

Evaluation of 3D T1-weighted imaging at 3 T across scanner vendors and models

Sjoerd B Vos^{1,2}, M Jorge Cardoso¹, Marzena Wylezinska-Arridge³, David L Thomas^{1,3}, Enrico De Vita^{3,4}, Marios C Yiannakas⁵, David Carmichael⁶, John S Thornton^{3,4}, Olga Ciccarelli⁵, John S Duncan^{2,7}, and Sebastien Ourselin¹

¹Translational Imaging Group, University College London, London, United Kingdom, ²MRI Unit, Epilepsy Society, Chalfont St Peter, United Kingdom, ³Neuroradiological Academic Unit, Department of Brain Repair and Rehabilitation, UCL Institute of Neurology, London, United Kingdom, ⁴Lysholm Department of Neuroradiology, National Hospital for Neurology and Neurosurgery, London, United Kingdom, ⁵Department of Neuroinflammation, UCL Institute of Neurology, London, United Kingdom, ⁶Institute of Child Health, University College London, London, United Kingdom, ⁷Department of Clinical and Experimental Epilepsy, UCL Institute of Neurology, London, United Kingdom

Synopsis

Volumetric analyses of 3D T1-weighted images has become an integral part of the clinical work-up and research studies. Variation between scanners, in both vendors and models, is a major confound in combining imaging-derived biomarkers across sites. In this work, we analyse test-retest data from different days on six 3 T scanners from three vendors to quantify this inter-scanner variability compared to intra-scanner variability. Contrast-to-noise ratios as well as volumetric analyses are performed showing between-scanner variation in total brain volumes – indicating different scanner calibrations – but also tissue-specific differences – possibly arising from different effective contrasts.

Introduction

Volumetric analyses using 3D T1-weighted acquisitions (3D-T1) has become an integral part of the clinical work-up for neurodegenerative diseases [1,2], epilepsy [3, <http://hipposeg.cs.ucl.ac.uk>], and many clinical trials with both cross-sectional and longitudinal designs [4]. A major confound in volumetric analysis is acquisition protocol variation – most notably between vendors but also between specific models, pulse sequences, and parameter choices – compromising the integration of imaging-derived biomarkers across sites and their use for the monitoring and management of patients scanned with different machines. In this work we evaluate the variations in T1 image quality and volumetry across six 3 T MRI scanners – from three main vendors – at a single site. This is a crucial step towards the standardisation of these sequences to improve their translation to clinical use and their harmonisation between centres.

Methods

Two healthy subjects (28yo male, 46yo female) were scanned on six 3 T scanners: a Philips Achieva, GE MR750, and Siemens Skyra, Trio, Prisma, and Biograph mMR (PET-MR). Each subject was scanned three times on all scanners. To discriminate potentially larger day-to-day variation from same-day scan-rescan changes [5], each subject underwent two scans on one day with a five-minute interval between scans and a further single scan on a separate day on each scanner. All protocols are the established standard sequences used in routine clinical practice and/or research studies, optimised for different dedicated applications and patient populations. Acquisition details for each scanner are given in Table 1. Scans were registered to an intermediate space to visualise contrast differences between protocols. Tissue segmentations were performed using Geodesic Image Flows framework (GIF [6]) in native space, which includes bias-field correction, from which total intracranial volume (TIV), WM, cortical (cGM) and deep (dGM) grey matter, and ventricular and non-ventricular CSF volumes were extracted. Signal-to-noise ratios were quantified within each tissue type by calculating mean and variance over the tissue, and the CNR between two tissue types was defined as the absolute difference in mean values between those two tissues divided by the weighted mean of the variances.

Results

Fig. 1 shows equivalent axial and coronal slices for all six scanners, in which contrast differences are evident between the six scanners. The clearest difference is in the GM-WM contrast between the scanners. The protocol on the Achieva has the highest visual GM-WM contrast, arguably at the cost of lower cGM-CSF contrast – as confirmed by the estimated CNR (Table 2). By contrast, images from the Siemens scanners tend to have higher GM-CSF contrast. Volumetric analyses are shown in Figs. 2 and 3. Maximum difference in TIV between scanners is around 60 ml, or just under 4%, where maximum intra-scanner TIV variation is well under 1%. Variation in TIV across scanners is consistent across subjects. The differences in contrast (Fig. 1) result in pronounced variations in segmented tissue volume – even when normalised by TIV. Most notably, the Biograph mMR yields WM volumes 1% higher than the other scanners, and the MR750 yields cGM estimates around 1-1.5% higher. In both cases there is no linked and equal decrease in adjacent tissues, indicating that it is not one tissue interface that is classified differently. Intra-scanner variation is low on all

scanners for WM, cGM, dGM, and non-ventricular CSF, with the lateral ventricles showing the largest inter-scan variability.

Discussion

Inter-scanner variation in volumetric brain analysis is larger than intra-scanner variation. Differences in TIV are possibly indicative of slight gradient calibration variations, leading to minor scaling differences between the true head dimensions and each scanner's representation thereof. However, estimated tissue volumes, normalised for TIV, still vary more across scanners than within. Slight inter-scanner imaging parameter variations may be the cause of this. Sequence implementation differences across vendors might also contribute to between-scanner variation, and the relatively close clustering of volume estimates across the Siemens Skyra, Trio, and Prisma systems, relative to those from the other scanners, suggests vendor-based differences – although the Siemens Biograph mMR shows a large deviation in WM volumes compared to all other scanners. With respect to CNR estimates, the vendors employ different inline or optional image reconstruction filters which might bias the comparison across vendors. Future work will extend this study to a larger subject group and investigate further metrics including localised analyses and cortical thickness.

Acknowledgements

SBV is funded by the UK National Institute for Health Research UCLH Biomedical Research Centre High Impact Initiative. Part of this work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme. We are grateful to the Wolfson Foundation and the Epilepsy Society for supporting the Epilepsy Society MRI scanner, and the UK MS Society for supporting the NMR Unit scanner. DLT is supported by the UCL Leonard Wolfson Experimental Neurology Centre (PR/ylr/18575).

References

[1] Heijer et al., Brain 2010; [2] Rohrer et al., The Lancet Neurology 2015; [3] Winston et al., Epilepsia 2013; [4] Salloway et al., N Engl J Med 2014; [5] Maclaren et al., Nature Scientific Data 2014; [6] Cardoso et al., IEEE TMI 2015.

Figures

Scanner	SE (mm)	TI (mm)	TR (mm)	FA (°)	FOV (mm)	Matrix APxRLxSI	ISZT	Coil (phantom)	Time (min)
Schnee	3.2	802	8.8	8	250x250x250	250x250x250	S2 AP	32	6:11
MR750	5.5	400P	7.4	11	250x250x250	250x250x250	G2 AP	32	4:20
Biograph	3.9	870	8.1	8	280x250x280	250x208x250	G2 AP	32	5:00
Prisma	3.7	800	7.4	8	240x250x250	240x250x250	G1 AP	64	5:25
Moya	3.0	800	6.3	8	250x208x250	250x208x250	G2 AP	32	4:53
Trio	3.8	800	6.3	8	250x208x250	250x208x250	-	32	9:28

Table 1: Scan parameters for the six scanners (IPAT=parallel imaging technique: S2=SENSE 2, G2=GRAPPA 2, AP=anterior-posterior, RL=right-left). *Note that the definition of TI is different for GE than for Siemens and Philips.

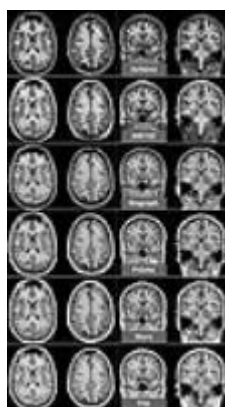


Fig. 1: Axial (left) and coronal (right) slices for one scan from one subject for all six scanners. Intensity scaling was done from 0 to 1.25 times the mean WM intensity per scan.

Scanner	Subject 1					Subject 2				
	WM-cGM	WM-dGM	WM-CSF	cGM-CSF	dGM-CSF	WM-cGM	WM-dGM	WM-CSF	cGM-CSF	dGM-CSF
Schnee	8.5	4.7	13.2	4.9	4.7	8.6	4.7	12.8	5.3	5.3
MR750	9.5	4.4	11.2	4.6	4.3	9.1	4.8	12.9	5.3	5.1
Biograph	5.5	3.9	11.6	5.5	4.2	5.1	3.8	11.6	5.5	5.7
Trio	5.7	4.0	13.3	5.2	4.9	5.3	3.7	13.5	5.9	5.8
Prisma	5.7	4.0	12.1	5.3	4.2	5.3	3.6	13.2	5.9	5.2
Moya	9.6	3.8	10.9	5.8	4.6	9.1	3.1	12.7	6.2	5.8

Table 2: Contrast-to-noise ratios (CNR) between different combinations of adjacent tissues. The cells highlighted in green are the highest CNR across scanners for that combination of tissues.

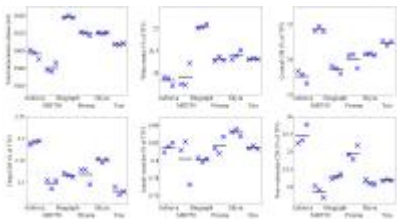


Fig. 2: Volume plots for the six segmented areas for subject 1. The blue crosses indicate the three scans per scanner, where the first two are the same-day test-retest and the third always the scan on another day. The black line indicates the average per scanner.

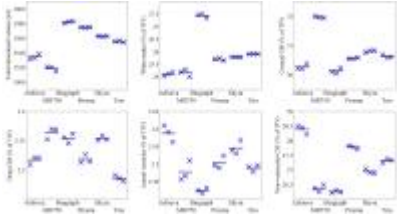


Fig. 3: Volume plots for the six segmented areas for subject 2. The blue crosses indicate the three scans per scanner, where the first two are the same-day test-retest and the third always the scan on another day. The black line indicates the average per scanner.