A retrospective multi-site, multi-manufacturer comparison of T1-weighted anatomical brain scans using tissue classification

Alain Pitiot[1], Olivier Mougin[2], Nikos Evangelou[3], and Paul S Morgan[4]

[1]School of Psychology, University of Nottingham, Nottingham, Nottinghamshire, United Kingdom, [2]Sir Peter Mansfield MR Centre, University of Nottingham, United Kingdom, [3]Division of Clinical Neurology, University of Nottingham, United Kingdom, [4]Medical Physics, University of Nottingham, United Kingdom

Target Audience

Scientists involved in designing MRI protocols for neuro anatomical multi-site trials.

Purpose

Multi-site neuroimaging MR studies are an effective methodological approach to increasing sample sizes while maintaining reasonable study timeframes. Larger samples make it possible to tease out finer details about a particular neurological condition or about differences between groups, and to generalise to the population level with greater confidence [1,2]. In terms of acquisition, most of the efforts focus on reducing variability in the imaging protocol parameters with the reasonable assumption that this will reduce variability in the acquired images, which, in turn, will reduce variability in the subsequent data analysis. However, MR acquisition is only the first stage of the overall acquisition/analysis procedure, and it is not clear which parameters in the processing pipeline will affect the overall variance the most. In particular, the algorithms of popular post-processing software packages such as SPM, FSL or Freesurfer are purposefully designed to correct for some of the scanner specific image characteristics, such as intensity non-uniformities that can arise from multi-element receive-only head coils. While standardising the MR acquisition parameters across sites is wise *a priori*, it is not clear how important this is to ensure the resultant image data can be combined successfully. In addition, producing near identical standardised MR protocols often results in protocols that are suboptimal for all scanners used.

Since systematically examining the effect of each settings of the relevant acquisition protocol variables would result in an impractically large amount of data collection, previous studies focused on specific acquisition parameters such as field strength, scanner upgrade and manufacturer [4], or pulse sequence, voxel geometry and parallel imaging [5]. They estimated on a group of healthy participants differences in the widely used cortical thickness measure induced by a discrete set of parameter settings (e.g. 2 field strengths and/or 2 voxel sizes and/or 2 scanner manufacturers and/or 2 sites and/or 2 sequences, etc.).

Here we investigate retrospectively the differences induced by sites and manufacturers on the tissue volumes estimated in a variety of brain regions of a single healthy control. We restricted our study to 1 mm^3 isotropic, whole brain, 3-D gradient echo scans at 3-T, since (a) these are typical of modern multi-site trials and (b) the significant effect of voxel sizes and field strengths has already been documented elsewhere ([6] for instance). In contrast to previous studies, the scans varied considerably in terms of sequence parameters. Additionally, using regional tissue volumes in a single participant makes it possible to alleviate the non-linear registration issues of multi-participants studies while retaining morphological variability (across regions). While the true volume of the individual's brain was not known, the study of this variability should provide insight into the robustness of the combined acquisition and analysis, and the extent to which protocols must be standardised across sites.

Methods

Within the context of a variety of neuroimaging studies, 41 different T_1-weighted, 3D gradient echo, 1 mm^3 isotropic whole brain scans of the same individual were acquired over a period of 9 years, all with approval from the relevant local ethics committees or IRBs. Scans from 3 T MR scanners manufactured by Philips Medical (Best, The Netherlands) and Siemens Medical (Erlangen, Germany) were acquired from three separate sites (two in the US, and one in the UK). All scan parameters were optimised to provide a visually acceptable contrast between white matter (WM) and gray matter (GM) with suppressed signal from CSF. However, they differed in a number of ways. Scans were acquired either axially or sagittally, using parallel imaging head coils with 8, 12 or 32 receive elements, all with a parallel imaging factor of 2 in the AP direction. The excitation flip angle varied between 8° and 15°, TE between 3.4 and 5.6 ms, TR per line between 8 and 15 ms, magnetisation preparation inversion pulse effective TI between 860 and 1100 ms (with a selection of both non-selective and selective inversions), and shot interval between 2200 and 3000 ms. Receiver bandwidth varied between 150 and 200 Hz/pixel. No interpolation or partial Fourier techniques were used. Voxels of 1x1x1 mm were acquired with a range of matrix dimensions from 224x224x160 to 256x256x192 to fit the head.

All scans were first re-oriented into MNI space, skull-stripped using FSL bet2 [3] and tissue classified in SPM8 using the New Segment approach [7] with standard parameters. We then computed the GM volumes for each of the 48 regions in the Harvard-Oxford cortical atlas, and the GM and WM volumes for each of the 9 regions in the MNI lobe atlas by integrating the modulated tissue probability values in standard space.

Results

Since nearly half of the volumetric measurements were non-normally distributed, as assessed both visually and using Levene's tests, we estimated differences across sites and across manufacturers using non-parametric independent-samples Kruskal-Wallis and Mann-Whitney U tests, respectively. All results were corrected for multiple comparison using the false discovery rate (FDR) at 5%.

Even though most acquisition parameters (e.g. flip angle, echo time, receiver bandwidth) were significantly different across sites and across manufacturers, we found no statistically significant differences between sites or between manufacturers for the MNI lobe volumes. We found only 4 (out of 48 possible) significant differences between sites and 5 (out of 48 possible) between manufacturers for the Harvard-Oxford cortical region volumes. Across sites and manufacturers, the relative differences between means were 1.5% and 3.5% on average across regions, respectively, and never exceeded 6%. We did not observe any significant variation of the volumes over time.

Discussion & Conclusion

The results suggest that, for a fixed voxel size and at the same field strength, differently optimised T_1-weighted 3D acquisitions of the brain from multiple MR scanners including two scanner manufacturers may be combined at analysis stage without undue concern to the exact acquisition parameters. Should those results hold in future studies with multiple subjects, it may be the case that a better compromise for multisite study would be for each site to use the best protocol for their scanner, independently from the others, rather than implement an inherently inferior, albeit standard, protocol.

References:

1. Jovicich J *et al*. NeuroImage, 30:436-443 (2006) | 2. Jack Jr CR *et al*. J Magn Reson Imaging, 27:685-691 (2008) | 3. Jenkinson M *et al*., Int. Conf. on HBM. (2005) | 4. Han X *et al.* NeuroImage, 32:180-194 (2006) | 5. Wonderlick J.S. *et al.* NeuroImage, 44:1324-1333 (2009) | 6. Lusebrink F *et al.* NeuroImage, 70:122-131 (2013) | 7. Ashburner J. Mag. Res. Imag. 27:1163 (2009)