Quality Assessment in a DTI Multicenter Study

A. Nayak[1], L. Walker[1], C. Pierpaoli[1], and .. the Brain Development Cooperative Group[2]
[1]NICHD, National Institutes of Health, Bethesda, MD, United States, [2]www.NIH-PediatricMRI.org

Introduction: Multicenter DTI studies are becoming increasingly popular for their ability to improve the statistical power of a study by recruiting a large number of subjects [1]. Processing and analyzing data originating from various centers, however, presents unique challenges due to the intrinsic higher heterogeneity of experimental procedures compared to a single center study. Previous studies, for example, have highlighted scanner and site contributions to variability in multicenter studies [1,2]. Here we report the quality assessment and artifact remediation strategy that we implemented on the DTI data of the NIH MRI study of normal brain development. This unique study scanned unsedated healthy children in the age range 0-18 years with the purpose of creating a database of normative MRI and neuropsychological data (www.NIH-pediatricMRI.org). The emphasis of the study was on structural MRI data and the DTI acquisition was added as an "ancillary" component without strict quality control requirements for data acquisition. The challenge during data processing has been to produce good quality tensor–derived quantities that would be suitable for inclusion into a database, despite the highly heterogeneous nature of the incoming raw data.

Method: We defined five levels of severity of confounds (i.e., protocol errors or artifacts): 0) Absent 1) Mild, 2) Moderate, 3) Severe, and 4) Unacceptable. The criterion for assigning confounds to each level takes into account the impact they are expected to have on the final computed tensor quantities. The evaluation of their impact was based on literature data and on specific tests or simulations developed when literature data was not available. This systematic quality assessment framework enabled us to classify potential confounds and identify modifications in the data processing pipeline that would make data more consistent across sites.

Results: We identified 11 different protocol errors (for example: improper slice thickness, improper in-plane resolution, incomplete acquisitions, improper gradient direction protocol, limited brain coverage, inconsistent echo time between series, zero-filling during image reconstruction, gap between slices, signal averaging at the scanner, oblique slice orientation, gain differences between series) and 7 main types of artifacts (cardiac pulsation, whole slice signal dropout, misregistration due to motion, misregistration due to eddy currents, ghosting, spike noise, and EPI distortion). Fig 1 shows the overall quality classification for incoming DTI data in our study using the quality assessment criteria we developed. Only 10% of the incoming scans were suitable for inclusion in the database without correction (green bar in the figure). The majority of datasets were not completely free from artifacts or protocol errors but the low severity of the problem or the possibility of performing an effective correction during post processing made it possible to classify them as "rescuable" datasets (orange/yellow bars). Unfortunately about 20% of datasets needed to be excluded from the database due to the presence of uncorrectable artifacts or severe protocol errors (red/pink bars). Rejection was more likely because of protocol errors rather than because of artifacts.

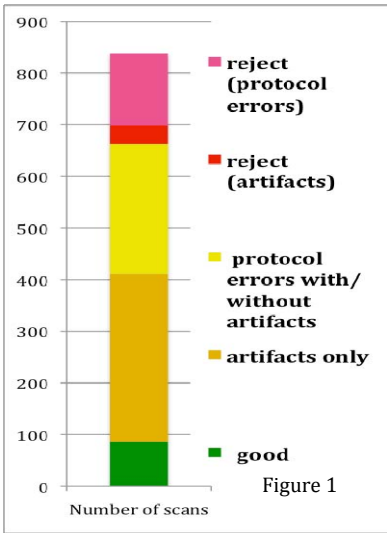

Figure 1

Discussion and Conclusion: In this work we aimed at defining a quality assessment criterion for DTI data that would have informed the decision of including datasets collected as part of a large multicenter pediatric study into a normative imaging database. We found that the lack of strict quality control requirements during acquisition resulted in a relatively large percentage of datasets having final quality unsuitable for inclusion in the database. Interestingly data were more likely to be rejected because of severe protocol errors rather than because of artifacts. In the presence of a severe protocol error, such as markedly incorrect in-plane resolution, there is little room for correction in post processing. Voxel size is known to have a marked effects on tensor-derived parameters, such as FA[3], and including data acquired at the lower resolution in the database will lead to unacceptable internal inconsistencies. On the other hand, data that has been acquired according to protocol and is only affected by artifacts can benefit from the presence of a centralized processing pipeline. Severe, moderate or mild artifacts can be successfully corrected with dedicated processing strategies [4] such as registration techniques and robust tensor estimation [5]. However, we found that remediation of artifacts was challenging if the acquisition lacked data redundancy and if there was the combined presence of both protocol errors and artifacts. In conclusion, our study suggests that the implementation of a well defined quality control strategy for data acquisition would be beneficial and that centralized processing is essential to achieve consistent quality of DTI data output in a multicenter DTI study.

References: [1] C.Vollmar et al 2010; Neuroimage 51 (2010), pages 1384-1394,[2]E.Pagani et al 2010; JMRI 31:pages1458-1468,[3]R.Gattu et al; ISMRM (2009),pages3579,[4]C.Pierpaoli et al; ISMRM,Stockholm,Sweden, (2010),[5]Chang L-C et al; MRM (2005).