Scan length, shrinkage and reliability of rsFC in the Human Connectome Project

A. F. Mejia\textsuperscript{1}, M. B. Nebel\textsuperscript{2,3}, A. Barber\textsuperscript{4}, A. Choe\textsuperscript{5,6}, B. Caffo\textsuperscript{1} and M. A. Lindquist\textsuperscript{1}

\textsuperscript{1}Department of Biostatistics, Johns Hopkins University, Baltimore, MD, USA, \textsuperscript{2}Kennedy Krieger Institute, Baltimore, MD, USA, \textsuperscript{3}Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA, \textsuperscript{4}Feinstein Institute for Medical Research, Manhasset, NY, USA, \textsuperscript{5}Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, MD, USA

\textbf{Background:} An important question in resting-state functional connectivity (rsFC) is the scan length required for reliable estimation. However, reliability can also be improved by adopting best practices in acquisition, processing and analysis. On such practice is shrinkage, in which subject-level estimates borrow strength from the group mean, and has been shown to result in 25-30\% higher reliability of rsFC within the motor cortex. Here, we examine the effect of both scan length and shrinkage on reliability of whole-brain rsFC.

\textbf{Methods:} The Human Connectome Project (N=461) includes four rsfMRI sessions each consisting of 1200 volumes (TR=0.72s) collected over two visits. After preprocessing and removal of structured artifacts with FIX, GICA was performed on all subjects to estimate a set of 300 spatial independent components (ICs). Time courses were estimated for each subject and IC through dual regression. The quantity of interest for each subject is $X_i(q, q')$, the true rsFC between ICs $q$ and $q'$, which we estimate using the first 300-2400 volumes (3.6-28.8 minutes) of the first visit. Letting $\overline{W}_\ell(q, q')$ denote the raw estimate of $X_i(q, q')$ and $\overline{W}_\ell(q, q')$ the average across subjects, the shrinkage estimate of $X_i(q, q')$ is

$$\overline{W}_\ell(q, q') = \lambda_\ell(q, q')\overline{W}_\ell(q, q') + \{1 - \lambda_\ell(q, q')\}W_\ell(q, q'),$$

where $\lambda_\ell(q, q') \in [0,1]$ is the shrinkage parameter, computed as the ratio of within-subject variance to total variance. Reliability of the raw and shrinkage estimates at each scan length is computed as absolute percent error (APE), relative to the raw visit 2 estimate.

\textbf{Results:} Figure 1 displays the APE of raw and shrinkage RSFC estimates for each seed region (median across connections). Shrinkage estimates exhibit greater reliability (lower APE) than the raw estimates, even with nearly 30 minutes (or 2400 volumes) of scan time. In fact, reliability of shrinkage estimates produced from 300 volumes is greater than that of raw estimates produced from 2400 volumes. Figure 2 displays the variance parameters and degree of shrinkage for each seed region (median across connections) and illustrates that as scan length increases, within-subject variance tends to decrease, while between-subject variance remains similar. As the degree of shrinkage is equal to the ratio of within-subject variance to total (within- plus between-subject) variance, it also tends to decrease, illustrating that with additional scan time, raw subject-level estimates become more reliable and hence require less shrinkage towards the group mean.

\textbf{Conclusions:} Shrinkage towards the group mean increases reliability of subject-level estimates of rsFC, even when produced from relatively long and high-quality rsfMRI scans.