Quantitative comparison of MP2RAGE skull-stripping strategies

Pavel Falkovskiy[1,2,3], Bénédicte Maréchal[1,2,3], Shuang Yan[4], Zhengyu Jin[4], Tianyi Qian[5], Kieran O'Brien[6,7], Reto Meuli[2], Jean-Philippe Thiran[2,3], Gunnar Krueger[2,3,8], Tobias Kober[1,2,3], and Alexis Roche[1,2,3]

[1]Advanced Clinical Imaging Technology (HC CMEA SUI DI BM PI), Siemens Healthcare AG, Lausanne, Switzerland, [2]Department of Radiology, University Hospital (CHUV), Lausanne, Switzerland, [3]LTS5, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, [4]Department of Radiology, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China, People's Republic of, [5]MR Collaborations NE Asia, Siemens Helathcare, Beijing, China, People's Republic of, [6]Centre for Advanced Imaging, University of Queensland, Brisbane, Australia, [7]Siemens Healthcare Pty Ltd., Brisbane, Australia, [8]Siemens Medical Solutions USA, Boston, MA, United States

Synopsis (100 words)

The MP2RAGE pulse sequence exhibits higher grey-matter/white-matter contrast compared to standard MPRAGE acquisitions and provides images with greatly reduced B1 bias. In theory, these qualities of MP2RAGE should lead to more accurate morphometric results. However, a major hindrance to MP2RAGE morphometric processing is the salt-and-pepper noise in the background and cavities. This poses a major problem for the skull-stripping stage of most automated morphometry algorithms. We investigated three skull-stripping strategies using the MorphoBox prototype and FreeSurfer automated-morphometry software packages and compared them to results obtained using the gold-standard MPRAGE contrast.

Introduction

Compared to standard MPRAGE, MP2RAGE acquisitions [1] provide improved grey/white-matter contrast [2] and exhibit greatly reduced B1 bias [1]. These properties should improve brain volumetric measurements based on the MP2RAGE contrast; however, a major difficulty in processing MP2RAGE data is the salt-and-pepper noise in the image background and cavities stemming from the intrinsic division of the two inversion contrasts acquired with the MP2RAGE. This poses a major problem for the skull-stripping stage of most automated-morphometry software packages. We evaluated three different skull-stripping strategies using the MorphoBox prototype [3] and FreeSurfer [4].

Materials and Methods

Experiments were performed on a clinical 3T MRI scanner (MAGNETOM Skyra, Siemens Healthcare, Germany) equipped with 20-channel head/neck and 32-channel head coil arrays. 12 healthy subjects were imaged after providing informed consent using the following protocols:

(a) 5:12-minute 3T MPRAGE protocol that resembles the protocol used in ADNI-2 [5]: TR/TI/BW/α= 2300ms/900ms/240Hz/px/9°;

(b) 8:22-minute 3T M2PRAGE protocol (TR/TI1/TI2/BW/α1/α2 = 5000ms/700ms/2500ms/240Hz/px/4°/5°, 1.00x1.00x1.00mm³) using a three-fold GRAPPA acceleration (iPAT=3)

For each subject, the measurement session consisted of a scan and rescan using MPRAGE and MP2RAGE protocols with both 32-channel and 20-channel coils. To note, the MP2RAGE outputs the 1st and 2nd inversion contrasts as well as the MP2RAGE "uniform" contrast as described in [1].

Subsequently, images were skull-stripped using both the MorphoBox prototype [3] and FreeSurfer [4] (version 5.3.0) by inputting auxiliary images with low background noise. We investigated 3 different types of auxiliary images:

- Method 1: Second inversion contrast
- Method 2: Product of the second inversion contrast and the uniform MP2RAGE image as in [6].
- Method 3: Denoised uniform MP2RAGE image as described in [7]. This strategy requires tuning a regularization parameter (denoising constant), which we hypothesise to be proportional to the noise variance estimated using the pseudo-replica method [8]. We tested a range of multiples of the variance as denoising constant (N*Var, N=1…30).

To quantify skull-stripping performance, we evaluated the across-subject root-mean-square errors (RMSE) between MP2RAGE-based and MPRAGE-based estimates of the total intracranial volume (TIV). We also evaluated the reproducibility of TIV estimation for each contrast by computing the RMSEs between scan and rescan.

Results and Discussion

Fig. 1 shows sample skull-stripped images using the above methods. Note that FreeSurfer brain masks exclude extra-ventricular CSF, and the TIV is estimated based on the Talairach transform without being explicitly segmented [9]. Visually, it is hard to distinguish between methods 2 and 3, whereas method 1 exhibits obvious errors in both MorphoBox and FreeSurfer results.

Fig. 2 illustrates the RMSE in TIV estimates as a function of the denoising constant using MorphoBox. Method 1 introduced large discrepancies in TIV compared to the MPRAGE acquisitions (4.10%/4.13% with 20-channel and 32-channel coils, respectively). Method 2 proved to be more similar to MPRAGE-based TIV extractions, yielding RMSEs of 0.94%/0.98%. With an optimal choice of the denoising constant, method 3 achieved 0.59%/1.38% difference to MPRAGE acquisitions. Scan-rescan RMSEs were comparable across the three methods: 0.172%/0.309% with method 1, 0.288%/0.127% with method 2, 0.164%/0.181% with method 3, which were not substantially different from the 0.264%/0.218% achieved using MPRAGE.

Using FreeSurfer, method 1 introduced considerable discrepancies in TIV estimates compared to MPRAGE as shown by the respective RMSEs of 9.35%/9.12% with 20-channel and 32-channel coils, respectively (Fig. 3). Method 2 yielded slightly smaller RMSEs of 7.00%/7.64% (Fig. 3), while method 3 achieved 1.89%/3.46% with optimally chosen denoising constants. Scan-rescan RMSEs for methods 1, 2, 3 and MPRAGE were respectively: 0.069%/0.134%, 0.167%/3.21%, 0.870%/1.32%, and 0.115%/0.127%. There is a noticeable decrease in scan-rescan stability of methods 2 and 3 compared to MPRAGE. In order to get a better insight into the large discrepancies between MPRAGE and MP2RAGE TIV estimates with FreeSurfer, we examined the brain masks as represented in Fig. 1. Fig. 4 illustrates the RMSE in brain mask volumes compared to MPRAGE. Both methods 2 and 3 produced considerably smaller discrepancies in brain mask volumes than in TIV and showed scan-rescan stability similar to MPRAGE.

Conclusion

Using MorphoBox, both methods 2 and 3 achieved similar TIV estimation performance compared to MPRAGE and clearly surpassed method 1. With FreeSurfer, method 3 performed considerably better than methods 2 and 1, however yielding rather poor correspondence with MPRAGE in addition to larger scan-rescan variability. When using MorphoBox, our results suggest in practice to estimate TIV from MP2RAGE data using method 2 since it can be applied retrospectively contrary to method 3. However, caution must be exercised when using FreeSurfer-based TIV measures computed from MP2RAGE data using the methods considered here (e.g., for volume normalization purposes). Future work will aim to investigate MP2RAGE-based volumetry for other brain structures.

References

[1] J. P. Marques, T. Kober, G. Krueger, W. van der Zwaag, P.-F. Van de Moortele, and R. Gruetter, "MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field.," *Neuroimage*, vol. 49, no. 2, pp. 1271–81, Jan. 2010.

[2] G. Okubo, T. Okada, A. Yamamoto, M. Kanagaki, Y. Fushimi, T. Okada, K. Murata, and K. Togashi, "MP2RAGE for deep gray matter measurement of the brain: A comparative study with MPRAGE," *J. Magn. Reson. Imaging*, p. n/a–n/a, 2015.

[3] D. Schmitter, A. Roche, B. Maréchal, D. Ribes, A. Abdulkadir, M. Bach-Cuadra, A. Daducci, C. Granziera, S. Klöppel, P. Maeder, R. Meuli, and G. Krueger, "An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease," *NeuroImage. Clin.*, vol. 7, pp. 7–17, Jan. 2015.

[4] B. Fischl, “FreeSurfer,” *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[5] C. R. Jack, M. A. Bernstein, B. J. Borowski, J. L. Gunter, N. C. Fox, P. M. Thompson, N. Schuff, G. Krueger, R. J. Killiany, C. S. Decarli, A. M. Dale, O. W. Carmichael, D. Tosun, and M. W. Weiner, “Update on the magnetic resonance imaging core of the Alzheimer’s disease neuroimaging initiative,” *Alzheimers. Dement.*, vol. 6, no. 3, pp. 212–20, May 2010.

[6] K. Fujimoto, J. R. Polimeni, A. J. W. van der Kouwe, M. Reuter, T. Kober, T. Benner, B. Fischl, and L. L. Wald, “Quantitative comparison of cortical surface reconstructions from MP2RAGE and multi-echo MPRAGE data at 3 and 7T,” *Neuroimage*, vol. 90, pp. 60–73, 2014.

[7] K. R. O’Brien, T. Kober, P. Hagmann, P. Maeder, J. Marques, F. Lazeyras, G. Krueger, and A. Roche, “Robust T1-weighted structural brain imaging and morphometry at 7T using MP2RAGE.,” *PLoS One*, vol. 9, no. 6, p. e99676, 2014.

[8] P. M. Robson, A. K. Grant, A. J. Madhuranthakam, R. Lattanzi, D. K. Sodickson, and C. a McKenzie, “Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions,” *Magn. Reson. Med.*, vol. 60, no. 4, pp. 895–907, Oct. 2008.

[9] R. L. Buckner, D. Head, J. Parker, A. F. Fotenos, D. Marcus, J. C. Morris, and A. Z. Snyder, “A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume.,” *Neuroimage*, vol. 23, no. 2, pp. 724–38, 2004.

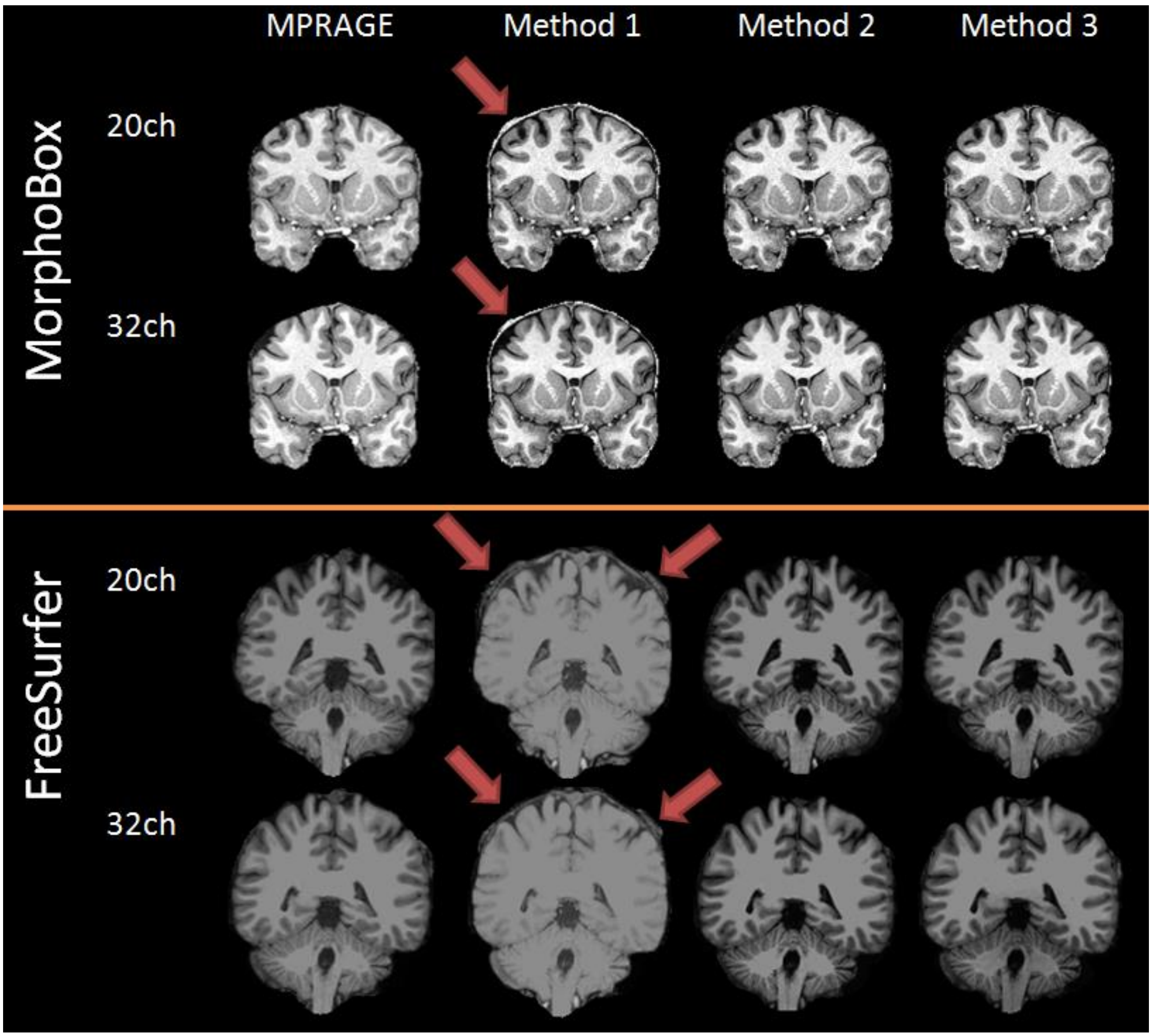

Figure 1: Exemplary skull-stripped brain volumes obtained with both MorphoBox and FreeSurfer using the different pre-processing methods. The red arrows indicate skull-stripping errors. Method 1: skull-stripping on second inversion contrast; Method 2: skull-stripping on product image; Method 3: skull-stripping on de-noised uniform MP2RAGE; 32ch: 32-channel head coil; 20ch: 20-channel head/neck coil.

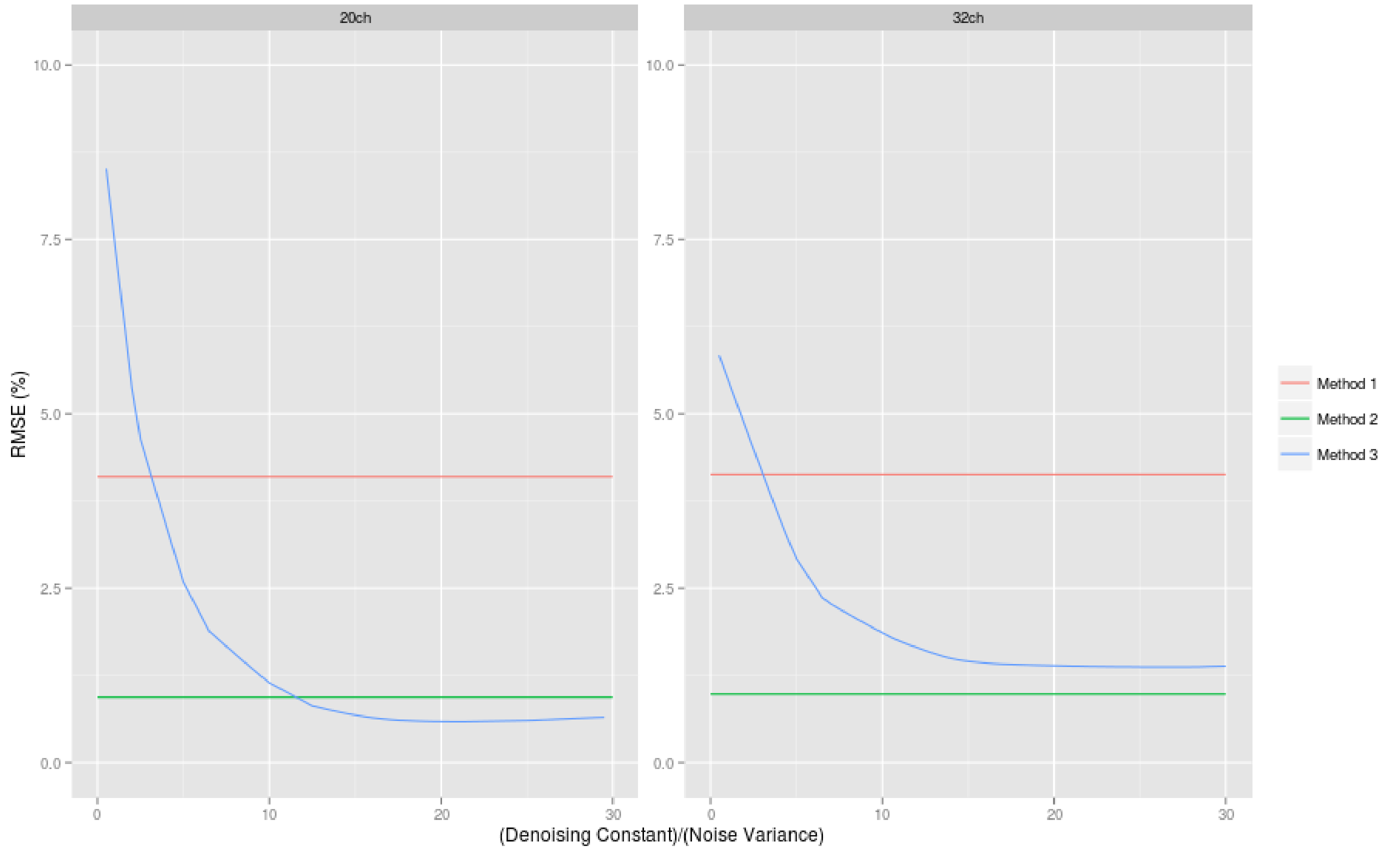

Figure 2: RMSE of TIV volumes processed with MorphoBox as a function of de-noising constant. Method 1: skull-stripping on second inversion contrast; Method 2: skull-stripping on product image; Method 3: skull-stripping on de-noised uniform MP2RAGE; 32ch: 32-channel head coil; 20ch: 20-channel head/neck coil.

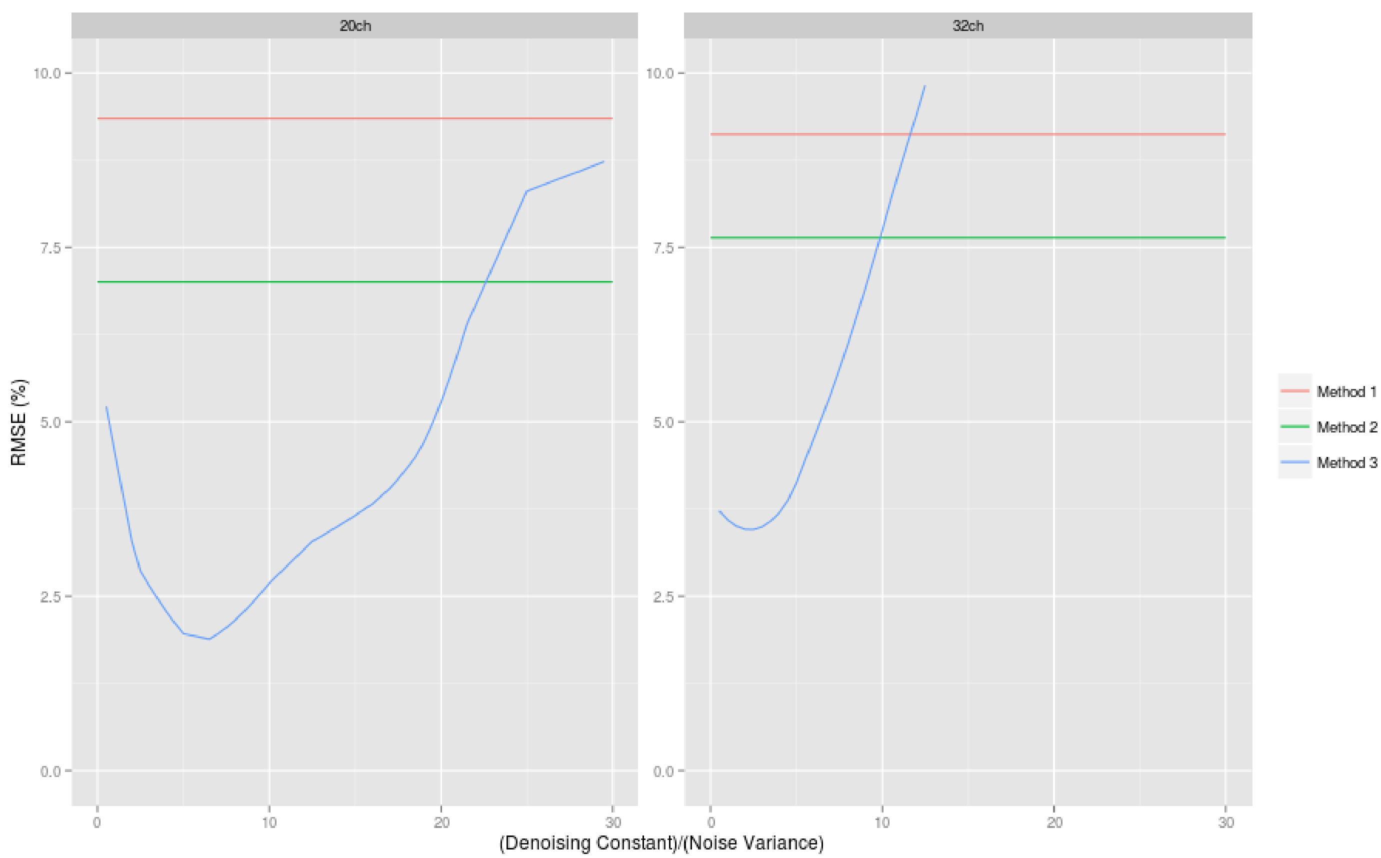

Figure 3: RMSE in TIV volumes processed with FreeSurfer as a function of de-noising constant. Method 1: skull-stripping on second inversion contrast; Method 2: skull-stripping on product image; Method 3: skull-stripping on de-noised uniform MP2RAGE; 32ch: 32-channel head coil; 20ch: 20-channel head/neck coil.

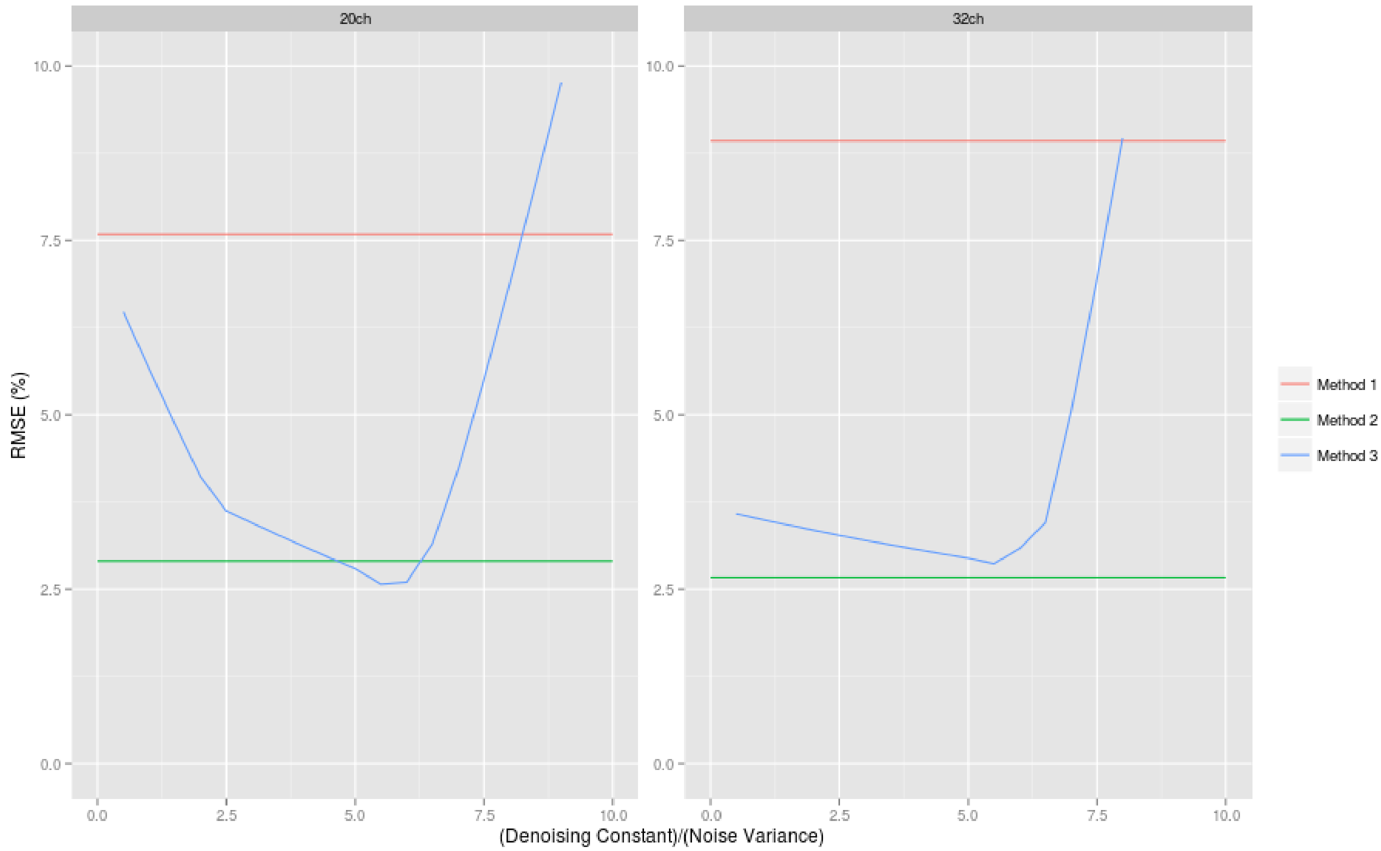

Figure 4: MP2RAGE vs. MP-RAGE RMSE in brain mask volumes processed with FreeSurfer as a function of de-noising constant. Method 1: skull-stripping on second inversion contrast; Method 2: skull-stripping on Inv2*uniform product image; Method 3: skull-stripping on de-noised uniform MP2RAGE. Results shown for both 20-channel head/neck and 32-channel head coils.