

Validation of a Fully Automated Hippocampal Segmentation Method on Patients With Dementia

Michael J. Firbank,* Robert Barber, Emma J. Burton, and John T. O'Brien

Institute for Ageing and Health, Newcastle University, Wolfson Research Centre, Westgate Road, Newcastle upon Tyne NE4 6BE, United Kingdom

Abstract: We describe a fully automated method for hippocampal segmentation. The method uses SPM5 (<http://www.fil.ion.ucl.ac.uk/spm/>) software to segment the brain into grey/white matter, and spatially normalize the images to standard space. Grey matter pixels within a predefined hippocampal region in standard space are identified to segment the hippocampi. The method was validated on 36 subjects (9 each of Alzheimer's disease, dementia with Lewy bodies, vascular dementia, and healthy controls). The mean absolute difference in volume compared with manual segmentation was 11% (SD 9%). Linear regression between manual and automated volume gave $V(\text{auto}) = V(\text{manual}) \times 0.83 + 401$ ml. The method provides an acceptable automated alternative to manual segmentation which may be of value in large studies. *Hum Brain Mapp* 29:1442–1449, 2008. © 2007 Wiley-Liss, Inc.

Key words: MRI, hippocampus, automated segmentation, atrophy

INTRODUCTION

Hippocampal atrophy is common in dementia and mild cognitive impairment [Barber et al., 2000; Chetelat and Baron, 2003] and can be useful in differential diagnosis. The gold standard for measuring hippocampal volume is manual segmentation using a well described protocol eg [Jack et al., 1995]. However, manual segmentation is both tedious and time consuming (~30 min per hippocampus), which limits its use in large scale studies or in routine clinical practice. Segmentation protocols vary between centers and between operators, making superficially simple comparisons between centers, such as magnitude of hippocam-

pal volumes in controls, impossible because of methodological differences.

The hippocampus is a difficult subject for automated segmentation because of a lack of clear boundaries apart from the cerebrospinal fluid (CSF) interface, and variability in its exact location and shape between individuals. This is particularly so in elderly individuals, where atrophy due to degenerative processes such as Alzheimer's disease (AD) can lead to considerable deviation in hippocampal geometry. A number of algorithms have been described for hippocampal segmentation, most of which are semiautomated in that they need an operator to provide some starting point, or correction during the segmentation ([Chupin et al., 2007] and references therein). Carmichael et al. [2005] described a comparison of fully automated segmentation procedures using image deformation to match the subjects image to an atlas image on which the hippocampi had been drawn.

Here, we describe a new approach in which the whole image is segmented into grey and white matter using the SPM5 software (<http://www.fil.ion.ucl.ac.uk/spm/>). The software also calculates a spatial transformation from the image to a template. Automatic segmentation of the hippocampus is then achieved by the identification of the

*Correspondence to: Michael J. Firbank, Institute for Ageing and Health, Wolfson Research Centre, Newcastle General Hospital, Westgate Road, Newcastle upon Tyne NE4 6BE, United Kingdom. E-mail: m.j.firbank@ncl.ac.uk

Received for publication 2 April 2007; Revised 18 June 2007; Accepted 8 August 2007

DOI: 10.1002/hbm.20480

Published online 2 November 2007 in Wiley InterScience (www.interscience.wiley.com).

grey matter pixels within a hippocampal region on the template. In this approach, the hippocampal region on the template is somewhat larger than the actual hippocampus, and the spatial transformation does not need to exactly match the subject's hippocampus to the template one, merely to ensure that the hippocampal grey matter is located within the region.

We describe the validation of the technique on a group of elderly subjects, including a substantial number with dementia.

METHODS

Subjects

Data were taken from a previously published study of atrophy in dementia [Barber et al., 2000], and we selected at random 36 subjects. Diagnoses of AD, vascular dementia (VaD), and dementia with Lewy bodies (DLB) were made in accordance with National Institute of Neurological and Communicative Disorders and Stroke – AD and Related Disorders Association (NINCDS-ADRDA) [McKhann et al., 1984], National Institute for Neurological Disorders and Stroke – Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN [Roman et al., 1993] and DLB consensus criteria [McKeith et al., 1996], and were arrived at by consensus between three experienced old age psychiatrists. Applying these criteria, nine patients had DLB, nine AD, nine VaD. There were also nine control subjects of comparable age who had no evidence of dementia. All subjects were assessed with a range of neuropsychological testing, including the mini-mental state exam (MMSE) [Folstein et al., 1975].

MRI

All scans were performed on a 1.0T Siemens Magnetom Impact Expert MRI Scanner (Erlangen, Germany). A T1 weighted 3D magnetization prepared rapid acquisition gradient-echo, turbo flash sagittal sequence was used to acquire whole brain images (repetition time 11.4 ms, echo time 4.4 ms, inversion time 400 ms, delay time, 50 ms, matrix 256×256 , field of view 256×256 mm², slice thickness 1 mm).

Manual Hippocampal Segmentation

Standard anatomic boundaries were used to define the hippocampus [Duvernoy, 1998; Jack et al., 1997]. The measurement included the hippocampus proper, dentate gyrus, subicular complex, alveus, and fimbria. The hippocampus was measured from the first slice identifying the head to the slice showing the longest length of fornix. All segmentations were performed by the same operator (RB). Intra rater reliability was assessed by measuring seven

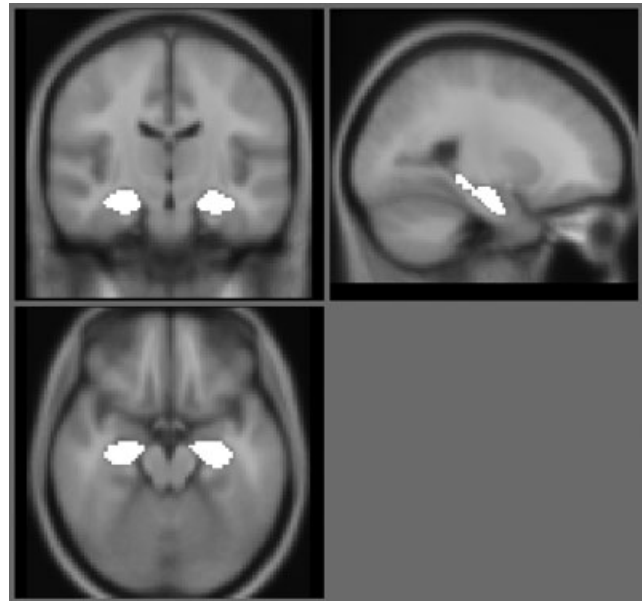


Figure 1.

The hippocampal regions overlaid onto the MNI 152 average brain.

subjects on three occasions. The mean coefficient of variation for hippocampal volume was 3%.

Automated Hippocampal Segmentation

These were undertaken blind to the results of the manual segmentation and by a different operator (MJF). First, we drew a region of interest on both the left and right hippocampus on the MNI (Montréal Neurological Institute) 152 average brain T1 weighted image. Regions were drawn so as to encompass all of the grey matter, and extended coronally from the first slice on which the hippocampus was visible to the crossing of the fornix. Figure 1 shows the hippocampal regions overlaid on the MNI template. This generated hippocampal regions of interest in standard space which were used for all further steps of the segmentation, and are referred to as template hippocampal regions.

The subject's images were segmented into grey and white matter and spatially normalized to MNI space using SPM5. This was performed using the default settings and the standard MNI template in SPM.

Using the spatial normalization transformation calculated for each subject, the left and right template hippocampal regions were transformed into the subject's native space image. Regions of hippocampal grey matter were identified by masking the grey matter segmentation with the transformed template hippocampal regions using the `imcalc` tool in SPM.

We compared two equations for this purpose

$$(\text{template ROI} > 0.5) \times (\text{greymatter} > 0.5) \quad (1)$$

$$(\text{template ROI} > 0.5) \times (\text{template ROI} + \text{greymatter} > 1.3). \quad (2)$$

The first is a straightforward identification of all pixels within the likely ROI which are likely to be grey matter. Visual inspection of the images showed that SPM correctly identifies with high probability the grey matter within the hippocampus. However, the SPM segmentation tends to incorrectly classify some pixels on WM/CSF boundary as GM, and the hippocampus contains some small regions which SPM correctly identifies as white matter. The second equation was an attempt to exclude the former and include the latter – the threshold for including voxels was set so that on the edge on the template (where templateROI was less than 1), a greater probability of being grey matter was required, whereas within the template, a less strict criteria applies. This might have the effect of including some CSF, though the SPM segmentation seemed quite good at identifying with high probability the CSF.

After this step, there remained by visual inspection on a subsample of images, isolated pixels not attached to the hippocampus. To remove these, and to fill in small gaps in the region, the region was smoothed with a 1.5 voxel Gaussian kernel, then thresholded at 50%. Figure 2 shows the steps in the automated segmentation. The template hippocampal regions are available from the author's website: (<http://www.staff.ncl.ac.uk/m.j.firbank/segmentation.html>).

Custom SPM a Priori Probabilities for Segmentation/Spatial Normalization

The templates and prior probability grey/white matter distributions provided with SPM are averages of the brains of healthy young adults. Older adults tend to have larger ventricles, and more atrophy, and experience with previous versions of SPM has found that creating a specific template can improve the segmentation and spatial normalization [Burton et al., 2002; Good et al., 2001]. We investigated the use of custom templates on the hippocampal segmentation. To do this, we took MRI scans [1.5T scanner with a 3D gradient echo T1-weighted sequence: repetition time 10 ms, echo time 4.6 ms, flip angle 20°, 1 mm cubic voxels] from 60 older individuals (different to the ones used for the hippocampus segmentation). Subject characteristics are described elsewhere [Firbank et al., 2003] but briefly, the subjects were community dwelling, nondemented and the average age was 73 years (SD 5, range 64–84 years). The MRI scans were segmented using the standard SPM5 parameters, and grey, white matter and CSF saved in subject's native space. The segmentations were visually inspected, and with the exception of two subjects who had excessive motion, and were removed, the segmentations were all acceptable. An affine spatial registration was used

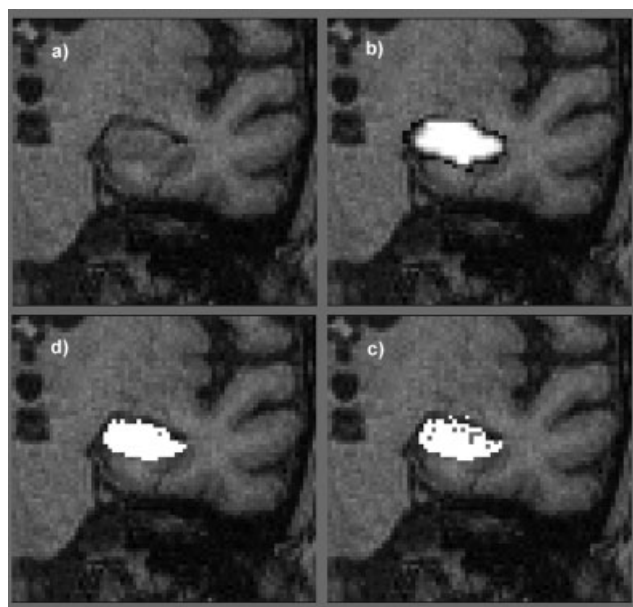


Figure 2.

Steps in the automated segmentation. Clockwise from top left (a) image (b) region transformed from MNI space (c) after multiplying by grey matter segmentation (d) final result after smoothing with Gaussian.

to match each subject's grey matter segmentation to the mean grey apriori SPM map. The grey matter affine registrations were then applied to each subject's grey, white matter and CSF maps, and after affine transformation to MNI space, the maps were averaged to create custom grey, white matter and CSF apriori maps, which were then used in the segmentation and normalization of the hippocampal subjects in the same manner as described earlier.

AIR (Automated Image Registration) Template Matching

To provide a further comparison, we used the automated image registration (AIR) package [Woods et al., 1998] to perform nonlinear warping of the scans to match an individual template. The individual template was of a 70-year-old healthy male who was scanned five times on the same day. The five scans were aligned together and averaged to create a high SNR image (see Fig. 3). On the template were drawn left and right hippocampi regions.

For each subject, the brain was automatically extracted using the BET tool [Smith, 2002], and then we used AIR v 5.2.5 (<http://bishopw.loni.ucla.edu/AIR5/>) with a 6th order polynomial to match the template and subject image. The resulting 'warp files' were applied to the template hippocampal ROIs to transform them to the individual subjects.

We also performed an additional step similar to above to identify the grey matter (from SPM segmentation) within the ROI.

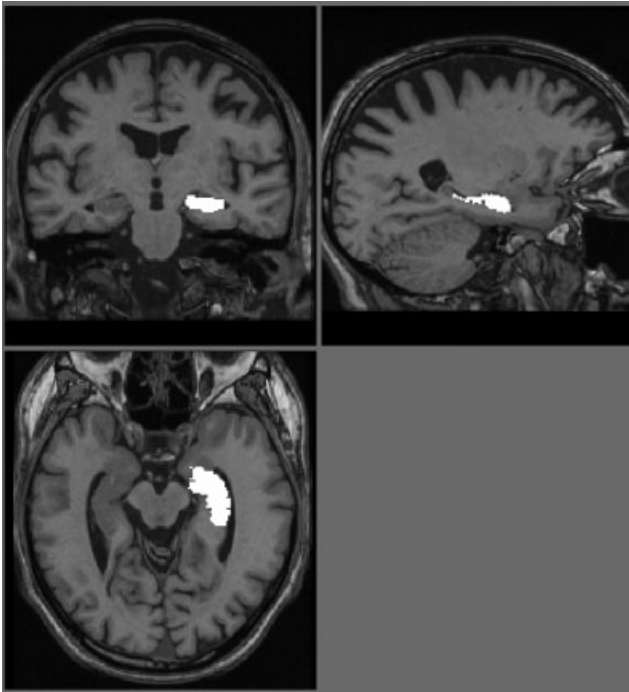


Figure 3.

Individual template used for AIR warp. Left hippocampus ROI overlaid.

Hippocampal Similarity Measures

We used the following quantitative measures to determine how well the automated segmentation performed.

- a. relative error (RE)

$$RE = 2 \frac{|V_m - V_a|}{V_m + V_a}$$

where V_m is the manually determined volume, and V_a the automated volume.

- b. Overlap ratio (OR)

$$OR = \frac{V(m \cap a)}{V(m \cup a)} = \frac{V_{ma}}{V_m + V_a - V_{ma}}$$

where V_{ma} is the volume of overlap ie the volume of those pixels in both V_m and V_a

- c. Similarity index (SI)

$$SI = \frac{V_{ma}}{(V_m + V_a)/2}$$

- d. Precision = $\frac{V_{ma}}{V_a}$

- e. Recall = $\frac{V_{ma}}{V_m}$

We also calculated the linear regression between the two volumes.

Reproducibility

Since this is a fully automatic segmentation method, repeating the analysis on the same scan produces exactly the same results. To determine the variability due to rescanning, one control subject was scanned five times on the same day (being repositioned in the scanner each time). The separate scans were segmented automatically using the SPM standard segmentation method, along with Eq. (2) above to identify grey matter followed by the smoothing step. The brains were then spatially registered together and the spatial registration parameters applied to the hippocampal segmentations, which were compared by determining the segmented pixel overlap.

RESULTS

Subject characteristics are shown in Table I. The hippocampal similarity values for the different segmentation methods are shown in Table II. As the relatively high recall figures show, the transformation of a ROI from atlas to subject either with SPM or AIR resulted in the inclusion of a reasonable proportion of the manually segmented hippocampus, but also much extraneous matter as shown by the relatively low Precision. As the regression figures show, this resulted in a poor relationship between the automated and manual hippocampal volume measurement. Inclusion of the grey matter segmentation greatly improved both the AIR and the SPM hippocampal segmentations in terms of the volume regression. We used two different equations to mask the hippocampal tissue (Eqs. (1) and (2) above). There was little difference between these two for most of the validation measures, though the regression was better for the 2nd equation. Further smoothing and thresholding the segmentation slightly improved the results.

We also investigated the use of age matched custom templates in the SPM segmentation and normalization. The results of this are shown in Table II. They were not particularly different to the results using the standard SPM templates. Table II shows that the optimum method in terms of the validation measures was the standard SPM segmentation with grey matter masking and smoothing and further results presented use this method. Table III shows the results subdivided by dementia group. A comparison between the groups showed the method produced somewhat less accurate results on the AD group.

TABLE I. Subject characteristics

	DLB	AD	VaD	Controls
Age	75 (6)	75 (6)	80 (6)	75 (4)
F:M	5:4	4:5	4:5	3:6
MMSE	9.8 (7.3)	14.6 (4.4)	17.9 (2.9)	27.0 (1.6)

TABLE II. Comparison of segmentations on the whole group

	Regression versus manual volume slope (SE); constant (SE)	Relative error mean (SD)	Overlap ratio mean (SD)	Similarity index mean (SD)	Precision mean (SD)	Recall mean (SD)
ROI transformed from standard with SPM but no grey matter masking	0.23 (0.08); 2919 (182)	0.49 (0.26)	0.40 (0.10)	0.57 (0.10)	0.47 (0.13)	0.75 (0.07)
Transformed ROI and masked grey matter with Eq. (1)	0.70 (0.06); 643 (123)	0.12 (0.09)	0.50 (0.06)	0.66 (0.06)	0.66 (0.08)	0.67 (0.07)
Transformed ROI and masked grey matter with Eq. (2)	0.78 (0.06); 572 (135)	0.12 (0.10)	0.50 (0.06)	0.67 (0.05)	0.65 (0.07)	0.69 (0.07)
Smoothed ROI masked grey matter Eq. (2)	0.83 (0.06); 409 (141)	0.11 (0.09)	0.54 (0.06)	0.70 (0.05)	0.69 (0.08)	0.71 (0.08)
AIR transformed hippocampus ROI	0.09 (0.08); 2362 (183)	0.30 (0.21)	0.38 (0.08)	0.54 (0.09)	0.50 (0.14)	0.61 (0.08)
AIR transformed ROI with masked grey matter	0.44 (0.06); 666 (124)	0.27 (0.16)	0.46 (0.06)	0.63 (0.06)	0.72 (0.10)	0.56 (0.08)
Using custom template in SPM. ROI masked with grey matter and smoothed	0.78 (0.07); 462 (152)	0.12 (0.09)	0.54 (0.06)	0.70 (0.05)	0.70 (0.07)	0.70 (0.08)

Figure 4 shows a comparison between manually and automatically determined regions. Visual inspection showed that there were two main sources of error – (a) inclusion of tissue wrongly classed as grey matter, and (b) problems due to the difficulty in determining the hippocampal/amygdala boundary. Figure 5 shows some examples of these. Linear regression for the right hippocampus gave $V_a = V_m \times 0.81 + 406$, for the left hippocampus, $V_a = V_m \times 0.84 + 417$, and for average L + R volume: $V_a = V_m \times 0.83 + 401$ ml.

Figure 6 shows a plot of manual versus automated volume for the left hippocampus.

On the subject who was scanned five times, the mean volumes were: Left 3110 (SD 56) mm³ and Right 2805 (SD 36) mm³. This gives a mean coefficient of variance (CoV) of 1.5% which compares favorably with the manual CoV of 3% on seven different subjects. Comparing the automated segmentations, 67% of pixels were identified as hippocampus on all five scans; 77% on 4 out of 5; 84% on 3 scans, and 91% on 2 out of 5 scans. Considering all possible pairs of scans, the mean overlap ratio was 0.83 (SD 0.012) and similarity index 0.91 (SD 0.007).

On a pentium pc with 2.2 GHz processor and 1 GB of RAM, the global grey/white matter segmentation and spa-

TABLE III. Comparison of automated and manual segmentations using the standard SPM segmentation and normalisation to identify grey matter voxels within the hippocampus region together with smoothing

	DLB ($n = 9$)	AD ($n = 9$)	VaD ($n = 9$)	Controls ($n = 9$)	Total ($n = 36$)
Av. manual volume (ml)	2202 (483)	1615 (546)	1838 (234)	2884 (368)	2135 (634)
Automated volume (ml)	2240 (441)	1869 (530)	1776 (409)	2769 (331)	2163 (573)
Left relative error	0.12 (0.9)	0.16 (0.11)	0.11 (0.12)	0.07 (0.05)	0.11 (0.10)
Right relative error	0.11 (0.08)	0.17 (0.11)	0.14 (0.12)	0.07 (0.06)	0.12 (0.10)
Total relative error	0.11 (0.06)	0.15 (0.12)	0.12 (0.11)	0.05 (0.05)	0.11 (0.09)
Left overlap ratio	0.53 (0.07)	0.50 (0.07)	0.54 (0.05)	0.57 (0.04)	0.53 (0.06)
Right overlap ratio	0.55 (0.10)	0.51 (0.06)	0.51 (0.07)	0.59 (0.03)	0.54 (0.08)
Total overlap ratio	0.54 (0.08)	0.50 (0.06)	0.53 (0.05)	0.58 (0.03)	0.54 (0.06)
Left similarity index	0.69 (0.06)	0.66 (0.06)	0.69 (0.04)	0.72 (0.03)	0.69 (0.05)
Right similarity index	0.71 (0.08)	0.67 (0.06)	0.67 (0.06)	0.75 (0.03)	0.70 (0.07)
Total similarity index	0.70 (0.07)	0.67 (0.05)	0.69 (0.05)	0.74 (0.03)	0.70 (0.05)
Left precision	0.67 (0.08)	0.61 (0.05)	0.71 (0.04)	0.73 (0.05)	0.68 (0.07)
Right precision	0.72 (0.08)	0.63 (0.06)	0.70 (0.05)	0.77 (0.03)	0.70 (0.08)
Total precision	0.70 (0.08)	0.62 (0.05)	0.68 (0.08)	0.75 (0.04)	0.69 (0.08)
Left recall	0.72 (0.06)	0.72 (0.10)	0.69 (0.09)	0.72 (0.04)	0.71 (0.07)
Right recall	0.70 (0.10)	0.73 (0.09)	0.65 (0.10)	0.72 (0.04)	0.70 (0.09)
Total recall	0.71 (0.08)	0.73 (0.09)	0.67 (0.09)	0.72 (0.03)	0.71 (0.08)

An ANOVA comparison between the total measures for the 4 groups gave total relative error $F_{3,35} = 2.0$, $P = 0.14$; total similarity index $F_{3,35} = 2.97$, $P = 0.046$ (Scheffe post hoc AD < control P 0.039); total overlap ratio $F_{3,35} = 3.08$, $P = 0.041$ (Scheffe post hoc AD < control P 0.036); total precision $F_{3,35} = 6.5$, $P = 0.001$ (Scheffe post hoc AD < control P 0.001); total recall $F_{3,35} = 0.9$, $P = 0.46$.

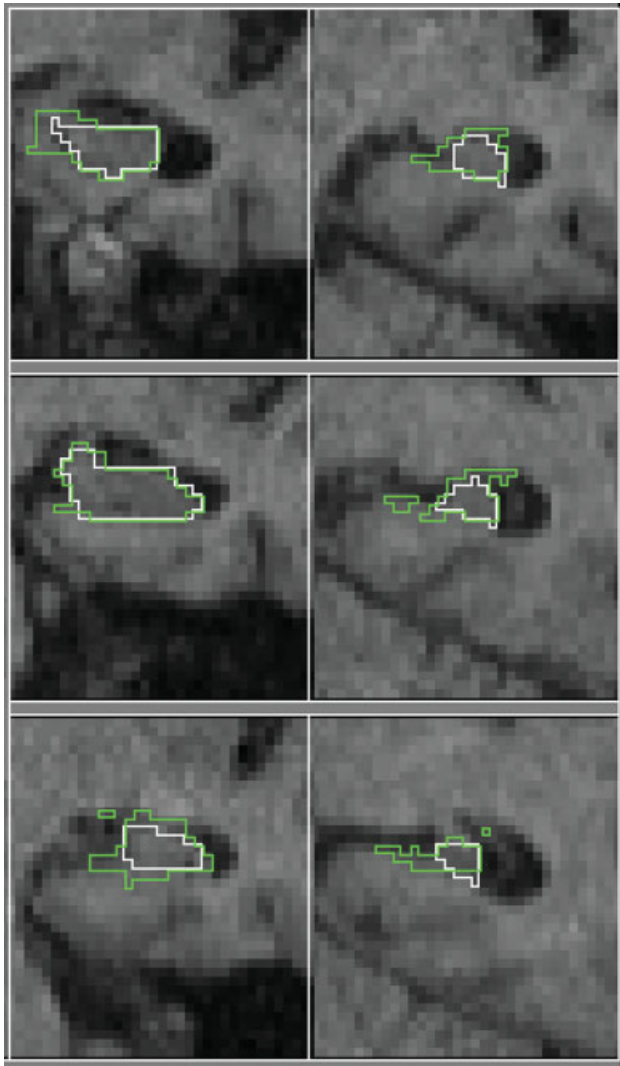


Figure 4.

Comparison of manual and automated segmentation. Coronal slices every 5 mm through a typical hippocampus. The manual segmentation is shown as a white contour, and the automated as a light grey (green online) contour. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

tial normalization took ~15 min, and the further segmentation of the hippocampus an additional minute.

DISCUSSION

The method produced good segmentations of the hippocampus with the widely used and freely available SPM5 software. The results are comparable with those of the fully automated segmentation of Carmichael [Carmichael et al., 2005] who obtained a overlap ratio of 0.55 using the

MNI atlas image with fully deformable registration and better than their results using SPM registration alone. They do not give timings for their segmentation, however fully deformable registration is a computer intensive procedure. We report in Table IV some accuracy values from other automated segmentation methods. Most of these require some operator input to indicate the location of the hippocampus. Our relative error value is comparable with these studies, though our overlap ratio is somewhat lower. Apart from Duchesne et al. (2002) they all have a relatively low number of subjects in the validation. A strength of our validation is that we included a large number of subjects picked at random from a population including mostly individuals with dementia.

The grey/white matter segmentation and spatial normalization in our method are likely to be useful steps in any further analysis, e.g. voxel based morphology [Good et al., 2002] and the additional processing to obtain hippocampal volumes from the normalized segmented images is very short (1 min).

We demonstrated that identifying grey matter within a hippocampal ROI improves the accuracy of the hippo-

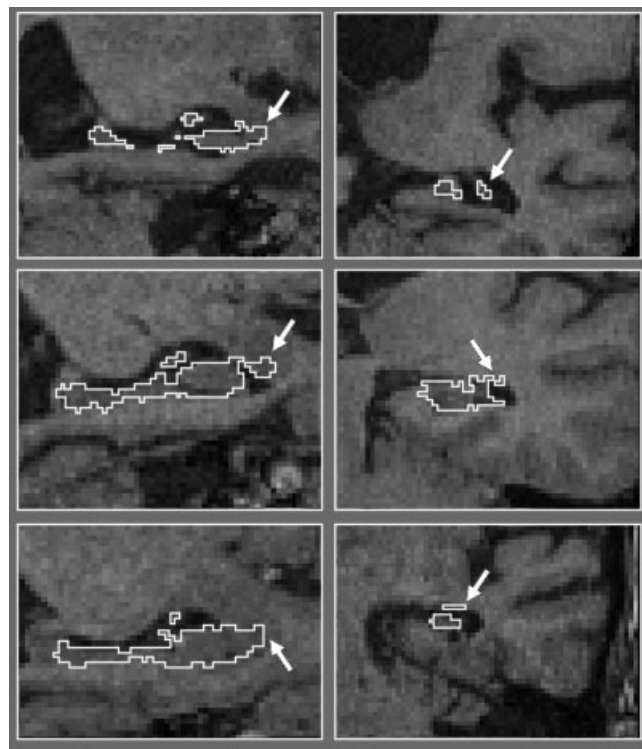


Figure 5.

Some examples of misclassified tissue. On the left, sagittal views showing amygdala pixels included in the hippocampus region. On the right, coronal views showing non grey matter tissue included in the hippocampus region. The region of misclassified tissue is shown with an arrow.

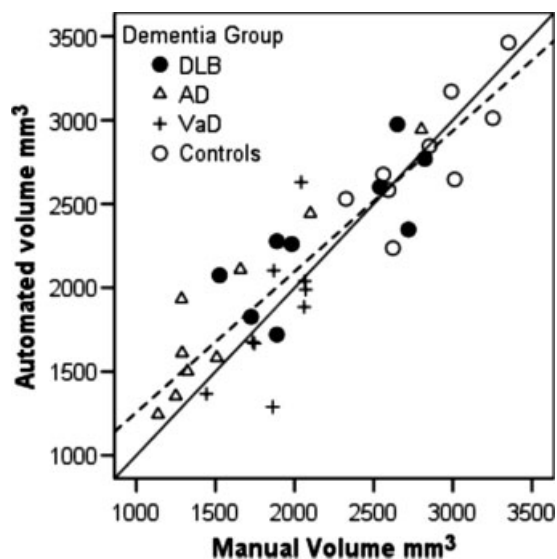


Figure 6.

Plot of manual versus automated left hippocampus volume. The diagonal solid line is the identity line, and the dashed line the regression $V_a = V_m \times 0.84 + 417$.

campal segmentation, particularly in terms of the accuracy of the volume measurement. The hippocampus does of course contain a small proportion of white matter, and our method may remove some of this from the hippocampal ROI. However, the improvement to the hippocampal segmentation from removing nonhippocampal white matter outweighs this small loss.

As would be expected, the segmentation was better in the control group, since controls have less atrophy. The MNI 152 average brain was generated from scans of young individuals who have little atrophy, and given this, the method works remarkably well in this population including mostly patients with degenerative dementias. An advantage of the MNI template is that the results are repeatable across different cohorts or centres. We investigated the use of a custom template from an age matched cohort, and did not find improvement. This may be

because although the hippocampus is atrophied in dementia, its location within the brain and hence relative to the MNI template is not changed. Also the atrophy process results in an increase of CSF, which is relatively easy to identify on T1 weighted scans.

The success of the method depends upon the global grey/white matter segmentation. Our data were collected at 1.0 Tesla. It is likely that the improved contrast to noise on more modern MRI scanners will lead to improved segmentation. In principle, this method could be used for the determination of any grey matter regional volume (e.g., basal ganglia) by having a suitable ROI in MNI space. Other approaches to segmentation, e.g., utilizing multiple image sequences or segmentation techniques which do not rely on prior probabilities, and hence are more generalizable to individuals with abnormal brains may bring about improvement in the method.

The Recall was 0.75 for the hippocampal ROI before masking with the grey matter segmentation, indicating that 75% of the manually drawn hippocampus ROI was within the ROI in MNI space. After extracting the grey matter from the ROI, the recall only decreased to 0.70 suggesting that not much of the true hippocampus was removed by the grey matter segmentation step. The precision for the Alzheimer's group was significantly lower than the controls, indicating that in that group a greater proportion of extra-hippocampal matter was falsely put into the ROI. This is not surprising given the tissue damage to the medial temporal region in AD. The lower Precision in the AD group is consistent with the regression slope implying that the automated segmentation overestimates volume for small hippocampi. This is likely to reduce somewhat the difference between groups in measured hippocampal volume and may decrease the statistical significance of studies comparing hippocampi with the method.

In summary, we have described and validated a fully automated and quick method using freely available software for determining hippocampal volume in a group of people with dementia. The method may be of utility in large scale studies of hippocampal atrophy in dementia as it provides a rapid technique with acceptable accuracy.

TABLE IV. Similarity index and relative error of manual and automated methods of hippocampus segmentation from different methods

Reference	Examined group	Operator interaction	Similarity index	Relative error
[Hogan et al., 2000]	5 young adults with temporal sclerosis	20 landmarks manually placed	75	10
[Shen et al., 2002]	10 subjects aged 55–85	50 landmarks manually placed	88	6
[Fischl et al., 2002]	7 young adults	None mentioned	80	10
[Duchesne et al., 2002]	70 normal subjects	None	68	—
[Chupin et al., 2007]	8 Alzheimer's disease	Box around structure and 2 seeds	84	9

REFERENCES

- Barber R, Ballard C, McKeith IG, Gholkar A, O'Brien JT. (2000): MRI volumetric study of dementia with Lewy bodies. *Neurology* 54:1304–1309.
- Burton EJ, Karas G, Paling SM, Barber R, Williams ED, Ballard CG, McKeith IG, Scheltens P, Barkhof F, O'Brien JT. (2002): Patterns of cerebral atrophy in dementia with Lewy bodies using voxel based morphometry. *Neuroimage* 17:618–630.
- Carmichael OT, Aizenstein HA, Davis SW, Becker JT, Thompson PM, Meltzer CC, Liu Y. (2005): Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. *NeuroImage* 27:979–990.
- Chetelat G, Baron JC. (2003): Early diagnosis of Alzheimer's disease: Contribution of structural neuroimaging. *Neuroimage* 18:525–541.
- Chupin M, Mukuna-Bantumbakulu AR, Hasboun D, Bardinet E, Baillet S, Kinkingnéhun S, Lemieux L, Dubois B, Garnero L. (2007): 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* 34:996–1019.
- Duchesne S, Pruessner JC, Collins DL. (2002): Appearance-based segmentation of medial temporal lobe structures. *Neuroimage* 17:515–531.
- Duvernoy HM. (1998): *The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections with MRI*. Berlin: Springer Verlag.
- Firbank MJ, Minett T, O'Brien JT. (2003): Changes in DWI and MRS associated with white matter hyperintensities in elderly subjects. *Neurology* 61:950–954.
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. (2002): Whole brain segmentation: Automated labeling of neuroanatomical structure in the human brain. *Neuron* 33:341–355.
- Folstein MF, Folstein SE, McHugh PR. (1975): A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12:189–198.
- Good CD, Johnsrude IS, Ashburner J, Henson RNA, Friston KJ, Frackowiak RSJ. (2001): A voxel based morphometric study of ageing in 465 normal adult human brains. *Neuroimage* 14:21–36.
- Good CD, Scahill RI, Fox NC, Ashburner J, Friston KJ, Chan D, Crum WR, Rosser MN, Frackowiak SJ. (2002): Automatic differentiation of anatomical patterns in the human brain: Validation with studies of degenerative dementias. *Neuroimage* 17:29–46.
- Hogan RE, Mark KE, Wang L, Joshi S, Miller MI, Bucholz RD. (2000): Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients. *Radiology* 216:291–297.
- Jack CR Jr, Theodore WH, Cook M, McCarthy G. (1995): MRI based hippocampal volumetrics: Data acquisition, normal ranges, and optimal protocol. *Magn Reson Imaging* 13:1057–1064.
- Jack CR Jr, Petersen RC, Xu YC, Waring SC, O'Brien PC, Tangalos EG, Smith GE, Ivnik RJ, Kokmen E. (1997): Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology* 49:786–794.
- McKeith IG, Galasko D, Kosaka K, Perry EK, Dickson DW, Hansen LA, Salmon DP, Lowe J, Mirra SS, Byrne EJ, Lennox G, Quinn NP, Edwardson JA, Ince PG, Bergeron C, Burns A, Miller BL, Lovestone S, Collerton D, Jansen ENH, Ballard C, deVos RAI, Wilcock GK, Jellinger KA, Perry RH. (1996): Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): Report of the consortium on DLB international workshop. *Neurology* 47:1113–1124.
- McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. (1984): Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology* 34:939–944.
- Roman GC, Tatemichi T, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, Amaducci L, Orgogozo JM, Brun A, Hofman A, Moody DM, O'Brien MD, Yamaguchi T, Grafman J, Drayer BP, Bennett DA, Fisher M, Ogata J, Kokmen E, Bermejo F, Wolf PA, Gorelick PB, Bick KL, Pajean AK, Bell MA, DeCarli C, Culebras A, Korczyn AD, Bogousslavsky J, Hartmann A, Scheinberg P. (1993): Vascular dementia—diagnostic criteria for research studies—report of the NINDS-AIREN international workshop. *Neurology* 43:250–260.
- Shen DG, Moffat S, Resnick SM, Davatzikos C. (2002): Measuring size and shape of the hippocampus in MR images using a deformable shape model. *Neuroimage* 15:422–434.
- Smith SM. (2002): Fast robust automated brain extraction. *Hum Brain Mapp* 17:143–155.
- Woods RP, Grafton ST, Holmes CJ, Cherry SR, Mazziotta JC. (1998): Automated image registration. I. General methods and intrasubject, intramodality validation. *J Comput Assist Tomogr* 22:139–152.