Journal of Neuroscience Methods 227 (2014) 35–46

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Computational Neuroscience

Automated algorithm to measure changes in medial temporal lobe volume in Alzheimer disease☆

Samaneh Kazemifar[a,b], John J. Drozd[a], Nagalingam Rajakumar[c], Michael J. Borrie[d,e], Robert Bartha[a,b,*], for the Alzheimer's Disease Neuroimaging Initiative

[a] Robarts Research Institute, University of Western Ontario, 1151 Richmond Street, London, Ontario, Canada N6A 3K7
[b] Department of Medical Biophysics, University of Western Ontario, 1151 Richmond Street, London, Ontario, Canada N6A 3K7
[c] Department of Anatomy and Cell Biology, University of Western Ontario, 1151 Richmond Street, London, Ontario, Canada N6A 3K7
[d] Department of Medicine, University of Western Ontario, 1151 Richmond Street, London, Ontario, Canada N6A 3K7
[e] Division of Aging, Rehabilitation and Geriatric Care, Lawson Health Research Institute, 268 Grosvenor Street, London, Ontario, Canada N6A 4V2

HIGHLIGHTS

- Fully automated medial temporal lobe (MTL) atrophy measurement.
- Significant MTL atrophy over 24 months in NEC, MCI, and AD groups.
- A significant difference in MTL atrophy detected between MCI and AD groups.
- A significant association between MTL atrophy and decline in cognition.
- Volume changes in MTL may help predict AD progression.

ARTICLE INFO

Article history:
Received 4 December 2013
Received in revised form 30 January 2014
Accepted 31 January 2014

Keywords:
Hippocampus
Medial temporal lobe segmentation
MRI
Shape
Multi-atlas
Alzheimer disease

ABSTRACT

Background: The change in volume of anatomic structures is as a sensitive indicator of Alzheimer disease (AD) progression. Although several methods are available to measure brain volumes, improvements in speed and automation are required. Our objective was to develop a fully automated, fast, and reliable approach to measure *change* in medial temporal lobe (MTL) volume, including primarily hippocampus.
Methods: The MTL volume defined in an atlas image was propagated onto each baseline image and a level set algorithm was applied to refine the shape and smooth the boundary. The MTL of the baseline image was then mapped onto the corresponding follow-up image to measure volume change (ΔMTL). Baseline and 24 months 3D T_1-weighted images from the Alzheimer Disease Neuroimaging Initiative (ADNI) were randomly selected for 50 normal elderly controls (NECs), 50 subjects with mild cognitive impairment (MCI) and 50 subjects with AD to test the algorithm. The method was compared to the FreeSurfer segmentation tools.
Results: The average ΔMTL (mean ± SEM) was $68 \pm 35\,mm^3$ in NEC, $187 \pm 38\,mm^3$ in MCI and $300 \pm 34\,mm^3$ in the AD group and was significantly different ($p < 0.0001$) between all three groups. The ΔMTL was correlated with cognitive decline.
Comparison with existing method(s): Results for the FreeSurfer software were similar but did not detect significant differences between the MCI and AD groups.
Conclusion: This novel segmentation approach is fully automated and provides a robust marker of brain atrophy that shows different rates of atrophy over 2 years between NEC, MCI, and AD groups.

☆ Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). 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 complete listing of ADNI investigators can be found at: http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

* Corresponding author at: Robarts Research Institute, University of Western Ontario, 1151 Richmond Street, London, Ontario, Canada N6A 3K7. Tel.: 1 519 663 5777x24039; fax: +1 519 931 5224.
E-mail address: rbartha@robarts.ca (R. Bartha).

0165-0270/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jneumeth.2014.01.033

1. Introduction

Alzheimer disease (AD) is the most common form of neurodegenerative dementia (Berchtold and Cotman, 1998) and progressively causes metabolic and structural changes in the brain that lead to symptoms of cognitive decline. In 2006, approximately 14.7% of people over the age of 85 years were living with AD worldwide. By 2050, the worldwide prevalence of AD is projected to quadruple compared to 2006 levels (Brookmeyer et al., 2007). The current diagnosis of AD is made clinically incorporating neuropsychological and neuroimaging assessments. The clinical progression of AD can be measured by cognitive assessments, most commonly the Mini-Mental-State Examination (Folstein et al., 1975) (MMSE), the Alzheimer Disease Assessment Scale-cognitive subscale (ADAS-cog) (Rosen et al., 1984) and more recently the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005). However, these assessments have only moderate sensitivity to disease progression (Anthony et al., 1982; Rosen et al., 1984), particularly in the early stages of the disease. Therefore, numerous non-invasive imaging techniques (Azari et al., 1993; Detoledomorrell and Morrell, 1993; Ikeda et al., 1994) are being developed as potential quantitative biomarkers for the detection and monitoring of small structural (Jack et al., 2004; Nestor et al., 2008), metabolic (Adalsteinsson et al., 2000; Dixon et al., 2002; Penner et al., 2010; Rupsingh et al., 2011), and functional (Miller et al., 1993) changes during the early stages of the disease. In addition, the development of amyloid imaging techniques using positron emission tomography (PET), now provides the ability to measure fibrillar amyloid accumulation in the brain (Furst and Lal, 2011; Klunk et al., 2004; Villain et al., 2012). These imaging tools could increase diagnostic accuracy and allow monitoring of disease progression and response to treatment (Nestor et al., 2008; Sperling et al., 2012; Whitwell et al., 2012).

Magnetic resonance imaging (MRI) is used to visualize anatomical changes in the brain with high resolution and is therefore considered the preferred neuroimaging method for early detection of AD (Kantarci and Jack, 2004). Progression of AD (Anderson et al., 2005; Atiya et al., 2003; Chetelat and Baron, 2003; Kantarci and Jack, 2003) is often associated with increased ventricle volume (Nestor et al., 2008) or atrophy of the medial temporal lobe (MTL) including the hippocampus (Atiya et al., 2003; Chetelat and Baron, 2003; Clerx et al., 2013; Kantarci and Jack, 2003). Although the hippocampus is considered one of the first brain structures affected by AD, accurate quantification of hippocampal atrophy is difficult because the structure has an inhomogeneous texture with poorly defined boundaries. Manual segmentation of the hippocampus on MRI is the gold standard (Geuze et al., 2005) but is time-consuming, subjective, and has high variability (Achten et al., 1998). In fact, only recently have attempts been made to standardize the definition of the hippocampus for volumetric measurements (Boccardi et al., 2011).

Several studies have proposed atlas-based segmentation techniques to find the optimal mappings between brain MRI and an atlas image using landmarks (Csernansky et al., 1998; Hogan et al., 2000; Hsu et al., 2002) or intensities (Chupin et al., 2009; Coupe et al., 2011; Gee et al., 1993; Gosche et al., 2001; Nestor et al., 2012; Wolz et al., 2010a). In addition, van der Lijn et al. (2008) combined atlas-based segmentation with graph cuts (Boykov et al., 2001; Greig et al., 1989) to define the hippocampus. Incorporating corresponding landmarks can also help to compute geometric transformations rather than relying on intensity-based registration. For example, an atlas-based registration method was applied by Wang et al. (2003) to segment the hippocampus; however, this method required the manual placement of a large number of landmarks in each scan. More recently, atlas-based registration has been proposed for the detection of pathological changes in addition to the segmentation of structures (Klein et al., 2009). Registration is treated as an optimization problem to find the spatial mapping between two images. For example, to segment the hippocampus and amygdala in images obtained from the ADNI dataset, Chupin et al. (2009) registered images to a probabilistic atlas with image deformation performed by minimization of an energy function (Chupin et al., 2007). In general, providing an accurate segmentation of a structure using atlas-based registration is important; however, the accuracy of the segmentation is heavily dependent on the atlas. A single atlas cannot represent a whole population, particularly in datasets with large variation and may lead to biased results. Therefore, the performance of atlas-based segmentation methods can often be improved using multi-atlas propagation (Heckemann et al., 2006; Rohlfing et al., 2004; Warfield et al., 2004).

A number of other methods have also been developed and applied to the segmentation of the hippocampus. Barnes et al. (2004) developed a semi-automated segmentation method to measure hippocampal volume changes by calculating the boundary shift integral (BSI) (Freeborough and Fox, 1997) in longitudinal MR images. Rusinek et al. (2003) applied the BSI to calculate the rate of MTL atrophy. Wolz et al. (2010a) combined a multi-atlas registration with an intensity refinement model (van der Lijn et al., 2008) to segment the hippocampus. This learning embedding for atlas propagation (LEAP) method was demonstrated by using images obtained from the ADNI database (Wolz et al., 2010a). Coupe et al. (2011) proposed a novel patch-based method for hippocampus and ventricle segmentation that incorporated nonlocal means label fusion (Buades et al., 2005). Fischl et al. (2002, 2004) have also proposed an automated tool (FreeSurfer) for segmentation of 37 neuroanatomical structures in the brain on MRI that includes 18 labels of subcortical structures. In this approach, a probabilistic atlas image is generated from a training set of 41 manually labeled brains. For each image to be segmented, the probability that a pixel belongs to a specific structure (label class) is computed. Preprocessing steps included an affine registration to Talairach space, intensity normalization, skull stripping, and nonlinear alignment to the Talairach atlas (Fischl et al., 2002, 2004). An initial segmentation is generated by assigning each pixel to the class for which the probability is greatest. Given this segmentation, a neighborhood function is used to recompute the class probabilities. The data set is then resegmented based on this new set of class probabilities. This process is repeated until the segmentation does not change. FreeSurfer is a fully automated segmentation pipeline (Fischl et al., 2002, 2004). Similarly, Shen et al. (2002) used a surface model for segmentation of the hippocampus. This model integrated rigid geometric alignment with shape variation as prior knowledge, but required manual placement of landmarks on the boundaries. The prior knowledge increased the accuracy of the boundaries and maintained the appropriate hippocampus shape, but manual placement of points within an individual MR image was needed.

Existing methods of hippocampal segmentation mostly require manual intervention at some point in the measurement to achieve accurate results (Csernansky et al., 2004; Hogan et al., 2000; Wang et al., 2003). However, in applying such methods to the monitoring of Alzheimer's disease progression, we postulate that segmentation accuracy is less important than sensitivity to change. Automation may be a more highly desirable feature because it increases reproducibility and aids in analyzing large datasets in a timely manner. Automated measurements of large volumes such as frontal, parietal, temporal, and occipital lobes as well as cerebellum have been successfully demonstrated by Andreasen et al. (1996). In addition, subcortical structures such as the thalamus, caudate nucleus, putamen and globus pallidus have also been measured (Iosifescu et al., 1997) using deformable shape models based on automated or semi automated image segmentation algorithms (Yang and Duncan,

2004). By formulating a shape-intensity model in terms of a level set function (Chan and Vese, 2001) it is possible to segment images with poor contrast and missing boundaries. Level sets are a type of geometric deformable model (Osher and Sethian, 1988). Intensity-based techniques are difficult to apply to the segmentation of the hippocampus because signal intensity varies and boundaries are ill-defined. Deformable models can be used to extract boundaries from objects with low contrast and indistinct edges. In this study, a geometric deformable model (level set algorithm (Sethian and Sethian, 1999)) was chosen that could automatically handle multiple objects or an object with unknown topology and sharp corners. The Fast Marching segmentation method is a type of level set method that can follow objects with topology changes when an object splits in two, develops holes, or the reverse of these operations. This method may offer an advantage when segmenting structures with low contrast boundaries such as the hippocampus.

Therefore, the goal of the present work was to develop a robust, reliable, and fully automated approach to quantify MTL atrophy using MRI. The novel segmentation approach presented here combined a multiple atlas-based registration method to define the tissue region of interest, a level set algorithm to refine the shape of the medial temporal lobe, and registration of baseline images to follow up images to increase measurement precision. The method was tested by comparing atrophy between a group of normal elderly controls (NECs), subjects with mild cognitive impairment (MCI), and subjects with AD. We hypothesized that the change in MTL (ΔMTL) volume would be robustly quantified by this fully automated segmentation method, and would be a sensitive marker of Alzheimer disease progression. We selected 24 months as the longitudinal interval, as this interval has been previously used for measuring disease progression in AD (Jack et al., 2011).

2. Methods

The automated segmentation approach developed in this study represents an expedient atrophy change in tissue (EXACT) measurement of the brain. Briefly, EXACT begins with the use of the expectation maximization (EM) algorithm to remove the skull and non-brain tissue from each image (Dempster et al., 1977), followed by a multiple atlas-based registration to identify the MTL volume of interest in a subject's baseline image, and a Fast Marching algorithm (Sethian and Sethian, 1999) (a type of level set algorithm) to refine the boundary of the MTL using image features such as the gradient. Our definition of the region of interest does not correspond to a single anatomical structure specially, but instead is guided by anatomical structures and refined using clearly identifiable tissue boundaries observed in T_1-weighted images. In contrast, trying to find the boundaries of the hippocampus, a structure with low contrast compared to surrounding tissue, is difficult (Chupin et al., 2007; Siadat et al., 2007). In addition, inclusion of other structures within and beyond the hippocampus may provide increased sensitive to early AD pathology (Dubois et al., 2007). Skull stripping and removal of non-brain tissue is repeated in the subject's follow-up images. Then, baseline images are registered to skull stripped follow-up images. Following registration, the transformation field is used to map the MTL volume from the baseline image to the follow-up image. The measured volume change in the MTL is the difference between the baseline and follow-up volume. A schematic representation of the entire automated segmentation procedure is provided in Fig. 1.

The EM segmentation is a model-based classification method that uses a Bayesian classifier (Williams and Barber, 1998). The EM segmentation algorithm classifies the brain image into cerebrospinal fluid (CSF), gray matter (GM) and white matter (WM) according to a probability distribution and the pixel intensity histogram (Pohl et al., 2007). It was chosen as a pre-processing step to extract the whole brain. Then the whole brain gray scale images of the atlas, baseline, and follow-up images were used in the registration step. The EM module from 3DSlicer (Gering et al., 1999; Pieper et al., 2004, 2006) (http://www.slicer.org) was used in this study. 3DSlicer is open-source software and based on the VTK (Visualization Toolkit: http://www.vtk.org) and ITK (Insight Segmentation and Registration Toolkit: http://www.itk.org) libraries. Although the EM segmentation method within 3DSlicer is one example of a skull stripping method, others could also be used. This tool is accessible to the research community and easy to use. We chose this method of skull stripping because it is automated and does not require parameter optimization. Other available methods (e.g. BET (Brain Extraction Tool) of FSL (Functional Magnetic Resonance Imaging of the Brain Software Library) (Smith, 2002)) require manual interaction to optimize parameters in some subjects if neck and non-brain structures remain.

The EXACT method requires the rough outline of the MTL region (using a bounding box) a priori in each *atlas* by an expert. This step is performed only once. Eight atlases were created using subject images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. These datasets were chosen visually because images had no visible rotation and medial temporal lobe atrophy spanned the range normally encountered in the brain in normal elderly controls ($N=3$), subjects with MCI ($N=2$) and subjects with AD ($N=3$). Our volume of interest (VOI) does not correspond to a specific anatomical structure. Rather the location is guided by the anatomical structure of interest but it is defined by clearly identifiable tissue boundaries that are observed in T_1-weighted images. The purpose of our registration step is simply to establish a bounding box around the MTL region that contains the hippocampus. Then, subsequently a Fast Marching method is used to identify a gray matter VOI. Accurate segmentation of the hippocampus was not the goal. Rather the goal is definition of a VOI with clearly defined boundaries. In this way, tracking atrophy over time is more precise.

To quantify the ΔMTL volume between two time points (e.g. a baseline and follow-up image) several steps were developed in ITK (Insight Segmentation and Registration Toolkit: http://www.itk.org) (Ibanez et al., 2005; Yoo et al., 2002). First the skull was removed in all eight atlases, as well as in the baseline and follow-up images using the EM segmentation algorithm. Then, the atlas that most closely matched the baseline image was selected by registering each atlas to the baseline image and choosing the atlas that maximized mutual information (MMI) (Studholme et al., 1999) within the local MTL region. The optimum atlas image was then aligned to the baseline brain image. The alignment was performed using affine transformations and the Demon's registration method (Thirion, 1998). The bounding box encompassing the MTL region defined within the optimum atlas image was propagated into the baseline image using the deformation matrix. A Fast Marching algorithm (Ibanez et al., 2005) was then applied to refine the shape and smooth the boundary of the MTL VOI. An initialization mask around the region of interest was delineated. A single seed point was initiated inside the region of interest that was propagated outwards until it reached the desired boundary. The second step involved mapping the MTL that was defined on the baseline image to the follow-up image. This propagation was accomplished by registering the baseline image to its corresponding follow-up image to produce the required transformation matrix and deformation field. The defined MTL in the baseline image was then propagated to the follow-up image using the deformation field. The difference between the volume of the MTL in the baseline image and follow-up image represented the atrophy that occurred between these time points. The technique was initially tested using a single image atlas,

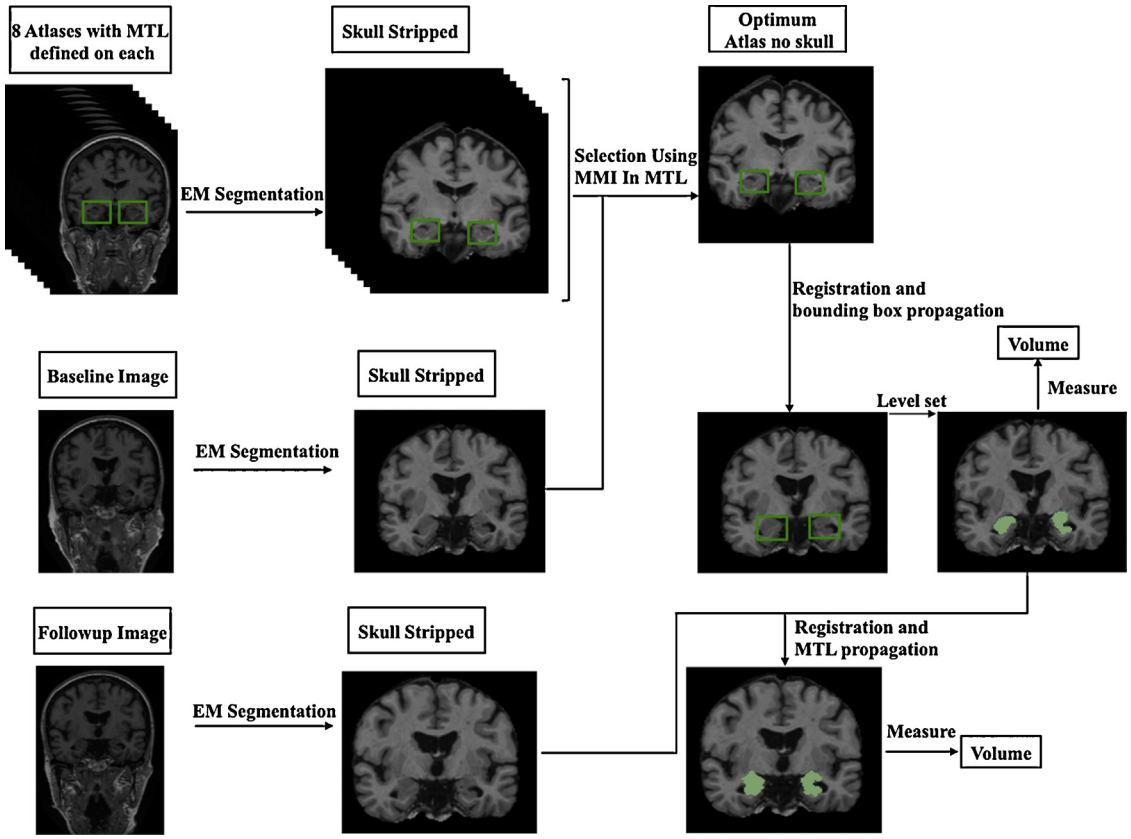

Fig. 1. Block diagram of the EXACT segmentation algorithm.

however we found that the bounding box failed to completely contain the hippocampus in some instances. Since public atlases are often created using the average of multiple healthy brains, such atlases are also not optimal for our application. Instead, we utilized eight image sets that showed a wide range of anatomical variance and atrophy (healthy brain, MCI, and AD). Using this approach we successfully labeled a relevant MTL volume in all subjects, which we confirmed by visual inspection. In a subset of NEC ($N = 10$), MCI ($N = 10$), and AD ($N = 10$) subjects a method of label fusion was also tested based on a majority voting strategy (Hansen and Salamon, 1990; Kittler, 1998). Eight label maps were independently produced using atlas-based registration to define the bounding box followed by Fast Marching segmentation to define the VOI. Finally the VOI was refined by label fusion where majority voting was used to determine whether a voxel should be included in the final VOI.

It should be noted that the smoothness of the edges of the VOI will depend on the objective function used to define the VOI, which incorporates a data term and a regularity term. The data term may include region-based and boundary-based terms. The regularization term is used to smooth the data term. The regularization term of the Diffeomorphic Demons registration (Vercauteren et al., 2009, Appendix A) is a convolution of a Gaussian kernel and spatial transformation (vector field) to smooth this vector field. In contrast, the Fast Marching method (a type of level-set method, Appendix B) does not use a regularization term. However, prior to application of the Fast Marching method, a filter is applied to the image (curvature anisotropic diffusion image filter (Ibanez et al., 2005)), which smoothes the noise and enhances the edges. This approach was used to increase sensitivity to minor volumetric changes over time.

2.1. Study subjects

Data used in the preparation of this article were obtained from the ADNI database (adni.loni.ucla.edu). The 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 primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment and early Alzheimer's disease. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, MD, 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 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55–90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

Table 1
Demographic information for study participants. M: male, SD: standard deviation, SEM: standard error of mean.

Diagnosis	Software	NEC	MCI	AD
Number of subject (*N*)		50	50	50
Age (Mean ± SD)		76.3 ± 5.0	74.5 ± 7.5	76.5 ± 6.4
Sex (M)		27	30	23
Baseline MMSE (mean ± SD)		29.1 ± 1.0	27.0 ± 1.7	23.3 ± 1.8
24 months follow-up MMSE (mean ± SD)		29.2 ± 0.9	25.1 ± 3.8	19.6 ± 5.4
ΔMMSE (mean ± SD)		−0.04 ± 1.1	1.8 ± 3.1	3.7 ± 5.0
Baseline ADAS (mean ± SD)		5.8 ± 3.3	11.2 ± 4.5	18.0 ± 5.7
24 months follow-up ADAS (mean ± SD)		5.6 ± 3.2	13.7 ± 6.7	25.0 ± 10.9
ΔADAS (mean ± SD)		0.19 ± 3.3	−2.5 ± 4.9	−7.1 ± 8.4
Baseline MTL volume (mean ± SEM) mm^3	EXACT	8569.8 ± 136.1	7773.6 ± 243.1	7161.5 ± 231.1
24 months MTL follow-up volume (mean ± SEM) mm^3	EXACT	8501.6 ± 135.1	7586.4 ± 245.3	6861.1 ± 226.0
Baseline hippocampus volume (mean ± SEM) mm^3	FreeSurfer	6599.7 ± 104.0	5896.1 ± 147.4	4993.3 ± 143.0
24 months hippocampus follow-up volume (mean ± SEM) mm^3	FreeSurfer	6441.2 ± 111.4	5481.0 ± 153.0	4589.9 ± 141.3

We randomly selected 1.5 T TT_1-weighted volumetric MRI images at baseline and 24 months in 50 NEC, 50 subjects with MCI, and 50 subjects with AD from the ADNI database to test the algorithm. All images were acquired using a sagittal 3D magnetization-prepared rapid acquisition with gradient echo MP-RAGE sequence with pixel size 0.94 mm × 0.94 mm × 1.2 mm; flip angle ∼8°; TE ∼4 ms; TR ∼9 ms; TI 1000 ms; matrix, 256 × 256; 166 slices (Jack et al., 2008). Additional image preprocessing included geometric distortion correction, bias field correction and geometrical scaling. All subjects had the MMSE, Logical Memory 1 (LM) exam (Wechsler, 1981) and ADAS-cog (https://ida.loni.ucla.edu/) cognitive assessments at baseline and 24 months interval. The MMSE scores were between 24 and 30 for NEC and MCI subjects while the MMSE scores for AD subjects were between 20 and 26. The measured change in MMSE and ADAS-cog is the difference between the baseline and follow-up scores. The major objective of ADNI is to provide a generally accessible data source for studying longitudinal changes in brain structure and metabolism.

2.2. MTL segmentation

Quantification of ΔMTL volume between baseline and 24 months was performed in each subject using the EXACT measurement summarized in Fig. 1. The results were compared to a measure of hippocampal atrophy provided in the ADNI database produced by FreeSurfer (http://surfer.nmr.mgh.harvard.edu/). The technical details of the FreeSurfer procedure were previously published (Fischl et al., 2002), and the complete process for hippocampus volume measurements using the longitudinal processing pipeline in ADNI can be found online (http://www.loni.ucla.edu/twiki/bin/view/ADNI/). The hippocampal volume measurements made using FreeSurfer version 4.4 were provided to ADNI by the UCSF medical center. An expert (N.R.) also manually segmented the hippocampus in a subset of data ($N = 10$ NEC, $N = 10$ MCI, and $N = 10$ AD) to compare hippocampal volumes to the volume of tissue included by EXACT. The Dice index (Dice, 1945) was calculated using the baseline images for each group.

2.3. Statistical analysis

One-way analysis of variance (ANOVA) was performed to compare ΔMTL volume between the NEC, MCI and AD groups. The association between the ΔMTL from the EXACT measurement and change in MMSE, LM and ADAS-cog scores were evaluated using linear regression. The Pearson correlation coefficients were calculated between the volume changes in MTL from the EXACT measurement and the volume changes in hippocampus using the FreeSurfer software. Prism GraphPad (Prism, version 5.00; GraphPad Software, San Diego, CA) was used for the statistical analyses. The receiver operating characteristic (ROC) was also used to calculate the classification rate between two groups using the EXACT measurement and the FreeSurfer analysis. The ROC curve analysis was performed using Prism GraphPad.

3. Results

Demographic characteristics are provided in Table 1 for all 150 subjects. There was no difference between group mean ages. The MTL volume (mean ± SEM, mm^3) for the baseline and the follow-up images in the NEC, MCI and AD groups are provided in Table 1. The segmentation of the MTL in one subject using the EXACT method is shown in Fig. 2 in the coronal, axial, and sagittal planes along with a 3D rendering. Similarly, the segmentation of the hippocampus in the same subject using FreeSurfer is shown in Fig. 3. The manual segmentation of the hippocampus (Fig. 4) showed that the hippocampus accounts for only about 40% of the total VOI included by

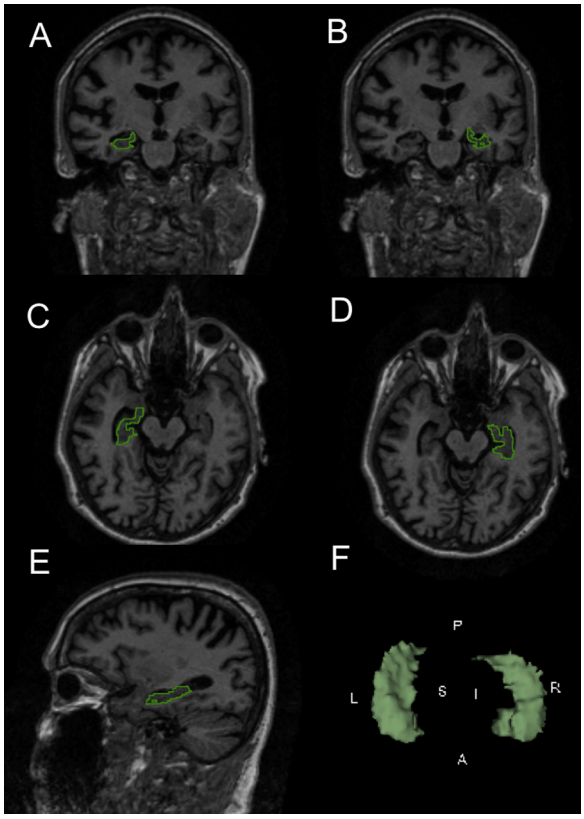

Fig. 2. An example segmentation of the MTL using the EXACT method in one subject shown in coronal (A/B), axial (C/D) and sagittal (E) orientations. The 3D rendering of the MTL is shown in (F).

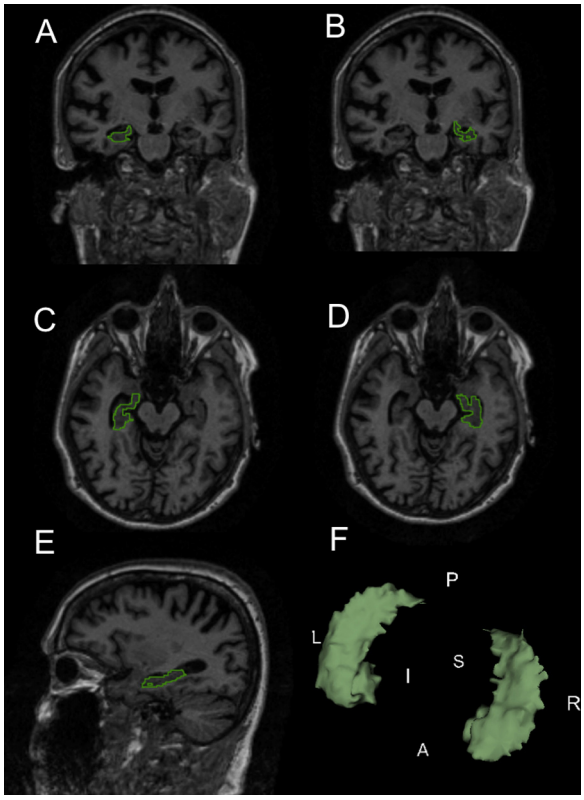

Fig. 3. An example segmentation of the hippocampus using the FreeSurfer method in same subject shown in coronal (A/B), axial (C/D) and sagittal (E) orientations. The 3D rendering of the hippocampus is shown in (F).

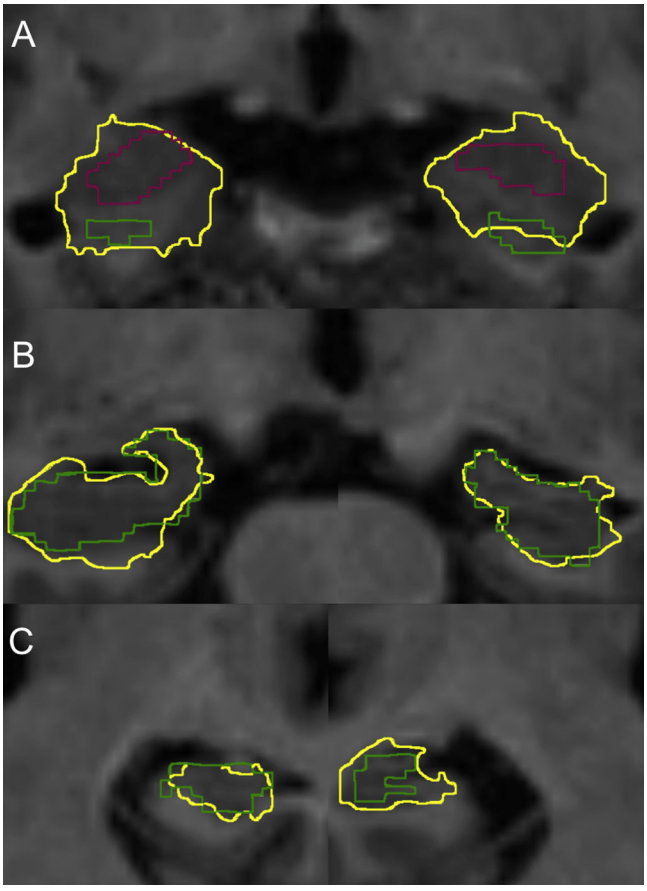

Fig. 4. An example manual segmentation of hippocampus, amygdala and corresponding MTL segmentation using the EXACT method in anterior (A), middle (B) and posterior views (C). Amygdala (purple), hippocampus (green), and MTL (yellow) are outlined separately. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

EXACT. Specifically, the Dice index was 0.43 in the AD group, 0.39 for the MCI group, and 0.45 for the NEC group. Using an Intel Core 2 Duo processor (3.06 GHz) on an iMac10.1 desktop computer, the total processing time to determine change in MTL volume was ~5 h using the EXACT method. Fig. 5 shows the average volume changes over 24 months for the MTL measured by the EXACT method and the change in the hippocampus measured by FreeSurfer in all groups. The average volume change measured using the EXACT method was significantly smaller than that measured using FreeSurfer in NEC ($p = 0.03$), MCI ($p = 0.001$), and AD ($p = 0.003$). The one-way ANOVA indicated that there was a significant difference between three groups using the EXACT method ($p < 0.0001$) and using the FreeSurfer software ($p < 0.0001$). An unpaired t-test showed that using the EXACT method, there was a significant difference in MTL volume changes between the NEC and MCI subject groups ($p = 0.02$), between the NEC and AD subject groups ($p < 0.0001$), and between the MCI and AD subject groups ($p = 0.02$). Similarly, an unpaired t-test of the FreeSurfer data also showed a significant difference in hippocampus between the NEC and MCI subject groups ($p = 0.0003$), and a significant difference between the NEC and AD subject groups ($p < 0.0001$). But no significant difference was detected between the MCI and AD subject groups.

Fig. 6 shows a significant correlation ($r = 0.43$, $p < 0.0001$) between the volume changes in the MTL using the EXACT method and volume change of the hippocampus using FreeSurfer including data for all groups. Fig. 7 shows a significant association between the volume changes in the MTL measured by the EXACT

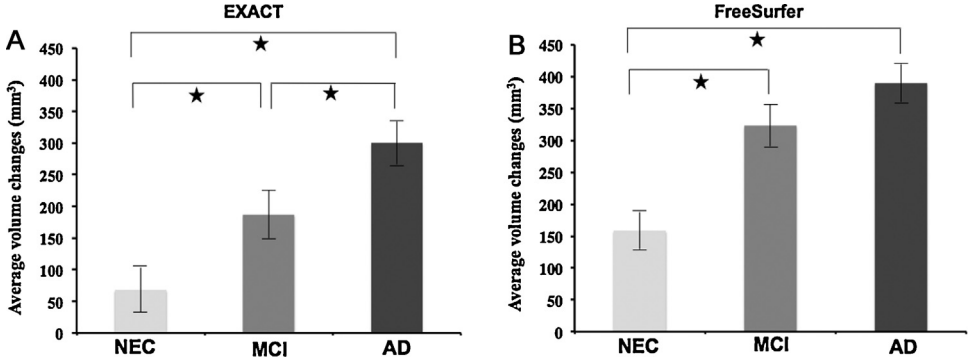

Fig. 5. Average volume changes in three groups were measured using the EXACT (A) and the FreeSurfer (B) methods. The error bars represent the standard error of the mean and asterisks show significant differences between groups ($p < 0.05$).

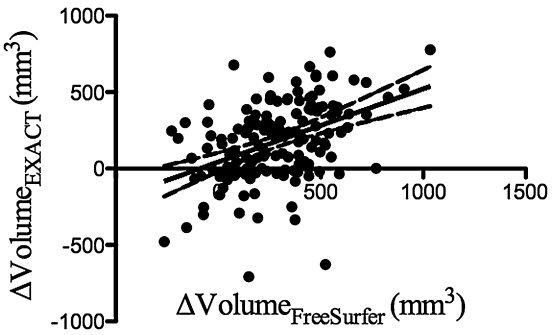

Fig. 6. Association between volume changes in MTL using the EXACT method and hippocampus volume change measured using the FreeSurfer method ($r = 0.43$, $p < 0.0001$). The 95% confidence intervals for the regressions are shown as dashed lines.

and the MMSE score ($r = 0.24$, $p = 0.003$), and between the hippocampus measured by FreeSurfer and the MMSE score ($r = 0.39$, $p < 0.0001$). Fig. 8 shows a significant association between the volume changes in the MTL measured by EXACT and the ADAS-cog changes ($r = -0.25$, $p = 0.002$), and between the hippocampus measured by FreeSurfer and the ADAS-cog changes ($r = -0.34$, $p < 0.0001$). A significant correlation between the volume changes in the MTL measured by EXACT and the 24 months LM score ($r = -0.39$, $p < 0.0001$), and between the hippocampus measured by FreeSurfer and the 24 months LM score ($r = -0.51$, $p < 0.0001$) is shown in Fig. 9.

A comparison of the VOI defined by the EXACT method and the VOI defined when incorporating label fusion by majority voting is shown in Fig. 10. With the label fusion approach the average MTL volume change was 68.0 mm^3 greater across all groups compared to the EXACT method ($N = 10$). With label fusion, the average

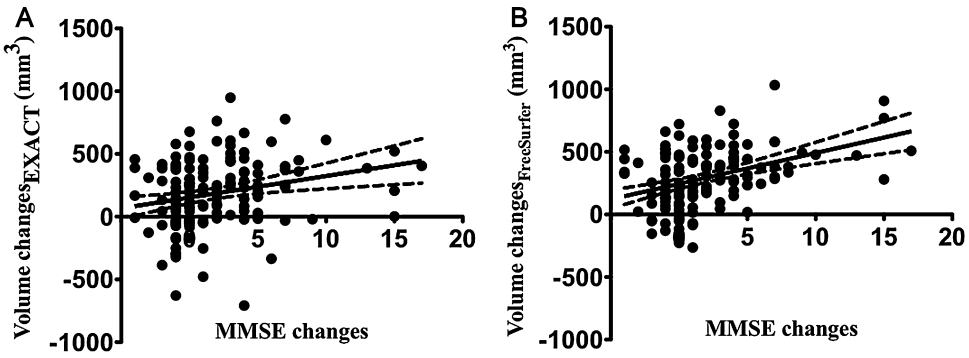

Fig. 7. Association between the volume changes in MTL using the EXACT (A) and the MMSE score ($r = 0.24$, $p = 0.003$). Relationship between the volume changes in hippocampus using FreeSurfer (B) and the MMSE changes score ($r = 0.39$, $p < 0.0001$). The 95% confidence intervals for the regressions are shown as dashed lines.

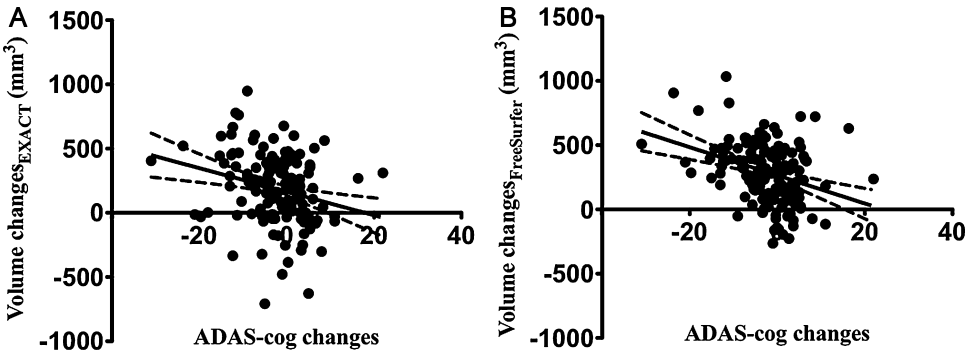

Fig. 8. Association between the volume changes in MTL using the EXACT (A) and the ADAS-cog changes ($r = -0.25$, $p = 0.002$). Relationship between the volume changes in hippocampus using FreeSurfer (B) and the ADAS-cog changes ($r = -0.34$, $p < 0.0001$). The 95% confidence intervals for the regressions are shown as dashed lines.

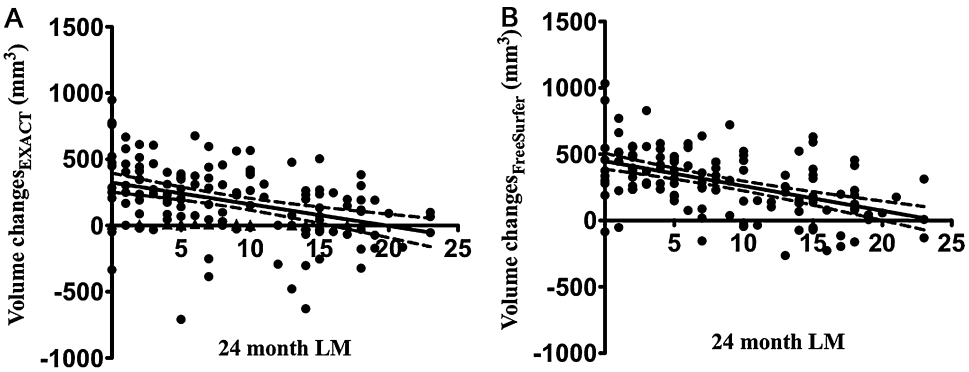

Fig. 9. Association between the volume changes in MTL using the EXACT (A) and the 24 month LM score ($r = -0.39$, $p < 0.0001$). Relationship between the volume changes in hippocampus using FreeSurfer (B) and the 24 month LM score ($r = -0.51$, $p < 0.0001$). The 95% confidence intervals for the regressions are shown as dashed lines.

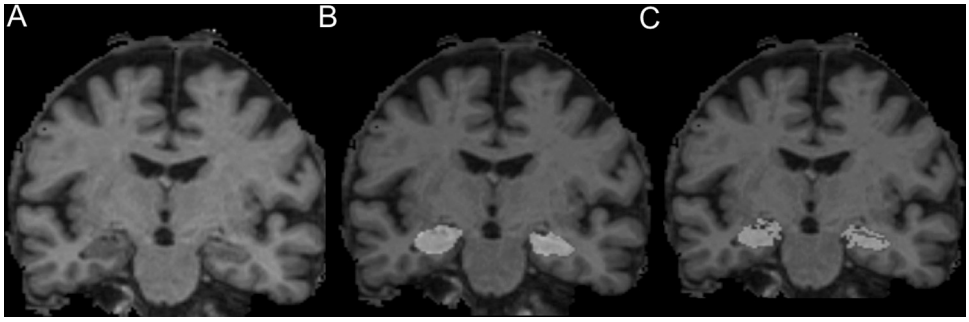

Fig. 10. An example T_1-weighted coronal MRI (A) following segmentation of the MTL using the EXACT method (B) and the label fusion method (C) in the same subject shown.

(±SEM) ΔMTL was 54.9 (±113.7) mm^3 in NEC, 380.0 (±64.3) mm^3 in MCI, and 384.0 (±74.3) mm^3 in the AD groups (ND = 10). Using the EXACT method in the same subject, the average ΔMTL was −49.5 (±75.2) mm^3 in NEC, 295.2 (±55.0) mm^3 in MCI, and 405.4 (±88.2) mm^3 in the AD groups. In the subset of individuals studied, the average volume change measured using the EXACT and label fusion methods were not significantly different in NEC ($p > 0.05$), MCI ($p > 0.05$), and AD ($p > 0.05$) groups. It should be noted that both the FreeSurfer and EXACT methods did measure an apparent increase in volume in some individuals (Figs. 6–9). A small increase in tissue volume may occur in some subjects due to an increase in tissue hydration (Streitburger et al., 2012), or more likely the small observed increases might be due to variability in image acquisition and post processing. A one-way ANOVA indicated that there was a significant difference between the three groups using the EXACT method ($p < 0.001$) and using the label fusion method ($p < 0.01$). An unpaired *t*-test showed that using the EXACT method, there was a significant difference in MTL volume change between the NEC and MCI groups ($p = 0.004$), and between the NEC and AD groups ($p = 0.001$). Similarly, an unpaired *t*-test of the label fusion volumes also showed a significant difference in MTL between the NEC and MCI groups ($p = 0.02$), and a significant difference between the NEC and AD groups ($p < 0.05$). In this subset of subjects neither method detected a difference between the MCI and AD groups.

A Bland–Altman plot was used to examine the agreement between the EXACT and FreeSurfer methods. The Bland–Altman plot in Fig. 11 shows the difference between the EXACT and FreeSurfer volume change measurements as a function of their average. The mean bias (108.4) and the 95% limit of agreement (265.8) computed as the mean bias ± SD of the difference between the two techniques are also shown. The results indicate that the FreeSurfer technique calculated larger changes in volume than the EXACT method. The ROC curve analysis between two groups using the EXACT method and the FreeSurfer software showed that the ability to classify subjects between NEC and AD groups with EXACT was slightly lower (76%) compared to FreeSurfer (80%). But, the ability to classify subjects between MCI and AD groups with EXACT was slightly higher (60%) compared to FreeSurfer (57%). The Z distribution indicates that there was no significant difference between the areas under the two ROC curves of NEC vs AD and MCI vs AD ($p > 0.05$).

4. Discussion

A novel fully automated method for segmentation of the MTL (EXACT) was developed that combined multiple atlas-based registration with a Fast Marching algorithm. The EXACT method was used to measure the change in MTL volume for 50 NEC, 50 subjects with MCI and 50 subjects with AD over a 24 months interval. The volume changes in the MTL correlated with changes in cognitive function measured using the MMSE and the ADAS-cog scores, as well as memory performance using the LM score. The volume of interest measured by EXACT does not correspond exclusively to the hippocampus. Rather it is a medial temporal lobe volume to which the hippocampus contributes ~40%. Comparison between manual tracings and the EXACT VOI indicated that EXACT also includes primarily the amygdala, fibers associated with the amygdala, and the parahippocampal gyrus.

The absolute change in MTL volume measured using the EXACT approach was slightly lower than the change measured by FreeSurfer in the hippocampus. However, this result was expected as FreeSurfer has been shown to significantly overestimate the rate of measured mean atrophy in NEC, MCI and AD groups compared to multi-template hippocampal segmentation (Nestor et al., 2012; Wang et al., 2011). When comparing the rate of atrophy measured over 2 years with EXACT to other methods, the values appear lower due in part to a larger MTL baseline measurement; the VOI defined by EXACT also includes tissue outside the hippocampus. Specifically, the mean MTL rate of atrophy measured over 2 years by EXACT was 0.7% in NEC, 2.7% in MCI and 4.2% in AD subjects while FreeSurfer measured hippocampal atrophy of 2.5% in NEC, 5.9% in MCI, and 8.0% in AD. Similarly, the 24-month mean rate of atrophy measured by EXACT was smaller than that recently reported by Nestor et al. (2012) in the hippocampus using multi-atlas based segmentation: 2.5% for NEC ($n = 173$), 5.0% for MCI ($n = 253$) and 7.9% for AD ($n = 111$). These differences may also be explained by differing sample sizes. More importantly, using the EXACT method,

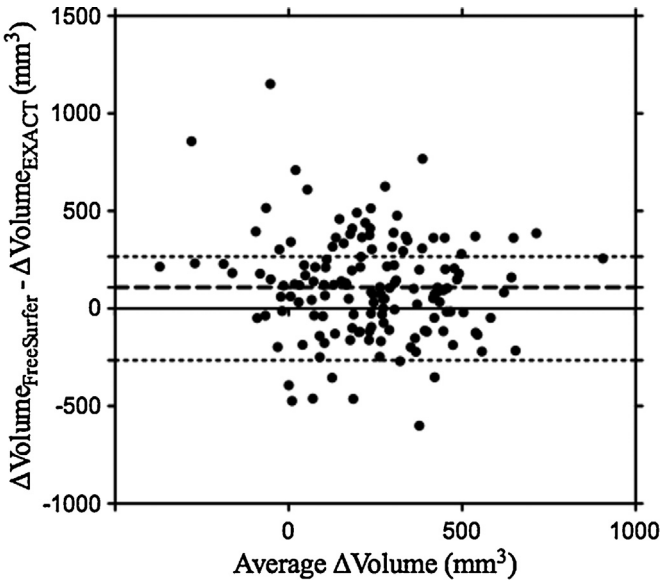

Fig. 11. Bland–Altman plot between FreeSurfer and EXACT method. The mean difference (±SD) between FreeSurfer and EXACT was 108.4 ± 265.8. Bold dashed line indicates the mean difference and dashed lines indicate the 95% limits of agreement.

differences were detected between all three groups (NEC, MCI, AD) while the FreeSurfer measurement only detected differences between NEC and MCI, as well as NEC and AD. The observed significant difference in ΔMTL volume between MCI and AD subjects suggests that ΔMTL volume may provide useful to differentiate between MCI and AD subjects. However both EXACT and FreeSurfer showed similar classification accuracy for NEC versus AD (76–80%) and for MCI versus AD (57–60%). Previous studies have also shown similar results. For example, Shen et al. (2012) performed a fully automated atlas-based segmentation using a template-based approach to obtain normalized hippocampal and amygdala volumes between subjects with AD, subjects with amnestic MCI (aMCI) and healthy control subjects. By using the area under the ROC curve on normalized hippocampus volume, they achieved a classification rate of 86–90% for the left hippocampus and of 85–87% for the right hippocampus for AD versus controls. However, they did not compare the classification rate between their subjects with aMCI and AD as well as examine the automated segmentation results with manual tracing. The EXACT method showed that the area under the ROC curve on volume changes was 78% in the left MTL and 75% in the right MTL between NEC and AD groups. However, a statistical comparison of the area under the ROC curves is not possible.

Both the EXACT and FreeSurfer techniques produced a significant relationship between change in volume and the change in MMSE and the ADAS-cog cognitive scores as well as 24 months LM score. With the EXACT method a positive correlation was observed between the volume changes in MTL and the change in MMSE score, and an inverse correlation was observed between the change in MTL and the change in ADAS-cog and LM scores, suggesting that MTL atrophy is related to cognitive function and memory performance. Previous studies using manual or semi-automated measures of MTL volumes have also shown a significant correlation between the rate of clinical decline and MTL atrophy in elderly patients over 12 months (Mungas et al., 2002, 2003). Arlt et al. (2012) showed correlations between hippocampus volume and neuropsychological tests for aMCI and AD groups. However, they did not investigate the classification rate of this fully automated technique or comparison with manual segmentation. Wolz et al. (2010b) proposed an automated method to measure hippocampal atrophy by segmenting longitudinal MR images from the ADNI database. They used an energy function based on Markov random fields (MRF) in combination with the graph cuts method (Boykov et al., 2001) and showed a significant correlation ($r = 0.30$, $p < 0.001$) between MMSE changes and hippocampal volume changes after 1 year. Similarly in the current study, we found a significant association between MTL volume changes after 2 years and the change in MMSE score ($r = 0.24$, $p = 0.003$). Although both studies demonstrated a similar association, there are several important differences between the approaches. First, we segmented the VOI in the baseline image and propagate this VOI to follow-up image. In contrast, Wolz et al. (2010b) proposed a method that segments hippocampus simultaneously in all longitudinal images. Second, their initial probabilistic atlas was created from healthy subjects, while EXACT selects the best atlas from a series of eight possibilities. Finally, they used the graph cuts method (Boykov et al., 2001) to segment longitudinal MR images, while we applied the Fast Marching method to the baseline images.

Several studies have performed manual or semi-automated segmentation of brain structures, which is time-consuming and is characterized by low intra- and inter-observer reproducibility depending on the level of experience of the observers (Pantel et al., 2004; Teipel et al., 2010; van der Flier et al., 2004). Although several automated hippocampus segmentation methods have been proposed (Friston et al., 1995; Jenkinson et al., 2002), these studies used single atlas-based segmentation potentially decreasing accuracy. In contrast, the current study used eight separate atlases to increase segmentation accuracy. Since EXACT requires significantly less processing time, it would be advantages in situations where large datasets must be analyzed.

Potential limitations of the algorithm are that the segmentation of the MTL in the baseline and follow up images are not independent (baseline images are registered to follow-up images). However this approach was taken to increase the precision of the atrophy measurement in this small structure. It should also be noted that the contour of the volume of interest is treated differently in the baseline and follow-up images. We expect that volumetric changes that occur between time points in one individual will be subtle. Therefore, we chose to use a regularization term in the diffeomorphic registration that was less severe than the smoothing filter used with the Fast Marching method so that the method would be sensitive to minor volumetric changes. Since the EXACT measurement is designed to be sensitive to atrophy, accurately defining the hippocampus is less important than sensitivity to change within the prescribed volume. Also, only a small number (eight) of atlases were used for selection, which should be increased for larger datasets. However the use of a small number of atlases is likely not a major limitation because the accuracy of the definition of the MTL is less important than sensitivity to change over time. Including more atlases may further increase the robustness of the technique in cases of extreme anatomical variability, however would also increase computation time. Optimizing the number of atlases will be the focus of future work. The number of atlases used can impact the quality of the final segmentation. The purpose of the atlas registration is to define a bounding box around the medial temporal lobe region of interest. As such, the demands on this registration step are light. However in cases of severe atrophy, the bounding box may fail to include the hippocampus if there are significant differences between the image and the atlas. This limitation was overcome in the current study by using multiple atlases with varying anatomy. Using eight atlases that spanned a wide range of atrophy conditions we successfully defined a MTL region label map in all subjects. Using an atlas fusion technique did not change the results and increased computation time. Finally, the method described in the current study is biased, since the baseline and follow-up volumes are measured using different regularization functions. Recently, several unbiased methods (Smith et al., 2002; Wolz et al., 2010b; Xue et al., 2006) have also been proposed for longitudinal studies.

5. Conclusion

In the present study, a fully automated longitudinal MTL atrophy measurement called EXACT was developed. The EXACT measurement showed significant differences in MTL atrophy over 2 years in normal elderly controls, subjects with mild cognitive impairment, and subjects with Alzheimer's disease. The MTL atrophy also correlated with decline in cognitive performance. This method is applicable to large datasets and could be incorporated in clinical evaluation in the future.

Funding

Ontario Research Fund.

Acknowledgements

This work was supported by Western University, Lawson Health Research Institute and the Alzheimer Society of Canada.

Data collection and sharing for this project was funded by ADNI (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute

of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129 and K01 AG030514.

Appendix A.

In Demons registration, an energy function is optimized to align two images by introducing correspondences between image pixels and applying a regularization term as prior knowledge to smooth the transformation of spatial points. The energy function can be formulated as shown in Eq. (A.1) and described in Vercauteren et al. (2009).

$$E(c,s) = \frac{1}{\sigma_i^2} Sim(F, M \circ c) + \frac{1}{\sigma_x^2} dist(s,c)^2 + \frac{1}{\sigma_r^2} Reg(s) \tag{A.1}$$

where σ_x represents the spatial uncertainly on the correspondences between image pixels, $\sigma_i = \left|F(p) - M \circ c(p)\right|$, $dist(s,c) = \|c - s\|$, the regularization term $Reg(s) = ||\nabla s||^2$ and Sim(.,.) is a similarity metric using the transformation of spatial points given by Eq. (A.2) in Vercauteren et al. (2009).

$$Sim(F, M \circ s) = \frac{1}{2}\left\|F - M \circ s\right\|^2 = \frac{1}{2\left|\Omega_p\right|} \sum_{p \in \Omega_p} \left|F(p) - M(s(p))\right|^2 \tag{A.2}$$

where Ω_p is the region of overlap between F and $M \circ s$. $M \circ s$ is a transformation of spatial points (s) belonging to the moving image M. To optimize Eq. (A.1) a spatial transformation is chosen. Given s, the vector field (u) is updated by minimizing Eq. (A.3) with respect to u using a Gauss–Newton iterative method from Vercauteren et al. (2009). Then the regularization is measured by convolution of a Gaussian kernel and vector field. This procedure performs iteratively till convergence.

$$E_s^{corr}(u) = \left\|F - M \circ (s+u)\right\|^2 + \frac{\sigma_i^2}{\sigma_x^2}\|u\|^2 \tag{A.3}$$

This Demons algorithm is made Diffeomorphic by computing an exponential of the vector field u with the spatial points s, prior to regularization.

Appendix B.

Details of the Fast Marching segmentation method can be found in (Ibanez et al., 2005; Sethian, 1996; Sethian and Sethian, 1999). Briefly, in the Fast Marching segmentation method, the goal is to solve the Eikonal equation $\left|\nabla T\right| F_1 = 1$. In its discrete form, it is given by Eq. (B.1) presented in Sethian (1996, 2001) and Tsitsiklis (1995).

$$\begin{aligned}\frac{1}{F_{1\{i,j,k\}}^2} = {} & \max\left(D^{-x}_{\{i,j,k\}}T, 0\right)^2 + \min\left(D^{+x}_{\{i,j,k\}}T, 0\right)^2 + \max\left(D^{-y}_{\{i,j,k\}}T, 0\right)^2 \\ & + \min\left(D^{+y}_{\{i,j,k\}}T, 0\right)^2 + \max\left(D^{-z}_{\{i,j,k\}}T, 0\right)^2 \\ & + \min\left(D^{+z}_{\{i,j,k\}}T, 0\right)^2 \end{aligned} \tag{B.1}$$

where $D^{-x}T(i,j,k) = T(i,j,k) - T(i-1,j,k)$, $D^{+x}T(i,j,k) = T(i+1,j,k) - T(i,j,k)$ and similarly for the y and z directions. In practice, the method propagates a contour curve in 2D or a contour surface in 3D with the speed function F_1. Where $T(i,j,k)$ is the time when the speed function reaches voxel (i,j,k).

The Fast Marching method has two phases: an initialization and propagation step. Each voxel has a value for T that is updated as the contour propagates. Voxels meeting a defined time threshold are included within the contour. The speed term in the differential equation is determined from the gradient magnitude. The curve propagates slowly near large image gradients but moves quickly when image gradients are small. Using this strategy, the contour propagates quickly until it nears the edge of an anatomical structure at which point it slows down. The contour propagates continuously over time and stops once a defined time threshold has been reached (Ibanez et al., 2005).

References

Achten E, Deblaere K, De Wagter C, Van Damme F, Boon P, De Reuck J, et al. Intra- and interobserver variability of MRI-based volume measurements of the hippocampus and amygdala using the manual ray tracing method. Neuroradiology 1998;40:558–66.

Adalsteinsson E, Sullivan EV, Kleinhans N, Spielman DM, Pfefferbaum A. Longitudinal decline of the neuronal marker N-acetyl aspartate in Alzheimer's disease. Lancet 2000;355:1696–7.

Anderson VC, Litvack ZN, Kaye JA. Magnetic resonance approaches to brain aging and Alzheimer disease-associated neuropathology. Topics in Magnetic Resonance Imaging: TMRI 2005;16:439–52.

Andreasen NC, Rajarethinam R, Cizadlo T, Arndt S, Swayze VW 2nd, Flashman LA, et al. Automatic atlas-based volume estimation of human brain regions from MR images. Journal of Computer Assisted Tomography 1996;20:98–106.

Anthony JC, Leresche L, Niaz U, Vonkorff MR, Folstein MF. Limits of the Mini-Mental State as a screening-test for dementia and delirium among hospital patients. Psychological Medicine 1982;12:397–408.

Arlt S, Buchert R, Spies L, Eichenlaub M, Lehmbeck JT, Jahn H. Association between fully automated MRI-based volumetry of different brain regions and neuropsychological test performance in patients with amnestic mild cognitive impairment and Alzheimer's disease. European Archives of Psychiatry and Clinical Neurosciences 2013;263:335–44.

Atiya M, Hyman BT, Albert MS, Killiany R. Structural magnetic resonance imaging in established and prodromal Alzheimer disease: a review. Alzheimer Disease and Associated Disorders 2003;17:177–95.

Azari NP, Pettigrew KD, Schapiro MB, Haxby JV, Grady CL, Pietrini P, et al. Early detection of Alzheimers-disease – a statistical approach using positron emission tomographic data. Journal of Cerebral Blood Flow & Metabolism 1993;13:438–47.

Barnes J, Scahill RI, Boyes RG, Frost C, Lewis EB, Rossor CL, et al. Differentiating AD from aging using semiautomated measurement of hippocampal atrophy rates. NeuroImage 2004;23:574–81.

Berchtold NC, Cotman CW. Evolution in the conceptualization of dementia and Alzheimer's disease: Greco-Roman period to the 1960s. Neurobiology of Aging 1998;19:173–89.

Boccardi M, Ganzola R, Bocchetta M, Pievani M, Redolfi A, Bartzokis G, et al. Survey of protocols for the manual segmentation of the hippocampus: preparatory steps towards a joint EADC-ADNI harmonized protocol. Journal of Alzheimers Disease 2011;26:61–75.

Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence 2001;23:1222–39.

Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. Alzheimer's & Dementia: The Journal of the Alzheimer's Association 2007;3:186–91.

Buades A, Coll B, Morel JM. A non-local algorithm for image denoising. In: 2005 IEEE computer society conference on computer vision and pattern recognition, vol. 2, proceedings; 2005. p. 60–5.

Chan TF, Vese LA. Active contours without edges. IEEE Transactions on Image Processing 2001;10:266–77.

Chetelat G, Baron JC. Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging. NeuroImage 2003;18:525–41.
Chupin M, Gerardin E, Cuingnet R, Boutet C, Lemieux L, Lehericy S, et al. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. Hippocampus 2009;19:579–87.
Chupin M, Mukuna-Bantumbakulu AR, Hasboun D, Bardinet E, Baillet S, Kinkingnehun S, et al. Anatomically constrained region deformation for the automated segmentation of the hippocampus and the amygdala: method and validation on controls and patients with Alzheimer's disease. NeuroImage 2007;34:996–1019.
Clerx L, van Rossum IA, Burns L, Knol DL, Scheltens P, Verhey F, et al. Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment. Neurobiology of Aging 2013;34:2003–13.
Coupe P, Manjon JV, Fonov V, Pruessner J, Robles M, Collins DL. Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. NeuroImage 2011;54:940–54.
Csernansky JG, Joshi S, Wang L, Haller JW, Gado M, Miller JP, et al. Hippocampal morphometry in schizophrenia by high dimensional brain mapping. Proceedings of the National Academy of Sciences of the United States of America 1998;95:11406–11.
Csernansky JG, Wang L, Joshi SC, Ratnanather JT, Miller MI. Computational anatomy and neuropsychiatric disease: probabilistic assessment of variation and statistical inference of group difference, hemispheric asymmetry, and time-dependent change. NeuroImage 2004;23(Suppl. 1):S56–68.
Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via EM algorithm. Journal of the Royal Statistical Society Series B: Methodological 1977;39:1–38.
Detoledomorrell L, Morrell F. Alzheimers-disease – new developments for noninvasive detection of early cases. Current Opinion in Neurology and Neurosurgery 1993;6:113–8.
Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26:297–302.
Dixon RM, Bradley KM, Budge MM, Styles P, Smith AD. Longitudinal quantitative proton magnetic resonance spectroscopy of the hippocampus in Alzheimer's disease. Brain: A Journal of Neurology 2002;125:2332–41.
Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurology 2007;6:734–46.
Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33:341–55.
Fischl B, Salat DH, van der Kouwe AJW, Makris N, Segonne F, Quinn BT, et al. Sequence-independent segmentation of magnetic resonance images. NeuroImage 2004;23:S69–84.
Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research 1975;12:189–98.
Freeborough PA, Fox NC. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. IEEE Transactions on Medical Imaging 1997;16:623–9.
Friston KJ, Ashburner J, Frith CD, Poline JB, Heather JD, Frackowiak RSJ. Spatial registration and normalization of images. Human Brain Mapping 1995;3:165–89.
Furst AJ, Lal RA. Amyloid-beta and glucose metabolism in Alzheimer's disease. Journal of Alzheimer's Disease: JAD 2011;26(Suppl. 3):105–16.
Gee JC, Reivich M, Bajcsy R. Elastically deforming 3D atlas to match anatomical brain images. Journal of Computer Assisted Tomography 1993;17:225–36.
Gering DT, Nabavi A, Kikinis R, Grimson WEL, Hata N, Everett P, et al. An integrated visualization system for surgical planning and guidance using image fusion and interventional imaging. In: Medical image computing and computer-assisted intervention, MICCAI'99, proceedings, vol. 1679; 1999. p. 809–19.
Geuze E, Vermetten E, Bremner JD. MR-based in vivo hippocampal volumetrics: 1. Review of methodologies currently employed. Molecular Psychiatry 2005;10:147–59.
Gosche KM, Mortimer JA, Smith CD, Markesbery WR, Snowdon DA. An automated technique for measuring hippocampal volumes from MR imaging studies. American Journal of Neuroradiology 2001;22:1686–9.
Greig DM, Porteous BT, Seheult AH. Exact maximum a posteriori estimation for binary images. Journal of the Royal Statistical Society Series B: Methodological 1989;51:271–9.
Hansen LK, Salamon P. Neural network ensembles. IEEE Transactions on Pattern Analysis and Machine Intelligence 1990;12:993–1001.
Heckemann RA, Hajnal JV, Aljabar P, Rueckert D, Hammers A. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. NeuroImage 2006;33:115–26.
Hogan RE, Mark KE, Wang L, Joshi S, Miller MI, Bucholz RD. Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients. Radiology 2000;216:291–7.
Hsu YY, Schuff N, Du AT, Mark K, Zhu XP, Weiner MW. Comparison of automated and manual MRI volumetry of hippocampus in normal aging and dementia. Journal of Magnetic Resonance Imaging 2002;16:305–10.
Ibanez L, Schroeder W, Ng L, Cates J. The ITK software guide. 2nd ed. Clifton Park, NY: Kitware Inc.; 2005.
Ikeda M, Tanabe H, Nakagawa Y, Kazui H, Oi H, Yamazaki H, et al. MRI-based quantitative assessment of the hippocampal region in very mild-to-moderate Alzheimers-disease. Neuroradiology 1994;36:7–10.
Iosifescu DV, Shenton ME, Warfield SK, Kikinis R, Dengler J, Jolesz FA, et al. An automated registration algorithm for measuring MRI subcortical brain structures. NeuroImage 1997;6:13–25.
Jack CR Jr, Barkhof F, Bernstein MA, Cantillon M, Cole PE, Decarli C, et al. Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. Alzheimer's & Dementia: The Journal of the Alzheimer's Association 2011;7:474–85.e4.
Jack CR Jr, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. Journal of Magnetic Resonance Imaging 2008;27:685–91.
Jack CR Jr, Shiung MM, Gunter JL, O'Brien PC, Weigand SD, Knopman DS, et al. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. Neurology 2004;62:591–600.
Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 2002;17:825–41.
Kantarci K, Jack CR Jr. Neuroimaging in Alzheimer disease: an evidence-based review. Neuroimaging Clinics of North America 2003;13:197–209.
Kantarci K, Jack CR Jr. Quantitative magnetic resonance techniques as surrogate markers of Alzheimer's disease. NeuroRx: The Journal of the American Society for Experimental NeuroTherapeutics 2004;1:196–205.
Kittler J. Combining classifiers: a theoretical framework. Pattern Analysis & Applications 1998;1:18–27.
Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage 2009;46:786–802.
Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Annals of Neurology 2004;55:306–19.
Miller BL, Moats RA, Shonk T, Ernst T, Woolley S, Ross BD. Alzheimer disease: depiction of increased cerebral myo-inositol with proton MR spectroscopy. Radiology 1993;187:433–7.
Mungas D, Reed BR, Jagust WJ, DeCarli C, Mack WJ, Kramer JH, et al. Volumetric MRI predicts rate of cognitive decline related to AD and cerebrovascular disease. Neurology 2002;59:867–73.
Mungas D, Reed BR, Kramer JH. Psychometrically matched measures of global cognition, memory, and executive function for assessment of cognitive decline in older persons. Neuropsychology 2003;17:380–92.
Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. Journal of the American Geriatrics Society 2005;53:695–9.
Nestor SM, Gibson E, Gao FQ, Kiss A, Black SE. A direct morphometric comparison of five labeling protocols for multi-atlas driven automatic segmentation of the hippocampus in Alzheimer's disease. NeuroImage 2012;66C:50–70.
Nestor SM, Rupsingh R, Borrie M, Smith M, Accomazzi V, Wells JL, et al. Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. Brain: A Journal of Neurology 2008;131:2443–54.
Osher S, Sethian JA. Fronts propagating with curvature-dependent speed – algorithms based on Hamilton–Jacobi formulations. Journal of Computational Physics 1988;79:12–49.
Pantel J, Schonknecht P, Essig M, Schroder J. Distribution of cerebral atrophy assessed by magnetic resonance imaging reflects patterns of neuropsychological deficits in Alzheimer's dementia. Neuroscience Letters 2004;361:17–20.
Penner J, Rupsingh R, Smith M, Wells JL, Borrie MJ, Bartha R. Increased glutamate in the hippocampus after galantamine treatment for Alzheimer disease. Progress in Neuro-Psychopharmacology & Biological Psychiatry 2010;34:104–10.
Pieper S, Halle M, Kikinis R. 3D Slicer. In: 2004 2nd IEEE international symposium on biomedical imaging: macro to nano, Vols. 1 and 2; 2004. p. 632–5.
Pieper S, Lorensen B, Schroeder W, Kikinis R. The NA-MIC Kit: ITK, VTK, pipelines, grids and 3D slicer as an open platform for the Medical Image Computing community. In: IEEE international symposium on biomedical imaging; 2006. p. 698–701.
Pohl KM, Bouix S, Nakamura M, Rohlfing T, McCarley RW, Kikinis R, et al. A hierarchical algorithm for MR brain image parcellation. IEEE Transactions on Medical Imaging 2007;26:1201–12.
Rohlfing T, Brandt R, Menzel R, Maurer CR Jr. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. NeuroImage 2004;21:1428–42.
Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. American Journal of Psychiatry 1984;141:1356–64.
Rupsingh R, Borrie M, Smith M, Wells JL, Bartha R. Reduced hippocampal glutamate in Alzheimer disease. Neurobiology of Aging 2011;32:802–10.
Rusinek H, De Santi S, Frid D, Tsui WH, Tarshish CY, Convit A, et al. Regional brain atrophy rate predicts future cognitive decline: 6-year longitudinal MR imaging study of normal aging. Radiology 2003;229:691–6.
Sethian JA. Evolution, implementation, and application of level set and fast marching methods for advancing fronts. Journal of Computational Physics 2001;169:503–55.
Sethian JA. A fast marching level set method for monotonically advancing fronts. Proceedings of the National Academy of Sciences of the United States of America 1996;93:1591–5.

Sethian JA, Sethian JA. Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science. 2nd ed. Cambridge, U.K., New York: Cambridge University Press; 1999.

Shen DG, Moffat S, Resnick SM, Davatzikos C. Measuring size and shape of the hippocampus in MR images using a deformable shape model. NeuroImage 2002;15:422–34.

Shen Q, Zhao W, Loewenstein DA, Potter E, Greig MT, Raj A, et al. Comparing new templates and atlas-based segmentations in the volumetric analysis of brain magnetic resonance images for diagnosing Alzheimer's disease. Alzheimer's & Dementia: The Journal of the Alzheimer's Association 2012;8:399–406.

Siadat MR, Soltanian-Zadeh H, Elisevich KV. Knowledge-based localization of hippocampus in human brain MRI. Computers in Biology and Medicine 2007;37:1342–60.

Smith SM. Fast robust automated brain extraction. Human Brain Mapping 2002;17:143–55.

Smith SM, Zhang YY, Jenkinson M, Chen J, Matthews PM, Federico A, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. NeuroImage 2002;17:479–89.

Sperling R, Salloway S, Brooks DJ, Tampieri D, Barakos J, Fox NC, et al. Amyloid-related imaging abnormalities in patients with Alzheimer's disease treated with bapineuzumab: a retrospective analysis. Lancet Neurology 2012;11:241–9.

Streitburger DP, Moller HE, Tittgemeyer M, Hund-Georgiadis M, Schroeter ML, Mueller K. Investigating structural brain changes of dehydration using voxel-based morphometry. PLoS ONE 2012;7:1–8.

Studholme C, Hill DLG, Hawkes DJ. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition 1999;32:71–86.

Teipel SJ, Ewers M, Wolf S, Jessen F, Kolsch H, Arlt S, et al. Multicentre variability of MRI-based medial temporal lobe volumetry in Alzheimer's disease. Psychiatry Research 2010;182:244–50.

Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. Medical Image Analysis 1998;2:243–60.

Tsitsiklis JN. Efficient algorithms for globally optimal trajectories. IEEE Transactions on Automatic Control 1995;40:1528–38.

van der Flier WM, van Buchem MA, Weverling-Rijnsburger AW, Mutsaers ER, Bollen EL, Admiraal-Behloul F, et al. Memory complaints in patients with normal cognition are associated with smaller hippocampal volumes. Journal of Neurology 2004;251:671–5.

van der Lijn F, den Heijer T, Breteler MM, Niessen WJ. Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. NeuroImage 2008;43:708–20.

Vercauteren T, Pennec X, Perchant A, Ayache N. Diffeomorphic demons: efficient non-parametric image registration. NeuroImage 2009;45:S61–72.

Villain N, Chetelat G, Grassiot B, Bourgeat P, Jones G, Ellis KA, et al. Regional dynamics of amyloid-beta deposition in healthy elderly, mild cognitive impairment and Alzheimer's disease: a voxelwise PiB-PET longitudinal study. Brain: A Journal of Neurology 2012;135:2126–39.

Wang H, Das SR, Suh JW, Altinay M, Pluta J, Craige C, et al. A learning-based wrapper method to correct systematic errors in automatic image segmentation: consistently improved performance in hippocampus, cortex and brain segmentation. NeuroImage 2011;55:968–85.

Wang L, Swank JS, Glick IE, Gado MH, Miller MI, Morris JC, et al. Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. NeuroImage 2003;20:667–82.

Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE Transactions on Medical Imaging 2004;23:903–21.

Wechsler D. The psychometric tradition – developing the Wechsler Adult Intelligence Scale. Contemporary Educational Psychology 1981;6:82–5.

Whitwell JL, Wiste HJ, Weigand SD, Rocca WA, Knopman DS, Roberts RO, et al. Comparison of imaging biomarkers in the Alzheimer disease neuroimaging initiative and the mayo clinic study of aging. Archives of Neurology 2012;69:614–22.

Williams CKI, Barber D. Bayesian classification with Gaussian processes. IEEE Transactions on Pattern Analysis and Machine Intelligence 1998;20:1342–51.

Wolz R, Aljabar P, Hajnal JV, Hammers A, Rueckert D. LEAP: learning embeddings for atlas propagation. NeuroImage 2010a;49:1316–25.

Wolz R, Heckemann RA, Aljabar P, Hajnal JV, Hammers A, Lotjonen J, et al. Measurement of hippocampal atrophy using 4D graph-cut segmentation: application to ADNI. NeuroImage 2010b;52:109–18.

Xue Z, Shen DG, Davatzikos C. CLASSIC: consistent longitudinal alignment and segmentation for serial image computing. NeuroImage 2006;30:388–99.

Yang J, Duncan JS. 3D image segmentation of deformable objects with joint shape-intensity prior models using level sets. Medical Image Analysis 2004;8:285–94.

Yoo TS, Ackerman MJ, Lorensen WE, Schroeder W, Chalana V, Aylward S, et al. Engineering and algorithm design for an image processing API: a technical report on ITK – the insight toolkit. Studies in Health Technology and Informatics 2002;85:586–92.