**Figure Legend**

**Figure 1:** Diagrammatic illustration of the ANTs longitudinal cortical thickness pipeline for a single subject with N time points. From the N original T1-weighted images (left column, yellow panel) and the group template and priors (bottom row, green panel), the single-subject template (SST) and auxiliary prior images are created (center, blue panel). These subject-specific template and other auxiliary images are used to generate the individual time-point cortical thickness maps, in the individual time point’s native space (denoted as “ANTs Native” in the text). Optionally, one can rigidly transform the time-point images prior to segmentation and cortical thickness estimation (right column, red panel). This alternative processing scheme is referred to as “ANTs SST”. For regional thickness values, regional labels are propagated to each image using a given atlas set (with cortical labels) and joint label fusion.

**Figure 2:** Top row: Canonical views of the template created from 52 randomly selected cognitively normal subjects of the ADNI-1 database. The prior probability mask for the whole brain (middle row) and the six tissue priors (bottom row) are used to “seed” each single-subject template for creation of a probabilistic brain mask and probabilistic tissues priors during longitudinal processing.

**Figure 3:** 95% credible intervals of the region-specific variance ratios are presented for each processing method. The ANTs SST method dominates the others across the majority of regions—its point estimates (posterior medians) are greater than those of the other processing methods except for the left and right EC values in FreeSurfer Long (although there is significant overlap in the credible intervals in those regions). These results also suggest that longitudinal processing is to be preferred for both packages.

**Figure 4:** Box plots showing the distribution of the residual variability, between subject variability, and ratio of the between-subject variability and residual variability for each of the 62 DKT regions. Note that the “better” measurement maximizes this latter ratio.

**Figure 5:** 3-D volumetric rendering of the regional variance ratio values on the generated ADNI template. The higher variance ratios indicate greater between-subject to residual variability.

**Figure 6:** Log-scaled *p*-values summarizing Tables 2 and 3 demonstrating performance differences across cross- sectional and longitudinal pipelines for the three diagnostic contrasts.