

Impact of the Analysis of Phantoms on Data Quality for the DTI Component of the NIH MRI Study of Normal Brain Development

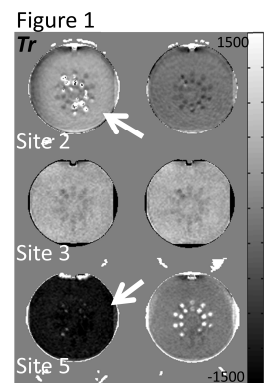
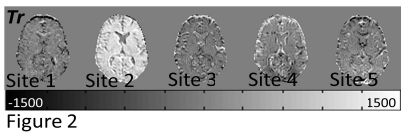
Lindsay Walker^{1,2}, Michael Curry¹, Amritha Nayak^{1,2}, Nicholas Lange³, Carlo Pierpaoli¹, and The Brain Development Cooperative Group⁴

¹PPITS/STBB/NICHD, NIH, Bethesda, MD, United States, ²CNRM, USUHS, Bethesda, MD, United States, ³Harvard Schools of Medicine and Public Health, Boston, MA, United States, ⁴www.NIH-PediatricMRI.org

Introduction: The NIH MRI Study of Normal Brain Development (PedsMRI)¹ aims to study human brain development (birth – 18 years) using MRI modalities including diffusion tensor imaging (DTI) with a representative sample matched to US census data, thus requiring the use of multiple imaging centers. Multicenter studies suffer from increased variability due to potential differences between sites. Previous studies found increased variability due to site effects and some recommend application of variance-based weights of tensor derived metrics in statistical analyses^{2,3,4}. However, increased noise in DTI data results in bias⁵ not only increased variance, and the impact of outlier datasets has been neglected. We define 3 sources of variability in DTI data: 1) occasional outlier datasets, 2) systematic differences between centers, and 3) intrinsic noise. Careful characterization of all 3 sources of variability is important for meaningful pooling of quantitative DTI data. The DTI component of the PedsMRI project is the largest prospective DTI study to date covering the key developmental stages; all data will be available to the public. Prior to the upcoming PedsMRI data release, including DTI data for the first time, we performed an analysis of variability on the PedsMRI phantom DTI data using our previously published framework⁶ with the aim of providing high quality DTI imaging data for the public data repository.

Methods: Imaging data: Phantom scans were acquired at 1.5T (GE or Siemens) at 5 participating centers, with b-values of 0s/mm² and 1000s/mm², 6 diffusion sensitization directions, repeated 4 times without averaging. Data were corrected for motion, eddy current and EPI distortions using TORTOISE⁷, and then registered to T2W images in MNI space. **Physical Phantom:** The ACR accreditation water phantom⁸ was scanned monthly at each center. Due to deviations from protocol, 69 scans from 3 of the 5 sites are analyzed here. **Living Phantom:** A healthy adult male aged 51 years at first scan was imaged at the beginning of years 1, 3 and 5 at each center. 20 out of 22 DTI scans were analyzed, 2 were rejected after visual QC for gross artifacts. **Analysis:** Data were analyzed according to our previously published framework⁶. Briefly, this is a two-step analysis framework applied to tensor derived metrics such as fractional anisotropy (FA) and mean diffusivity (*Tr*). Step 1: identification of outlier datasets by calculating the difference of each dataset from the median of all datasets, step 2: calculation of inter- and intra-site variability and inter- and intraclass correlation coefficients (ICC).

Results: Physical phantom: Outlier identification showed 3 patterns of outliers: increased *Tr* (Fig 1 site2), greatly reduced *Tr* (Fig 1 site 5), and slightly reduced *Tr* with increased FA. Site 3 showed a systematic increase in *Tr* at all time points. The inter-site contribution to variability (e.g. ICC_{inter,FA}=0.72) is much greater than the intra-site contribution (e.g. ICC_{intra,FA}=0.28), indicating a significant site effect. **Living Phantom:** Outlier identification showed 2 patterns of outliers: increased *Tr* with reduced FA (Fig 2 site 2), and elevated FA with normal *Tr*. All time points showed elevated or reduced FA and *Tr* at the edges of brain structures (Fig 2), indicating some mis-registration between time points, greater for inter- than intra-site. Inter-site variability was slightly higher on average than intra-site, indicating a small site effect. When the outlier datasets were removed, both steps of the pipeline were repeated, and surprisingly, inter-site variability increased. Careful inspection of outlier images indicated a small systematic site difference, with sites 2 and 4 (GE) having lower *Tr* values than sites 1, 3 and 5 (Siemens).



Discussion: Identification of outliers in the PedsMRI phantom data provided vital information for quality control, allowing for identification of and potential remediation or rejection of DTI data for the public data repository, thereby increasing the overall quality of DTI data released to the public. By removing outliers before performing parametric testing of variance, we discovered that extreme outliers had been masking a systematic site bias between GE and Siemens scanner sites, again highlighting the importance of identification of outlier datasets. The systematic site difference seen in the living phantom does not match the site difference seen in the physical phantom. Due to the composition of the physical phantom - water, which suffers from convective flow resulting in spurious anisotropy, and embedded plastic structures which create partial volume effects and susceptibility distortions at their edges - it is difficult to interpret these results with any confidence. If a physical phantom is used, a phantom designed specifically for DTI should be selected⁹ for clear interpretation. In conclusion, combining DTI data from multiple scanning centers is confounded by site biases and increased variance. Assessment of sources of variability in the PedsMRI DTI data, including outliers, increased the overall data quality, and indicates that differences between scanner manufacturers must be considered when performing analysis of this data.

References: [1]www.NIH-PediatricMRI.org, [2]Vollmar, C, et.al., NeuroImage 2010, [3]Pagani, E, et.al., JMRI 2010, [4]Zhu, T, et.al., NeuroImage 2011, [5]Pierpaoli, C, et.al., Radiology 1996, [6]Walker, L, et.al., 19th ann ISMRM #2004 [7]www.tortoisediti.org, [8] www.acr.org, [9]Pierpaoli, C, et.al., 17th ann ISMRM #1414