Evaluation of Brain Grey-White Ratios Using Automated Tissue Segmentation Packages

Kirt A. Schaper[1], Timothy R. Jarvis[1], Kristi Boesen[1], Joseph Gati[3], Ravi Menon[3], David A. Rottenberg[1,2]
Departments of [1]Neurology and [2]Radiology, UMN, USA and [3]Robarts Research Institute, London, ON

Introduction

Segmentation of MRI brain volumes into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) is an integral part of many analysis techniques including cortical surface extraction, the determination of cortical thickness and substructure volumes, inter-subject and cross-modality registration, and longitudinal studies of cerebral atrophy.

Historically quantitative measurements of the grey-white ratio (GWR) employed weighing paper cutouts of tissue boundaries [1] and *in vivo* xenon washout [2]. Given the importance of the GWR, we evaluated the performance of six brain-tissue segmentation algorithms with respect to the absolute value and reproducibility of the computed whole-brain GWR. We also examined the differences between "hard" segmentations (tissue labels assigned to each voxel) and "soft" segmentations (percentage of each tissue component defined at each voxel)

Methods

Six repeat T1-weighted MRI brain scans were acquired from six normal subjects on a 4T Siemens scanner. During the first scanning session three sequential 1xNEX scans at 1.0 mm^3 resolution and one 1xNEX scan at 1.5 mm^3 resolution were obtained; a single 1xNEX scan was acquired during subsequent sessions for a total of eight 1.0 mm^3 scans. Additionally, two 1xNEX volumes were acquired from each subject during six sessions on a 1.5T GE scanner.

For each subject the three 1xNEX scans acquired during the first scanning session were aligned and combined into three 2xNEX volumes and a single 3xNEX volume. Each scan was corrected for intensity nonuniformity using N3 [3], and non-brain tissues were removed using BET2 [4]. Six brain-tissue segmentation algorithms were applied to each MRI volume; four produced hard segmentations (FAST [5], INSECT [6], SEGM [7], and FANTASM [8]); four produced soft segmentations (FAST [5], PVS [9], SPM [10], and FANTASM [8]). For hard tissue segmentations GWRs were computed as the ratio of GM-to-WM voxels; for soft segmentations the GWR was defined as the ratio of the integral of fractional GM and WM volumes.

Results

Figure 1 illustrates the mean (filled dot) and standard deviation (SD, vertical line) of the GWRs averaged across all scans for each subject; a horizontal line indicates the average GWR across *all* subjects. The results of Hard and Soft segmentations are presented in separate panels. The absence of SDs for some algorithms (e.g., FANTASM and PVE) reflects a size smaller than the mean dot. The large SD for the PVS analysis of Subject 1 is due to a single outlier (one of the 1.5 mm^3 scans). For all algorithms the GWR for Subject 1 was markedly greater than that of the other subjects.

Two segmentation algorithms, INSECT (hard segmentation) and SPM (soft segmentation), produced GWRs that were significantly different from the GWRs produced by the other algorithms and had the largest within-subject GWR variation.

For those algorithms which produced both hard and soft segmentations (FAST and FANTASM), the hard segmentation resulted in a larger GWR.

In Figure 2 the data is broken down by volume type. For each algorithm, the mean GWR (filled dot) is plotted for each volume type, 1.5 mm^3 (4T), 1.5T (1 mm^3), and volumes with increasing NEX.

FAST (soft and hard segmentations) and SPM consistently produced greater GWRs with increasing NEX. With the exception of SEGM the GWR for the 1.5 mm^3 volumes was generally larger than that from 1.0 mm^3 volumes; INSECT and PVS produced inconsistent results across data sources.

No results are reported for SPM on the 1.5 mm^3 volumes because it failed to run on these volumes.

Figure 1. GWR Comparison Across Algorithms

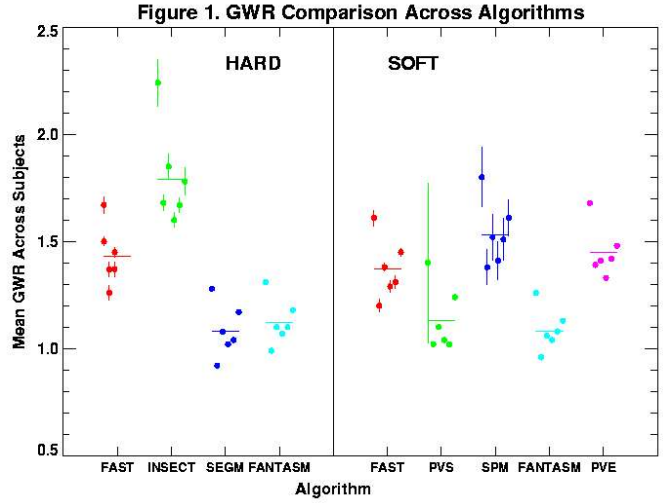


Figure 2a. GWR Sensitivity to Data Source

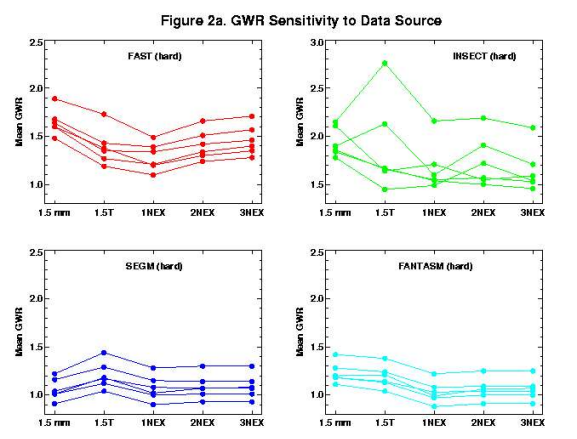
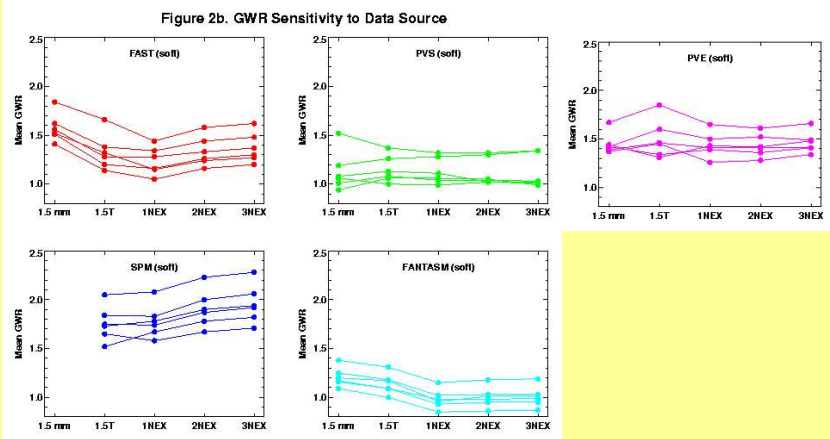


Figure 2b. GWR Sensitivity to Data Source



Discussion

FANTASM (hard and soft) and SEGM (hard) had GWRs that were most consistent with increasing NEX. PVS (soft) produced two tight clusters at high NEX which suggests an algorithmic bias rather than the expected intersubject variability.

Compared to other methods, SPM and INSECT significantly overestimated the GWR, which may be due to reliance on alignment to an anatomical template. The intra-subject variability of the GWRs produced by SPM and INSECT calls into question their suitability for longitudinal studies.

Compared to FAST, FANTASM produces GWRs that are closer to those produced by other methods, and based on this evaluation and results from reverse-segmentation (not presented here) FANTASM appears to be the algorithm of choice for brain-tissue segmentation.

References

1. Hennerberg RJ. Psychol. Neurol. 17:144. 1910. **2**. Hoedt-Rasmussen K, Skinhoj E. Neurology. 16:515-20. 1966. **3**. Sled JG, *et al*. IEEE TMI 17:87-97. 1998. **4**. Smith SM. HBM 17(3):143-155. 2002 **5**. Zhang Y, *et al*. IEEE TMI 20(1):45-57. 2001. **6**. Collins DL, *et al*. IPMI'99, LNCS 1613:210-23. 1999. **7**. Gur RC, *et al*. J Neurosci 19(10):4065-72. 1999. **8**. Pham DL, Prince JL. IEEE TMI 18(9):737-52. 1999. **9**. Shattuck DW, *et al*. NeuroImage 13(5):856-876. 2001. **10**. Ashburner J, Friston C. NeuroImage 11:805-821. 2000.

We want to thank Dzung Pham for the FANTASM software and David Shattuck for PVS.

This work is supported in part by NIH Grant P20 EB02013