Test retest reproducibility assessments for longitudinal studies: quantifying MRI system upgrade effects

J. Jovicich[1], S. Czanner[1], D. Greve[1], J. Pacheco[1], E. Busa[1], A. van der Kouwe[1], M. BIRN[2], B. Fischl[1,3]

[1]MGH/MIT/HMS Athinoula Martinos Center for Biomedical Imaging, Charlestown, MA, United States, [2]Morphometry BIRN, www.nbirn.net, BIRN, United States, [3]Computer Science & Artificial Intelligence Lab, Massachusetts Institute of Technology, Cambridge, MA, United States

INTRODUCTION

Neuroimaging longitudinal studies of change over time are becoming increasingly a standard element of clinical neuropsychiatric research [1,2]. One limitation of longitudinal studies is that major advances in technology, typically system upgrades, introduce technology related variability in the images that limits the power for following the progression of disease. This motivates the development and application of procedures both to standardize acquisition parameters over time and to estimate and correct for the error introduced by uncontrolled factors [3]. In this study we took the opportunity of an MRI system upgrade (1.5 T Siemens, from Sonata to Avanto) to investigate the reproducibility of a structural morphometry protocol before, after and across the upgrade. To make the results independent of brain morphometry tools, here we focus only on the reproducibility of image intensity.

MATERIALS AND METHODS

Upgrade changes: The upgrade from a Siemens Magnetom Sonata to a Magnetom Avanto involves major changes, which include: a) main magnet (both arre 1.5T, Avanto's length is 150cm, Sonata's is 160cm), b) gradient system (Avanto coils are more linear, Sonata 40 mT/m @ 200 T/m/s, Avanto 45 mTm @ 200 T/m/s), c) head RF coil (circularly polarized in Sonata, 12 channels in Avanto), and d) software. Even if we try to reproduce a structural MRI protocol (sequence and parameters) in both platforms it is of interest to know how the reproducibility of the images compare. To quantify this we scanned 4 healthy volunteers twice, in separate sessions (different head positions), before (Sonata) and after (Avanto) the upgrade.

Standardization of image acquisition protocol: The structural MRI protocol included acquisition of two 3D MP-RAGE scans (TR=2.73s, TE=3.44ms, TI=1s, 256x192, 1.33mm thick 128 sagittal slabs, flip angle= 7^0, 190 Hz/pixel, 8min 46s acquisition time). These parameters were kept for both platforms. Also, for both platforms brains were automatically aligned, in each scanning session, to an atlas for setting the slice prescription in aproximatelly AC-PC orientation [4,5]. The head RF coil set ups were different: circularly polarized for Sonata, and 4-channel configuration for Avanto.

Evaluation of test-retest reproducibility: Each scanning session was summarized by co-registering and averaging the two MP-RAGE scans. Each average was skull stripped [6], co-registered with a paired reference scan for reproducibility quantification, and intensity normalized (brain mean 100). For each subject we computed voxel-based relative error maps (image intensity difference) and error histograms for the following cases: Sonata test-retest (2 sessions on the Sonata, with and without 3D distortion corrections from gradient non-linearities [3]), Avanto test-retest (2 sessions on the Avanto), and Sonata-Avanto test-retest (one session on each). The mean image intensity errors for these cases were computed from the histograms for each of the subjects, and then averaged across subjects

RESULTS AND DISCUSSION

Preliminary results show that for our group of subjects the mean image intensity reproducibility error across the whole brain was 5.2% for Sonata test-retest (5.1% with 3D distortion correction), 3.1% for Avanto test-retest and 7.8% for Sonata-Avanto test-retest. We will continue to test the reproducibility effects when other sources of variability are corrected (like B1 inhomogeneities), other sequences are considered (like FLASH) on image intensity and morphometry estimates.

CONCLUSIONS

The quantification of the upgrade effects on image intensity reproducibility is not complete. Our first results show an improvement in the upgraded system relative to the older system, but a reproducibility deterioration when images from both platforms are combined. Additonal analyses will be done to test if the variability across platforms can be reduced. We propose that protocol reproducibility studies as the ones described in this abstract (not just phantom calibrations) be incorporated as part of longitudinal studies for understanding and quantifying system upgrade effects, which potentially allows for protocol tuning.

REFERENCES:

[1] Fox NC & Schott JM Lancet. 2004 Jan 31;363:392-4
[2] Ashburner J et al., Lancet Neurol. 2003 Feb,2:79-88
[3] Jovicich et al., ISMRM 2004
[4] Van der Kouwe, A. et. al. ISMRM 797, 2003
[5] Benner T., et al. ISMRM 2243, 2004
[6] Segonne F et al., Neuroimage 2004 22(3):1060-75

ACKNOWLEDGEMENTS: NCRR BIRN Morphometry Project (BIRN004).