Published in final edited form as:
J Comput Assist Tomogr. 2013 ; 37(2): . doi:10.1097/RCT.0b013e31828004ea.

Assessing the Performance of Atlas-Based Prefrontal Brain Parcellation in an Ageing Cohort

B.S. Aribisala, PhD[1,2,6,*], **S.R. Cox, MSc**[1,2,3,*], **K.J. Ferguson, PhD**[2,6], **S.E. MacPherson, PhD**[2,3], **A.M.J. MacLullich, PhD**[2,4,5], **N.A. Royle, MSc**[1,2,6], **M.C. Valdés Hernández, PhD**[1,2,6], **M.E. Bastin, DPhil**[1,2,6], **I.J. Deary, PhD**[2,3], and **J.M. Wardlaw, MD**[1,2,6]

[1]Brain Research Imaging Centre, University of Edinburgh, UK

[2]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, UK

[3]Department of Psychology, University of Edinburgh, UK

[4]Endocrinology Unit, University of Edinburgh, UK

[5]Geriatric Medicine, University of Edinburgh, UK

[6]Scottish Imaging Network, a Platform for Scientific Excellence (SINAPSE) Collaboration, Edinburgh, UK

Abstract

Objective—It is unclear whether atlas-based parcellation is suitable in ageing cohorts because age-related brain changes confound the performance of automatic methods. We assessed atlas-based parcellation of the prefrontal lobe in an ageing population using visual assessment, volumetric and spatial concordance.

Methods—We used atlas-based approach to parcellate brain MR images of 90 non-demented healthy adults, aged 72.7±0.7yrs and assed performance.

Results—Volumetric assessment showed that both single- and multi-atlas-based methods performed acceptably (Intraclass correlation coefficient, ICC:0.74 to 0.76). Spatial overlap measurements showed that multi- (Dice Coefficient, DC:0.84) offered an improvement over the single- (DC:0.75 to 0.78) atlas approach. Visual assessment also showed that multi-atlas outperformed single-atlas, and identified an additional post-processing step of CSF removal, enhancing concordance (ICC:0.86, DC:0.89).

Conclusions—Atlas-based parcellation performed reasonably well in the ageing population. Rigorous performance assessement aided method refinement, and emphasises the importance of age-matching and post-processing. Further work is required in more varied subjects.

Keywords

Prefrontal brain; multi-atlas; MRI; Ageing; skull thickening; brain atrophy

Author for Correspondence: Benjamin Aribisala, Brain Research Imaging Centre Department of Clinical Neurosciences, University of Edinburgh Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, United Kingdom Tel: +44 131 537 3093, Fax: +44 131 537 2661 benjamin.aribisala@ed.ac.uk.

*Joint First Authors

Europe PMC Funders Author Manuscripts

1. Introduction

Image segmentation plays an important role in region-based analysis of structural MR neuroimaging data, and can be broadly categorised into two types of approach. The first method is manual delineation of a particular anatomical structure or the 'region of interest' (ROI) drawn on the image to be analyzed [1]. This commonly-used approach is the reference standard, but it is time consuming because it requires ROIs to be drawn on every scan individually and it is also prone to user bias. The second approach is to use automatic parcellation methods based on image segmentation algorithms. This approach tends to require less user input, fewer person-hours, and is less susceptible to non-systematic bias. Automatic approaches can be summarised into four groups, namely atlas-based [2-5], supervised learning techniques [6], shape and appearance model approaches [7, 8] and energy-based techniques [9, 10]. Of these, atlas-based methods have been shown to perform the best [11-14], and have been proposed as the standard paradigm for exploiting spatial prior knowledge in brain MR image segmentation [1].

The atlas-based segmentation technique uses a standard anatomical atlas (e.g. Talairach or MNI atlas) to define the ROI in the atlas space or individual subject's native space, with the latter offering better accuracy due to reduction in partial volume effect errors [15]. A significant step in atlas-based methods is the registration of the atlas to each individual's native space which should potentially normalize variations in size and shape between the atlas and the individual's brain. Registration has most commonly used linear transformations [16]. However, both linear transformation and the use of a single atlas based on the brain of one individual do not adequately account for inter-subject variability in brain morphology. This results in relatively poor anatomical boundary matching. Non-linear registration has been demonstrated to improve boundary matching [17, 18]. Thus, a multi-atlas approach based on several brains rather than just one subject has been proposed [3]. The multi-atlas approach appears to outperform the single-atlas approach [1, 3, 4, 19].

Segmentation of ageing brains presents a significant challenge to any automatic segmentation technique because age-related changes [1] such as brain atrophy [20], skull thickening [21, 22], presence of white matter lesions [23] and infarcts [24] increase inter-individual variability [1]. Furthermore, the prefrontal lobe exhibits a particularly high degree of variation in sulcal folding patterns between subjects and is also highly susceptible to age-related atrophy and white matter lesions, both of which affect the performance of automated methods [25]. Combined with the considerable research attention it has received due to its involvement in complex cognition and its association with psychiatric, behavioural and neurological disorders [26], the prefrontal lobe is a highly relevant and challenging test-bed for shape-based automatic segmentation methods.

The performance of atlas-based techniques has been compared with manual methods in other studies [12, 27, 28], but most of these have used small sample sizes, typically 10 to 40 subjects [1], and analyses have been restricted to data acquired from children [12] or young adults [12, 27, 28] where age-related changes are not observable. In addition, measurements of comparison have tended to use either volumetric or spatial concordance, and rarely visual inspection. Although Desikan et al. [27] compared atlas-based methods with manual segmentation in an ageing population, their sample of 10 older adults is unlikely to be representative of the full range of premorbid and age-related structural differences, thus limiting inferential power. In addition, their comparison measure was limited to ICC consistency. To the best of our knowledge, no study has investigated the performance of atlas-based image segmentation techniques on the frontal lobes in a large sample of well-characterised older adults, using rigorous and multi-faceted comparison measures.

Here we compared the performance of single- and multi-atlas-based parcellation methods with manual segmentation using brain MRI data acquired from healthy older adults with a narrow age range (72.7 ± 0.7 years). The comparative analysis combined visual assessment, volumetric agreement using Bland-Altman and intraclass correlation analysis with measures of spatial concordance. We also compared the choice of atlas selection for both single- and multi-atlas approaches. Selection of atlases is an important step in multi-atlas-based techniques as the final outcome is significantly influenced by the choice of atlases, the number of atlases and the method for combining the atlases [3, 4, 11, 12]. These have been investigated [3, 4] and the use of image similarity metrics (e.g. normalised mutual information or correlation ratio) has been proposed as the most reliable atlas selection method. In addition, an optimum number of atlases has been suggested and classifier fusion based on majority vote rule has been shown to be an accurate atlas-combination method [4]. However, there have been no demonstrations of how these perform when applied to the ageing brain.

2. Materials and Methods

2.1 Subjects

Study data were selected from 700 members of the Lothian Birth Cohort 1936 (LBC1936; [29, 30]). The LBC1936 are surviving participants of the Scottish Mental Survey of 1947 [29, 31] living in the Lothian (Edinburgh) area of Scotland. At mean age 70 years they undertook a battery of tests including detailed cognitive and medical assessments [29, 30]. Three years later, as many of these subjects as possible underwent repeat cognitive and medical tests, and brain MRI [32]. Written informed consent was obtained from all participants under protocols approved by the Lothian (REC 07/MRE00/58) and Scottish Multicentre (MREC/01/0/56) Research Ethics Committee. Ninety males were selected based on the following criteria: not taking corticosterone medication or anti-depressants, no pathological MRI findings as identified by a consultant neuroradiologist (JMW), no severe cognitive impairment (Mini Mental Score Examination score of 24 or above) and non-depressed (Hospital Anxiety and Depression Scale – Depression score below 11).

2.2 Brain MRI Acquisition

Subjects were imaged with a GE Signa Horizon HDxt 1.5T clinical scanner (General Electric, Milwaukee, WI, USA) using a self-shielding gradient set with maximum gradient strength of 33 mT/m, and an 8-channel phased-array head coil. The imaging protocol included a coronal T_1-weighted (resolution 1 × 1 × 1.3 mm thickness), axial T_2*-weighted (1 × 1 × 2 mm thickness) and axial FLAIR (Fluid Attenuated Inversion Recovery, 1 × 1 × 4 mm thickness) whole brain scans; sequence details described in [32].

2.3 Manual Image Processing

Manual segmentation was performed using Analyze Software 8.1 (Mayo Clinic, Rochester, MN) [33]. T_1-weighted volumes were transformed so that the AC-PC line was horizontal at the midline in sagittal orientation, and the central fissure was vertical in both coronal and axial planes. Thresholding was then applied in order to remove dark grey elements such as meningeal tissue and signal noise from the image [34], which resulted in clearer grey matter-CSF boundaries.

The frontal lobe was then manually delineated on coronal slices of the transformed and thresholded image. This was achieved by first drawing a straight line between the depth of each adjacent major sulcus (superior and inferior frontal, lateral orbital, cingulate or paracingulate sulci) using a pen-driven cursor and tablet (Wacom Intuos 4, Wacom Co. Ltd., Saitama, Japan) to demarcate the internal extent of all major frontal gyri. An intensity-guided flood fill was then applied to enable automatic detection of the grey matter-CSF

boundaries. The posterior boundary of the superior frontal lobe was identified as the slice immediately anterior to the appearance of the pre-central gyrus (Kates et al., 2002; Figure 1). The selection of this boundary allows the simple and reliable identification of the frontal areas excluding pre-motor cortex using a common landmark that is easy to identify. The orbital aspect of the frontal lobe was identified using a coronal plane at the most posterior appearance of the lateral orbital sulcus, which allowed differentiation of the orbitofrontal cortex from insular cortex. Asymmetry of these boundaries between hemispheres of each individual was preserved. This process produced a library of 90 atlases with the anatomical scans.

2.4 Automated Image Processing

The brain was extracted from the surrounding tissue using a validated multispectral image processing tool, MCMxxxVI [35, 36]. This semi-automatic segmentation tool fuses pairs of MRI sequences (e.g. T_2*-weighted and FLAIR) in the red-green colour space to enhance signal differences between tissues, hence improving computational differentiation of signal differences and increasing accuracy of extraction of specific anatomical structure or ROI. T_2*-weighted and FLAIR volumes were fused to extract the brain as they provide good differentiation between brain-CSF and the inner skull table. After image fusion, the object extractor tool of Analyze 8.1 software [33] was used to extract the brain and the final result was visually inspected and manually edited to correct for any misclassification.

2.4.1 Segmentation using a Single Template—For single-atlas segmentation, we used atlases derived from four different representative brains in order to investigate how the choice of atlas selection affects performance. The first three were selected from the 90 subjects as the most representative in terms of intracranial volume, total brain tissue volume and frontal lobe volume. The fourth was a standard atlas developed from a right handed male young adult, available in MRIcro [37], included to compare the results of segmenting ageing brains using an atlas developed from a young adult with those achieved by applying an age-matched atlas. The frontal lobe volumes were computed from the manually segmented images described in Section 2.3.

Each of the four representative brains was transformed to individual subject's space using Automatic Registration Toolbox (ART, [17]), which was selected as it has previously been demonstrated as one of the most robust nonlinear image registration algorithms [38]. The registration process was in two stages. Firstly, the T_1-weighted volume from which the representative atlas was generated was registered to the T_1-weighted volume of each subject and the computed transformation matrix applied to the atlas using the nearest neighbour interpolation to preserve the binary nature of the atlas. Secondly, the atlas was applied to the T_1-weighted volume of each subject to extract the subject-specific prefrontal lobe.

2.4.2 Segmentation using Multiple Atlases—Multi-atlas segmentation consists of template selection, label propagation, label fusion and ROI definition [39]. First we used FSL (FMRIB Software Library, University of Oxford, UK) image registration tool (FLIRT [16]) to register all our library scans to each subject in a jack-knifing method [40] and computed the similarity between the source and the target image. Using this approach, each of the 90 subjects in our library became the target image and the remaining 89 were the source images. We used two image similarity metrics: the normalised cross correlation [41], and normalised mutual information [42] in order to investigate their accuracy in ageing population parcellation. For each subject, the library atlases were ranked based on the similarity metrics. The best-matched 20 atlases were selected as previously proposed [4], and were then registered to the target volume using ART [17], and combined using vote-rule-based decision

fusion at every voxel [4]. The resulting atlas was used to extract the participant's prefrontal region.

2.4.3 Comparative Analysis of Methods—The comparison between atlas-based and manual segmentation was conducted in three stages. Firstly, intra-class correlation coefficients (ICCs) [43] and Bland-Altman metrics [44] were calculated to examine the volumetric agreement between approaches. Secondly, as volumetric agreement does not give information about spatial agreement, we assessed spatial concordance using the Jaccard Index (JI) [45, 46] and Dice Coefficient (DC) [47]. The JI is defined as the ratio of the intersection of two images to their union, while the DC is the ratio of the intersection of two images to their mean value, where A and M represent automatically and manually generated segmentations respectively.

$$DC=2\left|\frac{A\cap M}{A+M}\right| \quad (1)$$

$$JI=\left|\frac{A\cap M}{A\cup M}\right| \quad (2)$$

Both measures have values ranging from 0 to 1 representing complete disjoint and complete overlap respectively. Paired t-tests were used to compare the spatial agreement between methods. Finally, results were visually assessed by inspecting consecutive sections displayed together on the same screen and three-dimensional rendering. Volumes which exhibited particularly high or low measures of spatial concordance were specially inspected in order to identify reasons for extreme values. Visual inspection typically showed that initial single- and multi-atlas outputs misclassified CSF voxels as brain matter, and that the output masks did not map well onto the gyral patterning of the target brain. In order to rectify this, the dark grey elements such as meningeal tissue and signal noise were removed from the T_1-weighted volumes using the threshold value identified during manual segmentation. The masks generated from this process were then applied to the results from single- and multi-atlas parcellation to produce CSF free prefrontal lobes.

3. Results

Table 1 presents the descriptive statistics for the study participants. Two subjects were excluded from the analysis because of registration failure, resulting in a sample of n=88 for analysis using the multi-atlas method. For the single atlas-based method, the agreement metrics for each of the study-based representative brains were excluded from the analysis, reducing the sample size to 87. Further registration failure after applying the young adult single-atlas method resulted in a cohort of n=77 for analysis using this approach.

3.1 Reproducibility of Manual Segmentation

In order to measure intra-rater reliability, 10 of the study cohort were randomly selected and segmented two weeks apart by the same rater (SRC) using the same protocol described in Section 2.3. The resultant similarity measures (ICC = 0.99, Bland-Altman mean = −0.07 %, 95% limits from 2.59 to −2.72 %; JI = 0.81, DC = 0.87) suggest that the manual method is highly reproducible.

3.2 Single-Atlas

All four representative single-atlases had comparably good spatial agreement with the manual method (range: JI = 0.61 to 0.64, DC =0.75 to 0.78), but there was significant divergence with respect to volumetric measures (Table 2, Figure 3). The three atlases selected from the target cohort gave acceptable volumetric agreement (ICCs > 0.74) with comparably wide Bland-Altman 95% confidence intervals (26.92 to 28.83 %, Table 2, Figure 2) and no apparent over- or under-estimation. However, where the young adult brain was used, it resulted in a very low ICC (0.31), systematic overestimation of frontal lobe volume (Bland-Altman mean = −18.70 %) and the widest Bland-Altman confidence interval spans of all approaches (34 %).

3.3 Multi-Atlas

Both multi-atlas methods (atlases selected using normalised mutual information and normalised correlation coefficients) initially gave highly similar volumetric (ICCs = 0.76) and spatial agreement (JI = 0.73 ± 0.02, DC = 0.84 ± 0.02) with the manual segmentation (Table 2, Figure 3). Although the volumetric agreement did not appear to improve with the use of multi-compared to single-atlases, the spatial agreement was significantly better, with JI improving by 0.09 and DC by 0.06; differences computed using atlases selected based on total brain tissue volume and normalised mutual information (paired t-test, $p < 0.001$). Removal of misclassified CSF voxels in a post-processing step identified via visual assessment further improved the performance of both multi-atlas methods (Figure 2, Table 3). Specifically, ICC agreement improved by 0.11, and there was a reduction in Bland-Altman confidence intervals of 6.94 (normalised correlation) and 8.97 % (mutual information). Furthermore, spatial overlap was significantly improved for both the JI (0.08) and DC (0.05) (paired t-test, $p < 0.001$). The distribution of both measures (Figure 3) also showed the multi-atlas based method performed better than the single-atlas method.

3.3.3 Visual Assessment—Preliminary visual assessment confirmed that boundary matching and grey matter-CSF boundaries were more accurately represented using the multi-atlas approach (the manual, single- and multi-atlas mask comparison is illustrated in Figure 4). The output masks from the single-atlas method did not map well onto the gyral patterning of the target brain. For the multi-atlas method, prior to the removal of CSF, estimation errors were primarily at medial and dorsolateral extents, and the posterior lobar boundary was often imprecise. This was particularly apparent amongst individuals who exhibited a higher degree of atrophy, although brain shape that deviates significantly from the atlas may also be a confounding factor for non-linear registration. However, introducing the post-processing step of CSF removal showed a marked improvement in the concordance between multi-atlas and manual methods. Following this step, variability between automated mask and reference was mainly confined to the posterior boundary, with medial and dorsal over-estimation showing improvement (Figure 4).

4. Discussion

We have presented an investigation of the performance of atlas-based brain parcellation techniques in an ageing population. As an exemplar, we parcellated the prefrontal lobe using both single- and multi-atlas-based approaches and compared their performance with that of manual parcellation. We also investigated the choice of atlas selection and their effects on parcellation accuracy. We found that, whereas single-atlas measures were inconsistent in terms of volumetric and spatial indices of comparison with the reference standard, both methods of multi-atlas parcellation gave consistently good results. Multi-atlas performed better than single-atlas on the majority of similarity measures, particularly those that consider spatial concordance. We also found that the additional post-processing step of CSF

removal improved the performance of the atlas-based methods. The spatial agreement measures show that the performance of multi-atlas methods in this ageing population compare very well with the implementation in young adults [3, 4] and with the state-of-the-art segmentation techniques (e.g. [2]) applied to young adults; therefore the current approach shows promise.

Our investigation on the choice of atlas selection revealed that the targeted selection of a study-based atlas for single-atlas methods of segmentation is of methodological importance. Using the atlas of a young healthy male resulted in much poorer performance than did the use of a study-based, older adult atlas. This supports an earlier finding [4] that application of an age-based atlas improves accuracy. Basing the single-atlas selection on average intracranial volume, brain tissue volume or frontal lobe size all seem to confer similar improvements on overall output but the selection based on the total brain tissue volume and frontal lobe volume gave the best performance. Both the multi-atlas methods showed an improvement in the spatial measures of similarity with the reference manual segmentation compared to the single-atlas approach. The multi-atlas method should now be tested in a much larger and more varied population. The current participants were comparatively healthy and cannot fully represent the extent of normal or pathological variation seen with advancing age. Any such approach should be used with caution until wider knowledge of the limitations is available.

On a further methodological note, we emphasise the importance of rigorous comparisons between automated and reference outputs. It is clear from the inconsistencies across measurement modalities that agreements in volume do not necessarily equate to accuracy; here the spatial improvement conferred by multi-atlas methods, as identified by both computational and visual assessment, is not reflected in the volumetric comparison. Furthermore, visual assessment is a crucial tool in the identification of errors and potential improvements; the potential to make significant improvements to the most promising methods by removing misclassified CSF voxels was only identified through visual comparison of the outputs. This implies that the performance of any automatic parcellation technique should not only be measured using ICC but also with spatial concordance and visual assessment by a trained rater.

Nevertheless, even when we used robust criteria for atlas selection, chose the best registration method and a large ROI, variance in brain morphology and particularly the effects of age-related atrophy continue to present challenges to purely automated methods. The implementation of post-processing largely rectified these discrepancies, although issues with posterior boundary positioning remained. An alternative approach is to combine automated methods with manual editing to correct discrepancies. This hybrid approach might confer advantages over automated methods in terms of accuracy, and at the same time reduce the amount of researcher-hours that are required for a purely manual method. Hence, future use of automatic parcellation methods could usefully be followed by visual assessment, cogent post-processing strategies and manual editing.

In addition to our robust assessments of performance, this study benefits from a large sample and narrow age range, making this the largest study to have investigated atlas-based parcellation performance in an ageing population. We also used a state-of-the-art non-linear registration tool and a well-validated semi-automatic tool for brain extraction. The present study is limited in that only healthy older male participants were used. We therefore do not offer an assessment of performance for healthy ageing female brains, or age-related degenerative disorders. There is increased skull-thickening in post-menopausal women, and this could affect the performance of the atlas-based method. Likewise, significantly higher degrees of atrophy in pathological ageing are likely to present increased, and possibly novel,

challenges to the present method. Further work could usefully test the effects of skull-thickening on automated parcellation and amongst subjects with significantly more brain atrophy than healthy cohorts, such as dementia. Moreover, this method has only been applied to a large cortical region that is predominantly bound by CSF. Future studies will assess performance on other brain regions, particularly sub-cortical areas that may not benefit from the post-processing step used here.

In conclusion, atlas-based parcellation methods performed reasonably well in our generally healthy ageing population, with the multi-atlas method performed better than the single-atlas, although this was only apparent using visual inspection and measures of spatial concordance. The performance of any parcellation scheme should be assessed by a robust and multi-faceted series of measures. However, brain shape and particularly the effects of age-related atrophy still challenge the performance of this, and other, computational methods. Visual assessment also allowed a targeted post-processing method to be identified and implemented, further augmenting the agreement of multi-atlas parcellation across all measures. We suggest that a high degree of volumetric and spatial concordance can be achieved when this method is combined with minimal manual editing, and may be a promising method for lobar parcellation in large ageing cohorts.

Acknowledgments

This work was funded by Age UK and the UK Medical Research Council as part of the Disconnected Mind (http://www.disconnectedmind.ed.ac.uk), The Centre for Cognitive Aging and Cognitive Epidemiology (CCACE; http://www.ccace.ed.ac.uk), The Row Fogo Charitable Trust and the Scottish Founding Council through the SINAPSE collaboration (http://www.sinapse.ac.uk). Funding (for CCACE; G0700704/84698) from the BBSRC, EPSRC, ESRC and MRC is gratefully acknowledged. The imaging was performed in the Brain Research Imaging Centre, University of Edinburgh (http://www.bric.ed.ac.uk), a SINAPSE Centre.

Conflicts of Interest and Sources of Funding:

The study was funded by: Medical Research Council (MRC), Row Fogo Charitable Trust, Wellcome Trust (088134), Age UK's The Disconnected Mind, Scottish Funding Council (SFC) through the SINAPSE Collaboration (Scottish Imaging Network. A Platform for Scientific Excellence, http://www.sinapse.ac.uk/), Centre for Cognitive Ageing and Cognitive Epidemiology, Lifelong Health and Wellbeing Initiative (G0700704/84698), Biotechnology and Biological Sciences Research Council (BBSRC), Engineering and Physical Sciences Research Council (EPSRC) and, Economic and Social Research Council (ESRC).

References

1. Cabezas M, Oliver A, Llado X, et al. A review of atlas-based segmentation for magnetic resonance brain images. Computer methods and programs in biomedicine. 2011; 104:e158–77. [PubMed: 21871688]
2. Fischl B, Van Der Kouwe A, Destrieux C, et al. Automatically parcellating the human cerebral cortex. Cerebral Cortex. 2004; 14:11–22. [PubMed: 14654453]
3. Heckemann RA, Keihaninejad S, Aljabar P, et al. Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation. Neuroimage. 2010; 51:221–227. [PubMed: 20114079]
4. Aljabar P, Heckemann RA, Hammers A, et al. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. Neuroimage. 2009; 46:726–738. [PubMed: 19245840]
5. Ashburner J, Friston K. Unified Segmentation. Neuroimage. 2005; 26:839. [PubMed: 15955494]
6. Moghaddam, MJ.; Soltanian-Zadeh, H. Information Processing in Medical Imaging, Proceedings. Prince, JLPDLMKJ., editor. 2009. p. 326-337.
7. Babalola, KO.; Cootes, TF.; Twining, CJ., et al. 3D brain segmentation using active appearance models and local regressors. Medical image computing and computer-assisted intervention : MICCAI; International Conference on Medical Image Computing and Computer-Assisted Intervention; 2008. p. 401-8.

8. Kelemen A, Szekely G, Gerig G. Elastic model-based segmentation of 3-D neuroradiological data sets. Ieee Transactions on Medical Imaging. 1999; 18:828–839. [PubMed: 10628943]
9. Leventon, ME.; Faugeras, O.; Grimson, WEL., et al. Level set based segmentation with intensity and curvature priors. 2000.
10. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. Neuroimage. 2006; 31:1116–1128. [PubMed: 16545965]
11. Ginneken, B.; Heimann, T.; Styner, M. 3D segmentation in the clinic: A grand challenge. MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge; 2007.
12. Igual L, Soliva JC, Hernandez-Vela A, et al. A fully-automatic caudate nucleus segmentation of brain MRI: Application in volumetric analysis of pediatric attention-deficit/hyperactivity disorder. Biomedical Engineering Online. 2011; 10
13. Babalola KO, Patenaude B, Aljabar P, et al. An evaluation of four automatic methods of segmenting the subcortical structures in the brain. Neuroimage. 2009; 47(4):1435–1447. [PubMed: 19463960]
14. Babalola, KO.; Patenaude, B.; Aljabar, P., et al. Comparison and evaluation of segmentation techniques for subcortical structures in brain MRI. Medical image computing and computer-assisted intervention : MICCAI; International Conference on Medical Image Computing and Computer-Assisted Intervention; 2008. p. 409-16.
15. Aribisala BS, He J, Blamire AM. Comparative Study of Standard Space and Real Space Analysis of Quantitative MR Brain Data. Journal of Magnetic Resonance Imaging. 2011; 33:1503–1509. [PubMed: 21591021]
16. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. Medical Image Analysis. 2001; 5:143–156. [PubMed: 11516708]
17. Ardekani BA, Guckemus S, Bachman A, et al. Quantitative comparison of algorithms for inter-subject registration of 3D volumetric brain MRI scans. Journal of Neuroscience Methods. 2005; 142:67–76. [PubMed: 15652618]
18. Avants BB, Tustison NJ, Song G, et al. A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage. 2011; 54:2033–2044. [PubMed: 20851191]
19. Rohlfing T, Brandt R, Menzel R, et al. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocol microscopy images of bee brains. Neuroimage. 2004; 21:1428–1442. [PubMed: 15050568]
20. Appelman APA, Exalto LG, Van Der Graaf Y, et al. White Matter Lesions and Brain Atrophy: More than Shared Risk Factors? A Systematic Review. Cerebrovascular Diseases. 2009; 28:227–242. [PubMed: 19571536]
21. Finby N, Kraft E. The aging skull: comparative roentgen study. 25 to 34 year interval. Clinical radiology. 1972; 23:410–4. [PubMed: 5083462]
22. May H, Peled N, Dar G, et al. Hyperostosis Frontalis Interna: What Does it Tell Us About our Health? American Journal of Human Biology. 2011; 23:392–397. [PubMed: 21387460]
23. Debette S, Markus HS. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. BMJ. 2010; 341
24. Appelman APA, Vincken KL, Van Der Graaf Y, et al. White Matter Lesions and Lacunar Infarcts Are Independently and Differently Associated with Brain Atrophy: The SMART-MR Study. Cerebrovascular Diseases. 2010; 29:28–35. [PubMed: 19893309]
25. Raz N, Lindenberger L, Rodrigue KM, et al. Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. Cerebral Cortex. 2005; 15(11):676–1689.
26. Steele JD, Lawrie SM. Segregation of cognitive and emotional function in the prefrontal cortex: a stereotactic meta-analysis. NeuroImage. 2004; 21:868–875. [PubMed: 15006653]
27. Desikan RS, Segonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage. 2006; 31:968–980. [PubMed: 16530430]
28. Destrieux C, Fischl B, Dale A, et al. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage. 2010; 53:1–15. [PubMed: 20547229]

29. Deary IJ, Gow AJ, Taylor MD, et al. The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. BMC Geriatr. 2007; 7:28. [PubMed: 18053258]
30. Deary IJ, Gow AJ, Pattie A, et al. Cohort profile: The Lothian Birth Cohorts of 1921 and 1936. International Journal of Epidemiology. 2012 In press.
31. Scottish Council for Research in Education. The Trend of Scottish Cognitive Ability: A Comparison of the 1947 and 1932 Surveys of the Cognitive Ability of Eleven-Year-Old Pupils. University Publishing Group; London: 1949.
32. Wardlaw JM, Bastin ME, Hernandez MCV, et al. Brain aging, cognition in youth and old age and vascular disease in the Lothian Birth Cohort 1936: rationale, design and methodology of the imaging protocol. International Journal of Stroke. 2011; 6(6):547–559. [PubMed: 22111801]
33. Mayo, C. Analyze 8.1. AnalyzeDirect, Inc. ; 2008. http://www.analyzedirect.com/Analyze/
34. Ferguson, KJ.; Wardlaw, JM. [Accessed 11th July 2012] Brain volume measurement in Analyze 8. 2012. http://www.sbirc.ed.ac.uk/documents/brain%20volume%20measurement%20in%20analyze%208.pdf
35. Hernández MDV, Ferguson KJ, Chappell FM, et al. New multispectral MRI data fusion technique for white matter lesion segmentation: method and comparison with thresholding in FLAIR images. European Radiology. 2010; 20:1684–1691. [PubMed: 20157814]
36. Hernández MC, Royle NA, Jackson NA, et al. Color Fusion of Magnetic Resonance Images Improves Intracrania Volume Measurement in Studies of Ageing. Open Journal of Radiology. 2012; 2:1–9.
37. Rorden C, Brett M. Stereotaxic display of brain lesions. Behavioural Neurology. 2000; 12:191–200. [PubMed: 11568431]
38. Klein A, Andersson J, Ardekani BA, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage. 2009; 46:786–802. [PubMed: 19195496]
39. Leung KK, Barnes J, Modat M, et al. Brain MAPS: An automated, accurate and robust brain extraction technique using a template library. Neuroimage. 2011; 55:1091–1108. [PubMed: 21195780]
40. Efron, B.; Tibshirani, RJ. An Introduction to Boostrap. Chapman and Hall; London: 1993.
41. Collignon, A.; Maes, F.; Delaere, D., et al. Automated multi-modality image registration based on information theory. 1995.
42. Studholme C, Hill DLG, Hawkes DJ. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition. 1999; 32:71–86.
43. Shrout PE, Fleiss JL. Intraclass correlations - used in assessing rater reliability. Psychological Bulletin. 1979; 86:420–428. [PubMed: 18839484]
44. Bland JM, Altman DG. Statistical Methods for Assessing Agreement between 2 methods of clinical measurement. Lancet. 1986; 1:307–310. [PubMed: 2868172]
45. Gee JC, Reivich M, Bajcsy R. Elastically deforming 3D Atlas to match anatomical brain images. Journal of Computer Assisted Tomography. 1993; 17:225–236. [PubMed: 8454749]
46. Jaccard P. The distribution of flora in the alpine zone. New Phytol. 1912; 11:37–50.
47. Zijdenbos AP, Dawant BM, Margolin RA, et al. Morphometric analysis of white-matter lesions in MR-images - method and validation. IEEE Transactions on Medical Imaging. 1994; 13:716–724. [PubMed: 18218550]

Europe PMC Funders Author Manuscripts

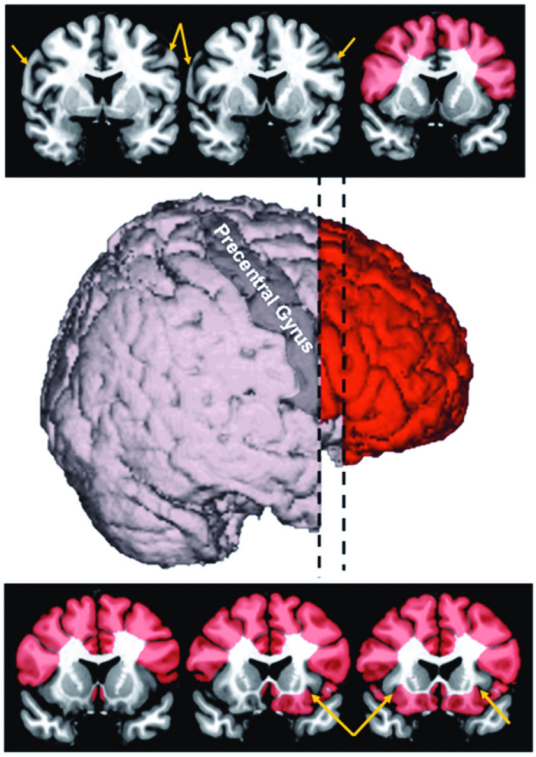

Figure 1. Manual definition of the ROI using coronal slices (left to right: posterior to anterior). Top row shows the identification of the most posterior coronal slice before the appearance of the precentral gyrus (PrCG; after Kates et al., 2002; PrCG indicated with arrows). Bottom row shows identification of the most posterior appearance of the lateral orbital sulcus (LOS; indicated with arrows). Middle image shows a 3D rendering of the prefrontal ROI in red. Dotted line indicates the anterior-most extent of the PrCG (left) and the most posterior coronal appearance of the LOS (right).

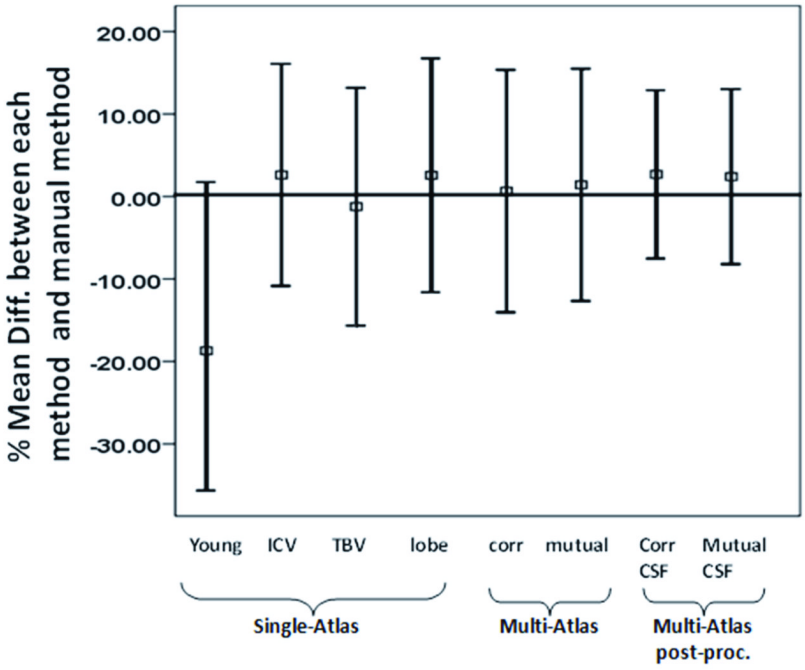

Figure 2. Bland-Altman analysis showing the variation of % mean difference against mean values between automated single-atlas and manual segmentation.

Tails denote upper and lower confidence limits. *Single-Atlas:* Young, ICV, TBV and lobe denote the representative brains selected based on young adult brain, ICV, total brain tissue volume and frontal lobe volume respectively. Multi-atlas: corr, mutual, corr CSF and mutual CSF denote selections of atlases using normalised correlation coefficient, normalised mutual information, normalised correlation coefficient followed by CSF removal and normalised mutual information followed by CSF removal respectively.

Europe PMC Funders Author Manuscripts

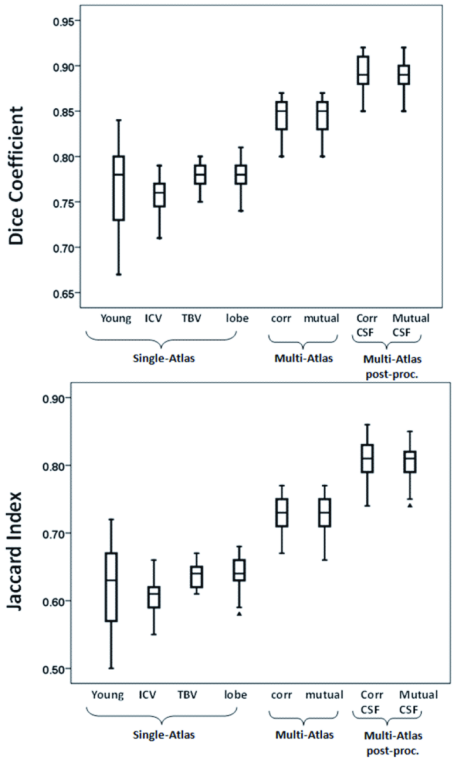

Figure 3. Boxplots of measures of agreements between manual and automatic atlas-based segmentation methods.

Single-Atlas: Young, ICV, TBV and lobe denote the representative brains selected based on young adult brain, ICV, total brain tissue volume and frontal lobe volume respectively. Multi-atlas: corr, mutual, corr CSF and mutual CSF denote selections of atlases using normalised correlation coefficient, normalised mutual information, normalised correlation coefficient followed by CSF removal and normalised mutual information followed by CSF removal respectively.

Europe PMC Funders Author Manuscripts

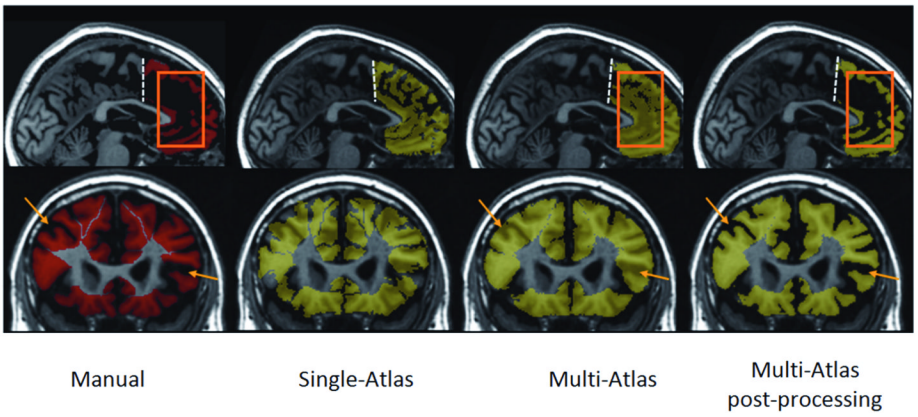

Figure 4. Sagittal (top row) and coronal (bottom row) planes for comparison of parcellation method on a representative brain from the cohort.

From left to right, manual, single atlas, multi-atlas, multi-atlas after post-processing to remove CSF. For single atlas, local patterns of gyrification were not well-matched. For multi-atlas, prior to CSF removal, voxels in lateral (orange arrows) and medial aspects (orange box) were classified as brain tissue. Broken white lines indicate posterior boundary. Single-atlas method shown was when representative brain was selected based on the total brain tissue volume. Multi-atlas method shown was when normalised correlation coefficient was used to select atlases.

Europe PMC Funders Author Manuscripts

Table 1

Descriptive Statistics

	Mean	SD	Range
Age at MRI (yrs)	73.30	0.37	72.41-74.22
MMSE	28.54	1.52	24-30
HADS-A	3.97	2.71	0-10
HADS-D	2.74	2.33	0-10

MMSE = Mini Mental State Exam (max=30), HADS-A = Hospital Anxiety and Depression Scale – Anxiety Subscale (max=21), HADS-D = Hospital Anxiety and Depression Scale – Depression Subscale (max=21).

Table 2

Volumetric and spatial agreement measures between manual and automatic parcellation.

Volumetric Comparison				
		ICC	**Bland-Altman**	
			Mean (%)	**Confidence Limits (%)**
Manual	Reference Standard	0.99	0.07	2.59 to −2.72
Single-atlas	Atlas selected based on ICV (N=87)	0.76	2.61	16.07 to −10.85
	Atlas selected based on TBV (N=87)	0.76	−1.23	13.17 to −15.66
	Atlas selected based on frontal lobe volume (N=87)	0.74	2.56	16.73 to −11.62
	Standard Atlas (N=77)	0.31	−18.70	1.74 to −35.65
Multi-atlas	Atlases selected based on correlation coefficient (N=88)	0.75	0.65	15.34 to −14.04
	Atlases selected based on mutual information (N=88)	0.75	1.41	15.49 to −12.67
Spatial Comparison				
		Frontal lobe Volume	Jaccard Index	Dice Coefficient
Manual	Reference Standard	193888 ± 23250	0.81±0.24	0.87±0.22
Single-atlas	Atlas selected based on ICV (N=87)	188330 ± 17361	0.61 ± 0.02	0.75 ± 0.02
	Atlas selected based on TBV (N=87)	195680 ± 18208	0.64 ± 0.02	0.78 ± 0.02
	Atlas selected based on frontal lobe volume (N=87)	188388 ± 17147	0.64 ± 0.02	0.78 ± 0.01
	Young Adult Standard Atlas (N=77)	235242 ± 27553	0.62 ± 0.06	0.76 ± 0.05
Multi-atlas	Atlases selected based on correlation coefficient (N=88)	211559 ± 19335	0.73 ± 0.03	0.84 ± 0.02
	Atlases selected based on mutual information (N=88)	209965 ± 18596	0.73 ± 0.03	0.84 ± 0.02

ICV = intracranial volume

TBV = total brain tissue volume

Europe PMC Funders Author Manuscripts

Table 3

Change in volumetric and spatial agreement measures for multi-atlas methods before and after CSF post-processing.

Volumetric Comparison				
		ICC	**Bland-Altman**	
		ICC	**Mean (%)**	**Confidence Limits (%)**
Initial	Atlases selected based on correlation coefficient (N=88)	0.75	0.65	15.34 to −14.04
	Atlases selected based on mutual information (N=88)	0.75	1.41	15.49 to −12.67
CSF Removed	Atlases selected based on correlation coefficient (N=88)	0.86	2.68	12.88 to −7.53
	Atlases selected based on mutual information (N=88)	0.86	2.40	13.01 to −8.21
Spatial Comparison				
		Volume	Jaccard Index	Dice Coefficient
Initial	Atlases selected based on correlation coefficient (N=88)	211559 ± 19335	0.73 ± 0.03	0.84 ± 0.02
	Atlases selected based on mutual information (N=88)	209965 ± 18596	0.73 ± 0.03	0.84 ± 0.02
CSF Removed	Atlases selected based on correlation coefficient (N=88)	189010 ± 19846	0.81 ± 0.03	0.89 ± 0.02
	Atlases selected based on mutual information (N=88)	188129 ± 19334	0.81 ± 0.03	0.89 ± 0.02