CORRECTION OF MRI SCALING PARAMETERS: A COMPARISON OF PHANTOM AND REGISTRATION METHODS ON THE ADNI DATASET

M. J. Clarkson[1,2], C. Nielsen[1], K. L. Leung[1,2], J. Barnes[1], J. L. Whitwell[3], J. Gunter[3], C. R. Jack[3], N. C. Fox[1], and S. Ourselin[1,2]

[1]Dementia Research Centre, University College London, London, United Kingdom, [2]Centre For Medical Image Computing, University College London, London, United Kingdom, [3]Mayo Clinic College Of Medicine, Rochester, MN, United States

Introduction: Rates of brain atrophy derived from serial magnetic resonance (MR) studies are beginning to be used to assess putative anti-dementia drugs. However, the rate of brain atrophy may be incorrectly measured due to changes in scanner voxel sizes, known as scanner or voxel drift. For this reason the Alzheimer's Disease Neuroimaging Initiavite (ADNI) included the imaging of a phantom with every scan to correct for this problem [1]. The logistical, practical and financial limitations of scanning a phantom at the time of all subject scans means that an alternative post processing correction would be highly desirable. In this study, we directly compare the scaling correction calculated using the phantom with scale factors calculated using a 9 degrees-of-freedom (DOF) registration algorithm [2], and measure the effects on the boundary shift integral (BSI), a measure of whole brain atrophy [3]. Previous work has demonstrated on smaller datasets, with larger scaling errors, that 9DOF registration can correct for scanner drift. This is the first study to directly compare registration and phantom scaling correction using data from a large, multi-site study.

Methods: From www.loni.ucla.edu/ADNI we downloaded 2 datasets, each containing 129 (79 control, 50 Alzheimer's disease (AD) subjects), pairs of baseline and repeat, T1 weighted volumetric scans with an interscan interval of approx. 1 year. The first dataset were corrected for gradient warping, B1 intensity inhomogeneity, and intensity inhomogeneity as appropriate [1], but not phantom scale corrected. The second dataset was additionally phantom scale corrected. For each scan, a brain mask was created by semi-automatically segmenting the brain and dilating 8 times. All registrations were performed using these brain masks. **Expt 1**. We took the phantom corrected dataset, and added 0.125%, 0.25%, 0.5% and 1.0% to the repeat image header X, Y and Z voxel size. This gave a volume change of 0.38%, 0.75%, 1.51% and 3.03% respectively. We registered each repeat scan to the corresponding baseline scan using a 9DOF registration algorithm [2]. From the final registration X,Y and Z scale factors, we calculated a volume scale factor (X * Y * Z), and compared that to the initial artificially added volume change. We compared the AD group to the Control group at each level of scale change using an unpaired T-Test. **Expt 2**. We took the phantom corrected dataset, and registered each of the repeat scans to its baseline scan. If the phantom correction worked perfectly, the registration algorithm should calculate a scale factor of 1 in the X, Y and Z direction. **Expt 3**. We took the images that were not phantom corrected, and registered them using 6DOF and 9DOF. After performing differential bias correction [4], we measured the whole brain BSI [3], and compared each method using a paired T-Test, and Pitman's test to assess the effects on the mean and variance of atrophy rates.

Results: Table 1 shows the results for Expt 1. Additionally, by graphing the results (not shown), we identified a single outlier for each set of artifical scale changes. We took the phantom corrected images, and registered them using 6DOF. Visual inspection revealed that the phantom correction was incorrect in the Y-direction (parallel to the scanner bore). For Expt 2. we registered phantom corrected images, using 9DOF and the registration algorithm gave a mean (SD) % volume change of 0.04(0.29)% for control and 0.04(0.32)% for AD subjects. The mean (SD) absolute % volume change was 0.20(0.20)% for control and 0.19(0.26)% for AD subjects.

Subject	Statistic	Artificial Applied Scale Change (% volume change)			
		0.38%	0.75%	1.51%	3.03%
Ctrl	Vol %	0.34(0.29)	0.72(0.29)	1.47(0.29)	2.99(0.30)
	Abs Diff %	0.20(0.21)	0.21(0.21)	0.21(0.21)	0.21(0.21)
AD	Vol %	0.34(0.32)	0.72(0.32)	1.47(0.32)	2.99(0.33)
	Abs Diff %	0.20(0.26)	0.20(0.26)	0.20(0.26)	0.20(0.26)
T-Test	AD c.f. Ctrl	0.96	0.96	0.95	0.95

Table 1. For each level of artificially added scale change, we show the mean (SD) percent volume change calculated by the 9DOF registration algorithm, and the mean (SD) absolute difference between the registration algorithm, and the target artificially added scale change.

Dataset	DOF	Mean (SD) BSI in ml	
		Control	AD
Uncorrected	6	5.83 (10.85)	14.59 (10.24)
Phantom Corrected	6	5.87 (10.18)	14.53 (9.59)
Uncorrected	9	5.55 (9.49)	14.10 (9.29)

Table 2. The effect of phantom and registration scaling correction on BSI measurements.

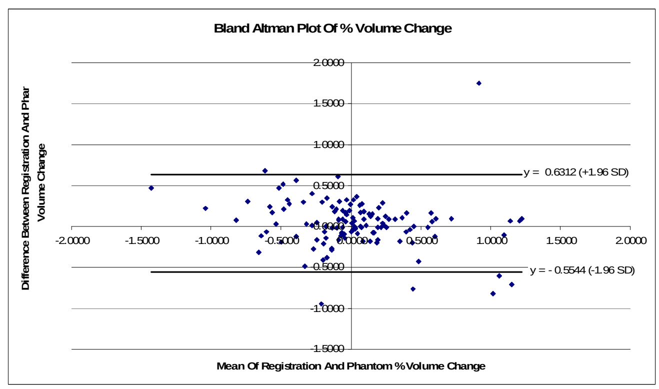

Figure 1. A Bland Altman plot comparing mean (X-axis) and difference (Y-axis) between the volume scale factor calculated using the phantom, and the 9DOF registration algorithm. Visual inspection of the 7 most outlier points revealed that the phantom scaling correction was most likely incorrect in that there was residual linear scaling after phantom correction that was removed by 9DOF registration.

Discussion: Expt 1 and Table 1 show that the registration algorithm can correct for scale changes ranging from 0.38% to 3.03%. The T-Tests reveal that there is no significant difference in performance between AD and control groups. This is important, as it demonstrates that we are not "registering away" atrophy when correcting for voxel scale change. The precision of the algorithm is 0.29% to 0.33%, giving a 95% confidence interval of 0.61%. Looking at the dispersion around the origin in Figure 1, we cannot claim that the algorithm is reliable under 0.61%. In Table 2 we see that 6DOF registration of phantom corrected images gives BSI values that are very close to the completely uncorrected measurements. This implies that in our dataset, the phantom is having little effect. In the baseline scans, the mean (SD) volume correction calculated by the phantom was 0.33 (0.36)% and a maximum of 1.66%, and similar for repeat scans. With 9DOF registration, there is a trend towards lower standard deviation of BSI measurements, but it did not reach significance. There was no significant difference in the mean BSI values between 9DOF registered images, implying that the registration scaling correction is comparable to the phantom scaling correction.

Conclusion: In this study we took a set of 129 pairs of images from the ADNI database, with and without phantom correction. The phantom corrections calculated were on average relatively small, resulting in no significant difference to the BSI. Subsequently, registering the un-corrected images using 9DOF enabled us to correct for scale changes to a comparable level, and indeed enabled us to detect cases where the phantom correction was incorrect. There was a trend suggesting that the 9DOF registration reduced the standard deviation of BSI measurements, which would improve group separation but it was not significant. Future work could investigate improving the level of precision of the registration algorithm. This work is of interest to anyone planning a long term clinical trial involving serial MR registration.

References

[1] Jack *et al.* J. Magn Reson Imaging, 2008, 27, 685-691
[2] Woods *et al.* JCAT, 1998, 22, 139-152
[3] Freeborough, Fox. IEEE TMI, 1997, 16, 623-629.
[4] Lewis, Fox. NeuroImage, 2004, 23, 75-83.