Semi-automated Hippocampal Segmentation in People With Cognitive Impairment Using an Age Appropriate Template for Registration

Bernd Merkel, MSc,[1,2]* Christopher Steward, PhD,[1,2] Lucy Vivash, PhD,[3] Charles B. Malpas, BA,[4] Pramit Phal, MBBS,[1,2] Bradford A. Moffat, PhD,[5] Kay L. Cox, PhD,[6] Kathryn A. Ellis, PhD,[7] David J. Ames, MD,[7,8] Elizabeth V. Cyarto, PhD,[8] Michelle M.Y. Lai, MPH, MMed,[8] Matthew J. Sharman, PhD,[9] Cassandra Szoeke, PhD,[3] Colin L. Masters, MD,[10] Nicola T. Lautenschlager, MD,[7,11] and Patricia Desmond, MSc, MD[1,2]

Background: To evaluate a new semi-automated segmentation method for calculating hippocampal volumes and to compare results with standard software tools in a cohort of people with subjective memory complaints (SMC) and mild cognitive impairment (MCI).
Methods: Data from 58 participants, 39 with SMC (17 male, 22 female, mean age 72.6) and 19 with MCI (6 male, 13 female, mean age 74.3), were analyzed. For each participant, T1-weighted images were acquired using an MPRAGE sequence on a 3 Tesla MRI system. Hippocampal volumes (left, right, and total) were calculated with a new, age appropriate registration template, based on older people and using the advanced software tool ANTs (Advanced Normalization Tools). The results were compared with manual tracing (seen as the reference standard) and two widely accepted automated software tools (FSL, FreeSurfer).
Results: The hippocampal volumes, calculated by using the age appropriate registration template were significantly ($P < 0.05$) more accurate (mean volume accuracy more than 90%) than those obtained with FreeSurfer and FSL (both less than 70%). Dice coefficients for the hippocampal segmentations with the new template method (75.3%) were slightly, but significantly ($P < 0.05$) higher than those from FreeSurfer (72.4%).
Conclusion: These results suggest that an age appropriate registration template might be a more accurate alternative to calculate hippocampal volumes when manual segmentation is not feasible.

J. MAGN. RESON. IMAGING 2015;42:1631–1638.

Dementia of the Alzheimer's type (DAT) and other dementias currently affect approximately 35.6 million people worldwide, which is predicted to increase to 115 million by 2050.[1] Mild cognitive impairment (MCI) is often seen as a potential prodromal syndrome of DAT and is therefore of high interest for clinical trials of potential

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.24966

Received Mar 8, 2015, Accepted for publication May 7, 2015.

*Address reprint requests to: B.M., Department of Radiology, Level 2, The Royal Melbourne Hospital, Grattan Street, Parkville, Victoria 3050, Australia. E-mail: bernd.merkel@mh.org.au

Drs. Lautenschlager and Desmond contributed equally to this work.

From the [1]Department of Radiology, The University of Melbourne, Melbourne, Victoria; [2]Department of Radiology, The Royal Melbourne Hospital, Melbourne, Victoria; [3]Department of Medicine, The University of Melbourne, Melbourne, Victoria; [4]Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Victoria; [5]Department of Anatomy and Neuroscience, The University of Melbourne, Melbourne, Victoria; [6]School of Medicine and Pharmacology, University of Western Australia, Perth, Western Australia; [7]Academic Unit for Psychiatry of Old Age, Department of Psychiatry, St. Vincent's Health, The University of Melbourne, Kew, Victoria; [8]National Ageing Research Institute, Parkville, Victoria; [9]School of Health Sciences, University of Tasmania, Hobart, Tasmania; [10]The Florey Institute, The University of Melbourne, Melbourne, Victoria; and [11]School of Psychiatry and Clinical Neurosciences & WA Centre for Health and Ageing, The University of Western Australia, Perth, Western Australia

therapies.[2] Similarly, individuals with subjective memory complaints (SMC) may also be at increased risk of developing DAT, although the predictive value and clinical utility of SMC is not entirely clear.[3,4]

The hippocampus in the medial temporal lobe is one of the brain structures, that is sensitive to aging effects and is affected early in the Alzheimer's disease (AD) process.[5] Although aging is associated with mild loss of volume in the hippocampus,[6] quantitative measures of hippocampal atrophy with MRI have been shown to be a strong predictor of AD diagnosis and disease progression. In comparison to healthy controls, previous work demonstrated reduced hippocampal volumes in people with MCI.[7,8] AD has been shown to be associated with a higher rate of volume loss in the hippocampus compared with normal elderly people, and in MCI patients, hippocampus atrophy was intermediate between healthy controls and AD patients.[9]

Therefore, hippocampal volume may be a useful metric to assess vulnerability to AD in participants with SMC and MCI. However, its quantification remains challenging; manual segmentation is considered the reference standard, but is subjective, time-consuming and requires anatomical training. A standardized protocol is being developed by an EADC-ADNI cooperation,[10] but currently there is no consensus on manual segmentation methods based on a statistical template of the human brain.

Different semi-automated and automated methods have been used to segment and analyze the volumes and shapes of hippocampi. For instance, FreeSurfer [11] and FSL (using the FIRST toolbox) [12] are two publicly available software packages with integrated tools for automated hippocampal segmentation.

Although previous studies have shown high correlations between automated methods and manual delineation,[13,14] reliability and validity of automated segmentation methods in older age groups remain suboptimal.[14,15] FreeSurfer has been observed to overestimate hippocampal volumes,[14,16] but was shown to perform better than FSL/FIRST in terms of volume overlap (Dice's coefficient), volume difference, and correlation [17] and has also shown similar patterns of atrophy in AD patients' brain structures including the hippocampus.[18]

The templates used for segmentation in both FSL and FreeSurfer are based on healthy young subjects with a diverse age range. A recent study in an epilepsy population has shown the importance of template selection in studies of pathology.[19] Following this, further studies have investigated the use of multiple atlases and templates to minimize atlas bias and improve registration.[20–22]

The aim of the current study was to create a study-specific template based on an older population for analysis of hippocampal volume. We hypothesized that in a study population of older participants with SMC and MCI, hippocampal volumes based on this template would yield results with greater agreement with manual hippocampal delineation when compared with the existing automated methods from popular image processing packages (FreeSurfer and FSL).

Materials and Methods

Participants

Data were analyzed from participants in the AIBL-Active study. This randomized clinical trial is a sub-study of the Australian Imaging, Biomarker and Lifestyle Flagship Study of Ageing (AIBL),[23] and AIBL-Active includes 108 participants from Melbourne, Victoria, Australia. AIBL-Active [24] is a study investigating the effect of physical activity on progression of white matter lesions in nondemented older adults with SMC and MCI. Cross-sectional baseline data from 58 older people, aged between 61 and 90 years (mean: 73.2 ± 5.6 years) were available for review at the time of this study, from 39 participants with SMC and 19 cases with MCI. Exclusion criteria and assessment details have been outlined previously.[24] Briefly, inclusion criteria for AIBL-Active were: male or female, community-dwelling, at least 60 years of age, presence of at least one vascular risk factor, diagnosis of SMC or MCI, and fluent in written and spoken English.

AIBL-Active has been approved by the Melbourne Health Human Research Ethics Committee (HREC 2011.014) and is fully compliant with the guidelines of the World Medical Association (WMA) Declaration of Helsinki. Written informed consent has been obtained from all participants and the trial registration number is Australia New Zealand Clinical Trials Registry ACTRN12611000612910

MRI Acquisition

For each participant, the following MRI sequences were acquired with a 12-channel phased array coil: (i) T1-W Magnetization Prepared Rapid Gradient Echo (MPRAGE), isotropic 1 mm voxel (repetition time [TR] = 1900 ms, echo time[TE] = 2.13 ms, flip angle = 9°, inversion time [TI] = 900 ms, matrix = 256 × 256, 176 sagittal slices); (ii) Sagittal 3D Fluid Attenuated Inversion Recovery (FLAIR), isotropic 1mm voxel (TR = 5000 ms, TE = 355 ms, flip angle = 120°, TI = 1800 ms). All subjects were scanned on the same machine (Siemens 3T Tim Trio) at the Royal Melbourne Hospital (Melbourne, Victoria).

Cognitive Measurements

The following assessments for determining eligibility and cognitive impairment were undertaken, as described elsewhere [24]: the 15-item Geriatric Depression Scale (GDS-15), the Alzheimer's Disease Assessment Scale – Cognitive Section (ADAS-cog), the Cambridge Contextual Reading Test (CCRT), the Standardized Mini-Mental State Examination (SMMSE), the Clinical Dementia Rating Scale (CDR), the n-Back Task, the Wisconsin Card Sorting Test (WCST-64), the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) battery, the Short-Form-36 version 2 (SF-36v2), Hospital Anxiety and Depression (HADS), Trails A & B/Trail Making Test (TMT), the Behavior Rating Inventory of Executive Function – Adult Version (BRIEF-A), Everyday Competence Questionnaire (ECQ), Memory Complaint Questionnaire (MAC-Q).

Participants had to have an MMSE at baseline of 24 or higher and a CDR total score 0.5 or lower.

Patients were classified as SMC if they: (a) endorsed the question "do you have any difficulty with your memory", but (b) did not demonstrate objective cognitive impairment. Absence of cognitive impairment was classified as scoring in the normal range (compared with age and sex matched healthy controls) on neuropsychological tests, specified in the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) neuropsychology assessment. Patients were classified as MCI if they: (a) endorsed a subjective memory complaint (as above), (b) demonstrated objective cognitive impairment (<1.5 SD on any CERAD subtest compared with controls), (c) had preserved activities of daily living (ADL), with only minimal impairment in complex instrumental ADLs, and (d) were not demented.

Manual Segmentation

Regions of interest (ROIs) for the hippocampal volume were manually delineated by one tracer (L.V., 6 years of expertise) on 58 datasets using the image analysis software Analyze (v10.0, Mayo Clinic, Rochester, MN). The tracer, who was blinded to the subject's identity, performed delineations of ROIs and defined their extents according to the method described by Cook et al.[25] First, MRIs were viewed in the coronal plane, and the posterior limit defined as the slice in which the greatest length of the fornix was observed. Delineation was performed by stepping anteriorly through to the amygdala. The alveus distinguished the hippocampus from the amygdala. The superior border was defined by the ambient cistern and the inferior border by the subicular white matter. The tracings were then verified and corrected in the sagittal plane, excluding the fornix. Final corrections were applied in the coronal plane, confirming the anterior and posterior limits of the hippocampus.

Finally, to evaluate the intra-operator agreement for manual segmentation, the hippocampal tracings were repeated on ten of the subjects.

Automatic Segmentation using FSL and FreeSurfer

To evaluate the performance of our own semi-automated segmentation, we compared it with two widely accepted software applications for image data analysis in the brain:

First, FSL (FMRIB Software Library, Oxford, UK) v5.0 [12]: FSL is an open source image library. After reorienting the MR images to match the standard template orientation in MNI152 space, we used the FIRST package (FMRIB's Integrated Registration and Segmentation Tool) in FSL to perform subcortical structure segmentation. For optimal tuning, the parameters for number of modes and boundary correction were not altered, as suggested by the software authors. Finally, the hippocampal volumes for corrected boundaries were calculated with fslstats, a tool for statistical summaries in FSL.

Second, FreeSurfer (Laboratory for Computational Neuroimaging, Charlestown, MA) Image Analysis Suite v5.3.0 [11]: Whole brain segmentation was calculated [26] with no manual intervention, using T1-weighted MPRAGE and FLAIR images. This permits a more accurate delineation between gray and white matter by allowing dura to be separated from gray matter to help refine the pial surfaces. Subcortical volume and surface files were automatically created comprising the following steps: motion correction, intensity normalization, skull-stripping, volumetric registrations, white and gray matter segmentation, spherical mapping, spherical registration and cortical parcellation. From the subsequent results of gray matter segmentation, the total intracranial volume (ICV) using the method of [27] and subcortical volumes were then obtained.

Semi-automated Segmentation Based on a Study-Specific Template

An unbiased, study-specific, age-matched template was generated from the T1-MPRAGE nifti images for 26 AIBL-Active participants (13 male, 13 female) using the ANTs (Advanced Normalization Tools) 1.9.v4 software package.[28–30] This approach has the advantage of preserving the anatomical features of the patient cohort more closely than with direct normalization to the standard SPM template MNI152.

Default settings were used for the template script (N4 bias field corrections of images, cross-correlation similarity metric, 60 $\times$ 100 $\times$ 30 nonlinear iterations using Gauss[3,0] for the regularization and a Greedy-SyN transformation model with step-size 0.25). All subject images were nonlinearly registered to this template with ANTs. This resulted in 58 different image transformation matrices R_i ($I = 1, \ldots, 58$) from subject space to template space and 58 inverse transformation matrices (R_i^{-1}) from template space to subject space.

The same tracer, who manually segmented the 58 participants' datasets, also delineated the hippocampi on the template (Fig. 1) to create a mask ROI of the hippocampi in template space. To segment the hippocampi in each individual, these hippocampi ROIs in the template space were projected back to each subject's native space using (R_i^{-1}) and nearest neighbor interpolation. This resulted in hippocampal ROIs for each subject, which were used to calculate hippocampal volume and for the statistical comparisons with the manual and automatic segmentations.

All image processing was performed on a 64-core Dell PowerEdge C6145 with 256GB of RAM using the Sun Grid Engine (SGE) from Oracle and Sun Microsystems (http://www.oracle.com/us/sun/index.htm) and was completed within 3 h.

Statistics

All the statistical analysis was performed using SPSS 21.0 (SPSS, Chicago, IL). To compare the different methods, we calculated Spearman correlations and analyzed both single left and right hippocampal volumes as well as the total hippocampal volume. Correlations were calculated between all segmentation methods. The correlation coefficients between the three methods and manual tracing were compared using a t-test procedure described elsewhere.[31]

Absolute differences between each method and manual tracing were compared using a one-way repeated-measures general linear model (GLM). Tracing method was entered as the within-subjects factor. Two orthogonal planned comparisons were conducted. The first compared the ANTs method to both FSL and FreeSurfer, and the second compared FSL with FreeSurfer. Effect sizes are reported as partial η^2, where a small effect = 0.01, medium effect = 0.06, and large effect = 0.14.

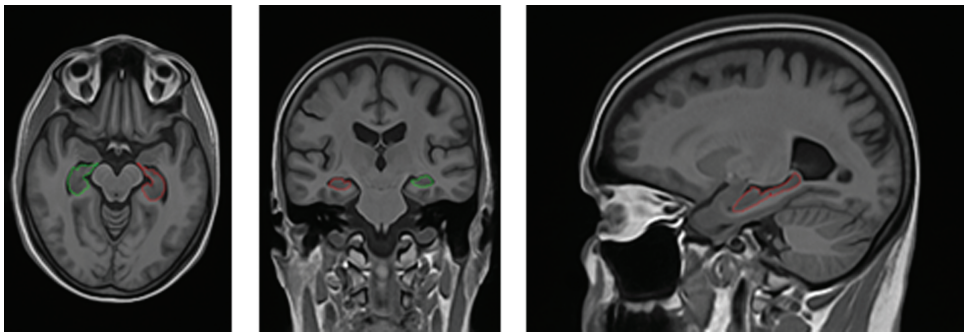


FIGURE 1: The age appropriate template in transversal, coronal and sagittal view (from left to right) with the manually segmented hippocampus.

Volume accuracy of a certain segmentation method S compared with manual tracing M was calculated for each patient by

$$1 - ((|\mathrm{Vol\,S} - \mathrm{Vol\,M}|) / \mathrm{Vol\,M}). \quad (1)$$

Dice coefficients between two different segmentations A and B were calculated as

$$2\,|A \cap B| / (|A| + |B|). \quad (2)$$

Results

Demographic and Clinical Findings

Table 1 shows the clinical and demographic characteristics of the two groups (SMC and MCI).

Absolute Hippocampal Volumes

The repeated manual hippocampal tracings on 10 subjects resulted in an intra-operator reproducibility error of less than 5%. The mean hippocampal volume estimates produced by each of the four methods are shown in Table 2. As shown in Figure 2, the FreeSurfer results were descriptively larger than those from the FSL analysis, which were in turn greater than the ANTs template and manually drawn ROIs. The ANTs based method produced the closest approximation to the manual tracing volumes.

Without correcting the volumes for IntraCranialVolume (ICV), the correlations between manually drawn ROIs in the left and right hippocampus and the different measurements were highly significant (all P-values < 0.01) at the level of $r \geq 0.37$ (Table 3). The total volumes from all four methods corresponded with $r \geq 0.48$ ($P < 0.01$) with each other (Table 4). Pearson correlations remained statistically significant after correcting for ICV (all r-values ≥ 0.70, $P < 0.01$).

The volumes calculated by all three methods were significantly correlated with manual tracing (Table 4). No one method correlated more highly than the others (all P values > 0.05, data not shown).

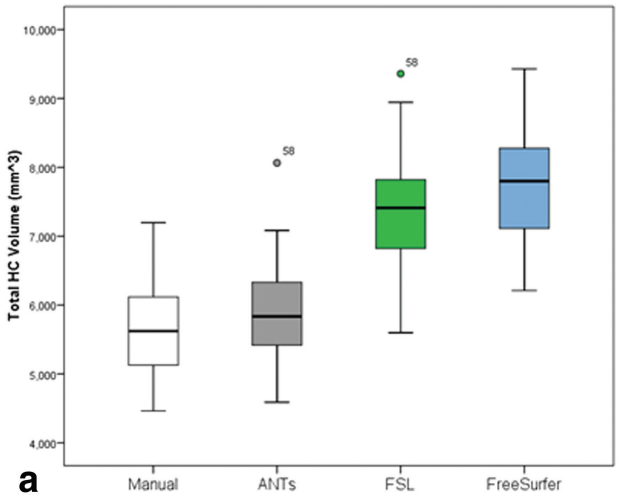
As shown in Figure 3, the differences between all three (semi-) automated methods and manual tracing were

TABLE 1. Demographical and Clinical Findings of the Dataset

	Participant group – Mean (SD)	
	SMC	**MCI**
N	39	19
Male/Female	17/22	6/13
Age [years]	72.6 (5.2)	74.3 (6.5)
MMSE	29.3 (0.9)	28.7 (0.7)

TABLE 2. Min., Max., and Mean (with Standard Deviation) for left, right, and total hippocampus volume, calculated by the four different methods (in mm^3)

	Estimated volume (mm^3)		
	Min.	**Max.**	**Mean (± SD)**
Left hippocampus:			
Manual	2210	3647	2824 (±361.9)
ANTs template	2212	4051	2896 (±365.1)
FSL (FIRST)	2543	4742	3656 (±450.4)
FreeSurfer	2807	4771	3774 (±410.6)
Right hippocampus:			
Manual	2200	3799	2863 (±365.8)
ANTs template	2367	4013	2992 (±351.7)
FSL (FIRST)	3055	4629	3733 (±367.7)
FreeSurfer	3307	4737	3959 (±368.1)
Total hippocampus:			
Manual	4463	7196	5687 (±660.7)
ANTs template	4590	8064	5889 (±686.4)
FSL (FIRST)	5598	9359	7390 (±731.5)
FreeSurfer	6210	9429	7733 (±732.3)



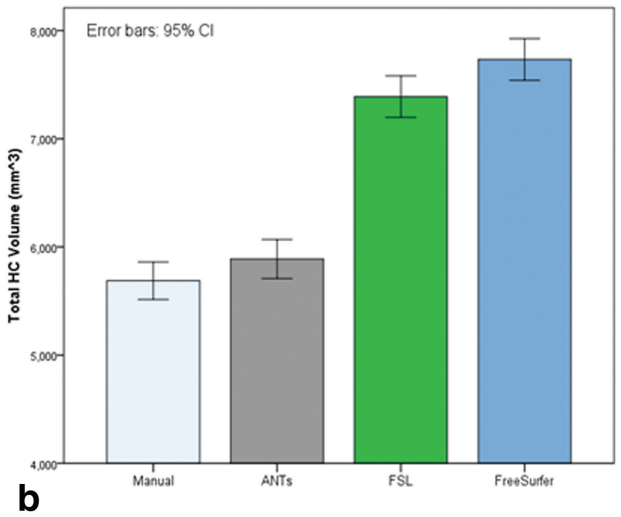


FIGURE 2: The absolute total hippocampal volumes in mm^3. Boxplots with median and interquartile ranges (a), and with 95% CI (b), both showing a significantly better estimation of the volumes with the ANTs template, compared with FSL and FreeSurfer.

significant. Mean accuracy of the volume estimates for the total hippocampus from ANTs were 0.901 (95% confidence interval [CI] = [0.878; 0.924]), compared with manual tracing. Mean FSL volume accuracy was 0.692 (95% CI = [0.658; 0.727]). The mean FreeSurfer estimates were 0.631 compared with manual tracing (95% CI = [0.597; 0.666]) (Fig. 3)

The absolute differences between tracing methods were statistically significant (Fig. 2). The assumption of sphericity was not violated, Mauchly's $W = 0.90$, $P = 0.06$. There was an overall difference between all tracing methods, $F(2,114) = 190.76$, $P < 0.001$, partial eta$^2 = 0.77$ (large effect). The difference between ANTs and the other two methods was statistically significant, $F(1,57) = 278.04$, $P < 0.001$, partial eta$^2 = 0.83$ (large effect). The difference between FSL and FreeSurfer also reached significance, $F(1,57) = 25.52$, $P < 0.001$, partial eta$^2 = 0.31$ (large effect).

TABLE 3. Spearman correlations between the four different measurements in the left HC (lower left part of the table, highlighted in blue) and the right HC (upper right part of the table, highlighted in orange)

	Manual	ANTs	FSL	FreeSurfer
Manual	-	0.372[a]	0.502[b]	0.473[b]
ANTs	0.505[b]	-	0.667[b]	0.581[b]
FSL	0.523[b]	0.559[b]	-	0.701[b]
FreeSurfer	0.663[b]	0.598[b]	0.615[b]	-

[a]Correlation is significant at the 0.01 level (two-tailed).
[b]Correlation is significant at the 0.001 level (two-tailed).

We also calculated the Dice coefficients (ranging from 0: no overlap, to 1: complete overlap) of our own age appropriate based method with the manual segmentation (mean: 0.753) as well as between FreeSurfer and the manual method (mean: 0.724). A comparison of the coefficients showed that the ANTs-based method was significantly 3% better than FreeSurfer (Table 5).

Discussion

In a recent study, Zhang and colleagues showed improved registration accuracy of a group-specific, whole-brain template compared with a single-subject, study-specific one with a volume-based template estimation.[32] They included three different populations: healthy young adults, patients with AD, and age-matched controls. There are also multi-atlas approaches for hippocampal segmentation.[20–22] Several manually labeled atlases are used, based on one single input template, to reduce atlas bias, registration errors and issues with resampling the images.

In general, registration templates are a powerful alternative to calculating brain structures based on an atlas.[19,27,32]

TABLE 4. Spearman correlations between the four different measurements for the total hippocampus volume: without normalization to the ICV lower left (highlighted in blue), normalized to the ICV upper right (highlighted in orange)

	Manual	ANTs	FSL	FreeSurfer
Manual	-	0.699[a]	0.768[a]	0.787[a]
ANTs	0.477[a]	-	0.639[a]	0.658[a]
FSL	0.573[a]	0.658[a]	-	0.833[a]
FreeSurfer	0.626[a]	0.636[a]	0.754[a]	-

[a]Correlation is significant at the 0.001 level (two-tailed).

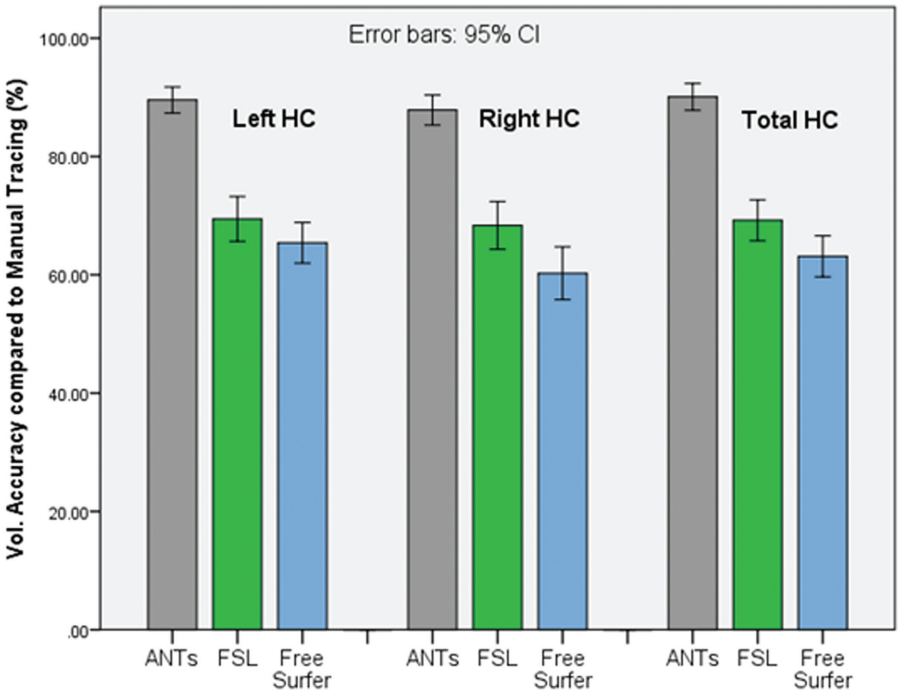


FIGURE 3: Mean volume accuracy (in percentage) between different segmentation methods and manual tracing. The ANTs template was the most accurate one (90.1% volume accuracy in the total HC, in average), followed by FSL (69.2%) and FreeSurfer (63.1%).

As Avants et al [19] suggested, the correct choice of the template is crucial when the study group is not a cognitively normal healthy group. As such, a cohort-specific subgroup was used to create a template for this study. The cohort included people with SMC and MCI. We used the ANTs software and MRI MPRAGE images from 26 subjects of our cohort to create an older adults' template. Our main finding was that a method based on this registration template performed significantly better in estimating hippocampal volumes than the popular automated software tools FreeSurfer and FSL. All three (semi-) automated methods were significantly correlated with the manually derived ROIs with FreeSurfer showing the highest correlation in terms of total hippocampal volume. Our own template, however, was found to be the most accurate method for calculating both hippocampal volumes and location (Dice coefficient).

By investigating left and right hippocampal volumes separately, significant correlations ($P < 0.05$) between all the methods were found with FSL yielding the highest correlation in the right hippocampus and with FreeSurfer in the left hippocampus.

Our results are in agreement with previous studies showing that FreeSurfer highly correlates with manually segmented ROIs, but also overestimates volumes. For example, in a similar study, Morey et al analyzed FreeSurfer and FSL-FIRST results of automated hippocampal quantifications in a cohort of 20 healthy controls.[17] They also found larger hippocampal volumes relative to the manual tracings as well as high correlation factors. In assessing total hippocampal volumes, FreeSurfer performed much better than FSL in terms of correlation whereas in our cohort, the two correlation coefficients for normalized volumes were similar. In keeping with previous studies, we also found all the four methods resulting in smaller volumes (in the mean) in the left hippocampus than in the right.[20,33] However, Agosta et al [8] reported smaller right hippocampal volumes in a cohort of amnestic MCI patients using FIRST.

TABLE 5. Mean Value, 95% CI, Min. and max. Values of the Dice Coefficients, Calculated for the Study-Specific Template (ANTs) and FreeSurfer in Comparison With Manual Tracing

Dice Coefficients	Mean ± SD	95% CI	Min.	Max.
ANTs vs. Manual	0.753 ± 0.038	[0.743, 0.763]	0.618	0.810
FreeSurfer vs. Manual	0.724 ± 0.040	[0.713, 0.734]	0.575	0.798
ANTs vs. FreeSurfer	0.733 ± 0.023	[0.727, 0.739]	0.654	0.771

An important issue is the right choice of the template. Frankó et al analyzed hippocampal atrophy over two time points.[5] They used a hippocampus which was manually segmented from the single-subject MNI-152 template in SPM. Marcus et al spatially warped images into an atlas which consisted of both old and young subjects without dementia.[34] However, these templates might not be suitable enough for different subject cohorts, due to partial volume effects and different degrees of atrophy.

We used a template based on a subgroup of our cohort to substantially minimize registration issues in mapping all images from the cohort into the same space. As described elsewhere,[19] the right choice of the template is even more crucial in people with disease compared with healthy subjects.

The template has been used to calculate hippocampal volumes in SMC and MCI participants more accurately than commonly used tools. However, it also has the potential to be applied in studies with different clinical questions because hippocampal morphometry has been proposed as a useful biomarker in other diseases, such as epilepsy,[20] depression,[35] and frontotemporal lobar degeneration (FTLD).[33]

This work shows that a study group based registration template may be better to use for hippocampal segmentation and volume measurement in elderly populations than templates based on healthy younger brains.[19,27,32]

There is an ongoing debate as to whether manual tracing should be seen as the reference standard for hippocampal segmentation because there is no empirical evidence to what extent it accurately reflects the true hippocampal volume and its subfields. A comparison of manually drawn, MRI-based hippocampal ROIs with post mortem measurements is rarely feasible. However, a first single-subject study using histology-based hippocampal subfield reconstruction on high-resolution MRI is a promising step for such an atlas.[36] Because we were only interested in volume and location accuracy in hippocampal segmentations in this study, we did not analyze any differences of the hippocampal subfields. Furthermore, different investigators may use different hippocampal tracing protocols. It is anticipated, that the EADC-ADNI joint effort to standardize manual hippocampus segmentations will be of significant benefit in this area.[10]

There are a diversity of options in creating a population-specific template for hippocampal studies and the impact of these different choices may be more critical in pathological cases than in healthy controls.[19] As with other similar methods our template might not be generalizable to other patient cohorts, which limits its use. For example, AD patients are reported to have higher hippocampal atrophy than people with MCI or healthy subjects. Therefore, in the presence of significantly atrophied brains, this template might perform not optimally due to considerable partial volume effects and more variability in hippocampal shapes.[37]

Furthermore, the results could be specific to the particular MRI acquisition parameters used in this study, which can vary slightly from study to study. In addition, it was not a comprehensive analysis of segmentation tools as there are also other software tools available for segmentation, such as HAMMER,[38] LocalInfo,[20] and SACHA.[39]

In conclusion, this study found that by using an age appropriate template of elderly people, taken as a subgroup of a study cohort, hippocampal volumes could be calculated more accurately in people with cognitive impairment (SMC and MCI) compared with previously published segmentation algorithms that rely on historical templates based on younger healthy volunteers. Furthermore, with the aid of modern computing resources, the use of a study specific template was entirely feasible with the use of freely available neuroimaging software packages. The next step will include the use and optimization of this template with other clinical samples and to create appropriate templates for other brain structures.

Acknowledgments

Contract grant sponsor: Australia's National Health and Medical Research Council; Contract grant number: 1005492.

We are most grateful to all volunteers taking part in the study and to the research staff from following institutes in Melbourne, Victoria: the National Ageing Research Institute (NARI), the Mental Health Research institute (MHRI) and the Department of Radiology at the Royal Melbourne Hospital. N.L., P.D., K.C., D.A., C.S., K.E., P.P., and C.M. were funded by a project grant from Australia's National Health and Medical Research Council. Funding for the AIBL study is provided by the CSIRO Flagship Collaboration Fund and the Science and Industry Endowment Fund (SIEF) in partnership with Edith Cowan University (ECU), Mental Health Research institute (MHRI), Alzheimer's Australia (AA), National Ageing Research Institute (NARI), Austin Health, CogState Ltd., Hollywood Private Hospital, Sir Charles Gardner Hospital. The study also receives funding from the National Health and Medical Research Council, the Dementia Collaborative Research Centres program (DCRC2) and The McCusker Alzheimer's Research Foundation and Operational Infrastructure Support from the Government of Victoria.

References

1. WHO, Alzheimer's Disease International. Dementia: a public health priority. Geneva: WHO; 2012. p 112.
2. Petersen RC. Mild cognitive impairment. N Engl J Med 2011;364: 2227–2234.
3. Jonker C, Geerlings MI, Schmand B. Are memory complaints predictive for dementia? A review of clinical and population-based studies. Int J Geriatr Psychiatry 2000;15:983–991.

4. Stewart R. Subjective cognitive impairment. Curr Opin Psychiatry 2012;25:445–450.
5. Frankó E, Joly O; Alzheimer's Disease Neuroimaging Initiative. Evaluating Alzheimer's disease progression using rate of regional hippocampal atrophy. PLoS One 2013;8:e71354.
6. Raz N, Lindenberger U, Rodrigue KM, et al. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. Cereb Cortex 2005;15:1676–1689.
7. Bozzali M, Filippi M, Magnani G, et al. The contribution of voxel-based morphometry in staging patients with mild cognitive impairment. Neurology 2006;67:453–460.
8. Agosta F, Rocca MA, Pagani E, et al. Sensorimotor network rewiring in mild cognitive impairment and Alzheimer's disease. Hum Brain Mapp 2010;31:515–525.
9. Schuff N, Zhu XP. Imaging of mild cognitive impairment and early dementia. Br J Radiol 2007;80:S109–S114.
10. Jack CR Jr, Barkhof F, Bernstein MA, et al. Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. Alzheimers Dement 2011;7:474–485.e4.
11. Fischl B. FreeSurfer. Neuroimage 2012;62:774–781.
12. Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. Neuroimage 2011;56:907–922.
13. Shen L, Saykin AJ, Kim S, et al. Comparison of manual and automated determination of hippocampal volumes in MCI and early AD. Brain Imaging Behav 2010;4:86–95.
14. Wenger E, Martensson J, Noack H, et al. Comparing manual and automatic segmentation of hippocampal volumes: reliability and validity issues in younger and older brains. Hum Brain Mapp 2014;35:4236–4248.
15. Sanchez-Benavides G, Gomez-Anson B, Sainz A, Vives Y, Delfino M, Pena-Casanova J. Manual validation of FreeSurfer's automated hippocampal segmentation in normal aging, mild cognitive impairment, and Alzheimer Disease subjects. Psychiatry Res 2010;181:219–225.
16. Cherbuin N, Anstey KJ, Reglade-Meslin C, Sachdev PS. In vivo hippocampal measurement and memory: a comparison of manual tracing and automated segmentation in a large community-based sample. PLoS One 2009;4:e5265.
17. Morey RA, Petty CM, Xu Y, et al. A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. Neuroimage 2009;45:855–866.
18. Lehmann M, Douiri A, Kim LG, et al. Atrophy patterns in Alzheimer's disease and semantic dementia: a comparison of FreeSurfer and manual volumetric measurements. Neuroimage 2010;49:2264–2274.
19. Avants BB, Yushkevich P, Pluta J, et al. The optimal template effect in hippocampus studies of diseased populations. Neuroimage 2010;49:2457–2466.
20. Akhondi-Asl A, Jafari-Khouzani K, Elisevich K, Soltanian-Zadeh H. Hippocampal volumetry for lateralization of temporal lobe epilepsy: automated versus manual methods. Neuroimage 2011;54(Suppl 1):S218–S226.
21. Chakravarty MM, Steadman P, van Eede MC, et al. Performing label-fusion-based segmentation using multiple automatically generated templates. Hum Brain Mapp 2013;34:2635–2654.
22. Pipitone J, Park MTM, Winterburn J, et al. Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. Neuroimage 2014;101:494–512.
23. Ellis KA, Bush AI, Darby D, et al. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. Int Psychogeriatr 2009;21:672–687.
24. Cyarto EV, Lautenschlager NT, Desmond PM, et al. Protocol for a randomized controlled trial evaluating the effect of physical activity on delaying the progression of white matter changes on MRI in older adults with memory complaints and mild cognitive impairment: The AIBL Active trial. BMC Psychiatry 2012;12.
25. Cook MJ, Fish DR, Shorvon SD, Straughan K, Stevens JM. Hippocampal volumetric and morphometric studies in frontal and temporal lobe epilepsy. Brain 1992;115:1001–1015.
26. Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 2010;53:1–15.
27. Buckner RL, Head D, Parker J, et al. A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. Neuroimage 2004;23:724–738.
28. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 2008;12:26–41.
29. Avants B, Gee JC. Geodesic estimation for large deformation anatomical shape averaging and interpolation. Neuroimage 2004;23(Suppl 1):S139–S150.
30. Klein A, Andersson J, Ardekani BA, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 2009;46:786–802.
31. Chen PY, Popovich PM. Correlation - parametric and nonparametric measures. Thousand Oaks, CA: SAGE Publications; 2002. 104 p.
32. Zhang YJ, Zhang JY, Hsu J, et al. Evaluation of group-specific, whole-brain atlas generation using Volume-based Template Estimation (VTE): application to normal and Alzheimer's populations. Neuroimage 2014;84:406–419.
33. Barnes J, Whitwell JL, Frost C, Josephs KA, Rossor M, Fox NC. Measurements of the amygdala and hippocampus in pathologically confirmed Alzheimer disease and frontotemporal lobar degeneration. Arch Neurol 2006;63:1434–1439.
34. Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL. Open access series of imaging studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J Cogn Neurosci 2007;19:1498–1507.
35. Bergouignan L, Chupin M, Czechowska Y, et al. Can voxel based morphometry, manual segmentation and automated segmentation equally detect hippocampal volume differences in acute depression? Neuroimage 2009;45:29–37.
36. Adler DH, Pluta J, Kadivar S, et al. Histology-derived volumetric annotation of the human hippocampal subfields in postmortem MRI. Neuroimage 2014;84:505–523.
37. Carmichael OT, Aizenstein HA, Davis SW, et al. Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. Neuroimage 2005;27:979–990.
38. Shen DG, Davatzikos C. HAMMER: hierarchical attribute matching mechanism for elastic registration. IEEE Trans Med Imaging 2002;21:1421–1439.
39. Chupin M, Mukuna-Bantumbakulu AR, Hasboun D, 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.