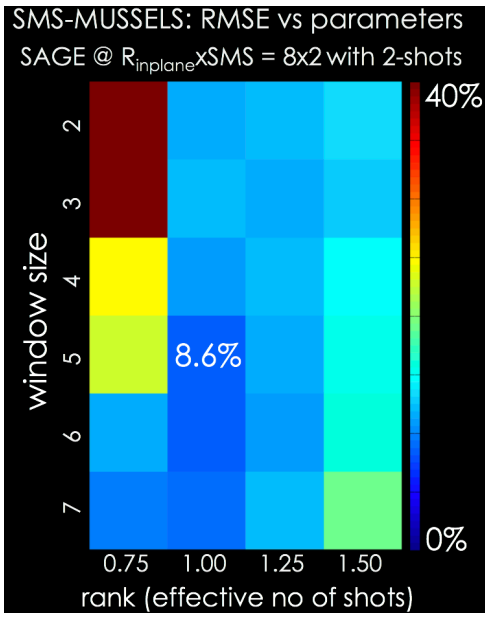
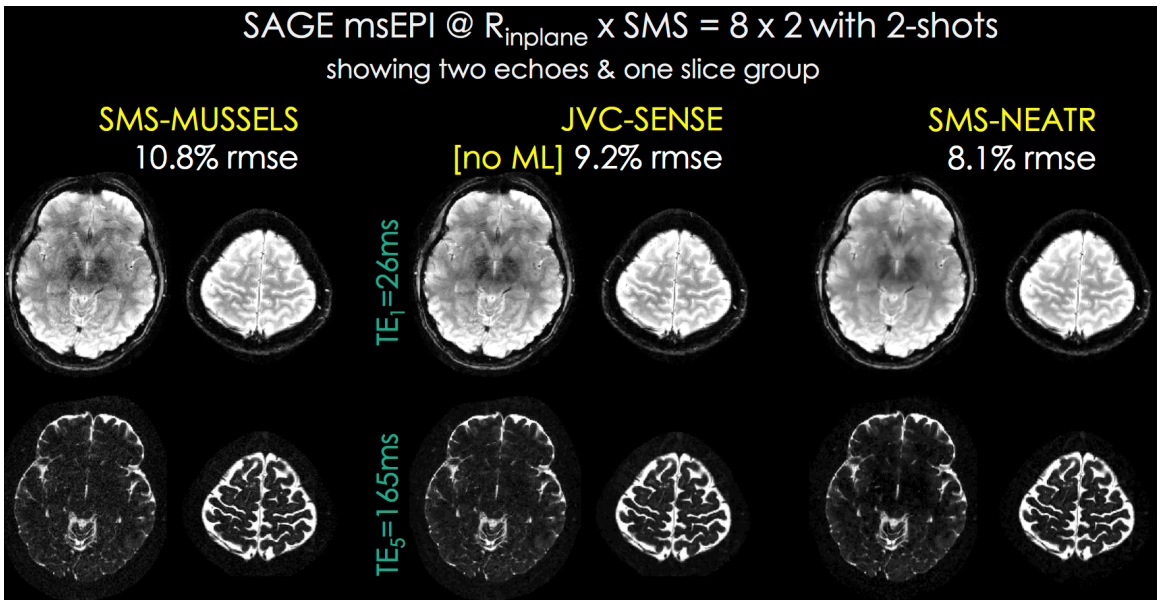
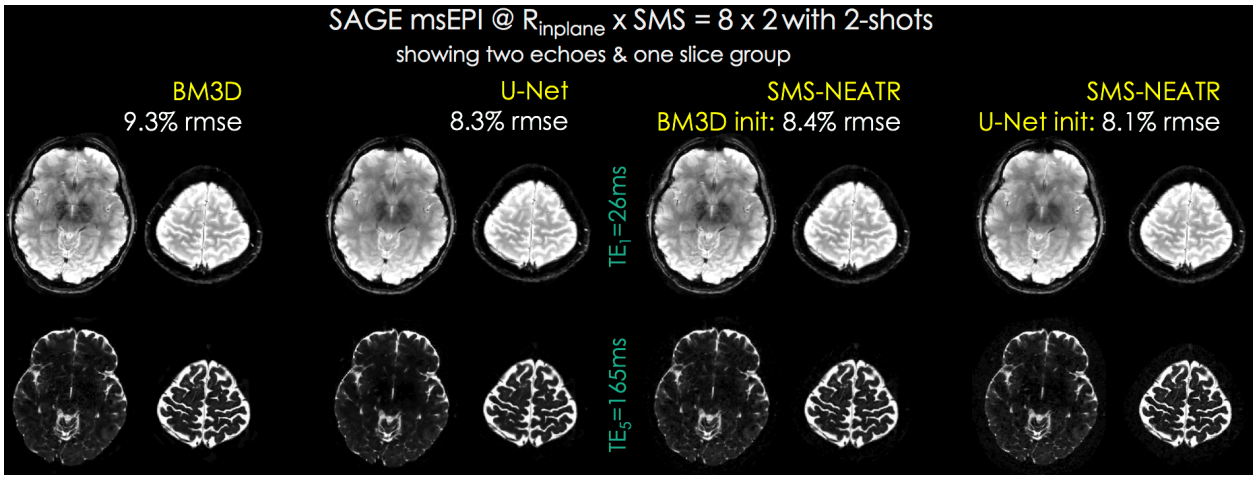
SUPPORTING INFORMATION FIGURES



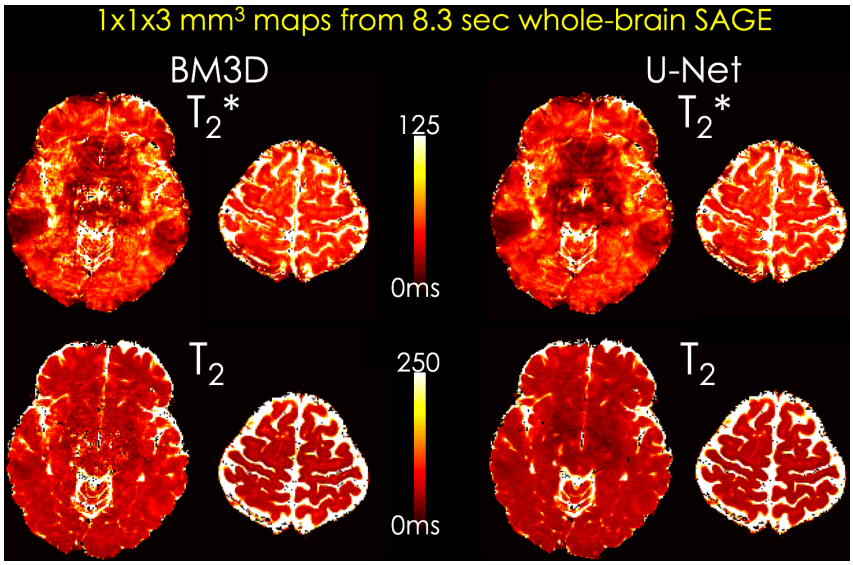
Supporting Information Fig S1. Dependence of the reconstruction performance of SMS-MUSSELS on the k-space window size r, and the rank constraint as represented by the effective number of shots (N_{eff}) for a slice group from a SAGE training dataset. The optimal parameter setting was $r = 5$ and $N_{eff} = 1$ for the 2-shot SAGE reconstruction at $R_{inplane}$xMB=8x2 acceleration.



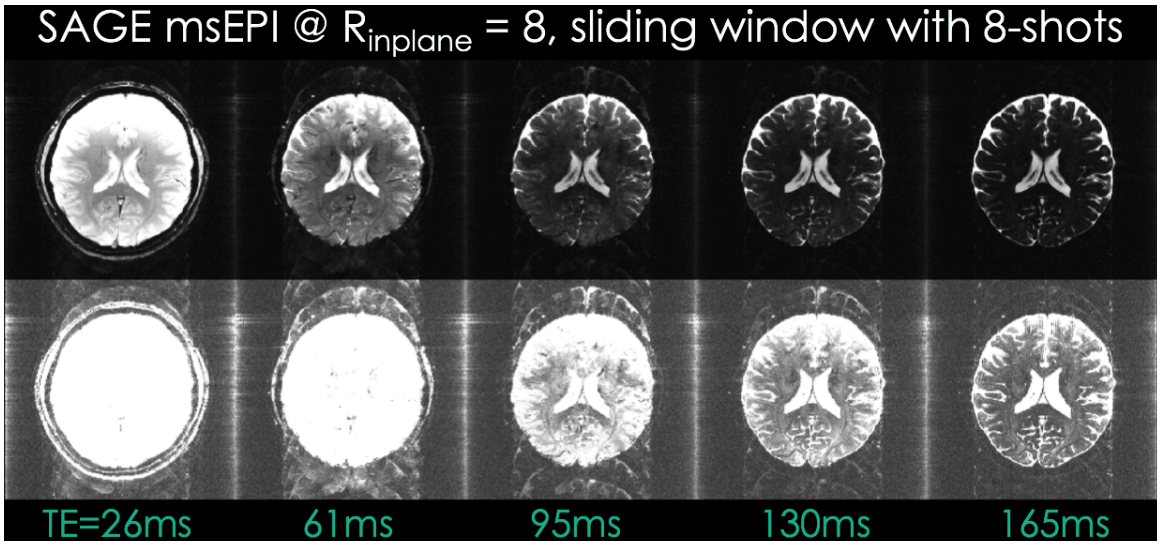
Supporting Information Fig S2. SAGE msEPI reconstruction results from the test dataset at $R_{inplane}$xMB=8x2 acceleration using 2-shots. The first and last echoes are displayed out of a total of five echoes belonging to this SMS slice group. SMS-MUSSELS yielded 10.8% RMSE (left), and was also used to initialize phase-cycling and JVC-SENSE reconstruction without machine learning (9.2% error, middle). Using U-Net to refine the SMS-MUSSELS result and jumpstart SMS-NEATR provided further improvement at 8.1% RMSE.



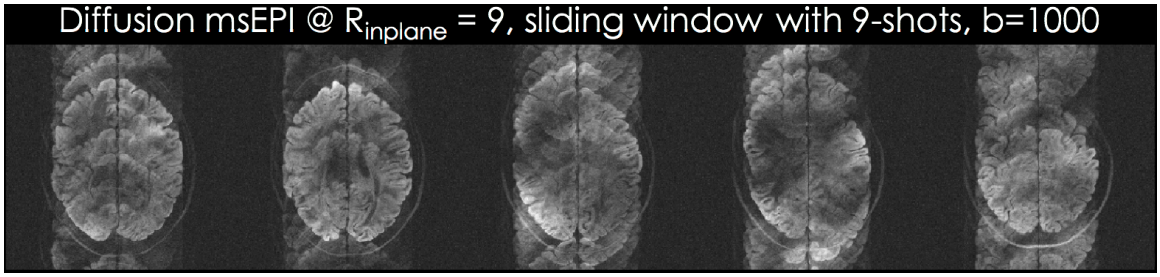
Supporting Information Fig S3. The SMS-MUSSELS reconstruction for the slice group in Fig S2 was denoised using BM3D and U-Net, where deep learning proved to be advantageous (BM3D: 9.3% versus U-Net: 8.3% error). Utilizing each of these denoised outputs to jumpstart SMS-NEATR led to similar reconstructions, and U-Net initialization had the best overall performance (8.1% RMSE).



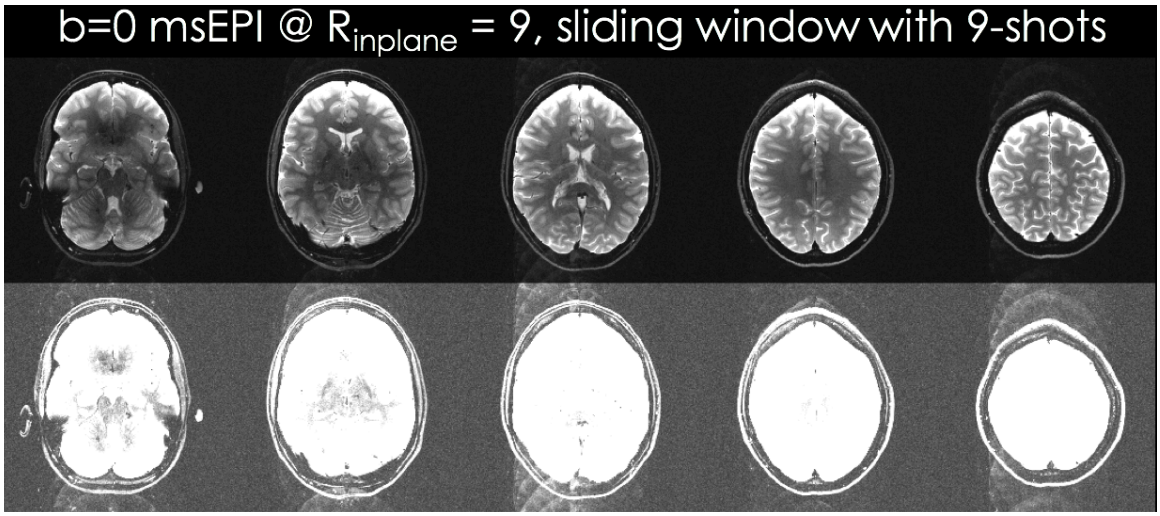
Supporting Information Fig S4. Bloch equation based signal modeling for the five echoes in the SAGE acquisition allows for T_2 and T_2* parameter mapping. Denoising the 2-shot SMS-MUSSELS reconstruction at $R_{inplane}$xMB=8x2 acceleration, corresponding to an 8.3sec whole-brain acquisition, using BM3D and U-Net led to improvements in the quality of these quantitative maps. U-Net appeared more successful in mitigating the noise amplification in the middle of the FOV than the conventional BM3D filtering.



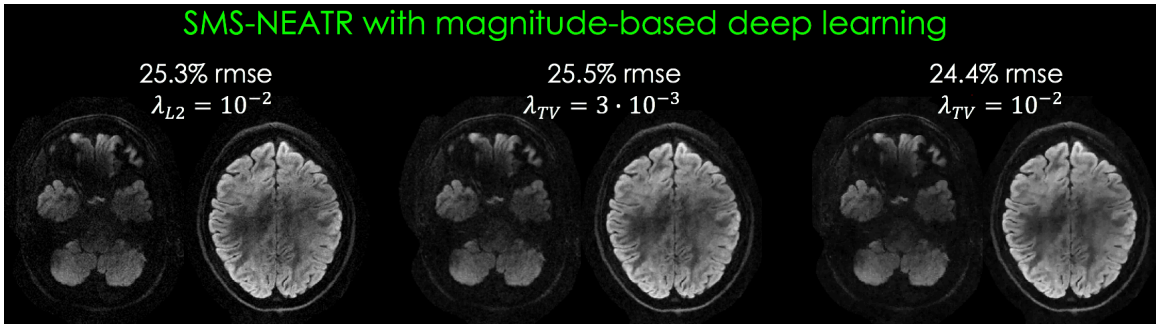
Supporting Information Fig S5. Sliding window (summation across shots in k-space) combination of 8-shots of SAGE data acquired at $R_{inplane}$=8 acceleration per shot. Due to physiological shot-to-shot phase variations, the combined fully-encoded images exhibit ghosting artifacts. This becomes more severe at later TEs, but is mitigated at the last echo, which is a spin echo image. The bottom row is scaled up to better demonstrate the artifacts.



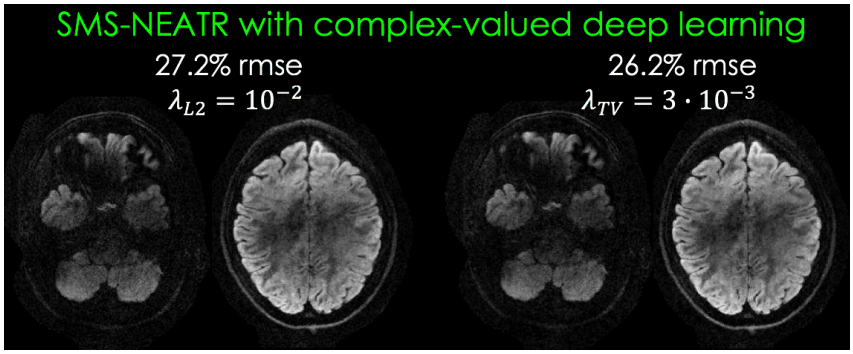
Supporting Information Fig S6. Sliding window combination of 9-shots of DWI data acquired at $R_{inplane}$=9 acceleration per shot. Five slices from a whole-brain acquisition are depicted. Due to shot-to-shot phase variations stemming from motion under the diffusion encoding gradients, there are severe ghosting artifacts in these otherwise fully-encoded images.



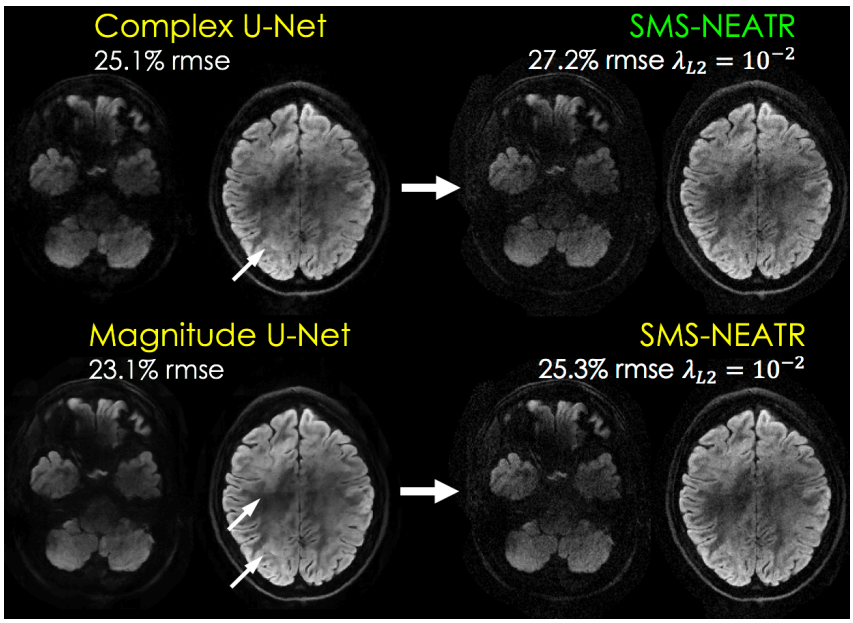
Supporting Information Fig S7. Sliding window combination of data from the same acquisition session as Fig S6, but this time the diffusion gradients have been switched off (b=0). These spin echo images look relatively devoid of ghosting artifacts, but the scaled-up images in the bottom row reveal that these errors persist. We anticipate that head motion contributed to these artifacts, since it took around 46 sec to sample 9-shots at TR=5.1 sec.



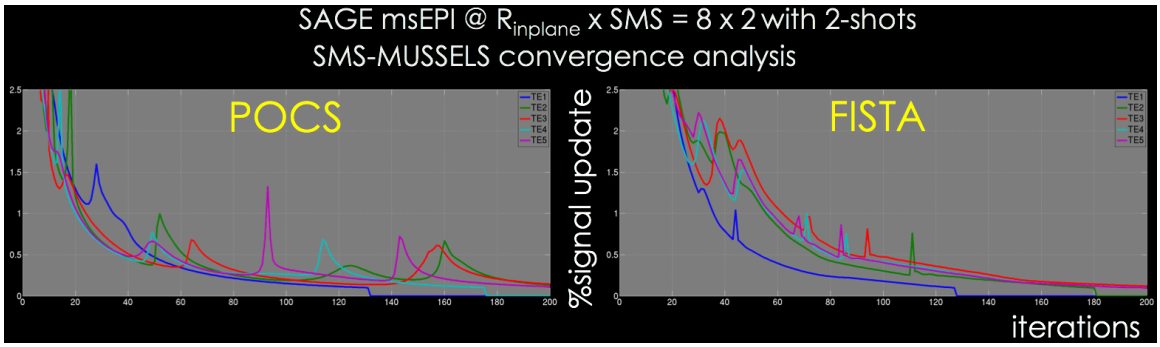
Supporting Information Fig S8. Employing TV-regularization in JVC-SENSE led to similar results as L2 penalty. Reconstruction obtained with the TV parameter value that yielded the optimal RMSE value ($\lambda_{TV} = 10^{-2}$) appeared over-smooth, hence reducing the regularization to $\lambda_{TV} = 3 \cdot 10^{-3}$ provided a better trade-off between RMSE performance and image sharpness. Magnitude-based U-Net was used to initialize SMS-NEATR in these reconstructions.



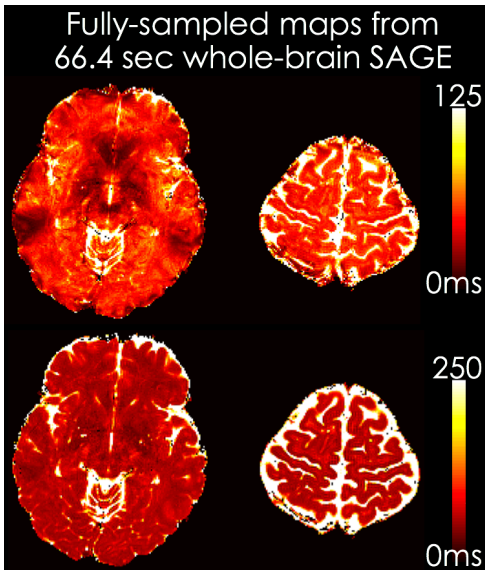
Supporting Information Fig S9. Using complex-valued deep learning to initialize JVC-SENSE yielded similar quality reconstructions as magnitude-based U-Net processing. JVC-SENSE could flexibly utilize L2 or TV regularizers to further stabilize the reconstruction with comparable results.



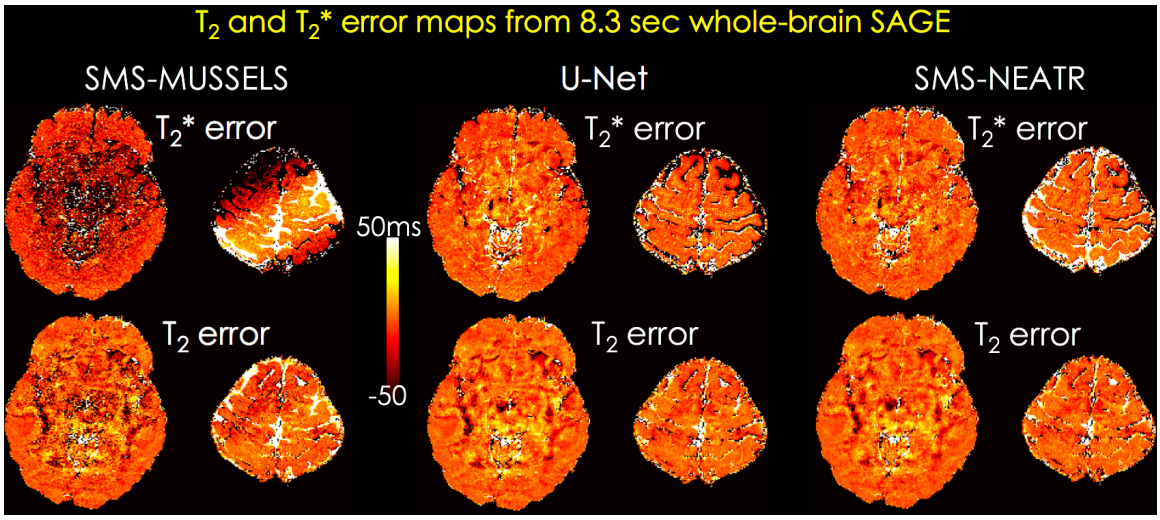
Supporting Information Fig S10. Both complex- and magnitude-valued U-Net processing could denoise diffusion images, albeit at the cost of remaining artifacts (arrows) and especially in the case of magnitude-valued network, over-smoothing. Using these to jumpstart SMS-NEATR led to crisp images with mitigated artifacts.



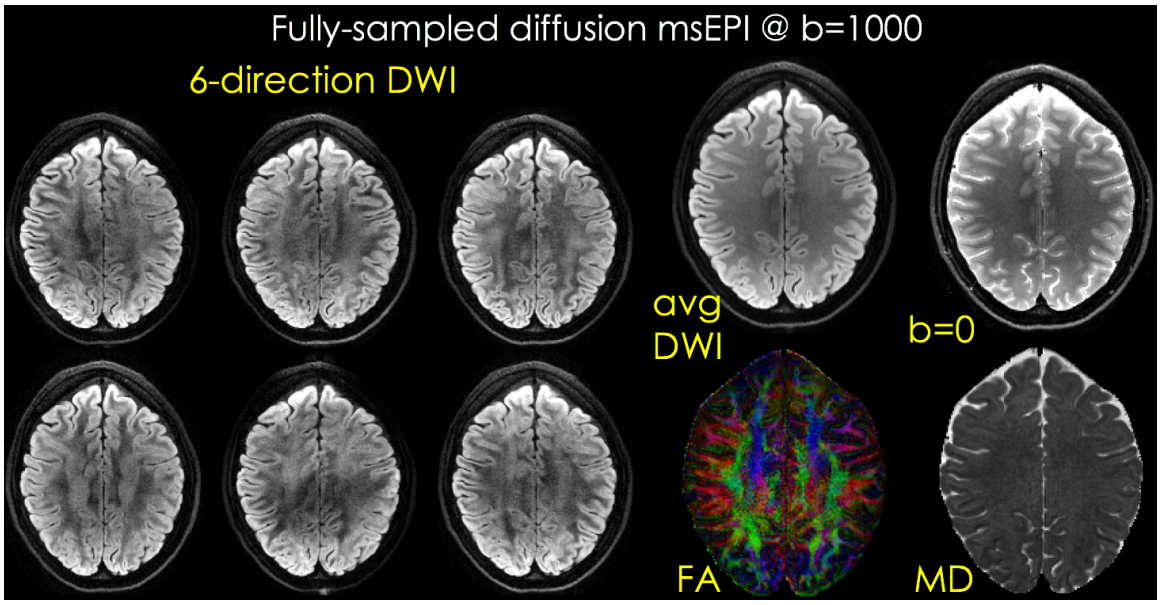
Supporting Information Fig S11. FISTA update rule helped stabilize SMS-MUSSELS reconstruction, especially in later iterations where POCS still experienced large signal updates. Different colors indicate different echo images. Termination criterion was reaching less than 0.1% change between successive image estimates and maximum iteration number was 200.



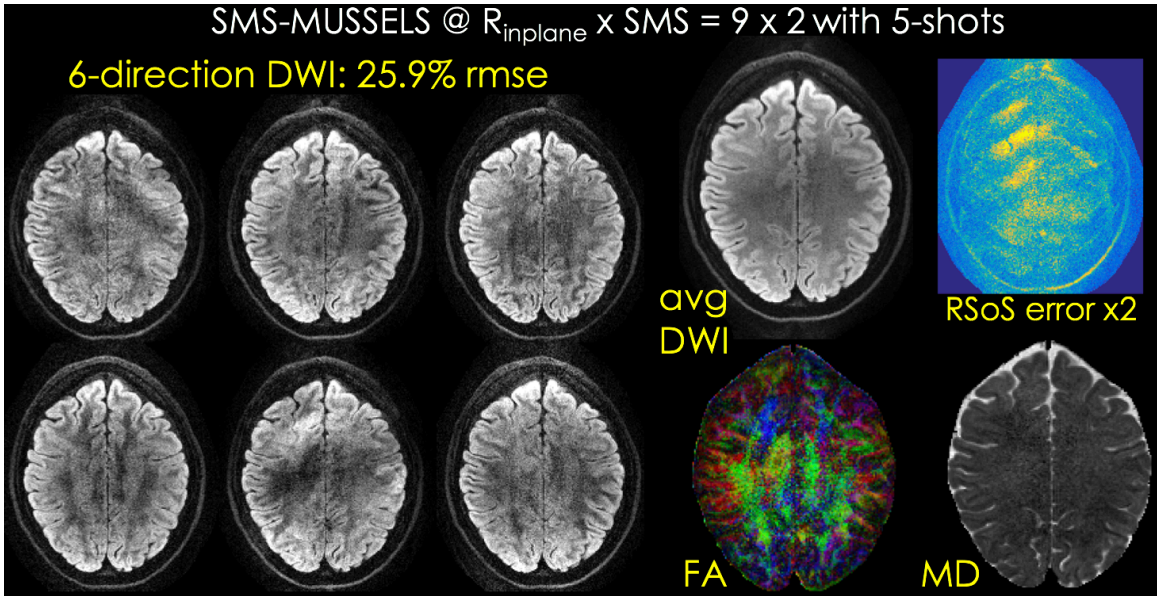
Supporting Information Fig S12. Parameter maps from fully-sampled MUSSELS reconstruction, corresponding to an 8-shot, 66.4 second SAGE scan.



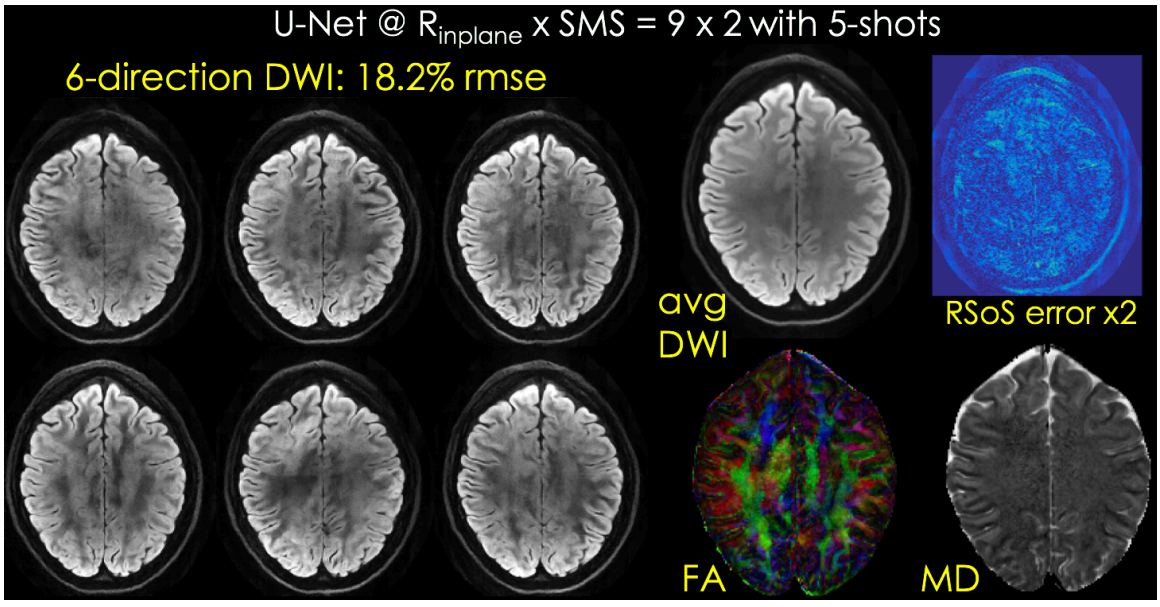
Supporting Information Fig S13. Parameter error maps from $R_{inplane}$ x SMS = 8 x 2-fold accelerated SMS-MUSSELS, U-Net and SMS-NEATR reconstructions relative to the fully-sampled data.



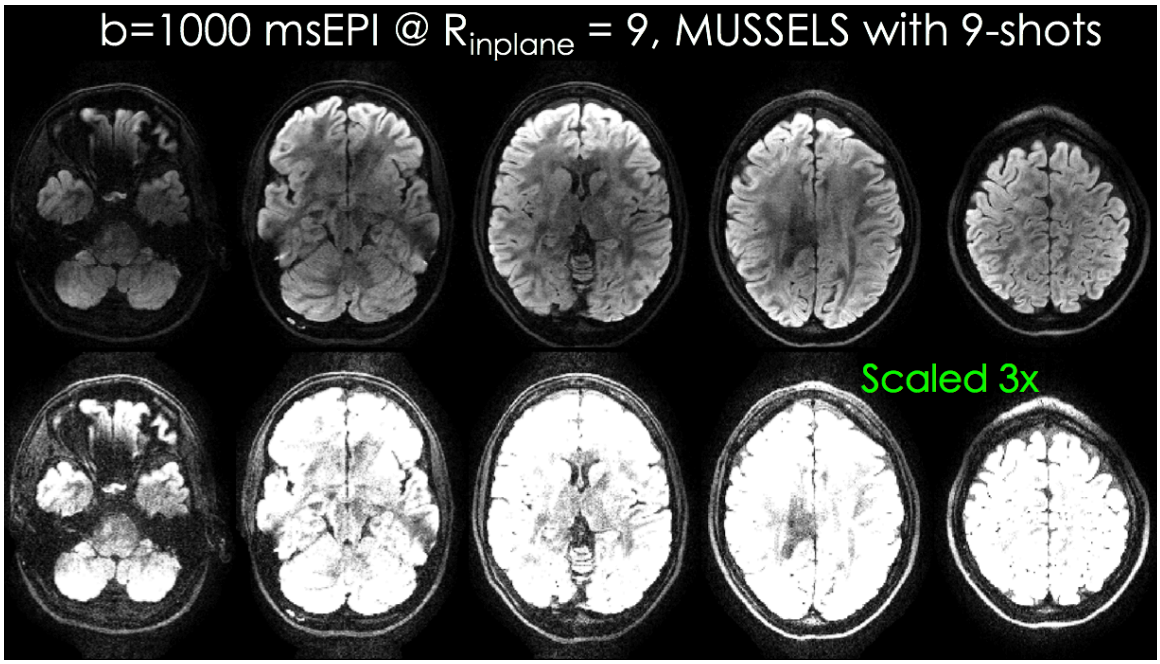
Supporting Information Fig S14. Diffusion images from 6-directions, average diffusion weighted image (DWI), b=0, color fractional anisotropy (FA) and mean diffusivity (MD) maps from the fully-sampled msEPI acquisition with MUSSELS reconstruction.



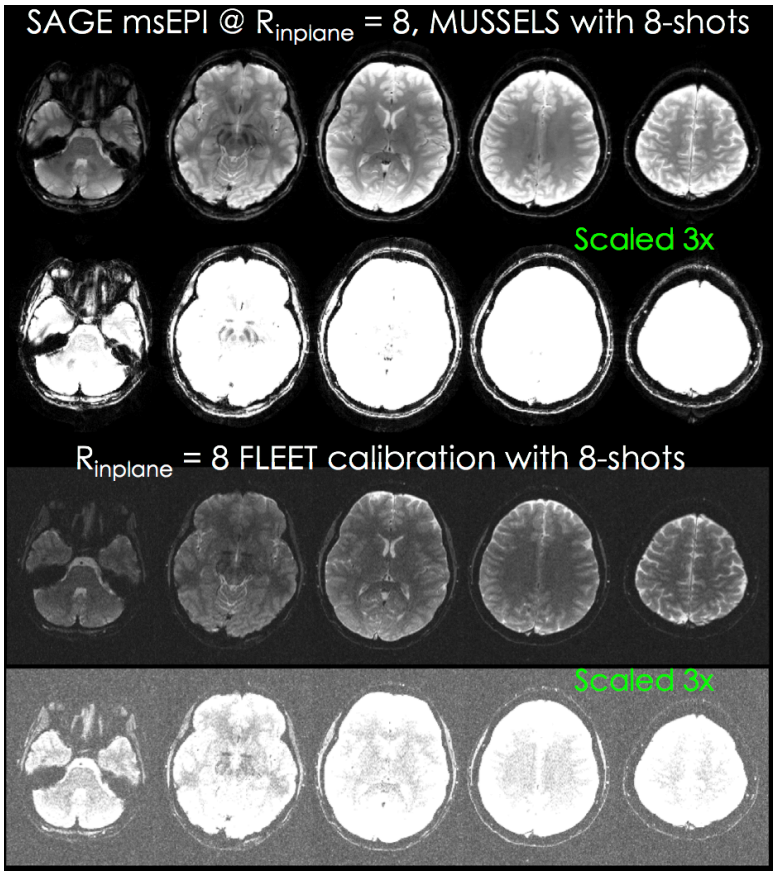
Supporting Information Fig S15. Diffusion images from 6-directions, root-sum-of-squares (RSoS) combination of error across the 6-directions, average DWI, color FA and MD maps from the accelerated SMS-MUSSELS reconstruction.



Supporting Information Fig S16. 6-direction diffusion images, RSoS error across the 6-directions, average DWI, color FA and MD maps from the magnitude-valued U-Net reconstruction.



Supporting Information Fig S17. Fully-sampled diffusion acquisition with MUSSELS reconstruction does not exhibit visible Nyquist ghost artifacts. Ghost-correction was performed using 1-dimensional navigators acquired for each slice and each shot individually.



Supporting Information Fig S18. Fully-sampled SAGE acquisition with MUSSELS reconstruction and multi-shot FLEET calibration data do not exhibit ghost artifacts. Ghost-correction was performed using 1-dimensional navigators acquired for each slice and each shot individually. 3-times scaled-up images are included to help the assessment of ghost level.