Effects of patient population specific atlases on automatic segmentation of subcortical structures in Freesurfer

I. Csapo[1], J. Price[1], T. Russell[1], J. Dewey[1], E. Sem[1], D. McCaffrey[1], C. R. Guttmann[1], B. Navia[2], and D. F. Tate[1]
[1]Radiology, Brigham and Women's Hospital, Boston, Massachusetts, United States, [2]Tufts New England Medical Center, Boston, Massachusetts, United States

Objective: To examine whether or not manually labeled training data from a specific patient cohort improves the automatic segmentation of novel images from the same cohort in Freesurfer (an automatic brain segmentation tool).

Background: Neuroimaging studies often require the accurate labeling of brain structures. A number of automatic segmentation procedures rely on probabilistic information obtained from a manually labeled training set. The probabilistic atlases obtained this way have been shown to capture the anatomical variations present in the training set, but might not generalize well to patient populations with different patterns of anatomical variations.

Freesurfer (FS) uses a probabilistic atlas for segmenting the brain into different tissue compartments, including over 37 subcortical structures [1]. In this study, we examined the overlap of FS segmentation results derived from several patient population specific atlases and manual tracings of several subcortical structures of interest.

Methods: The MRI scans from 106 HIV infected patients, collected as part of the multi-site MRS HIV Consortium study, were automatically segmented with FS using the default probabilistic atlas, followed by the manual tracing of 8 subcortical structures (left and right caudate, putamen, hippocampus, and amygdala). 64 of the manual segmentations formed a training set that was used to build 6 new atlases with 2, 4, 8, 16, 32, and 64 subjects each. The remaining 42 MRI images were then segmented with FS using each of the 6 new atlases.

The automatic results for the 8 subcortical structures were compared to the manual tracings of the same structures using the Dice coefficient, a similarity measure defined as

$$D(L_1, L_2) = \frac{2 * V(L_1 \cap L_2)}{V(L_1) + V(L_2)},$$

where V(L) is the volume of the label L.

Results: Figure 1 shows the mean Dice coefficients for the automated segmentation results using the default (d) and the new atlases (2-64). 6 structures showed no significant improvement when compared to the default atlas. However, the left hippocampus segmentations showed significant improvements for all but the 64-subject atlas, while the right hippocampus showed improvements for 3 of the new atlases.

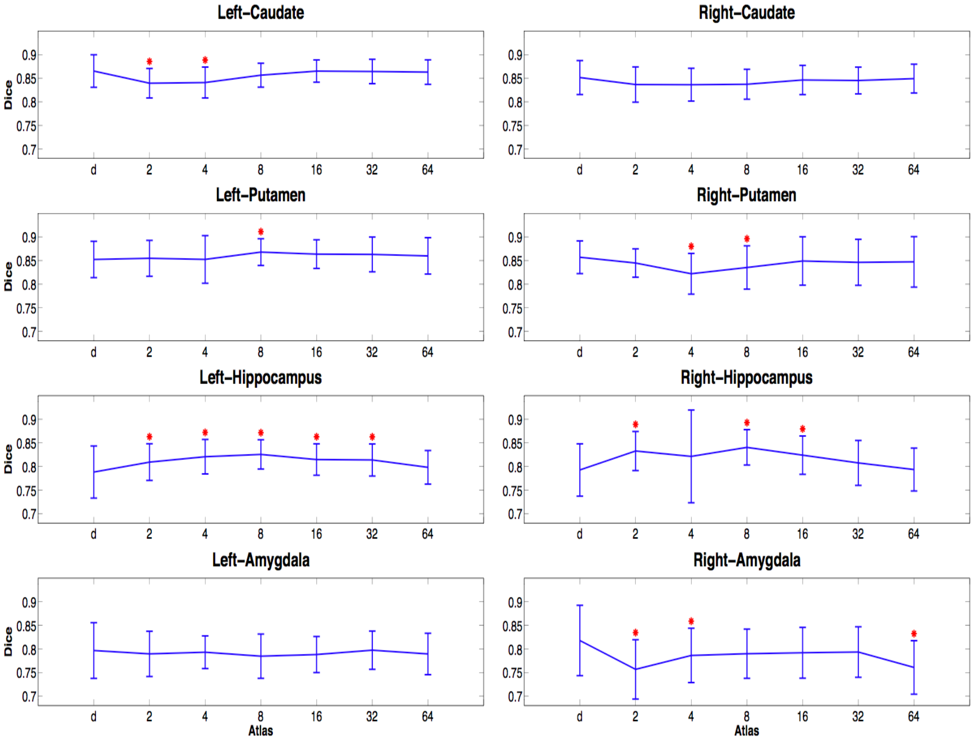

Fig 1.: Comparison of the automated segmentation results obtained using different atlases for each of the 8 subcortical structures. The x-axis represents the number of subjects used to build the atlas, "d" being the default Freesurfer atlas. The y-axis is the mean Dice coefficient (higher is better) for the 42 test subjects. The error bars are standard errors of the mean. The red asterix represents significant ($p<0.05$) difference from the mean Dice coefficient of the default atlas.

Note that the hippocampus segmentations significantly improved from the default atlas results for most of the population specific atlases.

Discussion: These results illustrate that 6 of the 8 structures did not improve significantly when using patient population specific atlases. However, the hippocampus showed significant improvement. The segmentation of the hippocampus is inherently difficult due to its convolved structure and the presence of both gray and white matter. The segmentation of such structures might benefit from a population specific atlas more than the well-delineated structures such as the caudate, putamen, and amygdala. Also, the anatomical effects of HIV on the hippocampus might be better captured with the population specific atlases.

References: 1. Fischl B, et al., *Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain*. Neuron, 2002. 33: 341-355.