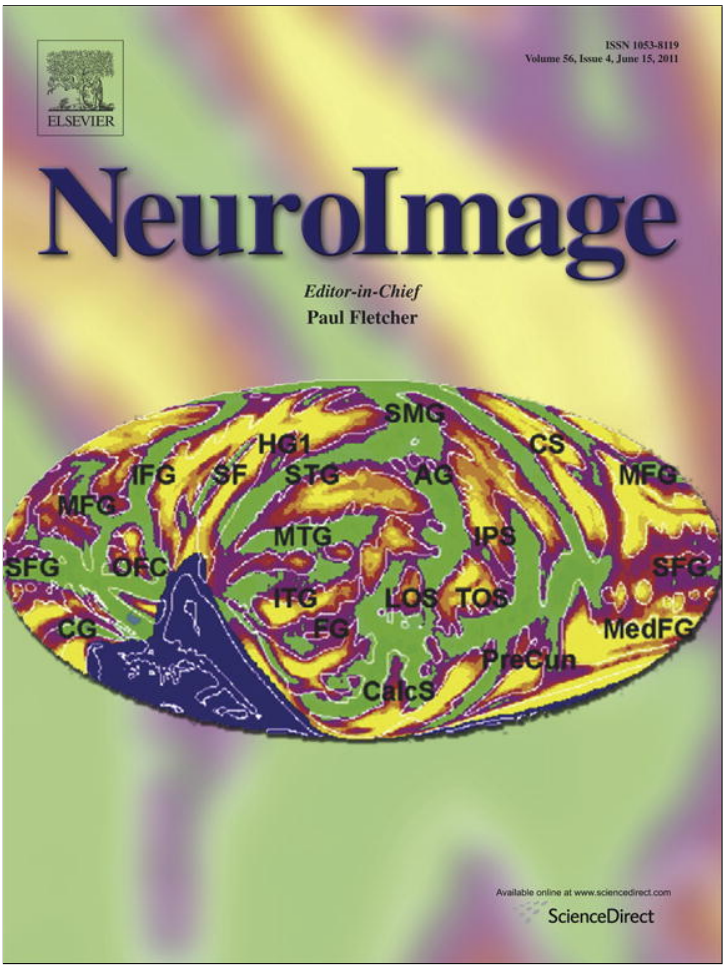
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

NeuroImage 56 (2011) 2024–2037

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Automatic morphometry in Alzheimer's disease and mild cognitive impairment ☆,☆☆

Rolf A. Heckemann [a,b,*,1], Shiva Keihaninejad [b,c], Paul Aljabar [d], Katherine R. Gray [d], Casper Nielsen [c], Daniel Rueckert [d], Joseph V. Hajnal [e], Alexander Hammers [a,b] and The Alzheimer's Disease Neuroimaging Initiative [2]

[a] The Neurodis Foundation (Fondation Neurodis), Lyon, France
[b] Centre for Neuroscience, Division of Experimental Medicine, Department of Medicine, Imperial College, London, UK
[c] Dementia Research Centre, UCL Institute of Neurology, London, UK
[d] Department of Computing, Imperial College, London, UK
[e] Imaging Sciences Department, MRC Clinical Science Centre, Imperial College, London, UK

ARTICLE INFO

Article history:
Received 21 November 2010
Revised 1 March 2011
Accepted 4 March 2011
Available online 11 March 2011

Keywords:
Image segmentation
Brain pathology
Brain atlas
Magnetic resonance imaging
Alzheimer's disease
Mild cognitive impairment

ABSTRACT

This paper presents a novel, publicly available repository of anatomically segmented brain images of healthy subjects as well as patients with mild cognitive impairment and Alzheimer's disease. The underlying magnetic resonance images have been obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. T1-weighted screening and baseline images (1.5 T and 3 T) have been processed with the multi-atlas based MAPER procedure, resulting in labels for 83 regions covering the whole brain in 816 subjects. Selected segmentations were subjected to visual assessment. The segmentations are self-consistent, as evidenced by strong agreement between segmentations of paired images acquired at different field strengths (Jaccard coefficient: 0.802 ± 0.0146). Morphometric comparisons between diagnostic groups (normal; stable mild cognitive impairment; mild cognitive impairment with progression to Alzheimer's disease; Alzheimer's disease) showed highly significant group differences for individual regions, the majority of which were located in the temporal lobe. Additionally, significant effects were seen in the parietal lobe. Increased left/right asymmetry was found in posterior cortical regions. An automatically derived white-matter hypointensities index was found to be a suitable means of quantifying white-matter disease. This repository of segmentations is a potentially valuable resource to researchers working with ADNI data.

© 2011 Elsevier Inc. All rights reserved.

Introduction

This paper presents results of a project that aims to provide anatomical labels based on automatic segmentation for magnetic resonance (MR) brain imaging data supplied by the Alzheimer's Disease Neuroimaging Initiative (ADNI). The result of this work is made available to the general scientific community via the same channels as the source ADNI data.

Anatomical segmentations of structural images of the human brain can be used for a plethora of purposes. A principal motivation is to understand the impact of neurodegeneration, trauma, epilepsy and other conditions on the brain's macroscopic structure. Such understanding leads to morphometric descriptors with the potential to serve as biomarkers for the diagnosis and monitoring of brain disease (Colliot et al., 2008; Duchesne et al., 2008; Heckemann et al., 2008; Klöppel et al., 2008). Beyond the realm of morphometric analysis, individual anatomical segmentation is frequently used in the analysis of functional imaging data, e.g. to precisely locate areas of hypo- or hypermetabolism within the subject's own anatomical reference frame. Anatomical segmentation also enables studies of

☆ The authors wish to thank Marc Modat, UCL, London for advice and support regarding the use of the Nifty Reg software.

☆☆ The Foundation for the National Institutes of Health (www.fnih.org) coordinates the private sector participation of the $60 million ADNI public–private partnership that was begun by the National Institute on Aging (NIA) and supported by the National Institutes of Health. To date, more than $27 million has been provided to the Foundation for NIH by Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson & Johnson, Eli Lilly and Co., Merck & Co., Inc., Novartis AG, Pfizer Inc., F. Hoffmann-La Roche, Schering-Plough, Synarc Inc., and Wyeth, as well as non-profit partners the Alzheimer's Association and the Institute for the Study of Aging.

* Corresponding author at: The Neurodis Foundation (Fondation Neurodis), Lyon, France.
E-mail address: soundray@imperial.ac.uk (R.A. Heckemann).

[1] RAH was supported by a research grant from the Dunhill Medical Trust, UK.

[2] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A listing of ADNI investigators is available at http://www.loni.ucla.edu/ADNI/Collaboration/ADNI_Manuscript_Citations.pdf.

1053-8119/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2011.03.014

Author's personal copy

regional connectivity based on diffusion tensor imaging [e.g. Traynor et al. (2010)].

ADNI MR imaging data have hitherto been provided with only minimal amounts of segmentation information. For a subset of ADNI images, labels of the left and right hippocampi are available. These labels have been generated using a semiautomatic tool (SNT, Medtronic Surgical Navigation Technologies, Louisville, CO) that relies on manual seed point placement. In work by Hsu et al. (2002), the SNT tool was claimed to yield hippocampal volume measurements equivalent to a manual delineation protocol, but the validation was not entirely convincing: Hsu et al. make reference to previous work by Watson et al. (1992), but the protocol described there finds distinctly larger volumes in normal adult hippocampi (Watson — right: 5264 ± 652 mm^3, left: 4903 ± 684 mm^3; Hsu — right: 3103 ± 505 mm^3, left: 2945 ± 503 mm^3). Furthermore, the SNT method yields volume measurements that are yet smaller than those of the manual reference (right: 2323 ± 326 mm^3, left: 2275 ± 253 mm^3). Both the validation and anatomical coverage of available ADNI segmentation data are thus limited.

Beyond the hippocampus, researchers requiring anatomical labels of ADNI data have three choices:

1. Normalize subject images to a reference space and apply one of a choice of anatomical volume or surface atlases available for this space [e.g. Talairach (Talairach and Tournoux, 1988), AAL (Tzourio-Mazoyer, 2002), Maximum Probability Brain Atlas (Hammers et al., 2003), The Whole Brain Atlas,[3] LPBA40 (Shattuck et al., 2008), PALS-B12 (Van Essen, 2005), the Freesurfer atlas Fischl et al. (2004), or a purpose-made atlas]. This can be a simple solution, in particular if other parts of the analysis already require spatial normalization. Since the segmentation process takes place in the common space, an inverse normalization has to be carried out in order to recover the volume and shape of segmented regions in native space. This approach is typically based on a single-subject atlas or maximum-probability atlas. The latter are generally preferable because they tend to eliminate idiosyncrasies due to anatomical variants in individual subjects. Success depends on the suitability of the chosen atlas, as well as the suitability and robustness of the chosen spatial normalization algorithm.
2. Carry out anatomical segmentation according to an existing or tailored protocol for manual region outlining in individual subject space. A full outlining protocol has been described by Hammers et al. (2003); other examples include the protocol by Shattuck et al. (2008) and another by Filipek et al. (1989). A further protocol for cortical labeling is under development as a collaborative project (brainCOLOR[4]). These methods require training of an operator in the chosen protocol and are expensive in terms of operator time and validation requirements, with costs rising approximately linearly with the number of images to be segmented and the number of regions labeled. The resulting segmentations are subject to intraobserver and interobserver variation.
3. Use one of a choice of semiautomatic approaches that require manual input, such as landmarks or seed points. Examples are SNT as noted above, Cardviews [Center for Morphometric Analysis, Massachusetts General Hospital, Boston, MA, USA, (Rademacher et al., 1992)], CARET [cortex only, Washington University School of Medicine, Saint Louis, MO, USA (Van Essen, 2005)] and LDDMM (Beg et al., 2004; Csernansky et al., 2004). Compared to manual outlining, interobserver variation is reduced, since even if the manual input varies within a certain range, the algorithms tend to arrive at the same results. These approaches are less labor-intensive, but the costs are still closely tied to the number of target images and regions.
4. Carry out anatomical segmentation in individual space using a fully automatic procedure. Software packages are available that implement the required functionality, but have limitations. For example, Mindboggle (Klein and Hirsch, 2005) and its extension using multiple atlases (Klein et al., 2005) are designed for cortical segmentation only, while the FS + LDDMM method achieves limited accuracy (Khan et al., 2008). Such approaches typically place a high demand on the computing infrastructure. An exception in this respect is the work by (Lötjönen et al., 2010), which is designed to reduce the computational demand sufficiently to make multi-atlas segmentation clinically feasible.

The present work is an instance of the fourth option. We have generated anatomical labels for ADNI MR images and provide them for download along with other ADNI data (http://www.loni.ucla.edu/ADNI). We present segmentations of 816 subjects' screening and baseline images into 83 regions, along with a statistical description of regional volumes.

To obtain automatic segmentations, we used multi-atlas propagation with enhanced registration [MAPER, Heckemann et al. (2010)]. This is a refined version of a previously validated approach (Heckemann et al., 2006). MAPER is the first automatic whole-brain multi-region segmentation method that has been shown to yield robust results in subjects with neurodegenerative disease. It uses training data ("atlases," images with reference segmentations) to segment T1-weighted brain MR images of any provenance into anatomical regions (Heckemann et al., 2010). We showed in previous work (Heckemann et al., 2006) that the accuracy achieved with MAPER is only slightly inferior to that of manual segmentation performed by a trained operator, and that the procedure is robust in the face of anatomical variation in the target subjects, specifically ventricular enlargement as seen in aging and neurodegeneration.

The implementation of MAPER used here relies on software tools sourced from the Image Registration Toolkit [IRTK,[5] Rueckert et al. (1999)] and from Nifty Reg,[6] (Modat et al., 2010). In a comparison of tools for intersubject registration of MR brain images, IRTK was recently found to be among the best-performing ones (Klein et al., 2009). Two other tools [SyN[7] (Avants et al., 2008) and ART[8]] (Ardekani et al., 2005) achieved more consistent results than IRTK in the comparison by Klein et al. Nevertheless, when working with heterogeneous data, we found IRTK to be more robust than ART and SyN, in particular when source (atlas) MR images had been acquired on different scanners than the target data for segmentation [e.g. ADNI images, Heckemann et al. (2010)]. ART and SyN have been shown to be suitable for registering pairs of images of identical provenance. MAPER is characterized by its robustness towards ventricular distension in the target subject. To achieve this, it relies on IRTK's ability to register multi-spectral tissue probability maps using cross correlation as the similarity measure, a feature that, to our knowledge, has not been implemented elsewhere. Our choice of IRTK rests on these two factors – robustness towards both intensity differences and typical pathology – although MAPER could in principle be implemented using other toolkits.

We validate the results using the volumes of the segmented regions as well as agreement measures between segmentations of images that have been serially acquired at different field strengths, and document limitations of the automatic procedure and the generated results for the benefit of future users of the data. We found that signal changes caused by white-matter disease can result in misclassification of tissues and lead to distortions in the segmentations. To quantify this influence, we describe and validate

[3] http://www.med.harvard.edu/AANLIB.
[4] http://www.braincolor.org/protocols.
[5] http://www.doc.ic.ac.uk/~dr/software.
[6] http://sourceforge.net/projects/niftyreg.
[7] http://www.picsl.upenn.edu/ANTS.
[8] http://www.nitrc.org/projects/art.

Author's personal copy

an automatically generated index. Finally, we show that statistical analyses of the automatically generated segmentations confirm previous observations of morphometric changes in Alzheimer's disease and mild cognitive impairment.

Materials and methods

MR data

Atlas data as required for MAPER consisted of 30 T1-weighted 3D image volumes acquired from healthy young adult volunteers at the National Society for Epilepsy at Chalfont, UK. Details of the acquisition are in Hammers et al. (2003). Hand-drawn segmentations of 83 structures had been previously prepared according to the protocols described in Hammers et al. (2003) and Gousias et al. (2008). Segmentation protocols are also available at http://www.brain-development.org.

MR images of patients with Alzheimer's disease and mild cognitive impairment as well as healthy elderly subjects were obtained from the ADNI database (www.loni.ucla.edu/ADNI).[9] The research presented here aligns with the primary goal of ADNI, which has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease. The full repository of ADNI images was accessed in February 2010. The clinical information was retrieved in August 2010. Each subject was assigned to one of five diagnosis groups: healthy subjects (HS), mild cognitive impairment with no conversion within the observation period (stable MCI, s-MCI), mild cognitive impairment at baseline, with progress to Alzheimer's disease within the observation period (p-MCI), Alzheimer's disease (AD), and Other (O). The latter assignment was used as a "catch-all" for subjects who did not fit the other categories, for example if ADNI noted a reversion from AD to MCI. The observation period was 24 11 months.

Preprocessing

As envisaged in the ADNI study design, images were obtained from the ADNI database in fully preprocessed versions. Depending on the scanner source, preprocessing included all or some of *GradWarp* geometric distortion correction (Jovicich et al., 2006), *B1 nonuniformity correction* to compensate for signal inhomogeneity (Jack et al., 2008), *N3* bias field correction (Sled et al., 1998) and *phantom scaling*. We chose the originally supplied linearly scaled images, irrespective of problems reported on a subset,[10] as linear scaling issues do not affect the segmentation procedure. Likewise, volume measurements, once normalized by intracranial volume as measured on the same source image, are unaffected by linear scaling.

To match the requirements of the MAPER procedure, we applied further preprocessing for brain extraction and tissue classification, as described in the following. Utilities used for these steps were taken from the Image Registration Toolkit [IRTK, Rueckert et al. (1999)], from the FSL suite (Smith et al., 2004) and from the ANTs toolkit (Avants et al., 2010).

For the brain extraction step, binary masks covering both intracranial white matter and gray matter (WM + GM) were available as the starting point. These had been generated as part of an earlier project using MIDAS, a semi-automatic procedure described elsewhere (Freeborough et al., 1997). Each mask was extended to cover the intracranial region generously by blurring (6 mm Gaussian kernel), thresholding at 27% and hole-filling. FSL FAST was applied to identify cerebrospinal fluid (CSF) within the pre-masked region. The original WM + GM mask was extended by the resulting CSF mask to obtain a complete intracranial mask that excluded meninges, sinuses and extracranial tissue. The original, semi-automatically created WM + GM mask is fully contained within the intracranial mask, reducing the impact of operator-dependent variability on the intracranial volume measurement.

Individual tissue probability maps for CSF, GM and WM obtained using FSL FAST were combined into a single multi-spectral volume. Fig. 1 shows a sample section from an image of a healthy subject. A binary maximum-probability gray matter mask was extracted from the discrete tissue class image generated by FAST.

T1-weighted screening (1.5 T) and baseline (3 T) images from the ADNI repository were obtained for all subjects for whom MIDAS-prepared brain masks were available. After removing data sets that had been withdrawn by ADNI after the download, a total of 996 images on 816 subjects (1.5 T: 811, 3 T: 185; of which paired: 180) were segmented and quality assessments carried out (cf. Statistical and visual analysis). For statistical analysis, only subjects in the HS, s-MCI, p-MCI, and AD groups (i.e. excluding the "Other" category), and of these only those who passed the outlier analysis (Outlier analysis section) were included (777 subjects, 953 images, 1.5 T: 772; 3 T: 181, of which paired: 176). The age distribution of included subjects is shown in Fig. 2. A breakdown by diagnostic group and gender is given in Table 1.

In the 176 included subjects for whom images had been acquired at both field strengths, the 3 T image was typically acquired within weeks after the 1.5 T image (median 22 days, range 0–112 days).

Segmentation

The MAPER procedure for robust, automatic segmentation of T1-weighted MR images of the human brain has been described and validated previously (Heckemann et al., 2010). Each target is paired with each of the atlases to generate an individual atlas-based segmentation. The steps are summarized in Table 2. In Steps 3 and 4, alignment of details in the image pair was achieved by optimizing a free-form deformation (FFD) represented by displacements on a grid of control points blended using cubic B-splines (Rueckert et al., 1999). These steps are carried out using each of the 30 atlases in turn, resulting in 30 segmentations, which are subsequently combined using vote-rule decision fusion (Rohlfing et al., 2004; Kittler et al., 1998).

In contrast to the approach discussed in Heckemann et al. (2010), where the entire registration was done in IRTK, we used Nifty Reg Version 1.3 (Modat et al., 2010) to carry out the detail-level registration (Step 4). The transformed and interpolated output from IRTK was used as the starting point. Nifty Reg is a particularly efficient implementation of the same FFD registration. To compare the accuracy of MAPER based on the combination (IRTK and Nifty Reg) with MAPER based on pure IRTK, we carried out a leave-one-out cross-comparison on the 30 atlas sets with both implementations, following the method described in Heckemann et al. (2010). Agreement between the generated and the manual label sets was measured using the mean Jaccard coefficient [*JC*; intersection divided by union (Jaccard, 1901)] across all 83 regions. The mean JC_m across the 30 atlas images was 0.691 for both methods [range

[9] ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public–private partnership. The Principal Investigator of this initiative is Michael W. Weiner, M.D., VA Medical Center and University of California San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research — approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years.

[10] http://www.loni.ucla.edu/twiki/bin/view/ADNI/ADNIMRICore.

Author's personal copy

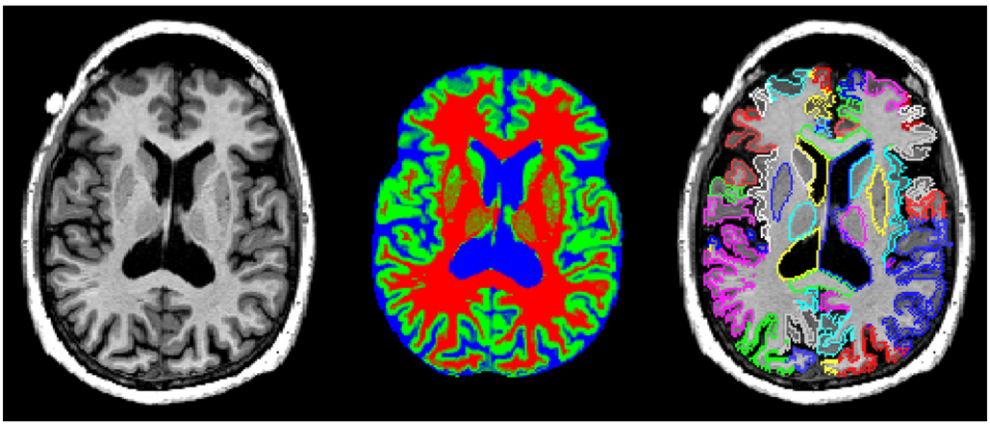


Fig. 1. Transverse section through the image of a healthy subject (73 year-old male, image ID I90026). Left panel: T1-weighted MR image. Middle panel: result of tissue classification with FAST; probability maps for each tissue class are combined into a single multi-spectral volume. Right panel: segmentation generated with MAPER. To create an accurate impression of the original resolution, the figure has been rendered without interpolation.

0.653–0.714, SD 0.0141 (IRTK), range 0.664–0.711, SD 0.0134 (IRTK and Nifty Reg)].

Quantifying white-matter disease

White-matter disease (WMD), characterized by diffusely hypointense regions within the white matter, is frequently seen in elderly subjects, and specifically in those with dementia (Black et al., 2009). Such regions can adversely affect the functioning of intensity-based methods. In the case of FSL FAST, they tend to be incorrectly labeled as gray matter, and this can impact subsequent processing — in the case of MAPER, resulting segmentations can be distorted. In particular, the lateral ventricle, the caudate nucleus, and the insula can be overestimated (cf. Outlier analysis). We developed a procedure to estimate the amount of white matter that is misclassified. The estimate is derived from a set of different label images derived from the T1-weighted image:

- A binary WM segmentation of the target by majority vote fusion of transformed atlas WM segmentations (each estimated from their intensities by FAST; note the atlas subjects are healthy young adults not affected by WMD): M_W
- A binary GM segmentation of the target generated with FAST: F_G
- A semi-automatically generated binary label covering white matter and gray matter of the target, as described in Preprocessing section: S_B
- A binary segmentation of both lateral ventricles in the target extracted from the fusion of the transformed atlas labels: M_V

An image with suspected WMD voxels is generated as

$$W = A \cup B$$

where

$$A = (M_W \cap F_G) \ominus E$$

$$B = (M_V \cap S_B) \ominus E$$

and E is a 3 3 3 structuring element used for eroding the intermediate images (this operation symbolized by $\ominus$).

In subjects where WMD leads to hypointensities that coincide with white matter, as identified by transforming atlas WM segmentations, such regions will be labeled by the intermediate Image *A*. In subjects where hypointensities border on the lateral ventricles, Image *B* will capture the affected regions. The volume of the resulting label *W*, normalized by the intracranial volume, provides an indication of the subject's WMD load. In the following, we refer to this measure as the white-matter hypointensities index (WMHI).

We assessed the validity of the WMHI by comparison with a semiquantitative rating. We adapted the rating scale described by Wahlund et al. (2001), which was designed for X-ray computed

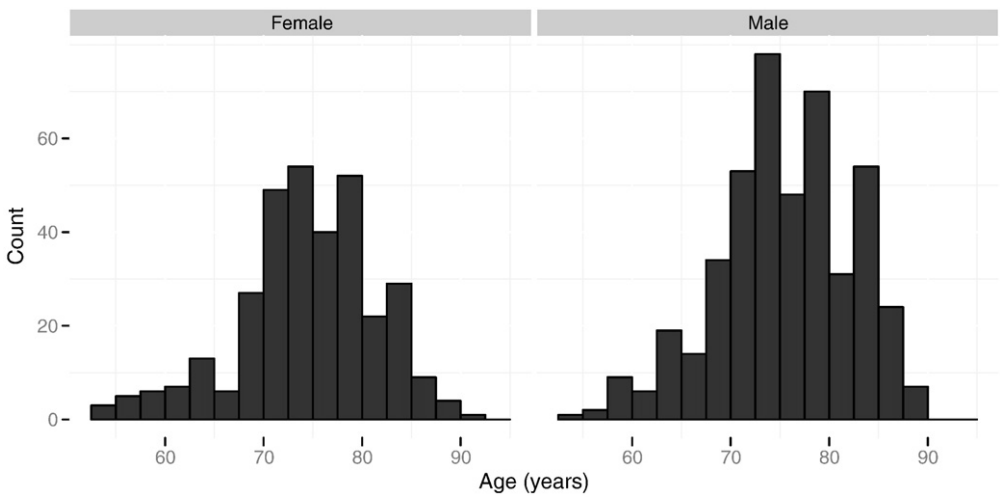


Fig. 2. Age distribution of included subjects.

Author's personal copy

Table 1
Numbers of subjects in each group.

	HS	s-MCI	p-MCI	AD	Total
Female	101	74	64	88	327
Male	110	144	99	97	450
Total	211	218	163	185	777

tomography and T2-weighted MR images, for use with T1-weighted images:

0: No hypointensities clearly identifiable as lesions
1: Focal lesions
2: Beginning confluence of lesions
3: Diffuse involvement of the entire region

The WMHI distribution was highly nonlinear with a small number of high values. To select a subset for visual scoring, we ranked the images according to WMHI, divided the sample into three equal parts, and sampled in a proportion of 42:21:7 from each group, yielding a total of 70 images for review.[11] An experienced rater (AH) who was blinded to WMHI, age, and diagnosis assigned the score after reviewing the T1-weighted images in three orthogonal planes. Where subjectively appropriate, based on comparisons within the sample, the rater assigned a tendency to the score, which was recorded as an addition or subtraction of 0.3 points to or from the integer score.

Masking based on tissue class

Depending on the application, it may be desirable to use segmented regions that have been multiplied with a binary tissue class label. In particular, since aging and Alzheimer's disease are characterized by cortical neuronal loss, the GM portion within each cortical label is often more relevant than the full label containing both GM and WM. We thus provide both raw segmentation data and masked versions. For the latter, regions with a substantial GM portion have been masked with a GM label (all except ventricles, central structures, cerebellum and brainstem), and the lateral ventricles have been masked with a CSF label. Unless otherwise noted, the analysis results reported in this work are based on the masked label sets.

Statistical and visual analysis

We assembled and analyzed the results of volumetry on all structures in all target subjects using standard statistical methods as provided by the R environment (http://www.r-project.org/).

Segmentation failures typically lead to grossly inaccurate estimations of the volume of individual regions. To detect outliers in the data, we grouped the images by diagnosis, gender and field strength, and determined per-group means and standard deviations of the regional volume (normalized by intracranial volume; masked by GM except for ventricles, central structures, brainstem and cerebellum). On this basis, all region volumes were converted to z scores. Regions where the z score was greater than 4 or less than -4 were flagged as outliers. Images containing outlier regions were visually assessed by an experienced reader (RAH). Label outlines were superimposed on the MRI image and the flagged region and its neighborhood viewed in the transverse, sagittal and coronal planes. The segmentation quality was rated on a visual analog scale from 1 to 5 (1: failed segmentation; 2: poor boundary matching, but correct indication of the relative position of neighboring regions; 3: fair; 4: good segmentation with minor boundary mismatches, 5: excellent segmentation with exact boundary matching). The likely cause of the outlying size of the region, based on the reader's subjective impression, was identified and recorded. The remainder of the image was searched in the transverse plane for obvious label mismatches beyond the flagged region and a note of the overall impression recorded.

[11] Originally, we had sampled 21 evenly from the ranked list. Based on the review of this original set, we decided to increase the sample size. The trisection approach allowed us to add to the existing sample, while maintaining even spacing within the parts and emphasizing the upper WMHI value range, where we expected the findings to be most relevant.

Table 2
MAPER steps for generating an individual segmentation. Sim: similarity measure, mstprob: multi-spectral tissue probability map, CC: cross correlation, NMI: normalized mutual information. Numbers indicate the control point spacing in millimeters.

	Type	Level	Image data	Sim	Toolkit	Tool
1	Global	Rigid	mstprob	CC	IRTK	rreg
2	Global	Affine	mstprob	CC	IRTK	areg
3	Coarse	Nonrigid 20	mstprob	CC	IRTK	hreg
4	Detailed	Nonrigid 10, 5, and 2.5	T1 signal	NMI	Nifty Reg	reg_f3d
5	Transform	Nonrigid 2.5	Atlas labels	n/a	Nifty Reg	reg_resample

Statistical analysis was carried out with a view to comparing diagnostic groups and determining potential volumetric criteria characteristic for Alzheimer's disease or impending progression from mild cognitive impairment. We also used MAPER measurements to determine balanced asymmetry indices for paired regions (A_r) as

$$A_r = \frac{2|V_R - V_L|}{V_R + V_L} \tag{1}$$

for right and left regional volumes, V_R and V_L. Unbalanced indices were generated by dividing V_L by V_R.

The volumetry and asymmetry studies were carried out using the images acquired at 1.5 T. The findings were compared with published knowledge as a consistency check for the correctness of the segmentation approach.

Comparison across field strengths

Where pairs of images acquired at 1.5 T and 3 T were available for individual subjects, the pair was rigidly registered and the unmasked label sets compared in the space of the 3 T image, using *JC*. A per-subject summary measure of agreement was obtained by calculating the mean *JC* across all 83 regions (JC_m). Key results are also provided as Dice similarity coefficients [DSC, intersection divided by average label volume, (Dice, 1945)].

Measuring precision by comparing independent segmentations

To derive a quantitative indicator of the precision of the segmentation of each target, we employed the following procedure. For each image in the ADNI set, we bisected the atlas set randomly into two subsets of 15 atlases each. From the pair of subsets, we generated a pair of independent segmentations using vote-rule decision fusion. The overall agreement between the unmasked segmentation pair was measured as the mean Jaccard coefficient across all 83 regions (JC_m).

Results and discussion

Segmentation results are available for download in NIfTI format as 3D label maps identifying 83 structures by spatial correspondence with the T1-weighted images as supplied by ADNI.[12]

[12] ADNI users can download the label maps from the image database (https://ida.loni.ucla.edu) via the "Advanced Search" feature by selecting "Post-processed" and entering "MAPER*" in the field "Series Description".

Author's personal copy

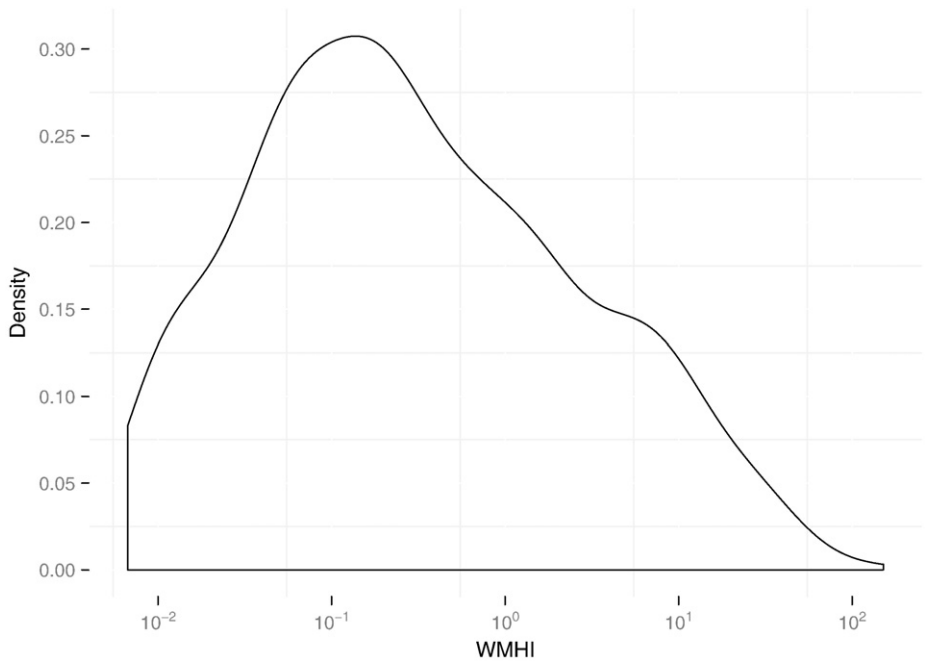


Fig. 3. Density plot of white-matter hypointensities index (WMHI) on 996 images.

Quality of intracranial masks

The mean intracranial volume (ICV) obtained by measuring the volume of the intracranial mask (cf. Preprocessing) was 1.41 ℓ (range 1.02–1.86 ℓ, SD 0.143 ℓ) on 1.5 T images. The images with the three largest (I72219, I40356, and I35499) and the three smallest (I63227, I82594, and I52799) ICVs were reviewed with the mask outline superimposed to search for visible under- and overestimations. All six masks were judged to adequately represent the intracranial volume after careful visual inspection.

In subjects for whom images had been acquired at both field strengths ($n = 176$), the measured ICV on 3 T was highly correlated with that of 1.5 T (Pearson's $r = 0.976$), giving smaller results on average than 1.5 T, but not significantly so (−1%, range −5%–+10%, SD 2 percentage points, $p = 0.32$). Similar observations have previously been made, when ICV measurements were compared on pairs of brain images of subjects who had been scanned serially at different field strengths (Keihaninejad et al., 2010). Automatic methods showed a tendency to underestimate ICV on 3 T and overestimate ICV on 1.5 T images. For the most consistent automatic method described by Keihaninejad et al., the difference was 0.7%.

Outlier analysis

Sixty regions in 42 subjects met the outlier criterion and were reviewed visually. Twelve subjects appeared twice in the list, two

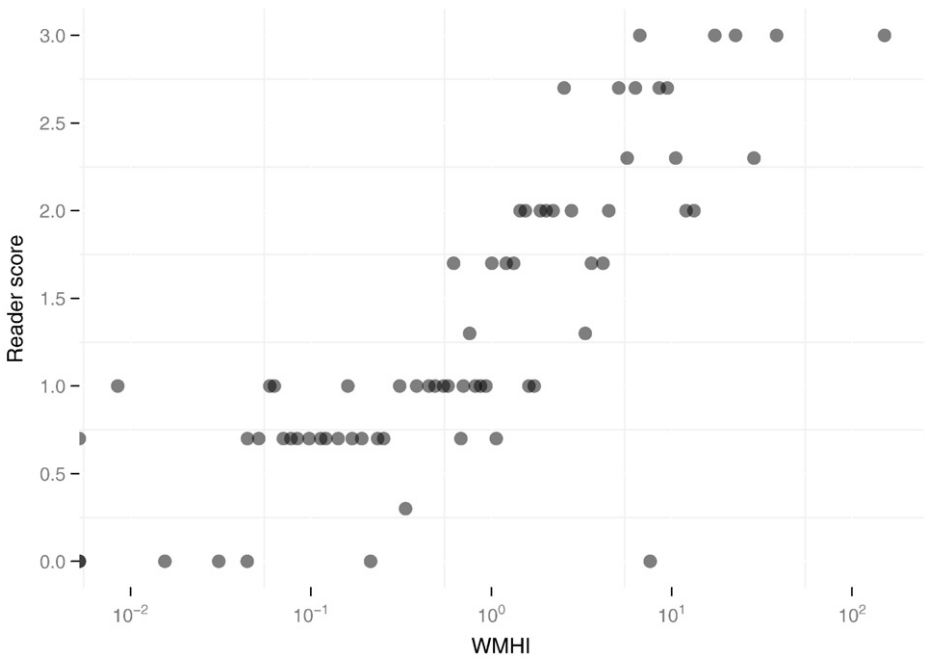


Fig. 4. Scatter plot of white-matter hypointensities index (WMHI, log-scaled axis) versus the Wahlund score assigned by the rater on 70 images. WMHI scores of zero are shown as half disks on the left edge of the plot.

Author's personal copy

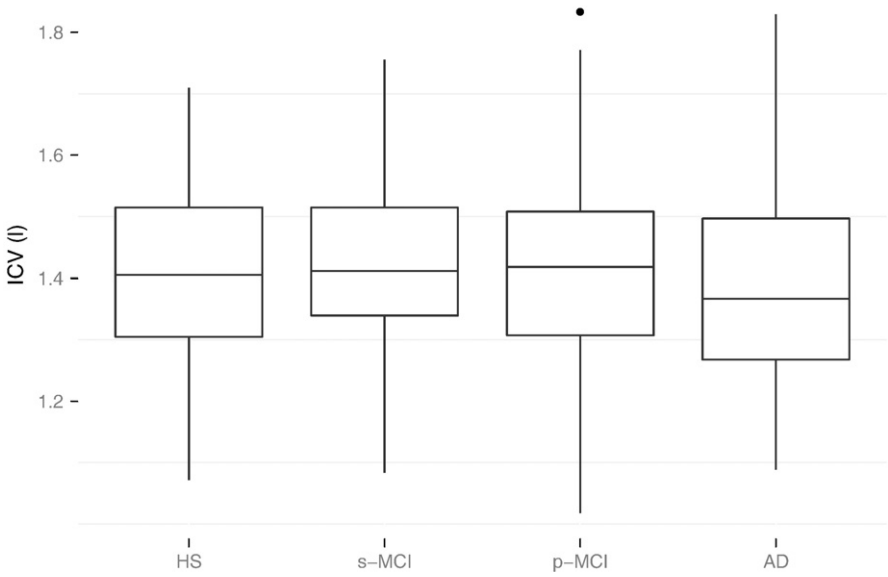


Fig. 5. Comparison of intracranial volumes by diagnosis groups, 1.5 T images. Center line shows median, boxes capture 25%–75% quantile range, whiskers indicate 1.5 interquartile range, dots denote outliers. ICV unit is ℓ.

subjects appeared three times and one appeared four times. The regions that appeared most frequently in the outlier list were the temporal horn of the lateral ventricle (8 right, 6 left), the caudate nucleus (7 right, 5 left), and the subcallosal area (2 right, 4 left).

On visual review, the flagged regions appeared to be affected by white-matter disease in a large number of cases (WMD: 24; other flawed segmentations: 19, correct: 17). In 13 of the 19 problematic segmentations that were not WMD-related, the flaw appeared to be limited to the region in question. No further segmentation problems were detected in these cases, and the extent of over- or undersegmentation was deemed to be mild or moderate (scoring 3 or 4 on the visual analog scale described in Statistical and visual analysis section). In the remaining six regions, more general problems were seen and the relevant four cases (I64189, I38944, I67210, and I91126) were excluded from further analysis (MR acquisition problems leading to lack of GM/WM contrast: four regions in three images; motion artifact: two regions in one image).

WMD is a highly prevalent feature in the subjects of this cohort, frequently leading to overestimation of the caudate nuclei and the insula regions. The gray-matter portion of labels of other cortical regions often included white-matter regions that had been mistaken for gray matter by the tissue classification. Subcortical regions other than the caudate nuclei appeared largely unaffected on visual review. We determined for each image an index (cf. Quantifying white matter disease section) that signals WMD load. This index correlates well with the visual appearance of distortion (cf. Measuring WMD using white-matter hypointensities index). It is provided with the label images as part of the metadata.

The raw MAPER-based label for the lateral ventricle is frequently overestimated, incorrectly including hypointense portions of white matter. We dealt with this issue by masking this region pair with the binary CSF label generated by FAST.

While MAPER is robust in the majority of cases, the limitations of automatic segmentation (and, indeed, manual segmentation) need to be considered in subjects whose anatomical configuration is severely abnormal and in those who show texture abnormalities such as white matter disease.

Measuring WMD using the white-matter hypointensities index

The WMHI ranged from 0 (seen in 135/996 images) to 151, with the distribution strongly skewed towards 0. The distribution is best visualized using a log scale as shown in Fig. 3.

Fig. 4 plots the rater score against WMHI. The measures are strongly correlated (Kendall rank correlation coefficient 0.71, significant at the limit of numeric precision), although there is some overlap of WMHI between adjacent score groups. One image (I79803) received a visual score of 0, although it scored high on WMHI (7.59). Review of the MR image along with the intermediate Image *A* showing hypointensities showed an artefactual step change of intensity in the MR along the vertical axis, which resulted in a hypointense white matter region in the brainstem. No other white matter regions where highlighted or visibly affected by white matter disease.

The WMHI is intended to alert users to possible WMD-related oversegmentation when susceptible region labels are used for analysis. Such regions include the lateral ventricles before CSF-masking, the caudate nucleus and the gray-matter masked cortical regions, especially the insula. The WMHI has some value as a metadatum indicating the reliability of the segmentation. With a view to the caudate nucleus, however, its value is limited due to the way the index is generated: hypointensities adjacent to the caudate nucleus tend to be included in the generated caudate label, in which case they are not identified as WMD. Thus, it is possible for a caudate nucleus to be oversegmented due to white matter disease, even when the WMHI is zero. Random visual

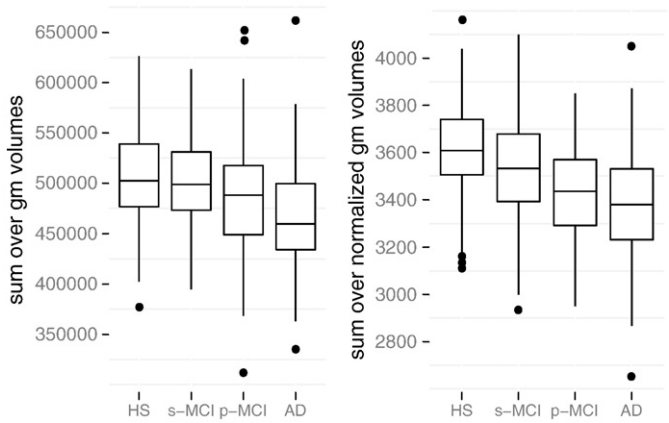


Fig. 6. Summed volume of gray matter portions of labeled regions, comparison by diagnosis groups, 1.5 T images. Left: absolute volumes in mm^3, right: normalized by ICV and scaled (arbitrarily) by 10^4.

Author's personal copy

reviews have revealed one image where this appears to be the case (I63489). In future work, we will seek to address the issue of WMD-related oversegmentation in a principled fashion by identifying affected subjects and regions a priori and counteracting the distorting effects at the registration step. We will also search for better criteria to indicate the overall veracity of the generated segmentations.

Volumetric analysis

Normalization

To reduce interindividual variation of region volumes, various measures have been proposed for normalization (Free et al., 1995). In particular, normalization of brain volume by intracranial volume was found to substantially reduce variation, and to remove gender-related differences (Whitwell et al., 2001). We found in previous work a correlation between hippocampal volume and ICV (Hammers et al., 2007), and this was confirmed in the present data (Pearson's $r = 0.56$ for the sum of both hippocampi in 1.5 T images of healthy subjects). Normalization by ICV also eliminates inaccuracies arising from problems with the phantom scaling, which have been reported for a subset of ADNI cases (Clarkson et al., 2009).

Our ICV measurements were stable across the diagnostic groups (cf. Fig. 5). Based on a two one-sided tests (TOST) procedure (Schuirmann, 1987), the null hypothesis of non-equivalence can be rejected for all paired comparisons of diagnosis groups, except (s-MCI, AD) where $p = 0.056$ ($\alpha = 0.05$; $\epsilon = 0.05\mu$). In the following, individual region sizes are expressed as a fraction of ICV, scaled by an arbitrary factor of 10^4.

The benefit of ICV normalization can be seen in group comparisons by diagnosis: the absolute total gray matter volume differs between

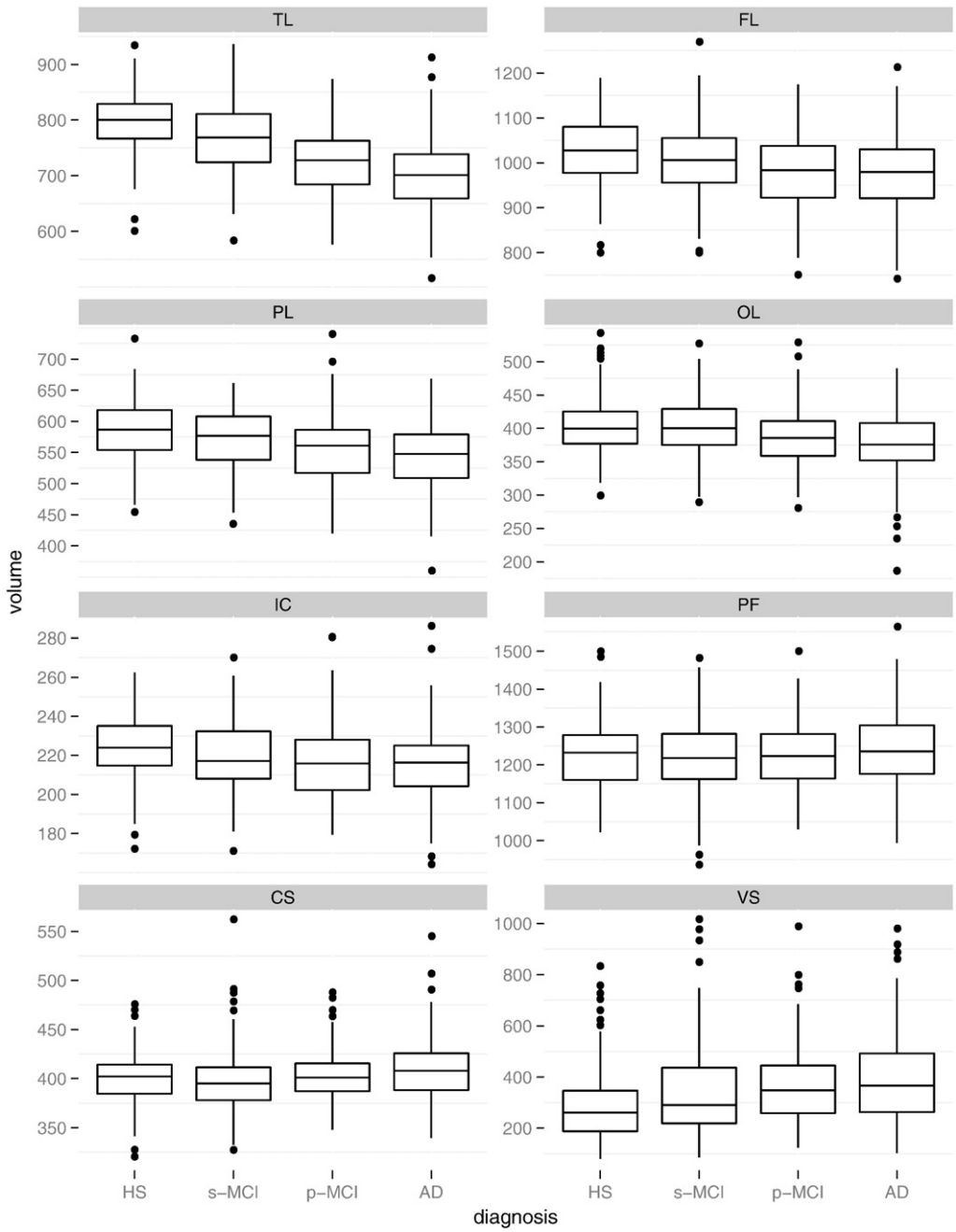


Fig. 7. Summed label volumes per superregion, normalized and scaled by 10^4. Gray matter masking applied where appropriate (all except CS, PF and VS). TL: temporal lobe, FL: frontal lobe, PL: parietal lobe, OL: occipital lobe, IC: insula and corpus callosum, PF: posterior fossa, CS: central structures, VS: ventricular system.

Author's personal copy

groups, but the distinction is comparatively weak (cf. Fig. 6, left panel). The right panel shows total gray matter volume with normalization, which results in larger group differences.

Aggregated regional analysis

For Fig. 7, volume results for individual regions have been aggregated into six superregions. The plots indicate that the temporal lobe is most distinctly different between diagnostic groups. Differences in the medians are also substantial for the ventricle regions, but the variance is greater in all groups, resulting in larger overlaps.

Individual regional analysis

The analysis of individual regional volumes reveals a pattern of increasing atrophy from the HS group via s-MCI and p-MCI to the AD group. Table 3 shows this for the 14 regions where the AD–HS difference is largest. An extended version of the table that includes all regions is provided as supplemental material. Most of the results match our expectations: ventricles are enlarged, especially the temporal horns; hippocampi are smaller, notably also when comparing HS with s-MCI (9% either side). The amygdala, the middle and inferior temporal gyrus and the fusiform gyrus are reduced in size, adding to the evidence that temporal lobe regions beyond the hippocampus are affected by the disease process. The amygdala is functionally connected with and spatially adjacent to the hippocampus, and its involvement in AD is well known from histopathology (Kromer Vogt et al., 1990; Scott et al., 1991) and imaging research (Cuénod et al., 1993; Jack et al., 1997; Lehericy et al., 1994). Other temporal lobe structures, notably the fusiform gyrus, the parahippocampal gyri, and the middle and inferior temporal gyri also have previously been found to be significantly affected (Chan et al., 2001).

In recent imaging studies, thalamic volumes have been found to be reduced in Alzheimer's disease (Cherubini et al., 2010; Zarei et al., 2010), in line with earlier post-mortem observations (Braak and Braak, 1991). de Jong et al. (2008) found reduced sizes of both putamen and thalamus. Our results confirm lower volumes of the thalamus, even when comparing the HS and s-MCI groups (5% either side, highly significant). For the putamen, the same comparison was marginally significant, while the difference between HS and AD was not. This finding may indicate a limitation of accuracy of the putamen segmentation in subjects with more advanced disease.

The heatmap in Fig. 8 indicates for each region and selected pairs of diagnostic categories the extent to which the measured volume can serve to distinguish the diagnosis groups. Red color indicates the "most significant" results in each column. Please note that p-values in this context are not used for the usual purpose of hypothesis testing, but for comparing regions; therefore we did not employ alpha thresholding or attempt correction for multiple comparisons. Regions in the mesial temporal lobe (hippocampus, amygdala, and parahippocampal gyri) are particularly prominent, along with the temporal horn of the lateral ventricle and the posterior temporal lobe. Outside of the temporal lobe, large posterior cortical regions (parietal lobe, occipital lobe) are highlighted. These observations align well with previously described AD patterns, specifically a posterior-to-anterior gradient in atrophy (Likeman et al., 2005).

Asymmetry

Generally, AD atrophy is described as a disseminated process with no lateral predilection. Regional counts of plaques and tangles in pathological specimens showed larger variability within one and the same region than between left and right counterparts (Janota and Mountjoy, 1988; Moossy et al., June, 1989; Wilcock and Esiri, 1987). Imaging studies comparing AD with other entities found that asymmetry indices may be a useful tool for differential diagnosis, as asymmetry of various regions frequently attends clinically similar conditions, specifically frontotemporal lobar degeneration (Barnes et al., 2006; Boccardi et al., 2003; Horínek et al., 2007; Likeman et al., 2005) and argyrophilic grain disease (Adachi et al., 2010). As a differential diagnostic criterion, asymmetry thus speaks against AD according to these studies.

For the hippocampus, a physiological right-larger-than-left asymmetry in healthy adults is well established [e.g. Pedraza et al., (2004)], but studies focussing on hippocampal asymmetry in AD have yielded varying results. Small lateral differences in atrophy rates between AD patients and controls were found by Barnes et al. (2005). Shi et al. (2007), focussing on shape characteristics rather than volume, also found small differences between AD and controls in the atrophy pattern. A metastudy on hippocampal volume found that right hippocampal volume was larger than left in all groups studied (AD, MCI and controls), with AD subjects showing smaller effect sizes due to larger variation (Shi et al., 2009). Similarly, Barber et al. (2001) report a loss of hippocampal asymmetry in AD patients versus controls. An increase in hippocampal asymmetry as a function of cognitive decline was seen in one study (Wolf et al., 2001).

In the present study, results for the hippocampal left/right volume ratio have a wide distribution. We therefore choose to report the median and the median absolute deviation (MD), which are more robust measures of central tendency and dispersion than means and standard deviations. For healthy subjects, we found the previously reported pattern of left<right hippocampal asymmetry (median L/R volume ratio 0.93; MD 0.073). In AD, the median of the volume ratio appears to be somewhat reduced (0.90; MD 0.12), but the difference is not significant. The balanced asymmetry index A_r is higher in AD than

Table 3
Regional volumes and volume differences. Column "HS vol" shows regional volume as a fraction of ICV, averaged across healthy subjects, scaled by 10^4. Columns labeled "d()" show the volume difference compared to HS as a percentage of HSvol, except "d(pMCI, sMCI)", which shows the volume difference between p-MCI and s-MCI as a percentage of the mean volume of the s-MCI group. The sort criterion is the modulus of the difference between AD and HS. Only the 14 regions ranking highest on the sort criterion are shown. "Code" is the numerical identifier for the region used in the label maps.

Code	Region	HS vol	d(sMCI, HS)	d(pMCI, HS)	d(pMCI, sMCI)	d(AD, HS)
47	Lat ventricle temp horn R	5.3	10	32	21	43
45	Lat ventricle main R	125.2	17	27	8	40
46	Lat ventricle main L	138.1	21	31	9	39
48	Lat ventricle temp horn L	4.8	6	21	14	34
2	Hippocampus L	13.3	−9	−15	−7	−20
3	Amygdala R	9.5	−8	−15	−7	−19
10	G parahippocamp/amb L	24.3	−9	−15	−7	−19
1	Hippocampus R	14.4	−9	−14	−6	−18
4	Amygdala L	9.9	−8	−14	−6	−18
9	G parahippocamp/amb R	23.4	−7	−13	−6	−18
14	Middle and inf temp gg L	59.9	−5	−14	−9	−17
15	Fusiform g R	20.4	−5	−11	−6	−16
16	Fusiform g L	21.4	−6	−12	−6	−16
13	Middle and inf temp gg R	62.3	−3	−10	−7	−14

Author's personal copy

	p(HS,AD)	p(HS,s–MCI)	p(s–MCI,p–MCI)	p(p–MCI,AD)
Hippocampus R TL	6.81e–27	1.38e–10	3.70e–04	2.86e–02
Hippocampus L TL	1.65e–32	1.99e–12	2.69e–05	8.86e–03
Amygdala R TL	8.84e–31	1.03e–09	1.03e–05	2.08e–02
Amygdala L TL	3.17e–28	6.17e–09	9.90e–05	2.52e–02
Ant temp lobe med R TL	3.92e–17	3.46e–02	2.74e–04	1.39e–02
Ant temp lobe med L TL	4.40e–16	2.87e–02	5.79e–04	2.27e–02
Ant temp lobe lat R TL	1.58e–12	3.03e–02	6.11e–04	1.60e–01
Ant temp lobe lat L TL	2.88e–14	1.05e–01	2.12e–04	4.36e–02
G parahippocamp/amb R TL	3.84e–33	2.94e–09	3.89e–05	2.92e–03
G parahippocamp/amb L TL	4.48e–39	3.78e–13	6.49e–06	9.06e–03
Sup temp gyrus post R TL	4.90e–13	1.75e–02	1.97e–03	5.92e–02
Sup temp gyrus post L TL	9.33e–14	3.89e–02	2.37e–04	5.02e–02
Middle & inf temp gg R TL	9.74e–25	1.24e–02	9.25e–07	1.24e–03
Middle & inf temp gg L TL	7.30e–33	1.04e–05	9.70e–10	2.50e–02
Fusiform g R TL	5.42e–23	2.00e–03	4.00e–04	1.10e–03
Fusiform g L TL	1.46e–23	1.06e–04	2.04e–04	8.58e–03
Post temp l L TL	9.79e–22	7.62e–04	1.08e–06	4.79e–02
Post temp l R TL	9.97e–23	4.56e–03	3.76e–09	7.01e–02
Sup temp g, ant part L TL	6.37e–03	1.26e–01	7.47e–01	3.26e–01
Sup temp g, ant part R TL	1.36e–03	6.32e–02	8.51e–01	2.43e–01
Middle frontal g L FL	5.49e–08	1.31e–03	8.33e–02	3.40e–01
Middle frontal g R FL	4.49e–07	1.25e–02	4.62e–02	4.63e–01
Precentral gyrus L FL	2.56e–01	6.41e–01	8.07e–01	6.46e–01
Precentral gyrus R FL	2.20e–01	4.59e–01	6.38e–01	3.42e–01
Straight gyrus L FL	1.72e–05	5.17e–03	2.36e–03	3.05e–01
Straight gyrus R FL	5.71e–07	8.85e–03	4.28e–03	7.42e–01
Ant orbital gyrus L FL	1.71e–03	1.43e–01	3.36e–03	2.61e–01
Ant orbital gyrus R FL	7.99e–04	4.60e–01	6.51e–04	3.59e–01
Inf frontal gyrus L FL	1.16e–04	1.39e–02	1.76e–01	8.69e–01
Inf frontal gyrus R FL	6.33e–04	4.00e–02	3.08e–02	7.44e–01
Sup frontal gyrus L FL	6.57e–04	7.50e–01	2.73e–03	9.96e–01
Sup frontal gyrus R FL	5.33e–04	1.04e–01	1.21e–01	6.95e–01
Med orbital gyrus L FL	8.40e–04	4.51e–03	9.52e–04	1.56e–02
Med orbital gyrus R FL	5.91e–08	3.07e–05	6.83e–03	1.17e–01
Lat orbital gyrus L FL	1.71e–01	8.76e–01	6.78e–02	7.02e–01
Lat orbital gyrus R FL	1.21e–02	1.49e–01	5.50e–01	6.44e–01
Post orbital gyrus L FL	5.29e–08	8.61e–04	9.14e–03	8.61e–01
Post orbital gyrus R FL	1.74e–07	1.69e–02	1.20e–03	8.58e–01
Subgenual frt cortex L FL	3.57e–02	7.29e–01	3.81e–01	4.07e–01
Subgenual frt cortex R FL	3.24e–01	4.91e–01	5.42e–02	7.50e–01
Subcallosal area L FL	9.32e–02	5.66e–01	3.10e–01	3.01e–01
Subcallosal area R FL	3.37e–01	3.89e–01	9.54e–02	9.90e–01
Pre–subgenual frt ct L FL	3.28e–04	6.82e–03	7.01e–01	2.05e–01
Pre–subgenual frt ct R FL	2.84e–03	5.28e–01	3.71e–02	7.99e–01
Cerebellum R PF	1.04e–01	9.65e–01	6.20e–01	3.03e–01
Cerebellum L PF	1.05e–01	9.92e–01	5.83e–01	3.13e–01
Brainstem PF	2.02e–01	3.00e–01	9.38e–02	5.39e–01
Insula L IC	6.59e–07	2.50e–05	1.72e–01	6.93e–01
Insula R IC	1.42e–06	2.83e–04	7.81e–02	7.71e–01
Cingulate gyrus, ant L IC	2.58e–01	3.89e–01	9.13e–02	8.87e–01
Cingulate gyrus, ant R IC	2.83e–01	8.84e–01	9.52e–01	3.57e–01
Cingulate gyrus, post L IC	2.62e–07	1.14e–02	2.96e–02	5.50e–01
Cingulate gyrus, post R IC	4.94e–06	3.27e–02	1.64e–02	8.14e–01
Occipital lobe L OL	6.97e–10	4.63e–01	1.03e–04	5.26e–02
Occipital lobe R OL	1.12e–09	3.68e–01	9.33e–04	9.34e–02
Lingual gyrus L OL	1.24e–01	2.62e–01	8.85e–02	3.60e–01
Lingual gyrus R OL	1.27e–03	9.13e–01	3.49e–01	4.54e–02
Cuneus L OL	1.22e–01	4.17e–01	8.53e–01	4.61e–02
Cuneus R OL	1.32e–02	9.58e–01	7.34e–01	4.04e–02
Parietal lobe (rem) L PL	2.82e–18	2.40e–03	1.04e–04	5.81e–02
Parietal lobe (rem) R PL	2.14e–14	2.33e–03	2.43e–03	4.54e–02
Postcentral gyrus L PL	1.26e–01	6.71e–01	3.21e–01	3.80e–01
Postcentral gyrus R PL	5.99e–01	6.81e–01	8.32e–01	9.42e–01
Sup parietal gyrus L PL	7.62e–09	3.50e–02	9.41e–04	4.25e–01
Sup parietal gyrus R PL	8.10e–09	2.96e–02	1.05e–03	5.56e–01
Caudate nucleus L CS	5.61e–06	1.94e–01	3.38e–01	2.72e–02
Caudate nucleus R CS	1.79e–07	1.27e–01	2.57e–01	1.13e–02
Nucleus accumbens L CS	3.83e–04	3.52e–03	2.68e–01	9.52e–01
Nucleus accumbens R CS	1.06e–02	8.98e–02	3.78e–02	3.58e–01
Putamen L CS	2.21e–01	2.57e–03	3.74e–01	5.44e–01
Putamen R CS	5.07e–01	1.82e–02	6.66e–01	3.32e–01
Thalamus L CS	3.21e–07	3.26e–06	6.90e–01	5.78e–01
Thalamus R CS	4.15e–08	7.07e–06	9.31e–01	2.04e–01
Pallidum L CS	1.84e–01	3.96e–01	1.96e–01	5.06e–01
Pallidum R CS	3.20e–02	5.00e–01	3.01e–02	7.24e–01
Corpus callosum CS	5.09e–07	1.66e–01	4.99e–02	9.09e–02
Substantia nigra L CS	5.14e–02	2.12e–02	9.52e–01	9.99e–01
Substantia nigra R CS	6.25e–02	4.47e–02	6.85e–01	6.44e–01
Lat ventricle main R VS	1.35e–11	1.55e–03	9.64e–02	3.68e–02
Lat ventricle main L VS	3.82e–10	3.31e–04	1.07e–01	2.29e–01
Lat ventricle temp R VS	4.81e–14	1.12e–02	9.52e–05	1.10e–01
Lat ventricle temp L VS	4.23e–13	5.90e–02	4.07e–04	1.59e–02
Third ventricle VS	6.98e–04	3.47e–01	3.19e–01	1.44e–01

Author's personal copy

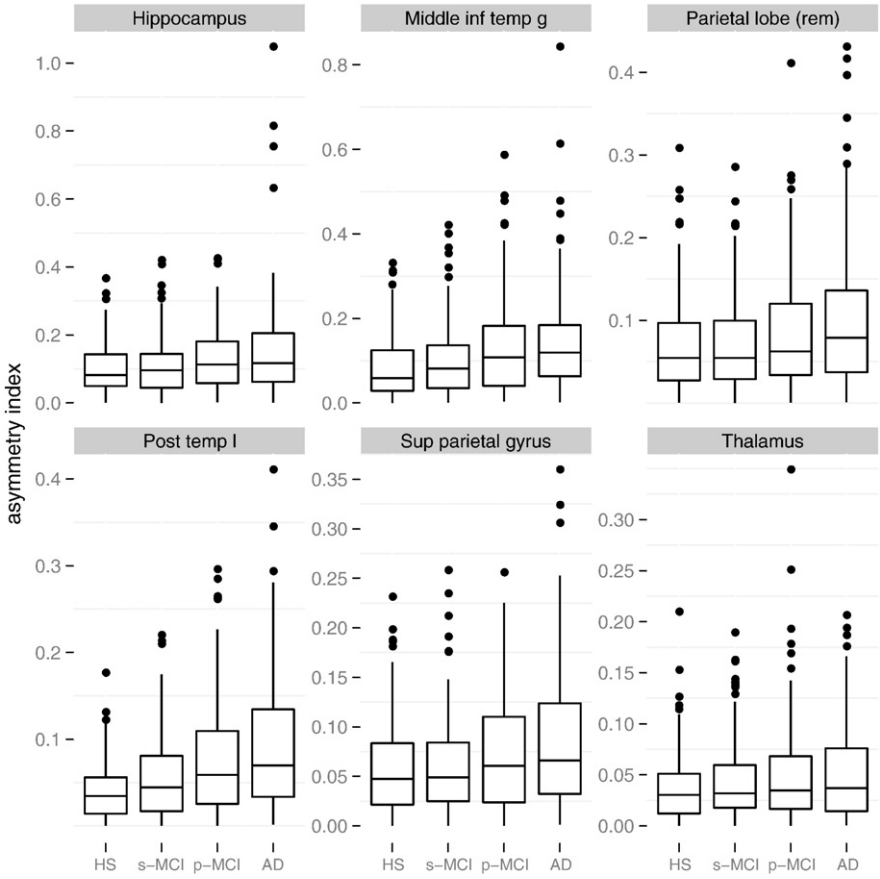


Fig. 9. Boxplot showing group differences of asymmetry index for selected regions. The difference between AD and HS is highly significant for each of the regions shown ($p<10^{-4}$); the difference between s-MCI and p-MCI is significant ($p<0.05$) for the shown regions except thalamus and parietal lobe. Cf. Fig. 8 for abbreviations.

in healthy subjects (0.12 versus 0.08), and this difference is significant ($p=1.3\times10^{-5}$).

Few studies have looked at asymmetry in AD beyond the hippocampus. The amygdala has been studied by Whitwell et al. (2005), who found a significant increase of asymmetry in AD patients. Thompson et al. (1998) found asymmetries in the Sylvian fissure in normal subjects, and these were significantly accentuated in AD.

We note that A_r is particularly large in AD compared to HS when considering large regions (posterior temporal lobe, HS: 0.041, AD: 0.092, $p=1.3\times10^{-15}$, parietal lobe, HS: 0.068, AD: 0.100, p at limit of precision, cf. Fig. 9). This is an area for future exploration.

Consistency across field strengths

High levels of agreement were seen when comparing the segmentation obtained on a 1.5 T image with that obtained on the same subject's 3 T image. The aggregate measure across all structures showed little variation between subjects (JC_m 0.802 ± 0.0146, range 0.749–0.829; corresponding to a DSC of 0.890). Results were equivalent for all four disease conditions as shown by a TOST procedure ($p<$1e-17 for $\alpha=0.05$ and $\epsilon=0.04$). In previous work, we used the JC_m measure to assess MAPER with reference to manual segmentation in normal adults (Heckemann et al., 2010), obtaining a mean of 0.691 (DSC 0.817). The fact that the MAPER method produces consistent results across field strengths indicates high precision and corroborates our previous findings showing the accuracy of the method.

Between individual regions, we note large differences in the standard deviation of the Jaccard coefficient (JC_σ). This standard deviation, here expressed as a percentage of the mean, ranges from 1.4% (brainstem) to 17.7% (left pallidum). Labels of regions that are well-defined by gray-scale intensity gradients in the T1 image are particularly consistent across field strengths. JC_σ for, e.g., the frontal horn and central part of the lateral ventricle is 3.0% (either side). For the precentral gyrus, which is a cortical region of average size within our set, JC_σ is 2.3% (left) and 2.5% (right). Small regions with weakly defined boundaries are naturally difficult to segment, both manually and by automatic procedures based on manual input. In the present results, this is reflected in large standard deviation values for the pallidum (JC_σ left: 17.7%, right: 14.3%) and the nucleus accumbens (left: 12.9%, right: 11.0%).

Precision based on atlas subsets

There was strong agreement between independent atlas-subset based segmentations of the target images (JC_m was 0.800 ± 0.0092, range 0.771–0.819, corresponding to a DSC of 0.889). Three outliers were seen out of 996, one of each in the HS (I64867, JC_m 0.765), s-MCI (I69360, JC_m 0.718) and p-MCI (I97327, JC_m 0.755) groups. On visual review of the corresponding segmentations, no substantial mislabelling was seen.

Fig. 8. Heatmap showing per-region results of unpaired two-tailed t-tests between selected pairings of diagnosis groups. P-values are mapped to colors so that the "most significant" results for each column are highlighted in red. Each column is color-scaled independently. R: right, L: left, ant: anterior, amb: ambiens, temp: temporal, med: medial, lat: lateral, sup: superior, post: posterior, inf: inferior, g: gyrus, gg: gyri, l: lobe, frt: frontal, rem: remainder; superregion abbreviations as in Fig. 7.

Author's personal copy

Future improvements

The choice of the segmentation approach employed in this work was guided by multiple considerations and represents a compromise that can be improved upon in future work. An important factor was proved robustness, as it helped to ensure that all segmentation results are traceable to a single procedure, and helped to avoid any individualized modification of the output data. As a consequence, a variety of recently published algorithmic developments have not been considered and may lead to even more accurate and detailed segmentations, once their robustness has been demonstrated. Promising developments are taking place in several areas. Registration algorithms are becoming more accurate and efficient as shown by Lötjönen et al. (2010) (albeit only for image pairs of identical provenance). The dependence on expensive expert input may in the future be reduced thanks to algorithms that uncover latent atlases (Riklin-Raviv et al., 2010). Procedures that select or weight atlas images from the outset (Aljabar et al., 2009) or at the segmentation combining stage (Artaechevarria et al., 2009) may yield more accurate results, especially when heterogeneous repositories are used as atlas data. Algorithms that revisit the image data after segmentation in order to refine the result are showing strong promise (Wolz et al., 2010b; van der Lijn et al., 2008).

A further compromise was made with regard to the choice of the atlas data. We settled on a set that has been segmented in high detail and with strong validation (Hammers et al., 2003), although it is based on young adults and is therefore demographically dissimilar from the ADNI target images. Work by Wolz et al. (2010a) shows for individual regions that the LEAP approach – propagating atlas labels indirectly via intermediate images – can yield improved results on such dissimilar targets. The question of how LEAP or a similar approach can be adapted for multiple regions is another promising research avenue.

Conclusion

In this work we present a repository of label data on healthy elderly subjects and patients with mild cognitive impairment or Alzheimer's disease. We offer segmentations of 996 screening and baseline images in 816 subjects. The data are publicly available as an accompaniment to the MRI data supplied by the ADNI project. We validated the segmentation results and presented results of statistical analysis that are congruent with established knowledge about atrophy progression in AD.

We are committed to maintaining and enhancing the repository with findings from future research. In particular, we envisage developing further indices of accuracy and adding them as metadata. As improvements to the segmentation algorithm are developed and validated, we are planning to add updated segmentations, using versioning to ensure that the original set described here remains available for reference. If members of the community should express an interest in segmentations of follow-up scans, such data will also be added.

For researchers working with ADNI data, the repository will provide reliable information on a large number of anatomical regions. We envisage that the segmentations be used to search for novel imaging biomarkers of Alzheimer's disease and progressive mild cognitive impairment, using regional volume, shape, and texture information that can be derived. Our data will also enable region-based analysis of the functional imaging data acquired using positron-emission tomography, and of the connectivity data acquired using diffusion tensor imaging in the respective subsets.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2011.03.014.

References

Adachi, T., Saito, Y., Hatsuta, H., Funabe, S., Tokumaru, A.M., Ishii, K., Arai, T., Sawabe, M., Kanemaru, K., Miyashita, A., Kuwano, R., Nakashima, K., Murayama, S., 2010. Neuropathological asymmetry in argyrophilic grain disease. Journal of Neuropathology and Experimental Neurology 69 (7), 737–744 URL http://dx.doi.org/10.1097/NEN.0b013e3181e5ae5c July.

Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V., Rueckert, D., 2009. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. NeuroImage 46 (3), 726–738 URL http://dx.doi.org/10.1016/j.neuroimage.2009.02.018 July.

Ardekani, B.A., Guckemus, S., Bachman, A., Hoptman, M.J., Wojtaszek, M., Nierenberg, J., 2005. Quantitative comparison of algorithms for inter-subject registration of 3D volumetric brain MRI scans. Journal of Neuroscience Methods 142 (1), 67–76 URL http://dx.doi.org/10.1016/j.jneumeth.2004.07.014 March.

Artaechevarria, X., Munoz-Barrutia, A., Oritz-de Solorzano, C., 2009. Combination strategies in multi-atlas image segmentation: application to brain MR data. IEEE transactions on medical imaging URL http://dx.doi.org/10.1109/TMI.2009.2014372 February.

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical Image Analysis 12 (1), 26–41 URL http://dx.doi.org/10.1016/j.media.2007.06.004 February.

Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C., 2010. A reproducible evaluation of ANTs similarity metric performance in brain image registration. NeuroImage URL http://dx.doi.org/10.1016/j.neuroimage.2010.09.025 September.

Barber, R., McKeith, I.G., Ballard, C., Gholkar, A., O'Brien, J.T., 2001. A comparison of medial and lateral temporal lobe atrophy in dementia with Lewy bodies and Alzheimer's disease: magnetic resonance imaging volumetric study. Dementia and Geriatric Cognitive Disorders 12 (3), 198–205 URL http://view.ncbi.nlm.nih.gov/pubmed/11244213.

Barnes, J., Scahill, R.I., Schott, J.M., Frost, C., Rossor, M.N., Fox, N.C., 2005. Does Alzheimer's disease affect hippocampal asymmetry? Evidence from a cross-sectional and longitudinal volumetric MRI study. Dementia and Geriatric Cognitive Disorders 19 (5–6), 338–344 URL http://dx.doi.org/10.1159/000084560.

Barnes, J., Whitwell, J.L., Frost, C., Josephs, K.A., Rossor, M., Fox, N.C., 2006. Measurements of the amygdala and hippocampus in pathologically confirmed Alzheimer disease and frontotemporal lobar degeneration. Archives of Neurology 63 (10), 1434–1439 URL http://dx.doi.org/10.1001/archneur.63.10.1434 October.

Beg, M.F.F., Helm, P.A., McVeigh, E., Miller, M.I., Winslow, R.L., 2004. Computational cardiac anatomy using MRI. Magnetic Resonance in Medicine 52 (5), 1167–1174 URL http://dx.doi.org/10.1002/mrm.20255 November.

Black, S., Gao, F., Bilbao, J., 2009. Understanding white matter disease: imaging–pathological correlations in vascular cognitive impairment. Stroke; A Journal of Cerebral Circulation 40 (3 Suppl) URL http://dx.doi.org/10.1161/STROKEAHA.108.537704 March.

Boccardi, M., Laakso, M.P., Bresciani, L., Galluzzi, S., Geroldi, C., Beltramello, A., Soininen, H., Frisoni, G.B., 2003. The MRI pattern of frontal and temporal brain atrophy in fronto-temporal dementia. Neurobiology of Aging 24 (1), 95–103 URL http://view.ncbi.nlm.nih.gov/pubmed/1249.

Braak, H., Braak, E., 1991. Alzheimer's disease affects limbic nuclei of the thalamus. Acta Neuropathol (Berl) 81 (3), 261–268 URL http://view.ncbi.nlm.nih.gov/pubmed/1711755.

Chan, D., Fox, N.C., Scahill, R.I., Crum, W.R., Whitwell, J.L., Leschziner, G., Rossor, A.M., Stevens, J.M., Cipolotti, L., Rossor, M.N., 2001. Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. Annals of Neurology 49 (4), 433–442 URL http://dx.doi.org/10.1002/ana.92.

Cherubini, A., Péran, P., Spoletini, I., Di Paola, M., Di Iulio, F., Hagberg, G.E.E., Sancesario, G., Gianni, W., Bossù, P., Caltagirone, C., Sabatini, U., Spalletta, G., 2010. Combined volumetry and DTI in subcortical structures of mild cognitive impairment and Alzheimer's disease patients. Journal of Alzheimer's Disease : JAD 19 (4), 1273–1282 URL http://dx.doi.org/10.3233/JAD-2010-091186 January.

Clarkson, M.J., Ourselin, S., Nielsen, C., Leung, K.K., Barnes, J., Whitwell, J.L., Gunter, J.L., Hill, D.L.G., Weiner, M.W., Jack, C.R., 2009. Comparison of phantom and registration scaling corrections using the ADNI cohort. NeuroImage 47 (4), 1506–1513 URL http://dx.doi.org/10.1016/j.neuroimage.2009.05.045 October.

Colliot, O., Chételat, G., Chupin, M., Desgranges, B., Magnin, B., Benali, H., Dubois, B., Garnero, L., Eustache, F., Lehéricy, S., 2008. Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. Radiology 248 (1), 194–201 URL http://dx.doi.org/10.1148/radiol.2481070876 July.

Csernansky, J.G., Wang, L., Joshi, S.C., Ratnanather, J.T., Miller, M.I., 2004. Computational anatomy and neuropsychiatric disease: probabilistic assessment of variation and statistical inference of group difference, hemispheric asymmetry, and time-dependent change. NeuroImage 23 (Suppl 1) URL http://dx.doi.org/10.1016/j.neuroimage.2004.07.025.

Cuénod, C.A., Denys, A., Michot, J.L., Jehenson, P., Forette, F., Kaplan, D., Syrota, A., Boller, F., 1993. Amygdala atrophy in Alzheimer's disease: an in vivo magnetic resonance imaging study. Archives of Neurology 50 (9), 941–945 URL http://dx.doi.org/10.1001/archneur.1993.00540090046009 September.

de Jong, L.W., van der Hiele, K., Veer, I.M., Houwing, J.J., Westendorp, R.G., Bollen, E.L., de Bruin, P.W., Middelkoop, H.A., van Buchem, M.A., van der Grond, J., 2008. Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. Brain : A Journal of Neurology 131 (Pt 12), 3277–3285 URL http://dx.doi.org/10.1093/brain/awn278 December.

Dice, L.R., 1945. Measures of the amount of ecologic association between species. Ecology 26 (3), 297–302 July.

Author's personal copy

Duchesne, S., Caroli, A., Geroldi, C., Barillot, C., Frisoni, G.B., Collins, D.L., 2008. MRI-based automated computer classification of probable AD versus normal controls. IEEE Transactions on Medical Imaging 27 (4), 509–520 URL http://dx.doi.org/10.1109/TMI.2007.908685 April.

Filipek, P.A., Kennedy, D.N., Caviness, V.S., Rossnick, S.L., Spraggins, T.A., Starewicz, P.M., 1989. Magnetic resonance imaging-based brain morphometry: development and application to normal subjects. Annals of Neurology 25 (1), 61–67 URL http://dx.doi.org/10.1002/ana.410250110 January.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cerebral Cortex 14 (1), 11–22 URL http://dx.doi.org/10.1093/cercor/bhg087 January.

Free, S.L., Bergin, P.S., Fish, D.R., Cook, M.J., Shorvon, S.D., Stevens, J.M., 1995. Methods for normalization of hippocampal volumes measured with MR. AJNR. American Journal of Neuroradiology 16 (4), 637–643 URL http://view.ncbi.nlm.nih.gov/pubmed/7611015 April.

Freeborough, P.A., Fox, N.C., Kitney, R.I., 1997. Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans. Computer Methods and Programs in Biomedicine 53 (1), 15–25 URL http://view.ncbi.nlm.nih.gov/pubmed/9113464 May.

Gousias, I.S., Rueckert, D., Heckemann, R.A., Dyet, L.E., Boardman, J.P., Edwards, A.D., Hammers, A., 2008. Automatic segmentation of brain MRIs of 2-year-olds into 83 regions of interest. NeuroImage 40 (2), 672–684 URL http://dx.doi.org/10.1016/j.neuroimage.2007.11.034 April.

Hammers, A., Allom, R., Koepp, M.J., Free, S.L., Myers, R., Lemieux, L., Mitchell, T.N., Brooks, D.J., Duncan, J.S., 2003. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. Human Brain Mapping 19 (4), 224–247 URL http://dx.doi.org/10.1002/hbm.10123 August.

Hammers, A., Heckemann, R.A., Koepp, M.J., Duncan, J.S., Hajnal, J.V., Rueckert, D., Aljabar, P., 2007. Automatic detection and quantification of hippocampal atrophy on MRI in temporal lobe epilepsy: a proof-of-principle study. NeuroImage 36 (1), 38–47 URL http://dx.doi.org/10.1016/j.neuroimage.2007.02.031 May.

Heckemann, R.A., Hajnal, J.V., Aljabar, P., Rueckert, D., Hammers, A., 2006. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. NeuroImage 33 (1), 115–126 URL http://dx.doi.org/10.1016/j.neuroimage.2006.05.061 October.

Heckemann, R.A., Hammers, A., Rueckert, D., Aviv, R.I., Harvey, C.J., Hajnal, J.V., 2008. Automatic volumetry on MR brain images can support diagnostic decision making. BMC Medical Imaging 8, 9 URL http://dx.doi.org/10.1186/1471-2342-8-9 May.

Heckemann, R.A., Keihaninejad, S., Aljabar, P., Rueckert, D., Hajnal, J.V., Hammers, A., 2010. Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation. NeuroImage 51 (1), 221–227 URL http://dx.doi.org/10.1016/j.neuroimage.2010.01.072 May.

Horínek, D., Varjassyová, A., Hort, J., 2007. Magnetic resonance analysis of amygdalar volume in Alzheimer's disease. Current Opinion in Psychiatry 20 (3), 273–277 URL http://dx.doi.org/10.1097/YCO.0b013e3280ebb613 May.

Hsu, Y.-Y., Schuff, N., Du, A.-T., Mark, K., Zhu, X., Hardin, D., Weiner, M.W., 2002. Comparison of automated and manual MRI volumetry of hippocampus in normal aging and dementia. Journal of Magnetic Resonance Imaging : JMRI 16 (3), 305–310 URL http://dx.doi.org/10.1002/jmri.10163 September.

Jaccard, P., 1901. Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. Bulletin de la Société Vaudoise des Sciences Naturelles 37, 241–272.

Jack, C.R., Petersen, R.C., Xu, Y.C., Waring, S.C., O'Brien, P.C., Tangalos, E.G., Smith, G.E., Ivnik, R.J., Kokmen, E., 1997. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology 49 (3), 786–794 URL http://view.ncbi.nlm.nih.gov/pubmed/9305341 September.

Jack, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L Whitwell, J., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. Journal of Magnetic Resonance Imaging : JMRI 27 (4), 685–691 URL http://dx.doi.org/10.1002/jmri.21049 April.

Janota, I., Mountjoy, C.Q., 1988. Asymmetry of pathology in Alzheimer's disease. Journal of Neurology, Neurosurgery, and Psychiatry 51 (7), 1011–1012 URL http://view.ncbi.nlm.nih.gov/pubmed/3204396 July.

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., Macfall, J., Fischl, B., Dale, A., 2006. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. NeuroImage 30 (2), 436–443 URL http://dx.doi.org/10.1016/j.neuroimage.2005.09.046 April.

Keihaninejad, S., Heckemann, R.A., Fagiolo, G., Symms, M.R., Hajnal, J.V., Hammers, A., Alzheimer's Disease Neuroimaging Initiative, 2010. A robust method to estimate the intracranial volume across MRI field strengths (1.5 T and 3 T). NeuroImage 50 (4), 1427–1437 URL http://dx.doi.org/10.1016/j.neuroimage.2010.01.064 May.

Khan, A.R., Wang, L., Beg, M.F.F., 2008. FreeSurfer-initiated fully-automated subcortical brain segmentation in MRI using Large Deformation Diffeomorphic Metric Mapping. NeuroImage 41 (3), 735–746 URL http://dx.doi.org/10.1016/j.neuroimage.2008.03.024 July.

Kittler, J., Hatef, M., Duin, R.P.W., Matas, J., 1998. On combining classifiers. IEEE Trans Pattern Analysis and Machine Intelligence 20 (3), 226–239 URL http://dx.doi.org/10.1109/34.667881.

Klein, A., Hirsch, J., 2005. Mindboggle: a scatterbrained approach to automate brain labeling. NeuroImage 24 (2), 261–280 URL http://dx.doi.org/10.1016/j.neuroimage.2004.09.016 January.

Klein, A., Mensh, B., Ghosh, S., Tourville, J., Hirsch, J., 2005. Mindboggle: automated brain labeling with multiple atlases. BMC Medical Imaging 5 (1), 7 URL http://dx.doi.org/10.1186/1471-2342-5-7 October.

Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage 46 (3), 786–802 URL http://dx.doi.org/10.1016/j.neuroimage.2008.12.037 July.

Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S.J., 2008. Automatic classification of MR scans in Alzheimer's disease. Brain 131 (3), 681–689 URL http://dx.doi.org/10.1093/brain/awm319 March.

Kromer Vogt, L.J., Hyman, B.T., Van Hoesen, G.W., Damasio, A.R., 1990. Pathological alterations in the amygdala in Alzheimer's disease. Neuroscience 37 (2), 377–385 URL http://view.ncbi.nlm.nih.gov/pubmed/2133349.

Lehericy, S., Baulac, M., Chiras, J., Pierot, L., Martin, N., Pillon, B., Deweer, B., Dubois, B., Marsault, C., 1994. Amygdalohippocampal MR volume measurements in the early stages of Alzheimer disease. AJNR. American Journal of Neuroradiology 15 (5), 929–937 URL http://www.ajnr.org/cgi/content/abstract/15/5/929 May.

Likeman, M., Anderson, V.M., Stevens, J.M., Waldman, A.D., Godbolt, A.K., Frost, C., Rossor, M.N., Fox, N.C., 2005. Visual assessment of atrophy on magnetic resonance imaging in the diagnosis of pathologically confirmed young-onset dementias. Archives of Neurology 62 (9), 1410–1415 URL http://dx.doi.org/10.1001/archneur.62.9.1410 September.

Lötjönen, J.M.M., Wolz, R., Koikkalainen, J.R., Thurfjell, L., Waldemar, G., Soininen, H., Rueckert, D., Alzheimer's Disease Neuroimaging Initiative, 2010. Fast and robust multi-atlas segmentation of brain magnetic resonance images. NeuroImage 49 (3), 2352–2365 URL http://dx.doi.org/10.1016/j.neuroimage.2009.10.026 February.

Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S., 2010. Fast free-form deformation using graphics processing units. Computer Methods and Programs in Biomedicine 98 (3), 278–284 URL http://dx.doi.org/10.1016/j.cmpb.2009.09.002 June.

Moossy, J., Zubenko, G.S., Martinez, A.J., Rao, G.R., Kopp, U., Hanin, I., 1989. Lateralization of brain morphologic and cholinergic abnormalities in Alzheimer's disease. Archives of Neurology 46 (6), 639–642 URL http://view.ncbi.nlm.nih.gov/pubmed/2730376 June.

Pedraza, O., Bowers, D., Gilmore, R., 2004. Asymmetry of the hippocampus and amygdala in MRI volumetric measurements of normal adults. Journal of the International Neuropsychological Society 10 (5), 664–678 URL http://dx.doi.org/10.1017/S1355617704105080 September.

Rademacher, J., Galaburda, A.M., Kennedy, D.N., Filipek, P.A., Caviness, V.S., 1992. Human cerebral cortex: localization, parcellation, and morphometry with magnetic resonance imaging. Journal of Cognitive Neuroscience 4 (4), 352–374 URL http://www.mitpressjournals.org/doi/abs/10.1162/jocn.1992.4.4.352.

Riklin-Raviv, T., Van Leemput, K., Menze, B.H., Wells, W.M., Golland, P., 2010. Segmentation of image ensembles via latent atlases. Medical Image Analysis 14 (5), 654–665 URL http://dx.doi.org/10.1016/j.media.2010.05.004 October.

Rohlfing, T., Brandt, R., Menzel, R., Maurer, C.R., 2004. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. NeuroImage 21 (4), 1428–1442 URL http://dx.doi.org/10.1016/j.neuroimage.2003.11.010 April.

Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Transactions on Medical Imaging 18 (8), 712–721 URL http://dx.doi.org/10.1109/42.796284 Aug.

Schuirmann, D.J., 1987. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. Journal of Pharmacokinetics and Biopharmaceutics 15 (6), 657–680 URL http://view.ncbi.nlm.nih.gov/pubmed/3450848 December.

Scott, S.A., DeKosky, S.T., Scheff, S.W., 1991. Volumetric atrophy of the amygdala in Alzheimer's disease: quantitative serial reconstruction. Neurology 41 (3), 351 URL http://www.neurology.org/cgi/content/abstract/41/3/351 March.

Shattuck, D.W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K.L., Poldrack, R.A., Bilder, R.M., Toga, A.W., 2008. Construction of a 3D probabilistic atlas of human cortical structures. NeuroImage 39 (3), 1064–1080 URL http://dx.doi.org/10.1016/j.neuroimage.2007.09.031 February.

Shi, Y., Thompson, P.M., de Zubicaray, G.I., Rose, S.E., Tu, Z., Dinov, I., Toga, A.W., 2007. Direct mapping of hippocampal surfaces with intrinsic shape context. NeuroImage 37 (3), 792–807 URL http://dx.doi.org/10.1016/j.neuroimage.2007.05.016 September.

Shi, F., Liu, B., Zhou, Y., Yu, C., Jiang, T., 2009. Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: meta-analyses of MRI studies. Hippocampus 19 (11), 1055–1064 URL http://dx.doi.org/10.1002/hipo.20573 November.

Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Transactions on Medical Imaging 17 (1), 87–97 URL http://dx.doi.org/10.1109/42.668698 February.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23 (Suppl 1), S208–S219 URL http://dx.doi.org/10.1016/j.neuroimage.2004.07.051.

Author's personal copy

Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System: An Approach to Cerebral Imaging. Thieme Medical Publishers. URL http://www.worldcat.org/isbn/0865772932. January.

Thompson, P.M., Moussai, J., Zohoori, S., Goldkorn, A., Khan, A.A., Mega, M.S., Small, G.W., Cummings, J.L., Toga, A.W., 1998. Cortical variability and asymmetry in normal aging and Alzheimer's disease. Cerebral Cortex 8 (6), 492–509 URL http://view.ncbi.nlm.nih.gov/pubmed/9758213 September.

Traynor, C., Heckemann, R.A., Hammers, A., O'Muircheartaigh, J., Crum, W.R., Barker, G.J., Richardson, M.P., 2010. Reproducibility of thalamic segmentation based on probabilistic tractography. NeuroImage URL http://dx.doi.org/10.1016/j.neuroimage.2010.04.024 April.

Tzourio-Mazoyer, N., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15 (1), 273–289 URL http://dx.doi.org/10.1006/nimg.2001.0978 January.

van der Lijn, F., den Heijer, T., Breteler, M.M., Niessen, W.J., 2008. Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. NeuroImage 43 (4), 708–720 URL http://dx.doi.org/10.1016/j.neuroimage.2008.07.058 December.

Van Essen, D.C., 2005. A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. NeuroImage 28 (3), 635–662 URL http://dx.doi.org/10.1016/j.neuroimage.2005.06.058 November.

Wahlund, L.O., Barkhof, F., Fazekas, F., Bronge, L., Augustin, M., Sjögren, M., Wallin, A., Ader, H., Leys, D., Pantoni, L., Pasquier, F., Erkinjuntti, T., Scheltens, P., European Task Force on Age-Related White Matter Changes, 2001. A new rating scale for age-related white matter changes applicable to MRI and CT. Stroke; A Journal of Cerebral Circulation 32 (6), 1318–1322 URL http://view.ncbi.nlm.nih.gov/pubmed/11387493 June.

Watson, C., Andermann, F., Gloor, P., Jones-Gotman, M., Peters, T., Evans, A., Olivier, A., Melanson, D., Leroux, G., 1992. Anatomic basis of amygdaloid and hippocampal volume measurement by magnetic resonance imaging. Neurology 42 (9), 1743–1750 URL http://view.ncbi.nlm.nih.gov/pubmed/1513464 September.

Whitwell, J.L., Crum, W.R., Watt, H.C., Fox, N.C., 2001. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. AJNR. American Journal of Neuroradiology 22 (8), 1483–1489 URL http://view.ncbi.nlm.nih.gov/pubmed/11559495 September.

Whitwell, J.L., Sampson, E.L., Watt, H.C., Harvey, R.J., Rossor, M.N., Fox, N.C., 2005. A volumetric magnetic resonance imaging study of the amygdala in frontotemporal lobar degeneration and Alzheimer's disease. Dementia and Geriatric Cognitive Disorders 20 (4), 238–244 URL http://dx.doi.org/10.1159/000087343.

Wilcock, G.K., Esiri, M.M., 1987. Asymmetry of pathology in Alzheimer's disease. Journal of Neurology, Neurosurgery, and Psychiatry 50 (10), 1384–1386 URL http://view.ncbi.nlm.nih.gov/pubmed/3681321 October.

Wolf, H., Grunwald, M., Kruggel, F., Riedel-Heller, S.G., Angerhöfer, S., Hojjatoleslami, A., Hensel, A., Arendt, T., Gertz, H., 2001. Hippocampal volume discriminates between normal cognition; questionable and mild dementia in the elderly. Neurobiology of Aging 22 (2), 177–186 URL http://view.ncbi.nlm.nih.gov/pubmed/11182467.

Wolz, R., Aljabar, P., Hajnal, J.V., Hammers, A., Rueckert, D., The Alzheimer's Disease Neuroimaging Initiative, 2010a. LEAP: learning embeddings for atlas propagation. NeuroImage 49 (2), 1316–1325 URL http://dx.doi.org/10.1016/j.neuroimage.2009.09.069 January.

Wolz, R., Heckemann, R.A., Aljabar, P., Hajnal, J.V., Hammers, A., Lötjönen, J., Rueckert, D., Alzheimer's Disease Neuroimaging Initiative, 2010b. Measurement of hippocampal atrophy using 4D graph-cut segmentation: application to ADNI. NeuroImage 52 (1), 109–118 URL http://dx.doi.org/10.1016/j.neuroimage.2010.04.006 August.

Zarei, M., Patenaude, B., Damoiseaux, J., Morgese, C., Smith, S., Matthews, P.M., Barkhof, F., Rombouts, S.A., Sanz-Arigita, E., Jenkinson, M., 2010. Combining shape and connectivity analysis: an MRI study of thalamic degeneration in Alzheimer's disease. NeuroImage 49 (1), 1–8 URL http://dx.doi.org/10.1016/j.neuroimage.2009.09.001 January.