Inline Supplementary Methods 1

Let I_{ia}, I_{im} be the indicators that voxel i is labeled to be in the brain mask for the automatic and manual masks, respectively.

A voxel i is labeled to be a true positive (TP) when $I_{ia} = 1$ and $I_{im} = 1$, false positive (FP) when $I_{ia} = 1$ and $I_{im} = 0$, false negative (FN) when $I_{ia} = 0$ and $I_{im} = 1$, and true negative (TN) when $I_{ia} = 0$ and $I_{im} = 0$. Let the total number of voxels be denoted by V. The number of true positive voxels is defined as:

$$\#\text{TP} = \sum_{i=1}^{V} (I_{ia} \times I_{im})$$

Sensitivity is defined as

$$\frac{\#\text{TP}}{\#\text{TP} + \text{FN}} = \frac{\sum_{i=1}^{V} (I_{ia} \times I_{im})}{\sum_{i=1}^{V} I_{im}},$$

specificity is defined as

$$\frac{\#\text{TN}}{\#\text{TN} + \text{FP}} = \frac{\sum_{i=1}^{V} \{(1 - I_{ia}) \times (1 - I_{im})\}}{\sum_{i=1}^{V} (1 - I_{im})},$$

overall accuracy is defined as:

$$\frac{\#\text{TN} + \text{TP}}{\#\text{TN} + \text{FN} + \text{TP} + \text{FP}} = \frac{\sum_{i=1}^{V} [(I_{ia} \times I_{im}) + \{(1 - I_{ia}) \times (1 - I_{im})\}]}{V},$$

and the Dice Similarity Index (DSI) is defined as

$$\frac{2 \times \#\text{TP}}{\#\text{TP} + \text{FN} + \text{TP} + \text{FP}} = \frac{2 \times \sum_{i=1}^{V} (I_{ia} \times I_{im})}{\sum_{i=1}^{V} I_{ia} + \sum_{i=1}^{V} I_{im}}.$$

Inline Supplementary Figure 1

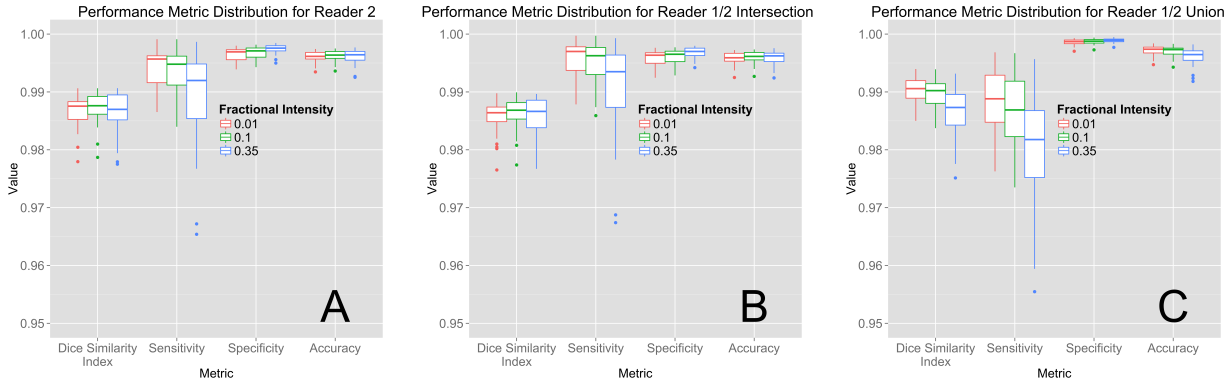

Figure 5: We display the boxplots for performance measures of the automated segmentation when using smoothed data with different fractional intensity (FI) with the gold standard being the manual segmentation from reader 2 (A), scan-wise intersection of the manual segmentation from reader 1 and reader 2 (B), or scan-wise union of the manual segmentation from reader 1 and reader 2 (C). Overall, using an FI of 0.01 and 0.1 perform high on all measures, regardless of manual segmentation used as the gold standard.

Inline Supplementary Figure 2

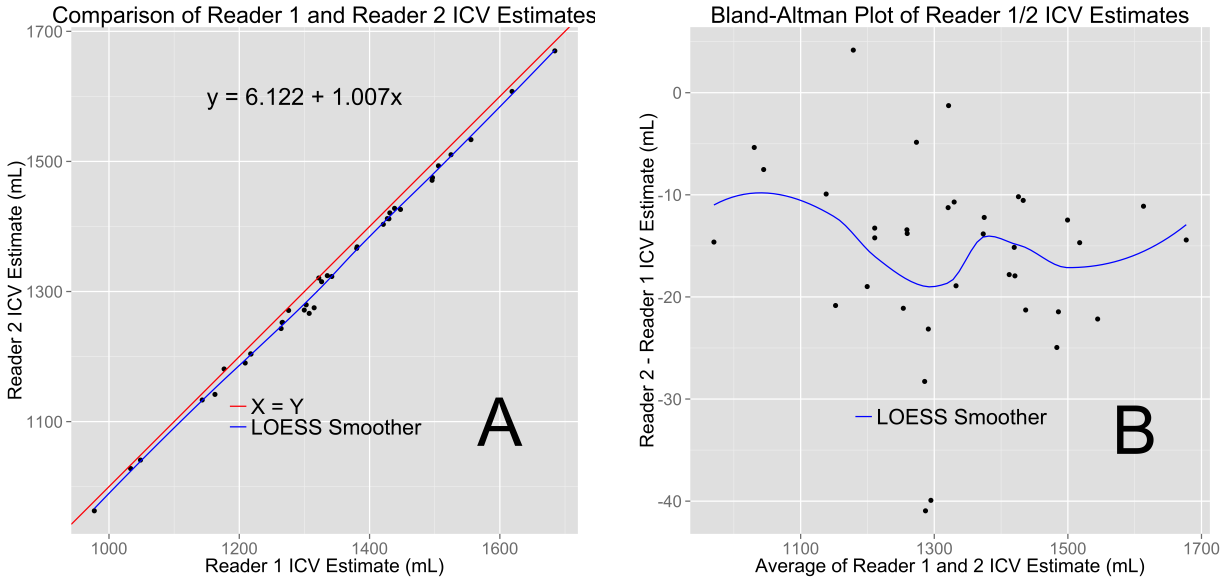

Figure 6: Panel A displays the intracranial volume (ICV) estimate from the manual segmentation of reader 1 versus reader 2. The blue line represents a LOESS scatterplot smoother of the data. The red line represents a linear fit. The slope is approximately 1 and the intercept is approximately 6 mL, indicating strong agreement of the estimates. The Bland-Altman plot in panel B denotes that there is no strong effect of the size of segmentation on the difference, but the ICV of reader 1 is higher on average than that of reader 2. These differences are small compared to the value of the ICV estimate, however.