SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Text S1 The specifics of the DLRPG-net processing data

1) The input data of the DLRPG-net is a pair of images acquired using reversed phase-encoding (UP/DOWN) with the EPI sequence, and the outputs are two field maps for UP and DOWN, respectively.
2) The input into the network first goes through a convolution, a batch normalization (BN) operation, and a max-pooling layer to accelerate the model training.
3) The output of the max-pooling layer is passed through two encoder-decoder blocks to obtain the feature maps of different resolutions.
4) Every feature map is passed through two different Residual Leaky ReLU modules to obtain two coefficient matrices with different sizes.
5) Each coefficient matrix is multiplied with a pre-generated basis set to obtain the corresponding final field map component based on the Back Projection module.

Text S2 Details of susceptibility artifacts

Here, we assume the susceptibility artifacts only exist in the PE direction, and the acquired signal for two-dimensional ssEPI in k-space can be given as

$$S\left(k_x,k_y\right)\propto\int_x\int_y\rho(x,y)e^{j(k_xx+k_y(y+\frac{\Delta B_0(x,y)t_{esp}}{G_yT_p}))}dxdy \quad \text{(S2-}$$

Where G_y and T_p denote gradient amplitude and duration, respectively. $\Delta B_0(x,y)$ is the local field inhomogeneity. $\rho(x,y)$ denotes a map of the density of excited protons, t_{esp} denotes echo spacing. k_x and k_y denote the frequency-encoding and slice-selection direction respectively.

From formula 1, we can note that the signal from the real space location (x, y) will be shifted to distorted location $(x, y+\delta y)$, where $\delta y = \frac{\Delta B_0(x,y) t_{esp}}{G_y T_p}$. And the shift in units of pixels (named displacement map) can be written as $u(x,y) = \frac{\delta y}{\Delta y} = \Delta B_0(x,y) t_{esp} N_y$, where $\Delta y = L_y / N_y$ indicates the resolution of one pixel. L_y, N_y denotes the field of view and phase encoding number, respectively. Moreover, the intensity of the distorted image will be proportional to the spin density, so if the distortion causes local compression, the signal strength will be higher through the compression. We can infer that the undistortion image needs to be obtained after the displacement correction and intensity correction of the distorted image.

Text S3 Pixel-wise Jacobian modulation

The distortion caused by susceptibility artifacts can be classified into geometric deformation, which is related to pixel shift, and intensity variation. Intensity variation manifests as unexpected bright or dark regions in the images, arising from pixel intensity accumulation or dispersion that comes along with pixel shift. In high-resolution imaging, with the aggravated image deformation due to lengthened echo spacing (ESP), intensity variation becomes prominent, especially in regions with highly inhomogeneous magnetic field, such as the frontal lobe and bilateral temporal lobes.

The traditional correction module will perform both geometric and intensity correction to obtain the undistorted image. From formula (S2-1), we can infer that susceptibility artifact correction can be divided into 2 steps: geometric and intensity correction. The geometric correction is performed in order to recover each pixel's real intensity (x, y) from the corresponding position $(x, y+u(x,y))$, where $u(x,y)$ is the shift in units of pixels. But as the image intensity will be proportional to the spin density, if distortion leads to local compression, the signal will be higher by the accumulation due to compression, and vice versa in case of local stretching. Thus, the intensity correction is necessary. The intensity can be directly scaled by the Jacobian J, which is the derivative of the offset point (displacement map), termed as J_{Field} in this paper.

$$J(x,y) = 1 + \frac{\partial u(x,y)}{\partial y} = 1 + T_{pe}\frac{\partial \Delta B(x,y)}{\partial y} \quad \text{(S3-}$$

Then, the corresponding geometric and intensity correction can be expressed as follows using the formula (S2-2).

$$I_{Undis}(x,y) = J(x,y) I_{Dis}(x, y+u) \quad \text{(S3-}$$

But because the displacement map generally is spatially smooth, the Jacobian matrix will also be smooth, so its accuracy will be affected by the partial volume effects in areas where the B0 field changes unevenly. Here, we use a new intensity correction method that is further smoothed in vivo experiments based on the two geometric corrected images in a pixel-wise manner, termed as J_{RPG}. The traditional way to calculate the Jacobian matrix is based on Numerical discrete differentiation as the field image is a digital image composed of discrete pixels. To overcome discrete numerical differentiation, the method that can derive the intensity of the undistorted image directly without the Jacobian calculation by combining the pair of corresponding pixels from the two EPI images with opposite PE directions is used, shown in Equation (1) of the main body of the paper.

Text S4 The procedure of simulation for evaluating SSIM metric

The whole procedure is shown in figure S1, and it contains four main steps:

1. Load T2 weighted images based on FSE and the field maps based on 3D GRE from the same mice;
2. T2 weighted image and Field map registration;
3. Perform Bloch Simulation with zero field map to obtain the reference images (Only for PSNR/SSIM metric);
4. Perform Bloch Simulation using field map to obtain the UP/Down images.

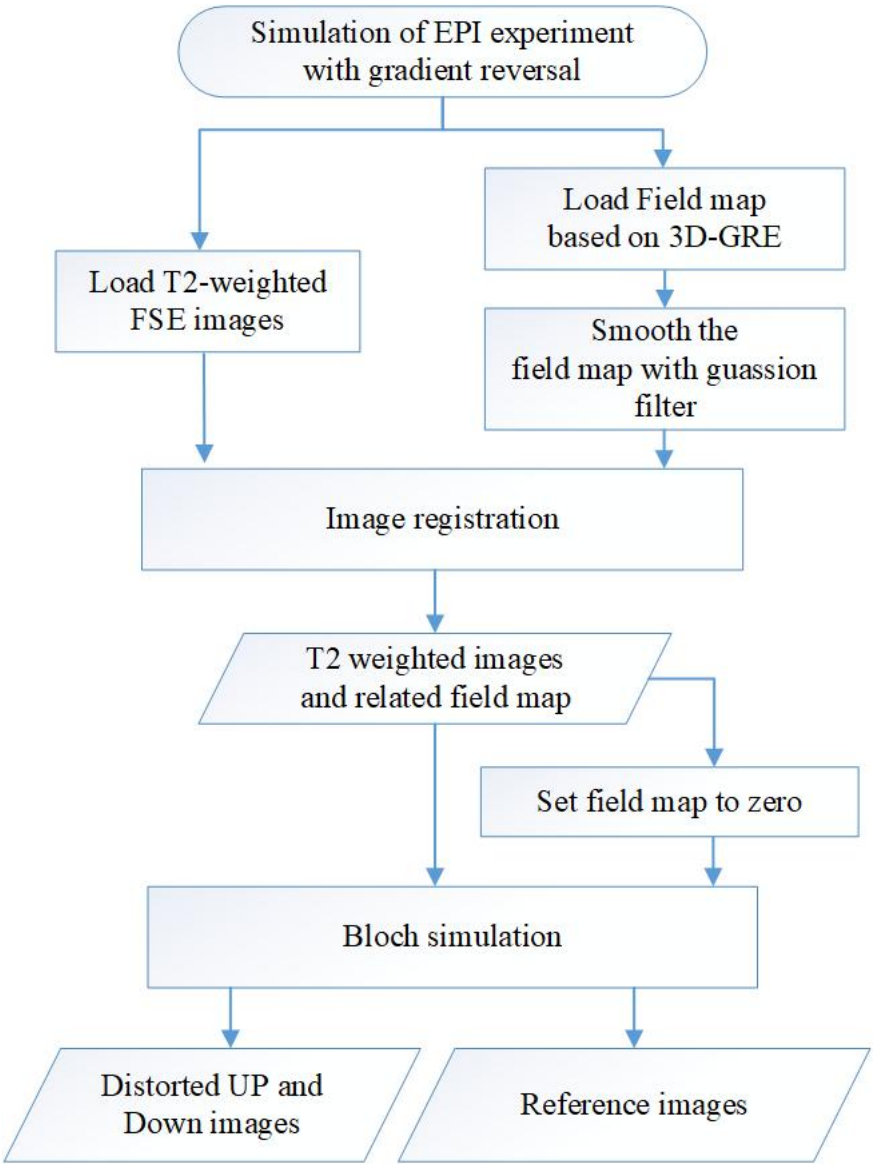

Figure S1. The procedure of simulation for data generation used for evaluating SSIM metric.

Additional clinical data results

In this section, the additional results for clinical data are demonstrated. Firstly, we show the statical distribution of correction for 1000 slices in Figure S2. Then, the results of fractional anisotropy (FA) maps and diffusion-encoded-color (DEC) maps obtained by different methods are shown in Figure S3.

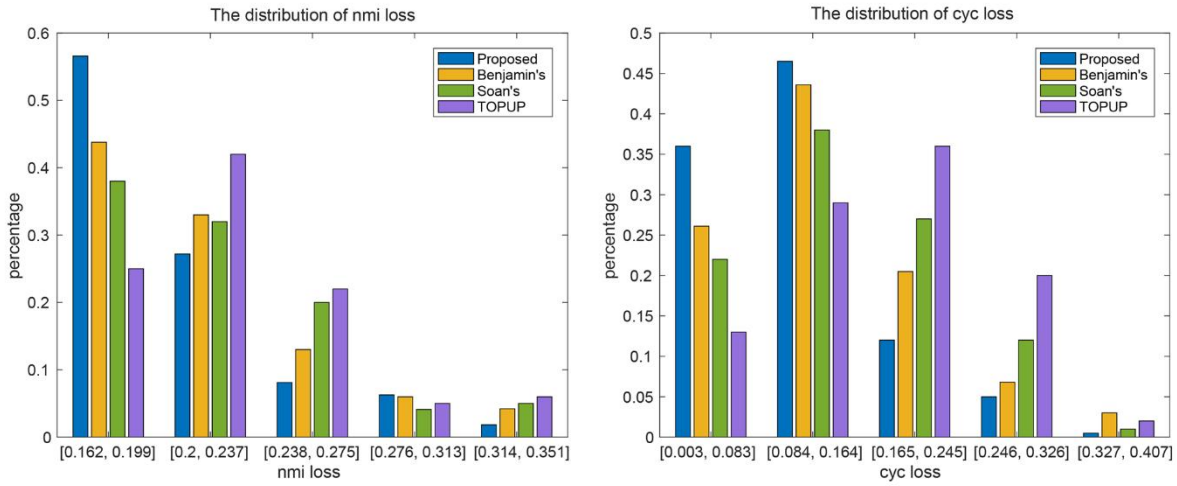

Figure S2. Statical distribution of correction for 1000 slices. The left panel shows the distribution of normalized mutual information (nmi) loss between corrected images and reference images, and the right panel shows the distribution of cycle consistency loss between the cycle back-calculated blip-UP/DOWN images and original UP/DOWN images.

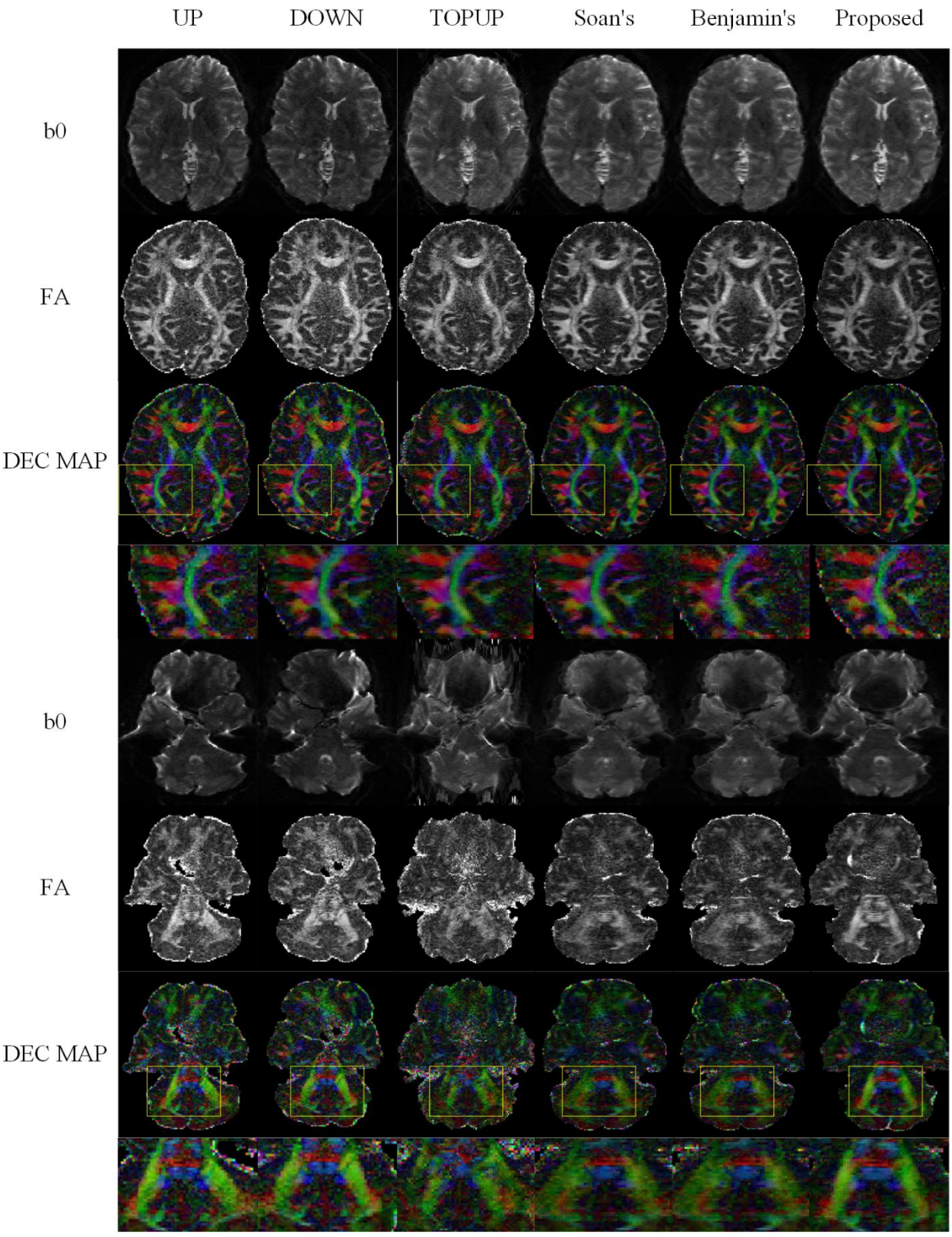

Figure S3. Comparison of different susceptibility artifact correction methods for two slices of the fractional anisotropy (FA) and diffusion-encoded-color (DEC) maps. The rows from up to down show the b0 images, diffusion-weighted images, the FA and DEC maps. The bottom four rows

show the b0 images, diffusion-weighted images, the FA and DEC maps for the slice that suffer from susceptibility artifacts close to the sinuses.

Additional preclinical data results

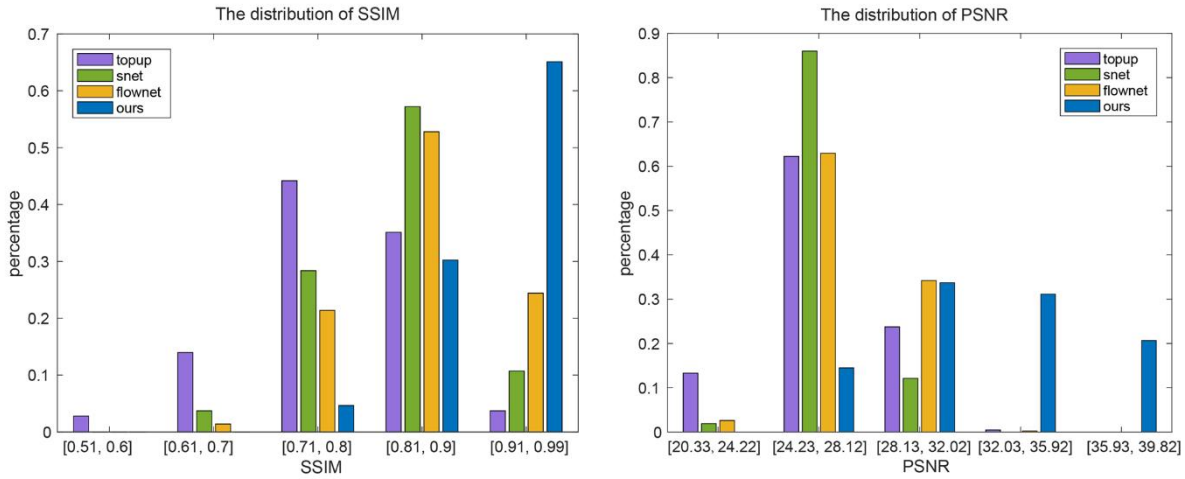

Figure S4. Statical distribution of correction for 500 simulated slices. The left panel shows the distribution of SSIM, and the right panel shows the distribution of PSNR.

Other ablation experiment results

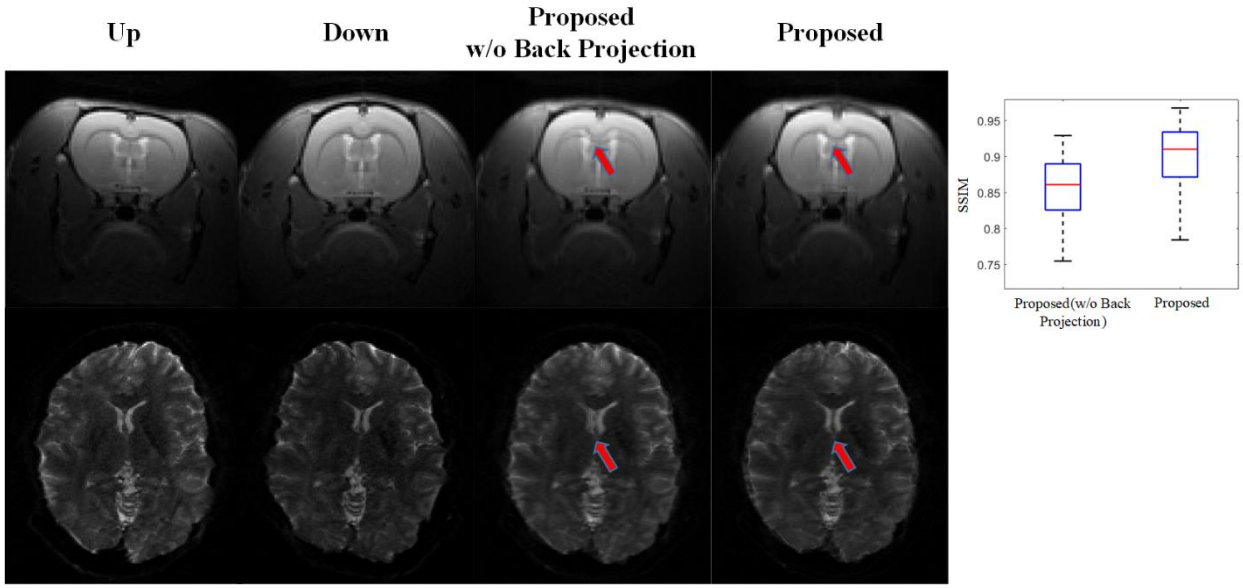

Figure S5. Comparison of our proposed method with and without Back Projection module for preclinical rat data (top row) and HCP data (bottom row).

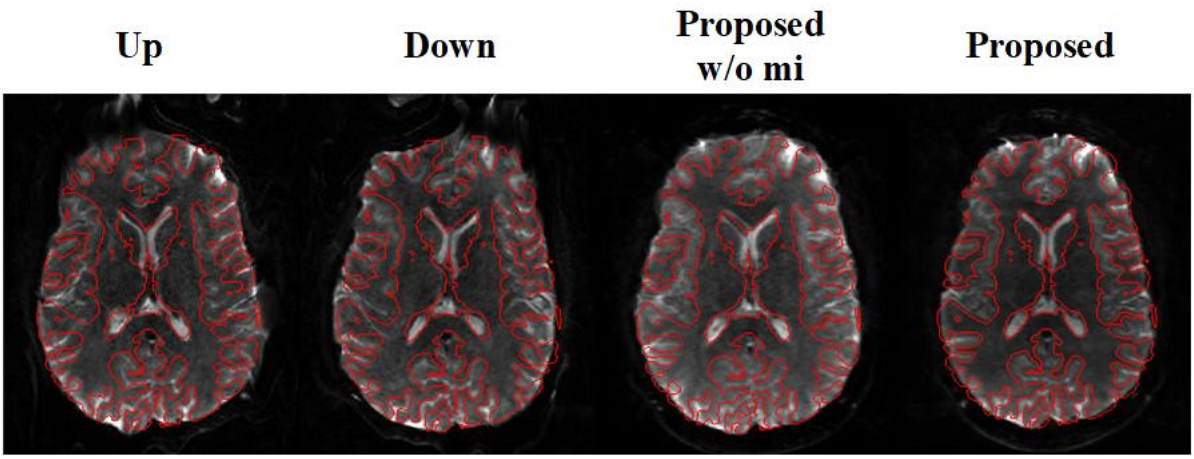

Figure S6. Comparison of performance of our proposed method with and without Multimodal similarity loss.

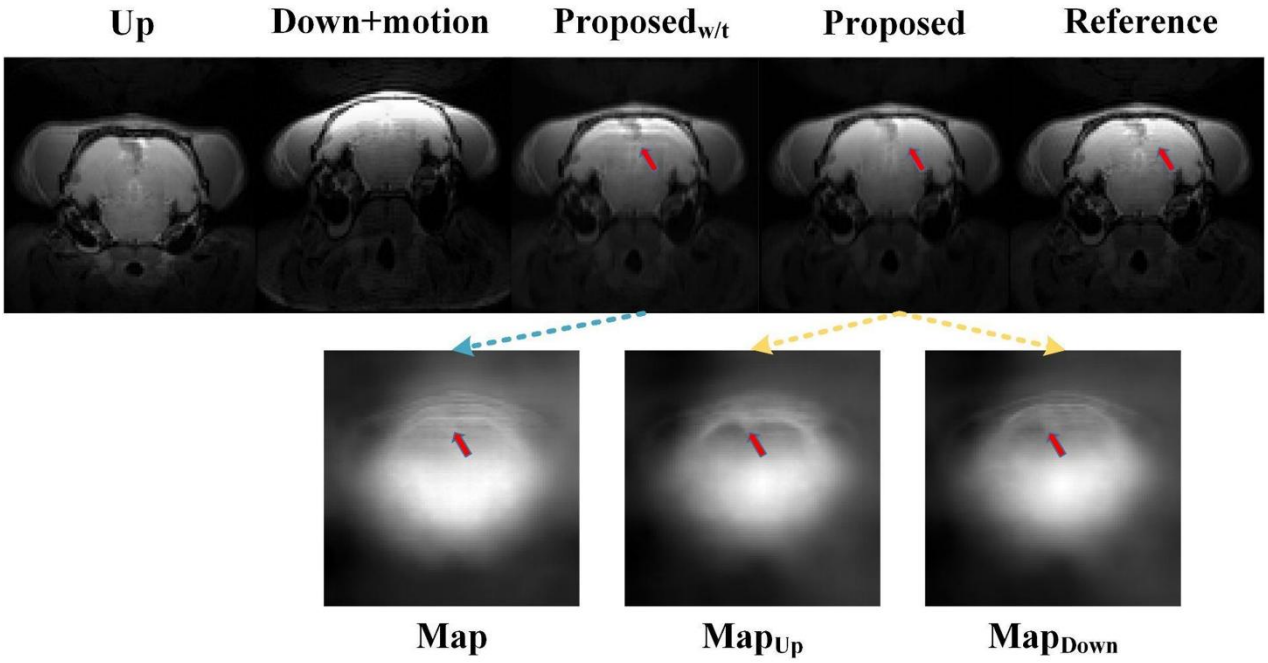

Figure S7. Comparison of performance of our proposed method with one displacement map (Proposed$_{w/t}$) and two displacement maps (Proposed).