# Figure Captions

**Figure 1:**

**Figure 2:** Canonical views of the mutlivariate, bilaterally symmetric template constructed from the MMRR data set1 (only shown are the FLAIR, T1, and T2 modalities--- the components relevant for this work). Template construction is detailed in2. These images are important for asymmetry-based features.

**Figure 3:** Representation of Stage 1 feature images for subject 01C1019. The FLAIR, T1-, and T2-weighted images are rigidly pre-aligned3 to the space of the T1 image. The three modality images are then preprocessed (N4 bias correction4 and adaptive denoising5) followed by application of standard ANTs brain extraction and -tissue segmentation protocols using the MMRR symmetric template and corresponding priors6 applied to the T1 image. The feature images are then generated for voxelwise input to the RF model which results in the voting maps illustrated on the right. This gives a probabilistic classification of tissue type. Not shown are the probability and voting images for the brain stem and cerebellum.

**Figure 4:** Sample FLAIR acquisition image slices showing both manual and random forest segmentations for both stages obtained during the leave-one-out evaluation. Manual segmentations were performed by one of the authors and provided the ground truth WMH labels for training the random forest models.

**Figure 5:**

**Figure 6:** Average *MeanDecreaseAccuracy* plots generated from the creation of all 24 random forest models for Stage 1 during the leave-one-out evaluation. These plots are useful in providing a quantitative assessment of the predictive importance of each feature. Features are ranked in descending order of importance. The horizontal error bars provide the percentile and illustrate the stability of the feature importance across the leave-one-out models. At this initial stage only 31 feature images are used.

**Figure 7:** Average *MeanDecreaseAccuracy* plots generated from the creation of all 24 random forest models for Stage 2 during the leave-one-out evaluation. These plots are useful in providing a quantitative assessment of the predictive importance of each feature. Features are ranked in descending order of importance. The horizontal error bars provide the percentile and illustrate the stability of the feature importance across the leave-one-out models. We augment the 31 feature images from the first stage by adding an additional seven voting maps and 7 segmentation posteriors from application of the Bayesian-based segmentation for a total of 45 images for the second stage.

**Figure 8:** (a) FLAIR image slice illustrating WMHs which have been manually delineated. The region around the WMHs is enlarged (b) in the original FLAIR and the (c) contralateral FLAIR difference image.

1. Landman BA, Huang AJ, Gifford A, Vikram DS, Lim IAL, Farrell JAD, Bogovic JA, Hua J, Chen M, Jarso S, et al. Multi-parametric neuroimaging reproducibility: A 3-T resource study. Neuroimage. 2011;54(4):2854–66.

2. Tustison NJ, Shrinidhi KL, Wintermark M, Durst CR, Kandel BM, Gee JC, Grossman MC, Avants BB. Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with aNTsR. Neuroinformatics. 2015;13(2):209–25.

3. Avants BB, Tustison NJ, Stauffer M, Song G, Wu B, Gee JC. The Insight ToolKit image registration framework. Front Neuroinform. 2014;8:44.

4. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC. N4ITK: Improved N3 bias correction. IEEE Trans Med Imaging. 2010;29(6):1310–20.

5. Manjón JV, Coupé P, Martí-Bonmatí L, Collins DL, Robles M. Adaptive non-local means denoising of mR images with spatially varying noise levels. J Magn Reson Imaging. 2010;31(1):192–203.

6. Tustison NJ, Cook PA, Klein A, Song G, Das SR, Duda JT, Kandel BM, Strien N van, Stone JR, Gee JC, et al. Large-scale evaluation of aNTs and freeSurfer cortical thickness measurements. Neuroimage. 2014;99:166–79.