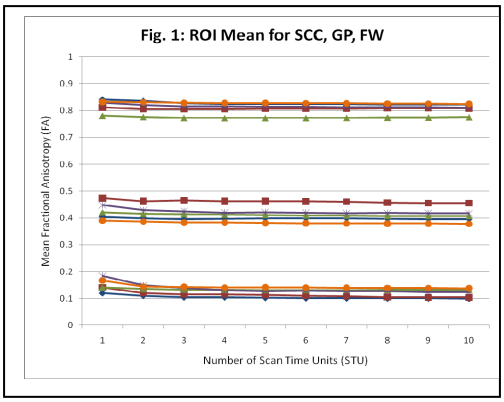
Multi-site Investigation of DTI Reproducibility

K. G. Helmer[1], M-C. Chou[2], A. Song[3], J. Turner[4], B. Gimi[5], and S. Mori[6]

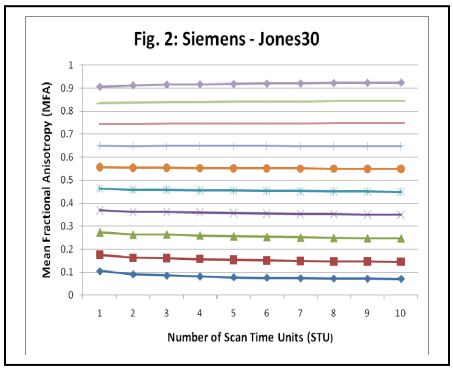
[1]Radiology, Massachusetts General Hospital, Charlestown, MA, United States, [2]Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan, [3]Duke University, Durham, NC, United States, [4]University of California, Irvine, Irvine, CA, United States, [5]UT Southwestern Medical Center at Dallas, Dallas, TX, United States, [6]Johns Hopkins University School of Medicine, Baltimore, MD, United States

Introduction: As researchers study ever more subtle disease effects, the need for an increasing number of subjects has made it necessary to collect data at multiple sites. To ensure that the data can be fruitfully combined, calibration studies must be undertaken to characterize the site effects and to determine whether they are less than the anticipated disease effect. There have been a number of studies undertaken to look at diffusion tensor imaging (DTI) metric reproducibility [1], but few studies have done so using multiple sites and scanner vendors. In this study, we collected DTI data at 5 sites representing 3 vendors. The goals were to 1) characterize the site/vendor effects on fractional anisotropy (FA) and 2) establish methods by which those wishing to include DTI in their multi-site study protocol can characterize their sites prior to the start of data acquisition and after hardware and/or software upgrades. The metric used was the reproducibility of FA as the amount of data used in the calculation was varied [2].

Materials and Methods: Five locally-recruited subjects were scanned at each of the five sites. Three scanner vendors were represented: Siemens (2 sites – Massachusetts General Hospital, UC-Irvine), GE (1 site - Duke), and Philips (2 sites – Johns Hopkins School of Medicine, UT Southwestern Medical Center, Dallas). All scanners had a field strength of 3.0T. Ten DTI scans were performed on each subject, using the Jones 30 set of diffusion-weighted directions (DWD) and 5 b=0 scans. Each Jones30/5 b=0 set was defined to be one "scan-time unit" (STU). Other protocol parameters include: b-value of 1000 s/mm^2, 2.5 mm^3 isotropic voxels, acquired matrix size: 96 x 96, full k-space coverage, FOV: 240 x 240 mm, slice thickness: 2.5 mm, number of slices: 25, parallel imaging: SENSE (p = 2) for Philips and GRAPPA for Siemens, 1 average, TR/TE (ms) were: Siemens = 4000/98 (MGH), 3800/98 (UCI); GE = 5200/69.8 (Duke); Philips = 4000/101.19 (Dallas), 4000/100.00. The achievable TR/TE is dictated by the achievable duty cycle of the scanner and the maximum gradient strength – the variations seen here were not expected to affect the results. Data sets with different numbers of STU were constructed by concatenating a sequentially increasing number of data sets together before calculation of the tensors and associated metrics, i.e., data set 1 (STU=1), data set 1 and 2 (STU=2), data sets, etc. Each frame within the concatenated data set was registered to the first b=0 frame using a 12 degree-of-freedom registration code. Tensors and tensor metrics were calculated using in-house code written in C. Noise and skull voxels were removed using a combination of Brain Extraction Tool (FSL, University of Oxford) and in-house code written in IDL (ITT-VIS).



Analysis: Whole-brain FA histograms were calculated and divided into 0.1-wide bin ranges from 0.0 to 1.0; the mean FA value was calculated for each bin range. The STU=10 data set, calculated using all 10 co-registered runs, was used as a "gold standard" and to identify the bin-range membership of each brain voxel. The corresponding bin means were then calculated at each STU value using those voxels identified as belonging in a given bin for the STU=10 data. This gave estimates for the effect of decreasing the amount of data used in the tensor calculation for each site. The data was sub-sampled to smaller numbers of directions (electrostatic model using 15, 10, 6 directions) and the same bin-mean vs. STU analysis was repeated to determine the effect of reducing the number of gradient directions on the reproducibility of the FA values as STU was changed. In addition, mean FA vs. STU plots for 6 brain structures (internal capsule {IC}, frontal white matter {FW}, centrum semiovale {CS}, globus pallidus {GP}, putamen {PUT}, and the splenium of the corpus callosum {SCC}) and for each site were generated from the Jones30 data.



Results and Discussion: Fig. 1 show examples of the mean ROI FA for three different brain structures that span the range of possible FA values (each trace is a different site). All three structures show evidence of the expected upward bias in the FA value at low STU. Note that amount of variance between different sites is maintained as STU increases. and was not statistically different for all of the 6 structures studied. Figure 2 shows the mean bin FA for a Siemens scanner that > 1 STU of data is needed to eliminate the upward FA bias at 3.0T using the Jones30 set of DWD's. There were statistically significant differences in mean bin FA at the lowest FA bins, but not at the higher FA values.

References: [1] Pfefferbaum, et al., JMRI, 18:427 (2003); [2] Farrell, et al., JMRI, 26:756 (2007).