignored.

Figure S4.2 depicts the effect of the coil-sensitivity profile on mSNR. After registering the brain acquired at orientation 2 to the brain at orientation 1, the two mSNR maps showed strong spatial differences, indicating that the mSNR maps are highly sensitive to the sensitivity-profile of the head coil.
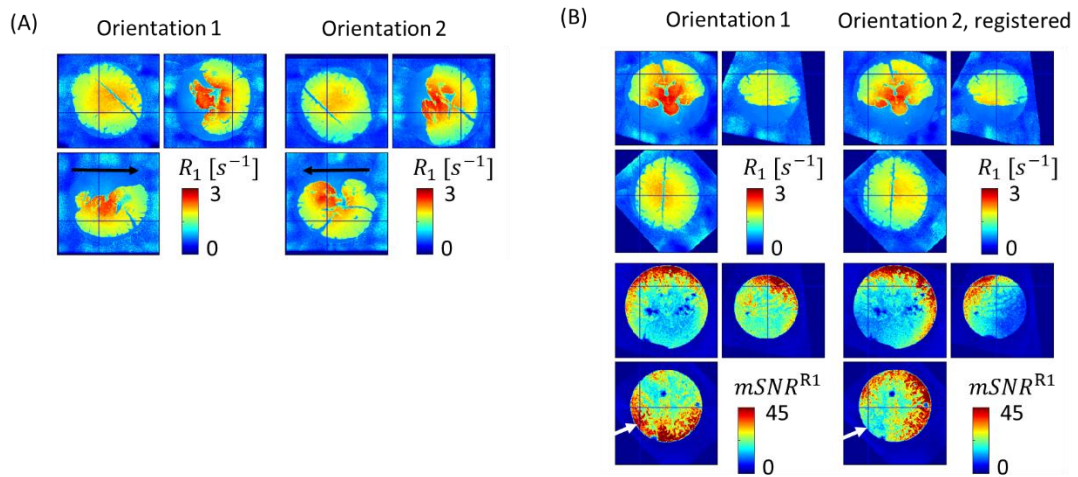
Figure S4.2: Effect of coil-sensitivity profile on mSNR – illustration for R1 parameter. A: Depicted is the R_1 map (top row) at two different head orientations (see arrow pointing along posterior-anterior direction, A) for the 32ch coil. B: Depicted are the R_1 maps (top row) and the associated $mSNR^{R1}$ maps (bottom row) after registering the brain at orientation 2 to orientation 1. While the registered R_1 maps show a similar contrast to the R_1 maps at orientation 1, the mSNR maps show local variation (e.g., region highlighted by white arrow).

S5: Relation between mSNR, error, and uncertainty maps

Figure S5.1 illustrates the relation between mSNR, error, and uncertainty maps for the R1 parameter across the three protocols P1-P3. Across all protocols, the mSNR was low where the error was high. The decline of mSNR was stronger towards the centre of the brain in left-right direction when using P1 as compared to P2 (dashed ellipsoids). This decline was accompanied by an increased error and an increased uncertainty.
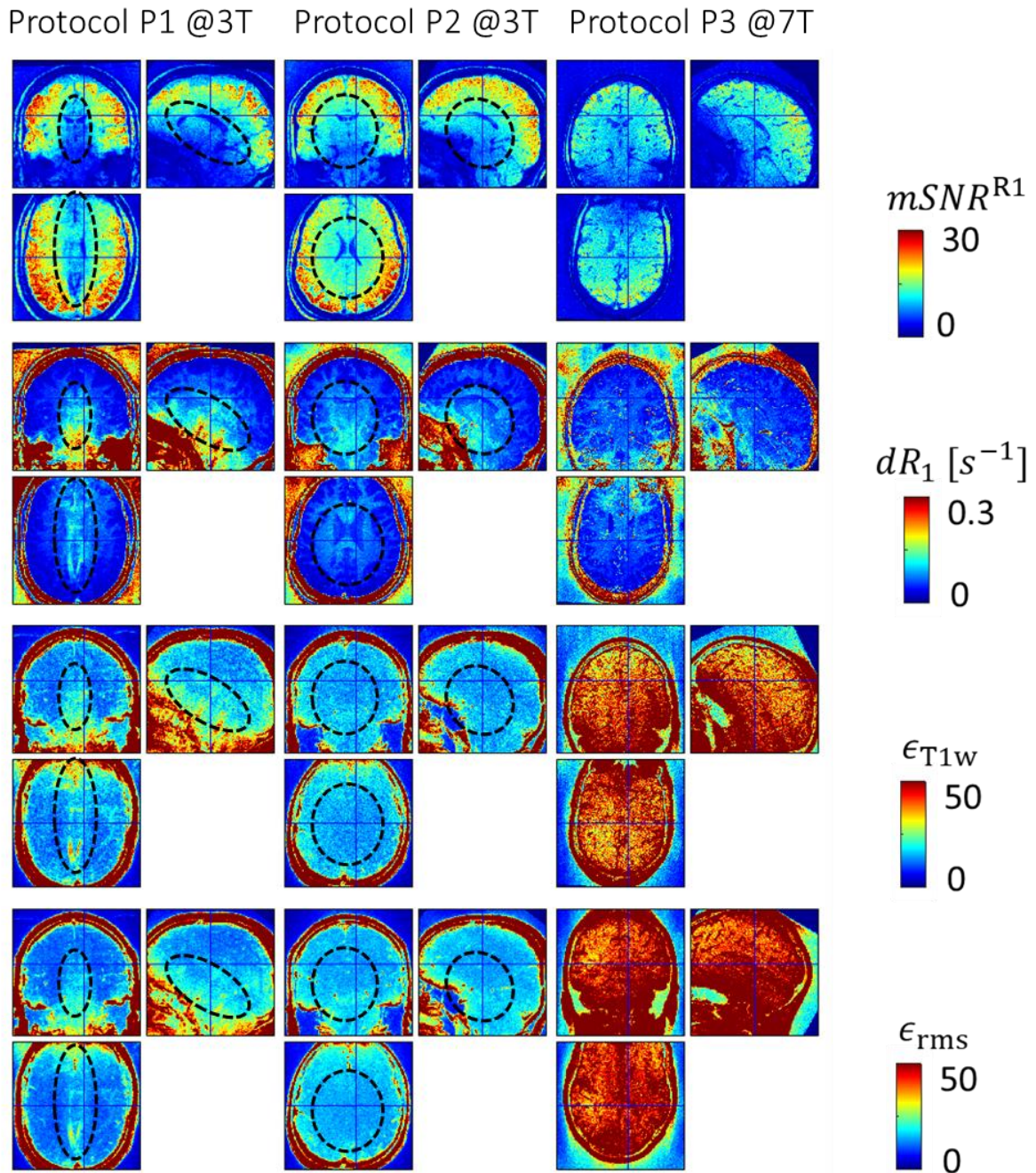

Figure S5.1: Relation between mSNR, error, and uncertainty maps – illustration for R1 parameter. Depicted is the $mSNR^{R1}$ maps (top row), the error maps dR_1 (second row from top), the uncertainty map for the T1w contrasts ϵ_{T1w} (third row from top), and the root-mean-square of the uncertainty maps across all three contrasts ϵ_{rms} (bottom row) for the three different protocols P1-P3. The decline of mSNR is stronger towards the centre of the brain in left-right direction when using P1 as compared to P2 (dashed ellipsoids) s.

S6: Illustration of error maps and robust combination in the presence of motion

This section is divided in two parts. In the first part, the analysis corresponding to Figure 5 is completed, i.e. the same information as in Figure 5 is depicted for the other two MPM parameters (i.e., PD and MT_{sat}) for the same subject. In the second part, a subset of the data from a previous study (Callaghan et al., 2015) is used to demonstrate that the error maps are sensitive to subject motion and that they can be used to mitigate the effect of subject motion on the MPM parameters using the robust combination. To this end, the subject was instructed to lay still in the first run. In the second run, the person was instructed to move. Additionally, to minimize movement artefacts in the first run, a prospective motion correction camera (PMC) was used as described in (Callaghan et al., 2015). In this analysis, the MPM parameters from the first run (i.e., lay still instruction and PMC on) were used as the reference because they are almost free of motion artefacts, whereas the second run is highly corrupted by motion artefacts.

Figures S6.1 and S6.2 complement Figure 5 in the main manuscript by illustrating regionally localized artefacts that were captured by the error maps for the other two MPM parameters (PD and MT_{sat}). They became less pronounced in the arithmetic mean and could be partly removed in the robust combination (see, e.g. red highlighted region).
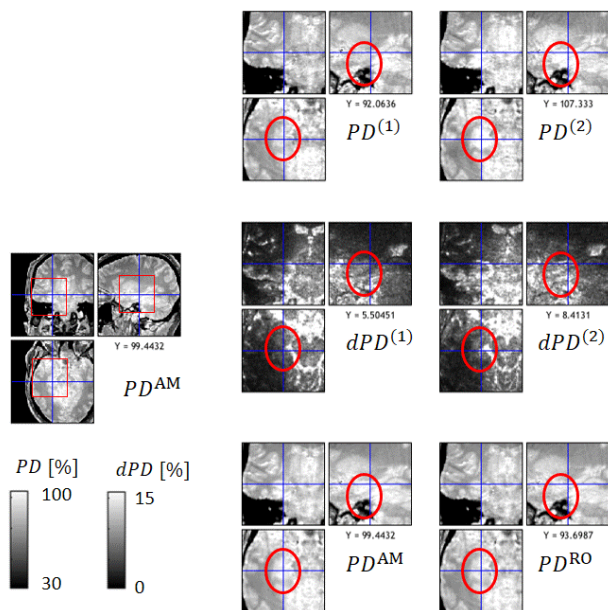

Figure S6.1: Reduced artefacts in robustly-combined proton density (PD) map. Depicted are: two successive repeats of the PD map using protocol 1 with superscript *(1)* and *(2)* (top row), the associated error maps for each repeat (middle row), and their arithmetic mean and robustly combined average with superscript *AM* and *RO* (bottom row). An area is magnified (red box, left column), where the error maps were sensitive to artefacts (hyper intensities) and the robust combined PD contained less artificially increased values than the arithmetic mean (circle) and single-repeat PD maps.
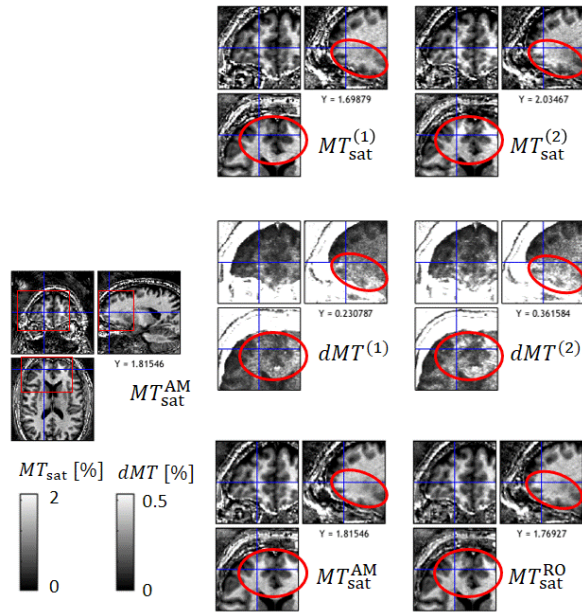
Figure S6.2: **Reduced artefacts in robustly-combined magnetization transfer saturation rate (MT_{sat}) parameter.** Depicted is the same information as in Figure S4.1 for the MT_{sat} parameter instead of the PD parameter.

Figure S6.3A shows the R_1 parameter and associated error maps for a person who moved in the first repeat and laid still in the second repeat. Figure S6.3B shows the arithmetic mean and robust combination of these two runs. The results demonstrated that the error map captures subject movement relate biases in the R_1 parameter: dR_1 was high where the motion-corrupted $R_1^{(1)}$ map was biased (red arrows in Fig. S6.3A) and dR_1 was low where $R_1^{(1)}$ was similar to gold-standard $R_1^{(2)}$ values (green arrows in Fig. S6.3A). When using the error maps in the robust combination R_1^{RO}, the reference $R_1^{(2)}$ values can be largely retrieved whereas the standard arithmetic mean combination R_1^{AM} remains largely biased (Fig. S6.3B).
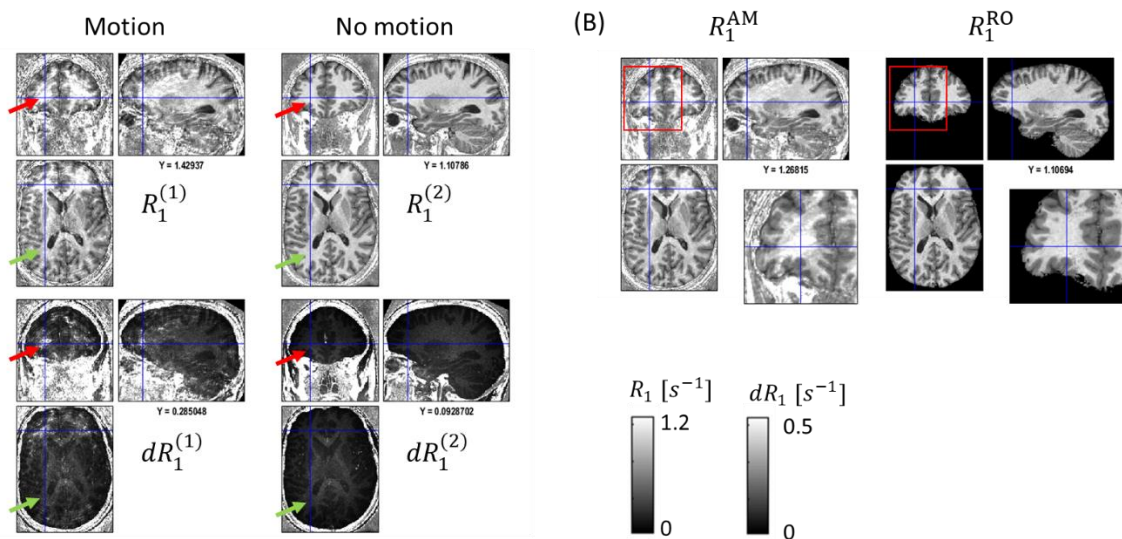

Figure S6.3: **Motion-induced artefacts captured by error maps and reduced by robust combination – illustrated for the R_1 parameter.** (A): Depicted are two longitudinal relaxation rate maps (R_1) and their

associated error maps (dR_1) acquired from two repeats of the MPM protocol. In the first run (superscript *(1)*), the subject was instructed to move and in the second run the person was instructed to lay still (superscript *(2)*). Additionally, to minimize movement artefacts in the second run, a prospective motion correction camera was used as described in (Callaghan et al., 2015). (B): Depicted is the whole-brain view of the arithmetic mean of both runs (R_1^{AM}) as well as their robust combination (R_1^{RO}). In addition, the highlighted region is magnified to better appreciate the reduced artefact level in R_1^{RO} as compared to R_1^{AM}. Note that the image-pixel value at the cross-hair is depicted below the sagittal view.

S7: Reducing artefactual variation at the group level

In this supplementary Figure S7.1, we show how artefactual variations at the group level vary for different combinations of a two-repeat MPM acquisition in the grey matter (GM). Variability at the group level was assessed by the standard-error-of-the-mean (SEM). The SEM showed higher values toward the outer edge, potentially caused by residual inaccuracies in spatial registration.
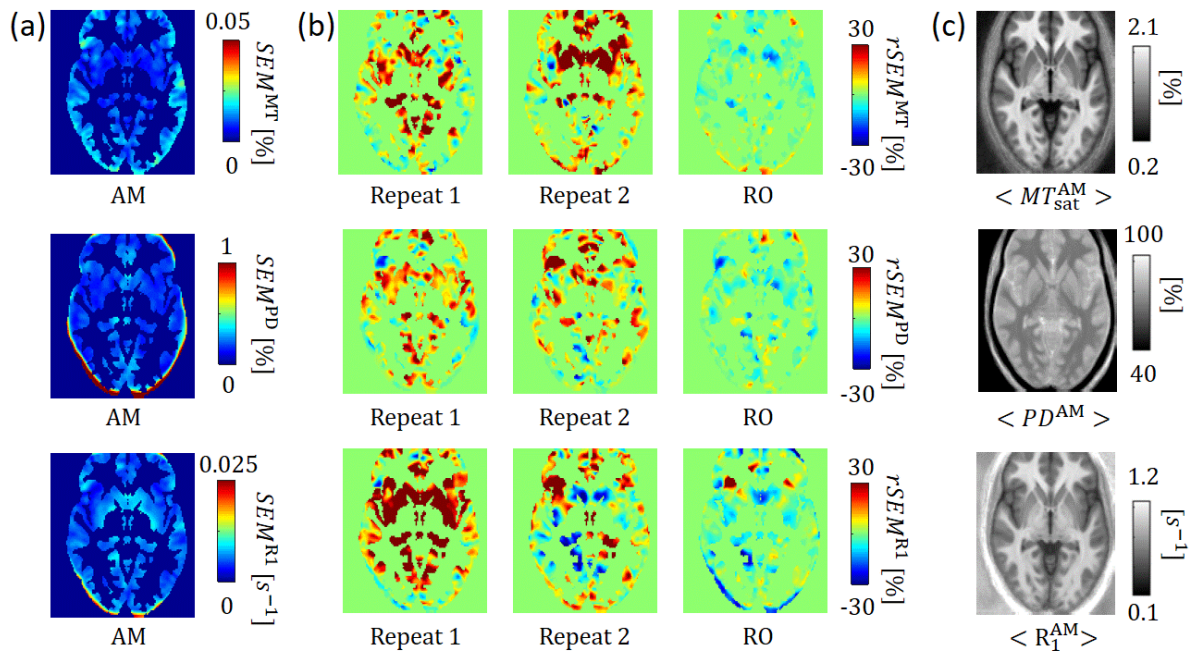

Figure S7.1: Variability across subjects for grey matter. Depicted is the same information as in Fig. 8 but for grey matter instead of white matter.

S8: Across-group spatial variability of mSNR maps

To assess the spatial variability of the error (Fig. S8.1) and mSNR (Fig. S8.2) maps for protocol 1 across the group of healthy subjects, we calculated the group average and standard deviation of the respective maps after non-linear spatial transformation into MNI space. The spatial processing steps were described in the methods section "Map creation and spatial processing".

We found the same overall spatial pattern in the group average of the error and mSNR maps as depicted for the one-subject example in Figure 4b and 4c (see Results section, Analysis I): a strong left-right gradient which appeared as an increase in the error map towards the centre and as a decrease in the mSNR maps. The standard deviation of the error maps across the group was high in the cerebellum for all three contrasts (white ellipse). Additionally, it was high in the CSF for the PD parameter, and in focal posterior regions of the brain for the R_1 parameter.
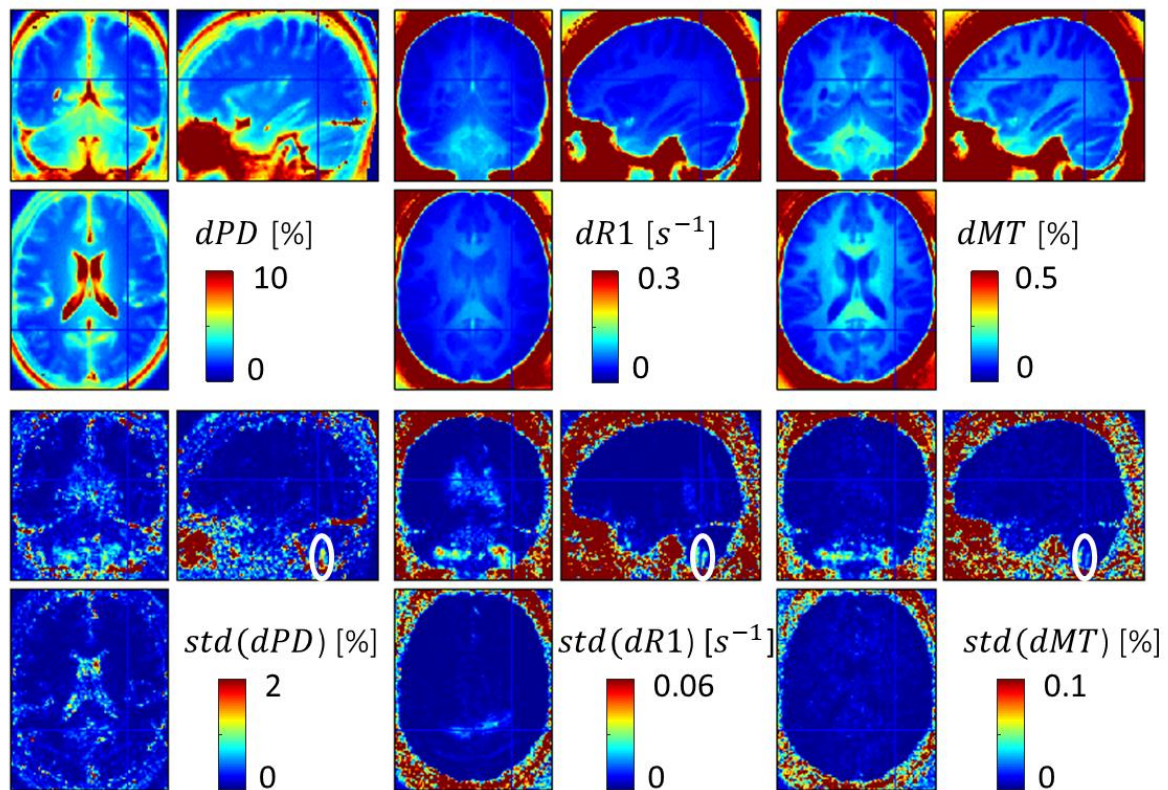

Figure S8.1: Spatial variation of error maps across group. Depicted are the group average (top row) and standard deviation (bottom) of the error maps for the three MPM parameters: PD, R_1, and MT_{sat}. The group average shows the same spatial pattern as seen in the one-subject example in Figure 4. In the standard deviation map, however, artefacts are revealed that appear inconsistently across the group. For example, all errors show high standard deviation in the cerebellum (highlighted with white ellipse).
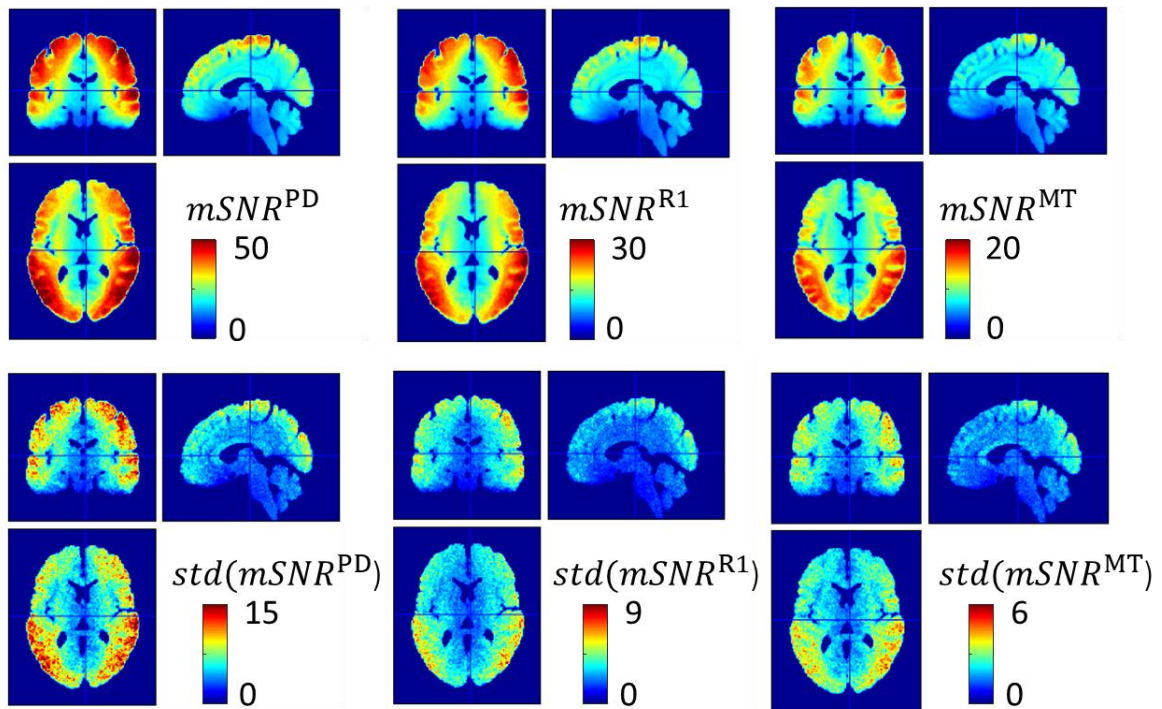
Figure S8.2: Spatial variation of mSNR maps across the group. Depicted are the group average (top row) and standard deviation (bottom row) of the mSNR maps for the three MPM parameters: PD, R_1, and MT_{sat}. The same spatial pattern as in the one-subject example in Figures 3a and 4 can be seen in the group average and in the standard deviation across the group.

References

Callaghan MF, Josephs O, Herbst M, Zaitsev M, Todd N, Weiskopf N (2015) An evaluation of prospective motion correction (PMC) for high resolution quantitative MRI. Front Neurosci 9:97.

Koay CG, Chang L-C, Carew JD, Pierpaoli C, Basser PJ (2006) A unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging. J Magn Reson San Diego Calif 1997 182:115–125.

Mohammadi S, Nagy Z, Hutton C, Josephs O, Weiskopf N (2012) Correction of vibration artifacts in DTI using phase-encoding reversal (COVIPER). Magn Reson Med 68:882–889.

Zwiers MP (2010) Patching cardiac and head motion artefacts in diffusion-weighted images. NeuroImage 53:565–575.