



OASIS is Automated Statistical Inference for Segmentation with applications to multiple sclerosis lesion segmentation in MRI

Elizabeth M. Sweeney^{1,2}, Russell T. Shinohara^{2,3}, Navid Shiee⁴, Farah J. Mateen⁵, Avni A. Chudgar⁶, Jennifer L. Cuzzocreo⁷, Peter A. Calabresi⁸, Dzung L. Pham⁴, Daniel S. Reich^{1,2,7,8}, Ciprian M. Crainiceanu¹
¹Department of Biostatistics, The Johns Hopkins University, Baltimore, MD 21205; ²Translational Neuroradiology Unit, Neuroimmunology Branch, National Institute of Neurological Disease and Stroke, National Institute of Health, Bethesda, MD 20892; ³Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; ⁴Center for Neuroscience and Regenerative Medicine, Henry M. Jackson Foundation, Bethesda, MD 20892; ⁵Department of International Health, The Johns Hopkins University, Baltimore, MD 21205; ⁶Department of Radiology, Division of Emergency Radiology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115; ⁷Department of Radiology, The Johns Hopkins University School of Medicine, Baltimore, MD 21208; ⁸Department of Neurology, The Johns Hopkins University School of Medicine, Baltimore, MD 21208

Introduction

Magnetic resonance imaging (MRI) can be used to detect lesions in the brains of multiple sclerosis (MS) patients and is essential for evaluating disease-modifying therapies and monitoring disease progression. In practice, lesion load is often quantified by manual segmentation of MRI, which is time - consuming, costly, and associated with large inter- and intra- observer variability. We propose OASIS is Automated Statistical Inference for Segmentation (OASIS), an automated statistical method for segmenting MS lesions in MRI studies. The OASIS model uses intensity-normalized T1-weighted (T1), T2-weighted (T2), fluid-attenuated inversion recovery (FLAIR) and proton density (PD) MRI volumes.

Methods

Image Preprocessing

Image preprocessing was performed in MIPAV (McAuliffe, et al., 2001) with JIST (Lucas et al., 2010):

- 1) N3 inhomogeneity correction (Sled et al., 1998)
- 2) Skull strip (SPECTRE) (Carass et al. 2011)
- 3) Rigid registration to MNI standard space

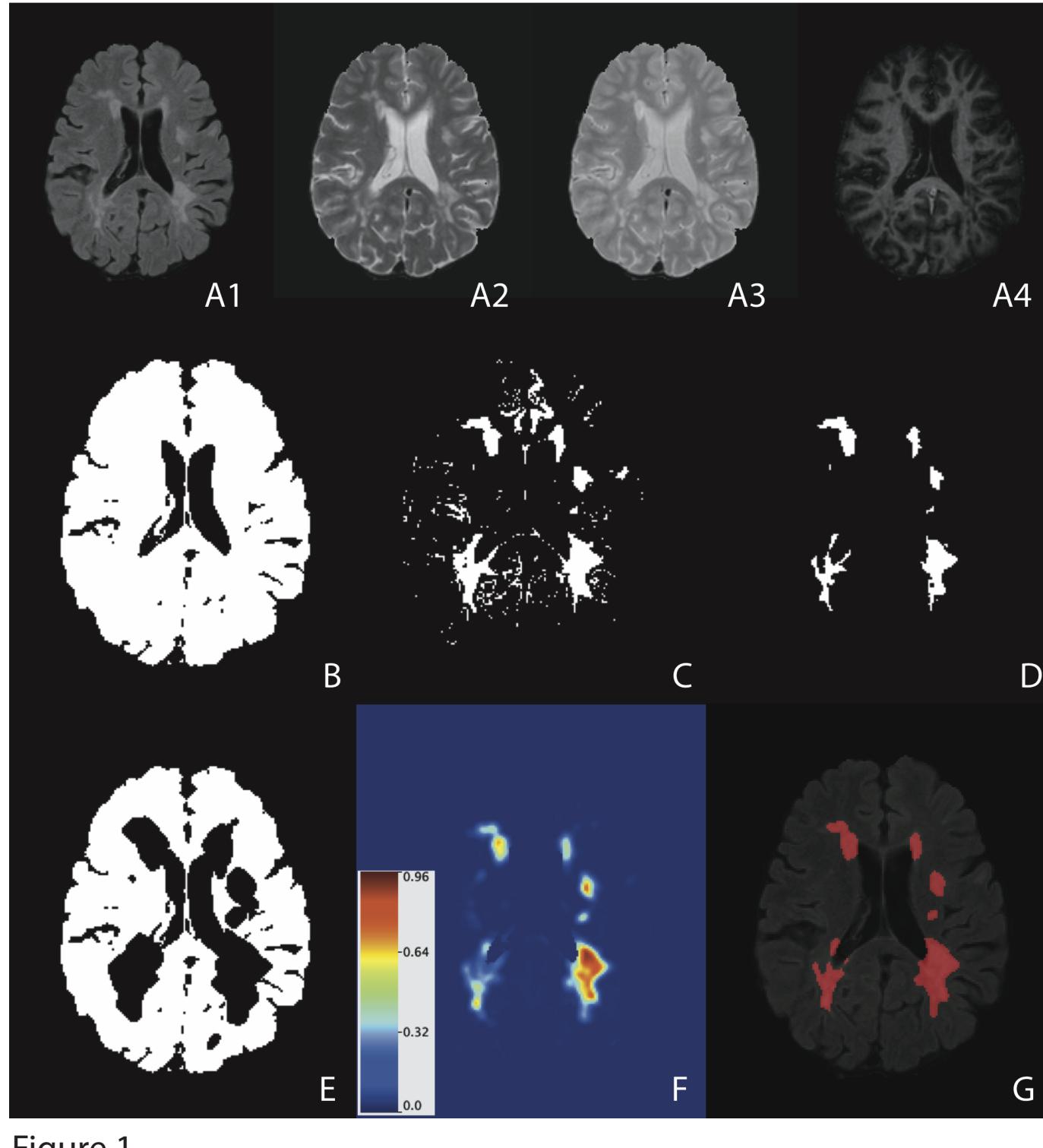


Figure 1

Intensity Normalization

The intensities in a voxel v from the T1, T2, FLAIR and PD volumes for a subject i are expressed with the following notation:

$$M_i^0(v), M = \text{FLAIR}, \text{PD}, \text{T2}, \text{T1}$$

We adapt the normalization method described in Shinohara et al. (2011) to normalization over the brain tissue mask shown in Figure 1 B:

$$M_i^N(v) = \frac{M_i^0(v) - \mu_{i,M}^0}{\sigma_{i,M}^0}$$

where $\mu_{i,M}^0$ and $\sigma_{i,M}^0$ are the mean and standard deviation of the observed voxel intensities in the brain tissue mask of subject i , from volume M . The normalized volumes are shown in Figure 1 A. (A1. FLAIR; A2. T2; A3; PD; A4. T1)

BLUR is a Local Uniform Representation (BLUR) Volumes

We found that even after N3 inhomogeneity correction, there remain intensity variations in the same tissue class within an MRI volume. To account for this intensity inhomogeneity we use a sequence of multiresolution BLUR volumes, obtained using different levels of three dimensional smoothing of the normalized volume from each modality over the brain tissue mask. Gaussian blurs with kernel window sizes of 10 and 20 are used to smooth over the features in the brain and capture the pattern of the remaining inhomogeneity. The notation for the BLUR volume intensity for subject i in voxel v is: $M_i^k(v)$, where M is the imaging modality and k is the kernel window size. The BLUR volumes for the four modalities (FLAIR, T2, PD, and T1) with both kernel window size (10 and 20) are shown in Figure 2.

OASIS is Automated Statistical Inference for Segmentation

OASIS uses logistic regression to model the probability that a voxel is part of a lesion. The model must be trained on gold standard manual lesion

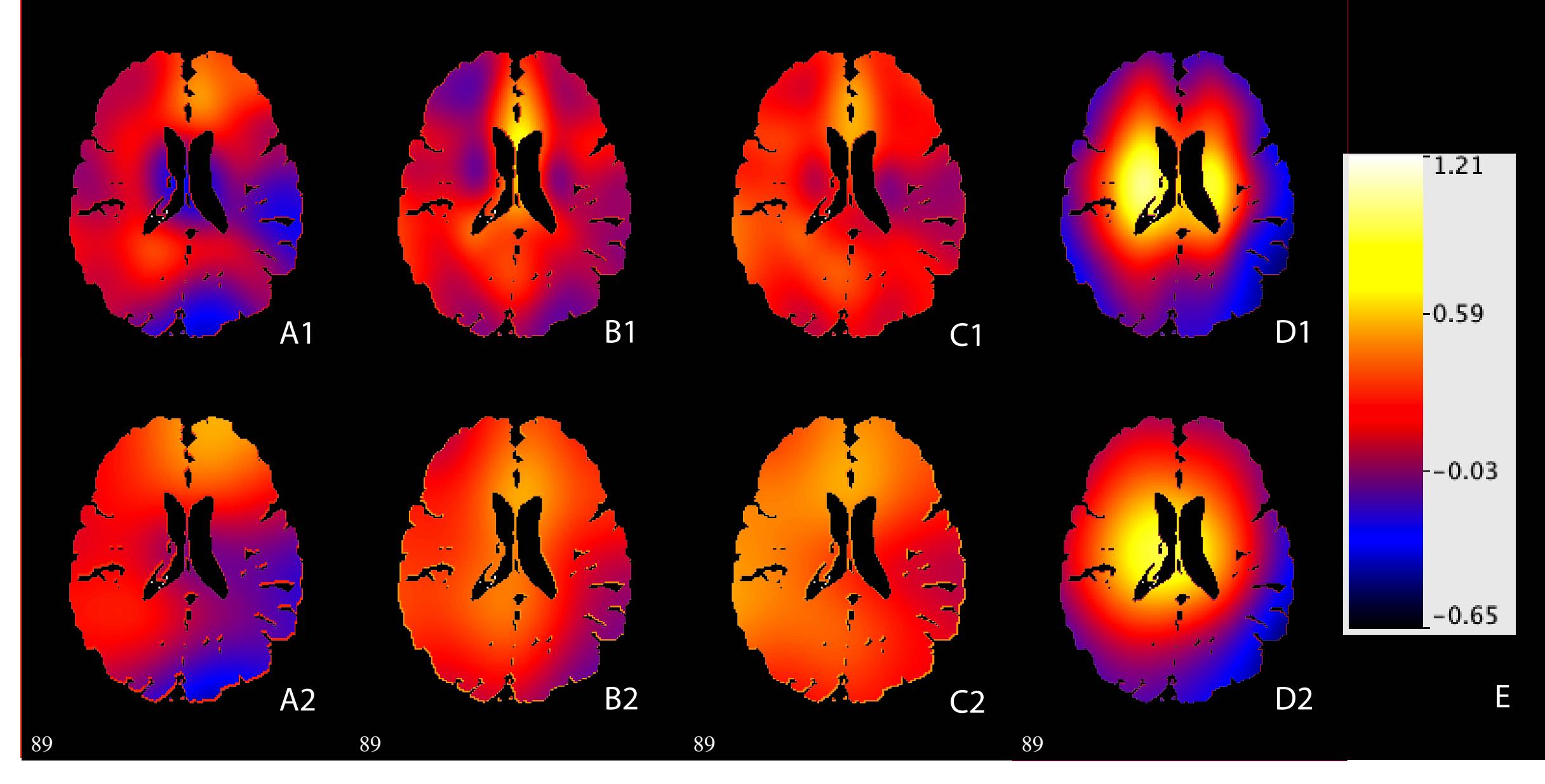


Figure 2

Methods Continued

segmentations (Figure 1D). The model is fit only over candidate voxels, the 85th percentile and above of intensities of voxels in the brain tissue mask of the FLAIR volume (Figure 1C). There are two iterations of the OASIS model fitting. This is done to reduce the influence of lesions in the BLUR volumes. Lesions are removed after the first fit of the model (Figure 1E). Then a second set of BLUR volumes are created and the area with lesions are inpainted with the expected value of healthy tissue. Let $L_i(v)$ be a random variable denoting lesion presence ($L_i(v) = 1$ if there is a lesion in voxel v of subject i , 0 otherwise) The OASIS model is shown below:

$$\begin{aligned} \logit[P\{L_i(v) = 1\}] = & \beta_0 \\ \beta_1 FLAIR_i^N(v) + & \beta_2 GFLAIR_i^N(v, 10) \\ \beta_4 PD_i^N(v) + & \beta_5 GPD_i^N(v, 10) \\ \beta_7 T2_i^N(v) + & \beta_8 GT2_i^N(v, 10) \\ \beta_{10} T1_i^N(v) + & \beta_{11} GT1_i^N(v, 10) \\ & + \beta_3 GFLAIR_i^N(v, 20) \\ & + \beta_6 GPD_i^N(v, 20) \\ & + \beta_9 GT2_i^N(v, 20) \\ & + \beta_{12} GT1_i^N(v, 20) \\ & + \beta_{13} FLAIR_i^N(v) * GFLAIR_i^N(v, 10) \\ & + \beta_{14} FLAIR_i^N(v) * GFLAIR_i^N(v, 20) \\ & + \beta_{15} PD_i^N(v) * GPD_i^N(v, 10) \\ & + \beta_{16} PD_i^N(v) * GPD_i^N(v, 20) \\ & + \beta_{17} T2_i^N(v) * GT2_i^N(v, 10) \\ & + \beta_{18} T2_i^N(v) * GT2_i^N(v, 20) \\ & + \beta_{20} T1_i^N(v) * GT1_i^N(v, 10) \\ & + \beta_{21} T1_i^N(v) * GT1_i^N(v, 20) \end{aligned}$$

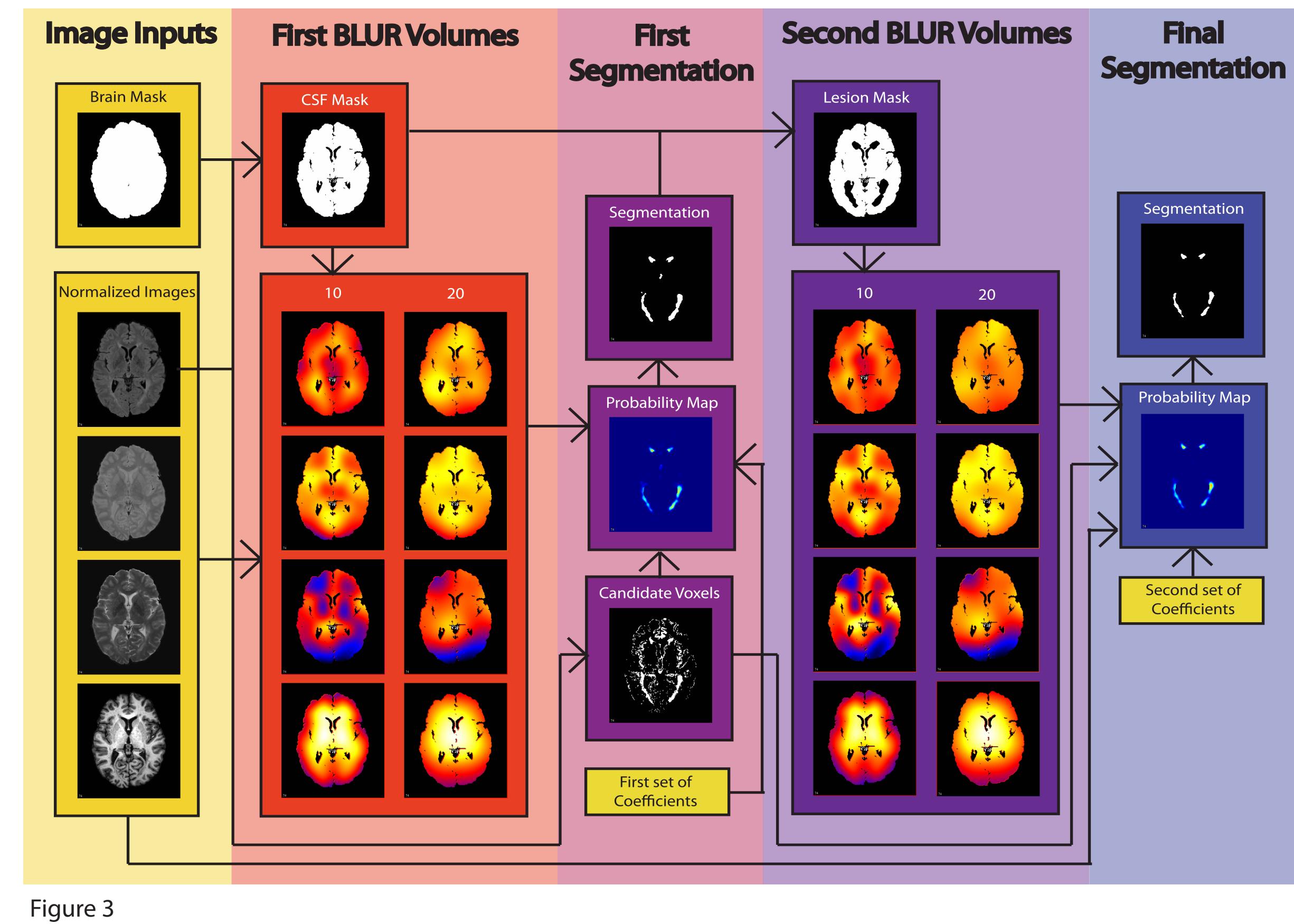


Figure 3

The OASIS model produces a set of coefficients that can be used to create probability maps of lesion presence for new MRI studies (Figure 1F). Figure 3 is a flow chart describing using the fitted OASIS model to generate a probability map from a set of new images. To make a probability map for a new study, the two sets of regression coefficients, a brain mask, and the intensity normalized FLAIR, PD, T2-weighted, and T1-weighted volumes are required. The probability maps can be thresholded at a population level to create binary segmentations (Figure 1G).

Results: Validation Set 1 and 2

We evaluate the performance of OASIS on MRI volumes acquired at two different imaging centers.

Validation Set 1

This set consists of 131 MRI studies (98 MS subjects, 33 healthy controls) with manual segmentations. The model is fit on 20 (15 MS, 5 healthy) randomly selected subjects and tested on the remaining 111. Figure 4 shows the partial ROC curve for the voxel-level detection of lesions in the testing set of Validation Set 1 with bootstrapped 95% confidence intervals. Table 1 shows threshold values, sensitivity, and Dice similarity coefficient for four different false positive rates for the model fit over all of the studies in Validation Set 1.

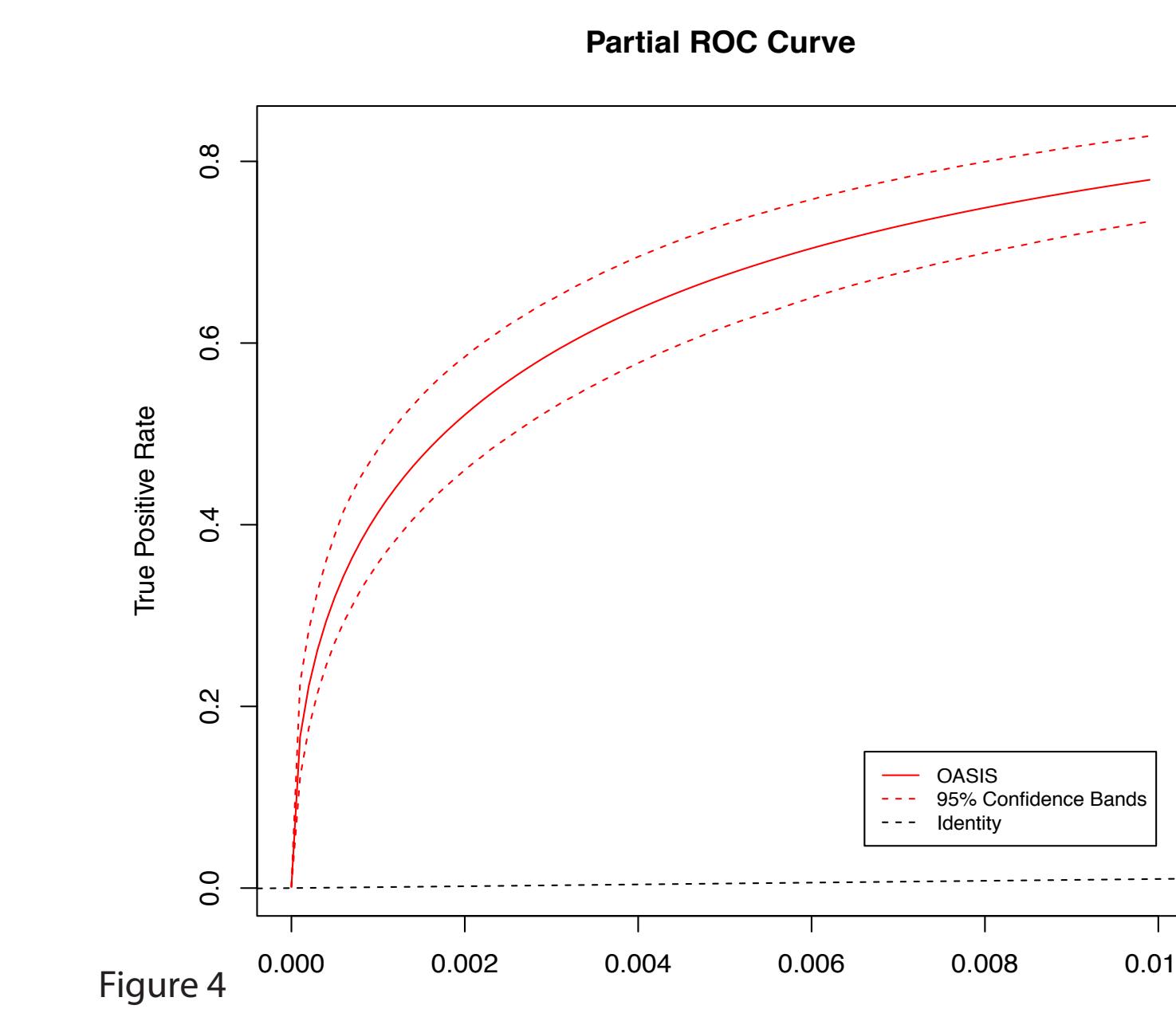


Table 1: Binary segmentation thresholds with false positive rate, sensitivity and DSC for Validation Set 1

False Positive Rate	Sensitivity	Threshold Value	DSC
1 %	80%	0.10	0.55
0.75%	76%	0.12	0.58
0.5 %	69%	0.16	0.61
0.25%	58%	0.23	0.59

Validation Set 2

This set consists of 45 MRI studies. We used the coefficients from the model fit on Validation Set 1 to produce segmentations for Validation Set 2. The only tuning parameter in the OASIS model when moving to a new population is the threshold value for binary segmentations. For Validation Set 2, manual segmentations were not available. Instead, expert evaluations of the segmentations were used for validation. We compared the OASIS segmentations at the population level threshold value from Validation Set 1 and the population level empirically adjusted threshold to the segmentations produced by the open source software LesionTOADS (<http://www.nitrc.org/projects/toads-cruise/>). Five of the 45 MRI studies were repeated to assess rater reliability, for a total of 50 MRI studies. A neuroradiologist, radiologist and neurologist were blinded to the segmentation method and scored the performance of each of the segmentations on a continuous scale from 0 to 100, with 0 being an unusable lesion segmentation and 100 being a perfect segmentation. Figure 5 shows an example of the rating system presented to the experts. Figure 6 and Table 2 show the results from the ratings.

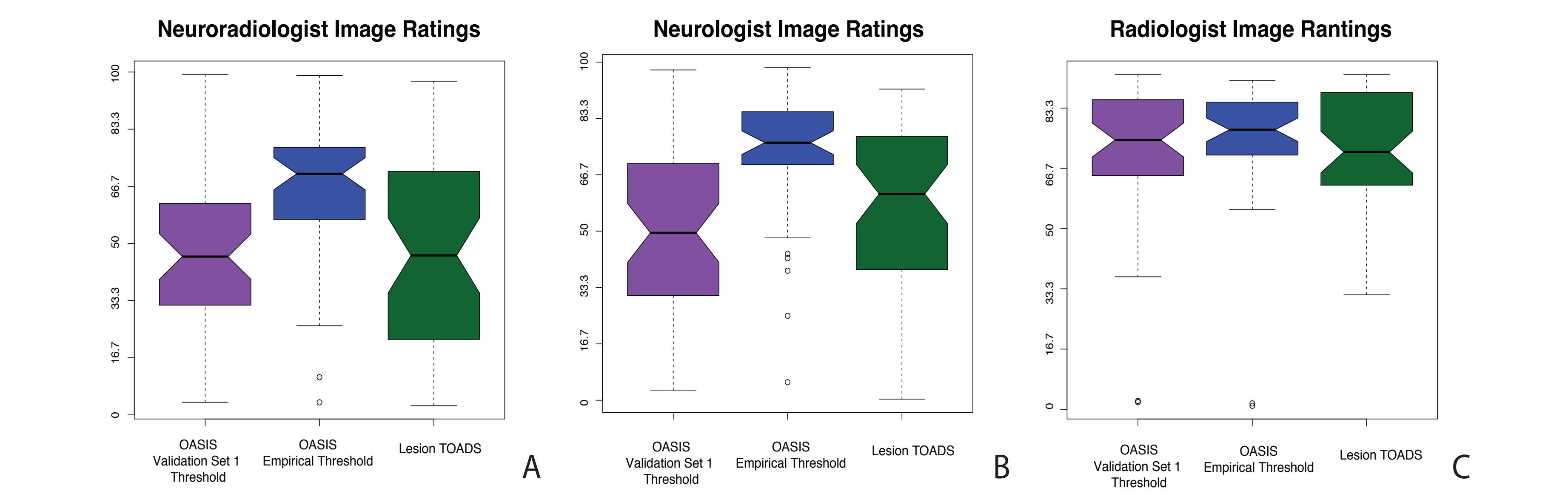
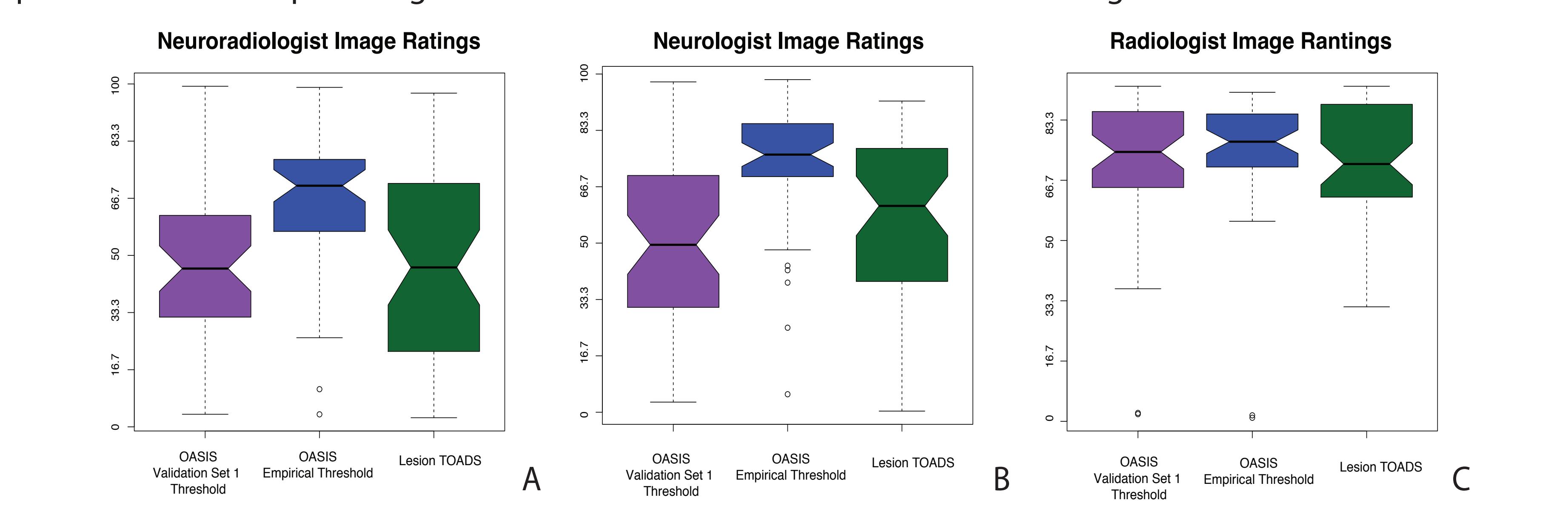


Figure 5

Table 2: Mean and standard deviation of the rating from the neuroradiologist, neurologist, and radiologist for OASIS Validation Set 1 threshold, OASIS empirical threshold and LesionTOADS on 50 studies from Validation Set 2; mean difference between OASIS empirical threshold and LesionTOADS and percentage of times OASIS was ranked higher than LesionTOADS on these images

	OASIS Validation Set 1	OASIS Empirical	LesionTOADS	Mean Difference (95% CI)	Percentage Rank
Neuroradiologist	46.3 (22.0)	66.1 (20.2)	47.3 (27.2)	18.7 (11.2, 26.3)	76% (64%, 88%)
Neurologist	48.7 (24.3)	73.1 (18.5)	50.6 (26.0)	63.5 (7.0, 25.9)	63% (52%, 78%)
Radiologist	71.6 (19.6)	74.1 (17.9)	71.8 (16.5)	2.3 (-4.2, 8.8)	52% (38%, 66%)



A B C

Conclusion

OASIS may be used to assist or even replace manual segmentation of MS lesions in the brain. After training, our fully automatic method does not require human input and avoids the variability introduced by manual segmentation. Using the explicit form of the statistical model, OASIS can easily be adapted and trained for cases where more or fewer imaging sequences are available. The potential practical utility of OASIS is high; OASIS may assist in standardizing scoring of MS lesion load in clinical trials with multiple sites and scanners. OASIS is likely to be both time- and cost-saving in large-scale studies employing neuroimaging endpoints in MS.

References

- Carass, A., Cuzzocreo, J., Wheeler, M.B., Bazin, P.L., Resnick, S.M., and Prince, J.L., 2011. Simple paradigm for extra-cerebral tissue removal: Algorithm and analysis. *NeuroImage*. 56 (4), 1982 - 1992.
- Lucas B.C., Bogovic J.A., Carass A., Bazin P.L., Prince D.L., Landman B.A., 2010. The Java Image Science Toolkit (JIST) for rapid prototyping and publishing of neuroimaging software. *Neuroinformatics*. 8 (1), 5 - 17.
- McAuliffe, M., Lalonde, F., McGarry, D., Gandler, W., Csaky, K., Trus, B., 2001. Medical image processing, analysis and visualization in clinical research. Proceedings of the 14th IEEE Symposium on Computer-Based Medical Systems (CBMS 2001), 381 - 386.
- Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity non-uniformity in MRI. *IEEE transactions on medical imaging*. 17(1), 87-97.
- Shinohara, R.T., Crainiceanu, C.M., Caffo, B.S., Gaitan, M.I., Reich, D.S., 2011. Population-wide principal component-based quantification of blood-brain-barrier dynamics in multiple sclerosis. *NeuroImage*. 57(4),1430 - 1446.