

Validation of hippocampal volumes measured using a manual method and two automated methods (FreeSurfer and IBASPM) in chronic major depressive disorder

Woo Suk Tae · Sam Soo Kim · Kang Uk Lee · Eui-Cheol Nam · Keun Woo Kim

Received: 30 October 2007 / Accepted: 5 March 2008
© Springer-Verlag 2008

Abstract

Introduction To validate the usefulness of the packages available for automated hippocampal volumetry, we measured hippocampal volumes using one manual and two recently developed automated volumetric methods.

Methods The study included T1-weighted magnetic resonance imaging (MRI) of 21 patients with chronic major depressive disorder (MDD) and 20 normal controls. Using coronal turbo field echo (TFE) MRI with a slice thickness of 1.3 mm, the hippocampal volumes were measured using three methods: manual volumetry, surface-based parcellation using FreeSurfer, and individual atlas-based volumetry using IBASPM. In addition, the intracranial cavity volume (ICV) was measured manually.

Results The absolute left hippocampal volume of the patients with MDD measured using all three methods was significantly smaller than the left hippocampal volume of the normal controls (manual $P=0.029$, FreeSurfer $P=0.035$, IBASPM $P=0.018$). After controlling for the ICV, except for the right hippocampal volume measured using Free Surfer, both measured hippocampal volumes of the patients with MDD were significantly smaller than the measured hippocampal volumes of the normal controls (*right* manual

$P=0.019$, IBASPM $P=0.012$; *left* manual $P=0.003$, Free Surfer $P=0.010$, IBASPM $P=0.002$). In the intrarater reliability test, the intraclass correlation coefficients (ICCs) were all excellent (*manual* right 0.947, *left* 0.934; *Free Surfer* right 1.000, *left* 1.000; *IBASPM* right 1.000, *left* 1.000). In the test of agreement between the volumetric methods, the ICCs were right 0.846 and *left* 0.848 (manual and FreeSurfer), and right 0.654 and *left* 0.717 (manual and IBASPM).

Conclusion The automated hippocampal volumetric methods showed good agreement with manual hippocampal volumetry, but the volume measured using FreeSurfer was 35% larger and the agreement was questionable with IBASPM. Although the automated methods could detect hippocampal atrophy in the patients with MDD, the results indicate that manual hippocampal volumetry is still the gold standard, while the automated volumetric methods need to be improved.

Keywords Magnetic resonance imaging · Hippocampus · Volumetry · Validation · Major depressive disorder

Introduction

MRI-based hippocampal volumetry has been widely used in various neuropsychiatric disorders. In the past 20 years, the number of clinical studies using hippocampal volumetric assessment has increased rapidly because the hippocampus plays an important role in forming memory and regulating emotion, and is involved in various neuropsychiatric diseases, such as depression, schizophrenia, epilepsy, dementia, and sleep disorders [1]. To optimize the determination of hippocampal volume, various manual [2] and automated protocols have been described [3–11]. Taking advantage of

W. S. Tae · S. S. Kim · K. U. Lee · E.-C. Nam · K. W. Kim
Neuroscience Research Institute,
Kangwon National University College of Medicine,
Kangwon, Korea

S. S. Kim (✉)
Department of Radiology, Kangwon National University Hospital,
17-1 Hyoja 3-dong, Chuncheon,
Kangwon-do 200-947, Korea
e-mail: samskim@kangwon.ac.kr

advanced image-processing techniques, a precise manual hippocampal volumetry method was introduced in the late 1980s, and its precision was validated with a phantom [12–17]. The advent of thin-slice three-dimensional (3-D) MRI with slice thicknesses of 1 to 2 mm and advanced measurement techniques for brain volumetry has enabled the detection of mild hippocampal atrophy and the prediction of Alzheimer disease, mild cognitive impairment, and epilepsy [18–20]. In addition, although 1.5-T MRI is not able to resolve the details of the hippocampal cellular anatomy reliably, the borders of the hippocampal formation can be distinguished sufficiently for the entire hippocampus [21], and 3-T MRI is the current standard, rather than very-high-field MRI.

Despite the usefulness of manual hippocampal volumetry, it is time-consuming and dependent on rater experience, and automated methods with high reproducibility and accuracy may potentially be more efficient than manual volumetry. Several automated volumetric methods have been developed. The FreeSurfer [8] software package (<http://surfer.nmr.mgh.harvard.edu>) provides completely automated parcellation of the brain cortex and subcortical structures. FreeSurfer calculates brain subvolumes by assigning a neuroanatomical label to each voxel in a MRI volume based on probabilistic information estimated automatically from a manually labeled training set [22]. Individual Brain Atlases using Statistical Parametric Mapping (IBASPM) [3] is a fully automated, easy volumetric method (<http://www.thomaskoenig.ch/Lester/ibaspm.htm>) based on the SPM software package (<http://www.fil.ion.ucl.ac.uk/spm>). Volume measurement with IBASPM software is based on an individual brain atlas masked using a predefined Automated Anatomical Labeling (AAL) atlas [23]. Free Surfer and IBASPM are freely available via the Internet, and the processes for volumetric measurements are automated, whereas other automated methods are not available publicly or the processes are complicated [4–7, 9–11]. The automated volumetric methods that we tested here are at the cutting edge of medical imaging technology. These packages implement hundreds of recent algorithms, such as linear and nonlinear spatial normalization, surface base registration, intensity nonuniformity correction, segmentation, and skull stripping, although the users of FreeSurfer and IBASPM can perform automated hippocampal volumetry without knowing all of them.

If these automated methods provide brain subvolumes with high correlation, investigation of brain morphology would be very easy and simple. Although each development group has validated the various automated volumetric methods, users must revalidate the methods using their own MRI images because image quality is affected by noise, the acquisition parameters, or the specific MRI machine [24]. Therefore, in this study, we compared two recently

developed automated volumetric methods using FreeSurfer and IBASPM with a precisely defined manual method regarded as the gold standard. In addition, we validated the usefulness of these automated methods prior to their application in other hippocampal volumetric studies.

Materials and methods

Subjects

A group of 21 female patients aged 18–60 years with major depressive disorder (MDD) were recruited from the Department of Psychiatry at our hospital. In addition, 20 female normal control subjects, matched to the patients in terms of age, handedness, and parental socioeconomic status, were recruited from the community. This study was approved by the hospital institutional review board, and all of the subjects provided written informed consent before participating. After an initial psychiatric interview, the subjects underwent a physical examination and screening tests that included a complete blood count, plasma electrolytes, liver function tests, thyroid function tests, and routine urine analysis.

The patients were included if they met the criteria for MDD based on the structured clinical interview (SCID) for DSM-IV. Depressive symptoms were assessed using the Beck depression inventory [25] and the 17-item Hamilton depression rating scale (HDRS) [26]. Both the SCID and HDRS were administered by an experienced psychiatrist (K.U.L.). Patients were excluded from the study if they had a history of childhood trauma or other major axis I disorder, including bipolar disorder, schizophrenia, schizoaffective disorder, claustrophobia, and a current or past history of alcohol or substance abuse or dependence. Patients with an underlying dysthymic disorder, that is double depression, or panic attacks in the context of MDD were included. All patients were on antidepressant medication and none was taking antipsychotics or mood stabilizers. The control subjects had no history or family history of axis I disorders. Neither the patients nor the control subjects had ever been traumatized enough to cause posttraumatic stress disorder. Subjects were also excluded if they had a major medical or neurological illness, history of significant head trauma or electroconvulsive therapy, or an IQ of less than 80. Subjects in whom MRI was contraindicated and those who had taken oral or intravenous steroids were also excluded.

As a group, the patients with depression had a mean \pm SD age of 41.7 ± 11.00 years (range 21–57 years) versus 41.9 ± 10.26 years (range 24–58 years) for the controls. All but one of the patients and all of the controls were strongly right-handed as assessed by the Edinburgh handedness inventory [27]. The mean age at the first depressive episode

was 33.2 ± 13.0 years; the mean number of life-time depressive episodes was 3.9 ± 3.3 , and the mean illness duration was 80.0 ± 67.0 months. Five patients presented with a first episode of depression, while 16 subjects had recurrent MDD. None of the patients showed atypical, catatonic features or had a postpartum onset. No patient reported psychotic symptoms, such as delusions and hallucinations. Four patients reported a family history of major depressive disorder.

MRI acquisition

All of the subjects were scanned with a 1.5-T MRI scanner (Gyroscan ACS-NT; Philips Medical Systems, Best, The Netherlands). Coronal 3-D T1-weighted turbo field echo (TFE) MRI was obtained with the following scanning variables: 1.3 mm thickness; no gap; 160 slices; scanning time 10 min 13 s; repetition time/echo time 10/4.3 ms; number of signal averages 1; matrix 256×256 ; field of view 22×22 cm; and flip angle 8° . Coronal slices were obtained perpendicular to the long axis of the anterior commissure (AC) to the posterior commissure (PC) in the midsagittal plane. The final voxel size was $0.86 \times 0.86 \times 1.30$ mm ($x \times y \times z$). The Philips data format PAR, REC, was converted to 16-bit Analyze format using the software MRIcro (<http://www.sph.sc.edu/comd/orden>), and was not converted to 8 bits to preserve the original intensity scale.

Volume measurements

Intracranial cavity volume

The T1-W MRI and manual volume measurements of the intracranial cavity volume (ICV) and hippocampus were preprocessed using a Unix-based Sun Ultra 1 Creator workstation (Sun Microsystems, San Diego, CA) and Analyze 7.5 (Biomedical Imaging Resource, Mayo Foundation, Rochester, MN). The ICV was used to normalize the regional brain volumes. Skull growth occurs along the suture lines and is determined by brain expansion, which takes place during normal growth of the brain [28], so the ICV represents the size of the brain at maximal maturity unaffected by age or disease-related atrophy. To measure the ICV, the original TFE MRI data were reconstructed into 5-mm-thick sagittal images, which were then magnified twofold. The cerebrum, cerebellum, and midbrain were included in the ICV volume [29–31]. The lateral limits of the ICV were defined as the right- and leftmost slices of the brain parenchyma on the sagittal images, while the lower tip of the cerebellum was the lower limit. We increased the brightness of the image to improve the visual clarity of the boundary of the dura mater for the outer boundary. Using the established measurement criteria [31, 32], the dura

mater of the cerebrum, the cerebellum, and the midbrain, except for the inferior boundary, were traced manually by a skilled technical staff member (W.S.T.), who was blind to all clinical information. A horizontal line was drawn across the midbrain to include the lower tip of the cerebellum to establish the inferior boundary on the head tilt-corrected sagittal images [30, 31]. Details of the ICV measurement are described in a previous paper [33].

To test the reproducibility of the ICV measurement, the ICVs of ten subjects (five normal controls and five patients with MDD) were remeasured.

Manual hippocampal volumetry

To reduce manual tracing errors, the coronal T1-W MR images were magnified four times and trilinear interpolation was performed. The entire hippocampus was measured from the anterior-most head to the posterior tail, including the cornu ammonis, gyrus dentatus, hippocampus, and subiculum [34]. The anterior-most boundary was identified with alveus and simultaneous 3D position markings in the coronal and sagittal views (Fig. 1). The investigator identified suspected boundary pixels from the coronal view, while the orthogonal view displayed the pixel in the same position in the sagittal view. On sagittal MR images, the hippocampus was easily differentiated from the amygdala [35]. The lateral border of the hippocampus was delineated against the entorhinal cortex by the upper margin of the white matter of the subiculum. The posterior-most hippocampus was measured to the end of the tail of the hippocampus. Sectors CA-1 through CA-4 of the hippocampus proper, the dentate gyrus, and the subiculum were included in the in-plane boundaries. The detailed in-plane boundary criteria used were as described previously [12]. According to the defined hippocampus boundary criteria, the rater outlined the hippocampus on the alveus manually with a track-ball mouse, and the tracing line was included as the hippocampal volume (Fig. 2). The ICV and hippocampal volume were calculated by multiplying the total voxel count by the volume of one voxel [36].

Automated hippocampal volumetry using FreeSurfer

The volume of the hippocampus was measured automatically by FreeSurfer 3.04 installed on a Fedora core 3 (<http://fedoraproject.org>) Linux workstation. The coronal MRI was reformatted to an axial image, and then Analyze format was converted to FreeSurfer's mgz format. The automated procedures for hippocampal parcellation have been described previously [8]. This procedure automatically assigns a neuroanatomical label to each voxel in a MRI volume based on probabilistic information estimated automatically from a manually labeled training set. The optimal linear transform

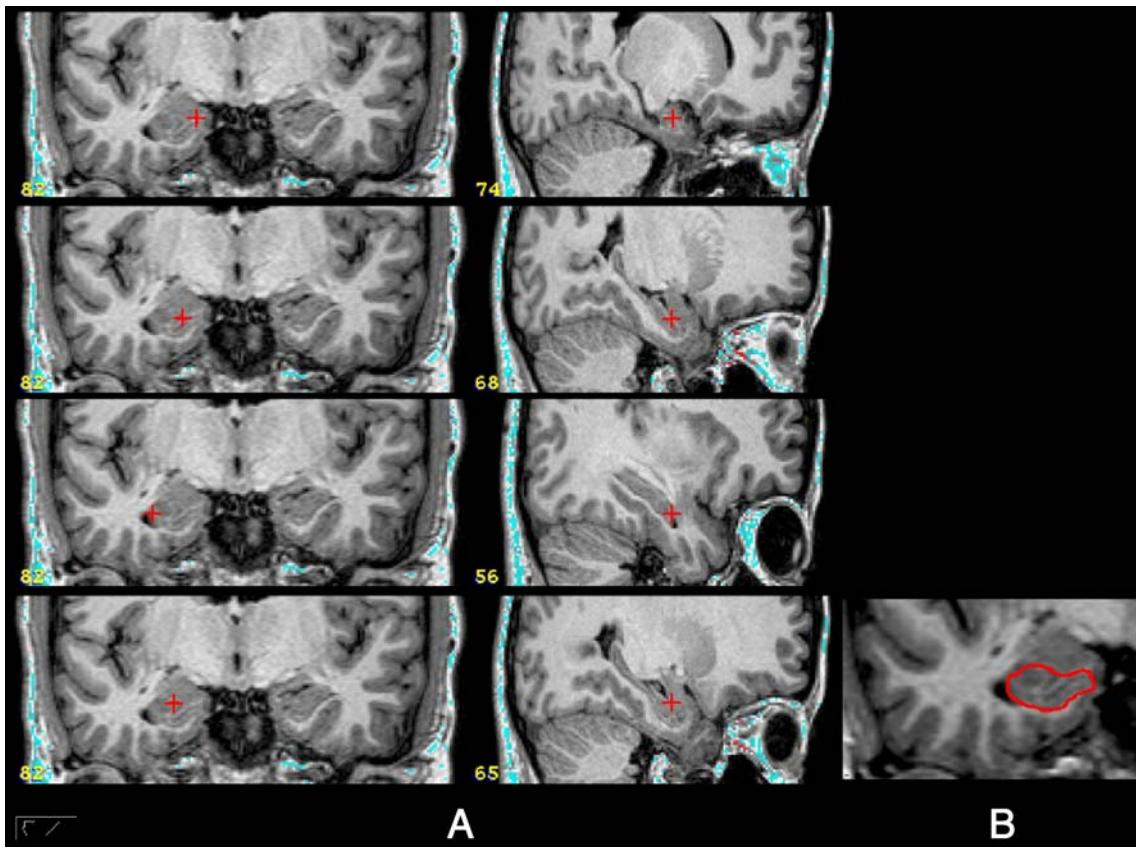


Fig. 1 The boundary separating the hippocampal head from the amygdala. The obscure boundary of the hippocampal head was identified with the alveus and simultaneous 3-D position markings in the coronal and sagittal views (**a**), and the tracing line was confirmed

(**b**). The red crosshairs in the coronal and sagittal images of the same row are at the same position. The tracing lines were redrawn manually with a thick line for review purposes only

was calculated, maximizing the likelihood of the input image using an atlas constructed from manually labeled images of 14 young and middle-aged subjects, and then the nonlinear transformation was computed. Bayesian segmentation was performed, and then the maximum a posteriori estimate of the labeling was determined [22].

Automated hippocampal volumetry using IBASPM

Using the IBASPM toolbox implemented in MATLAB 7.0 (MathWorks, Natick, MA), individual brain atlas-based volumetry was performed. IBASPM uses the spatial normalization and segmentation routines of SPM2 (Wellcome Department of Cognitive Neurology, Institute of Neurology, University College London, London, UK). This procedure consisted of the following five processes.

1. *Segmentation*: The MR image is segmented into gray matter, white matter, and cerebral spinal fluid in native space.
2. *Normalization*: The MR image is normalized to the ICBM 152 T1 template MNI space to obtain the spatial

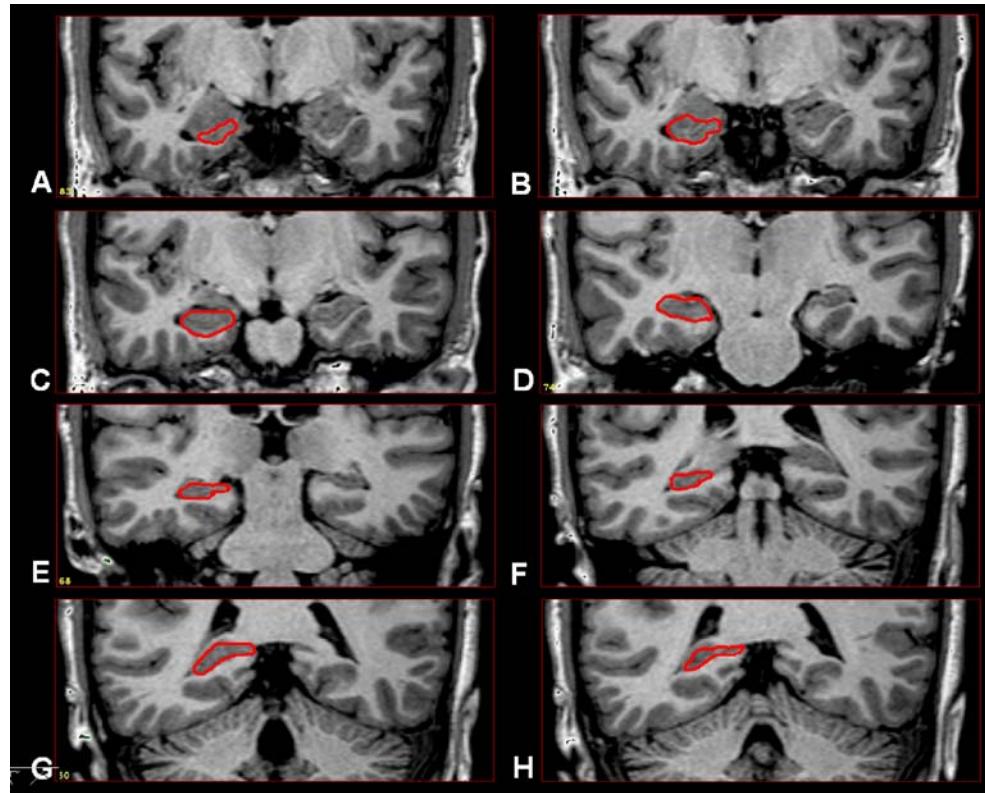
transformation matrix; using this transformation matrix, the gray matter image is transformed to MNI space.

3. *Automatic labeling*: Each normalized individual gray matter voxel is labeled based on an AAL atlas [23] of predefined 116 segmented structures.
4. *Atlasing*: Individual atlases are created from each subject's MR image. The individual atlases are reversely normalized, and the gray matter image in native space is masked using its individual 116 brain structures.
5. *Volume statistic*: For all individual atlases, the volumes of 116 predefined brain structures are calculated. When the voxel values of the gray matter image exceed the voxel values of the white matter and CSF images, the voxel is included for volume calculation.

All MRI data from the normal controls and patients with MDD were mixed in the same directory without clinical information before the measurements. For the intrarater reliability test, each volumetric method was performed twice for five normal controls and five patients.

The ICV was expressed in cubic centimeters, and hippocampal volume was expressed in cubic millimeters.

Fig. 2 The region of interest (ROI) definition of the hippocampus. The ROI of the entire hippocampus was traced manually from the anterior-most head (**a**) to the posterior-most tail (**h**) of the hippocampus. The MR images were magnified four times, and the tracing line was included in the total hippocampal volume. The alveus was used for the upper boundary, and the gray and white matter junction was used for the lower boundary. The tracing lines were redrawn manually with a thick line for review purposes only



Statistical analysis

Statistical analysis was performed with SPSS 11.5 (SPSS, Chicago, IL). The normality of the distribution was tested using the Kolmogorov-Smirnov test and normal plot. The significance of differences in the ICV and hippocampal volumes between the controls and patients were tested using the *t*-test. To compare the hippocampal volume, analysis of covariance (ANCOVA) was performed with the covariate ICV. To determine the degree of correlation between measurements simple regression analysis was used. The asymmetry between the right and left hippocampus in the normal controls and patients with MDD measured using the three volumetric methods was tested using a paired *t*-test.

Since the degree of simple correlation depends on the number of samples, a high correlation does not mean high agreement between the two methods, and the use of correlation could be misleading. Therefore, the reliability of the intraclass correlation coefficients (ICCs) for the intrarater reliability and the agreement between volumetric methods was tested using Cronbach's alpha [37]. A Bland-Altman plot was also used to examine the agreement between volumetric methods. A Bland-Altman plot is a method of data plotting used in comparing two different measurement methods that was popularized in medical statistics. One common application of the Bland-Altman plot is to compare a new measurement technique with a

gold standard based on graphical techniques and simple calculations [38].

The differences in the mean volumes of the hippocampus measured using the three volumetric methods were tested using one-way repeated ANOVA, and the differences in the variances of the hippocampal volume measured using the three volumetric methods were tested manually using Pitman's *t*-test (<http://core.ecu.edu/psyc/wuenschk/StatHelp/Pitman.doc>) [39]. The details of Pitman's *t* calculation are as follows:

1. Compute F as the ratio of the larger variance to the smaller variance.
2. Compute $t = \frac{(F-1)\sqrt{n-2}}{2\sqrt{F(1-r^2)}}$, where n is the number of pairs of scores and r is the correlation between the scores in sample 1 and the scores in sample 2.
3. Evaluate this t on $n-2$ degrees of freedom.

All tests were two-tailed, and the level of significance was $P<0.05$.

Results

The age distribution of the controls and patients did not differ ($P=0.956$), and the ICV was similar between the two groups ($P=0.942$). The absolute left hippocampal volumes as determined with all three volumetric methods were smaller in the patients with MDD than in the normal

Table 1 Volumes (means \pm SD) measured with one manual and two automated volumetric methods (intracranial cavity volumes are in cubic centimeters; hippocampal volumes are in cubic millimeters).

Method		Normal controls	MDD patients	P values	
				t-test	Ancova ^a
Manual	Intracranial cavity volume	1397.0 \pm 123.52	1393.6 \pm 170.84	0.942	
	Right hippocampus	2871.9 \pm 278.21	2658.5 \pm 412.31	0.061	0.019
	Left hippocampus	2811.1 \pm 226.51	2589.3 \pm 377.24	0.029	0.003
FreeSurfer	Right hippocampus	4017.7 \pm 414.56	3797.7 \pm 459.75	0.116	0.055
	Left hippocampus	3802.9 \pm 327.23	3558.2 \pm 386.82	0.035	0.010
IBASPM	Right hippocampus	2981.0 \pm 363.12	2714.4 \pm 475.89	0.051	0.012
	Left hippocampus	3456.3 \pm 388.75	3115.4 \pm 485.62	0.018	0.002

^a Ancova statistical test covariate with intracranial cavity volume.

controls (Table 1), whereas the right hippocampal volumes were at the borderline of significance. After correcting for ICV, both hippocampal volumes as determined with the manual and IBASPM volumetric methods were smaller in the patients with MDD than in the normal controls. Although the left hippocampal volumes measured using FreeSurfer were smaller than those in the normal controls, the right hippocampal volumes measured using FreeSurfer did not differ significantly.

The asymmetry between the right and left hippocampal volumes in the normal controls and patients was not significant. However, FreeSurfer showed the right hippocampal volume to be significantly larger than the left, and IBASPM showed the left hippocampal volume to be larger than the right (Table 2).

In the analysis of reliability between raters, the ICCs (Cronbach's alpha) for manual volumetry were 0.947 (right) and 0.934 (left), for FreeSurfer were 1.000 (right) and 1.000 (left), and for IBASPM were 1.000 (right) and 1.000 (left) (Fig. 3); and the ICC of the ICV was 0.998.

In the analysis of the agreement between the volumetric methods, the ICCs between manual volumetry and Free Surfer were 0.846 (right) and 0.848 (left), and between manual volumetry and IBASPM were 0.654 (right) and 0.717 (left). The correlations for the right hippocampus between manual volumetry and FreeSurfer were $r=0.749$ and $P<0.000001$, and between manual volumetry and IBASPM were $r=0.494$ and $P=0.001$. For the left hippocampus, the correlations were $r=0.743$ and $P<0.000001$ (manual vs. FreeSurfer), and $r=0.594$ and $P=0.000043$ (manual vs. IBASPM) (Fig. 4).

In the Bland-Altman plot, the mean differences (d_m) and limits of agreement ($d_m\pm 2SD$) for the volumetric methods were as follows: manual minus FreeSurfer measurements $-1142\pm 588 \text{ mm}^3$ (right) and $-980\pm 507 \text{ mm}^3$ (left); and manual minus IBASPM measurements $-82\pm 811 \text{ mm}^3$ (right) and $-584\pm 751 \text{ mm}^3$ (left) (Fig. 5).

The mean hippocampal volumes measured using the three volumetric methods (manual, FreeSurfer, and IBASPM) differed significantly (right hippocampus $P<0.000001$, left hippocampus $P<0.000001$; repeated ANOVA). In the post hoc analysis with Bonferroni's correction, the right hippocampal volumes differed significantly between the manual and FreeSurfer methods ($P<0.000001$) and between the FreeSurfer and IBASPM methods ($P<0.000001$), while the right hippocampal volumes did not differ significantly between the manual and IBASPM methods ($P=0.627$). The left hippocampal volumes differed significantly between the manual and FreeSurfer methods ($P<0.000001$), between the FreeSurfer and IBASPM methods ($P<0.000001$) and between the manual and IBASPM methods ($P<0.000001$).

The variance of each hippocampal volumetric method differed significantly for the right hippocampus (manual vs. FreeSurfer, $P=0.0008$; manual vs. IBASPM, $P=0.00520$), and for the left hippocampus (manual vs. FreeSurfer, $P=0.007$; manual vs. IBASPM, $P=0.0002$; FreeSurfer vs.

Table 2 Right-left hippocampal asymmetry (values are volumes in cubic millimeters).

	Manual		FreeSurfer		IBASPM	
	Controls	MDD	Controls	MDD	Controls	MDD
Mean difference ^a	60.9	69.1	214.8	239.4	-475.3	-401
SD	161.7	209.4	243.7	209.9	272.3	228.8
SE	36.2	45.7	54.5	45.8	60.8	49.9
P value ^b	0.109	0.146	0.001	<0.0001	<0.0001	<0.0001

^a Right-left hippocampal volume.

^b Paired t-test.

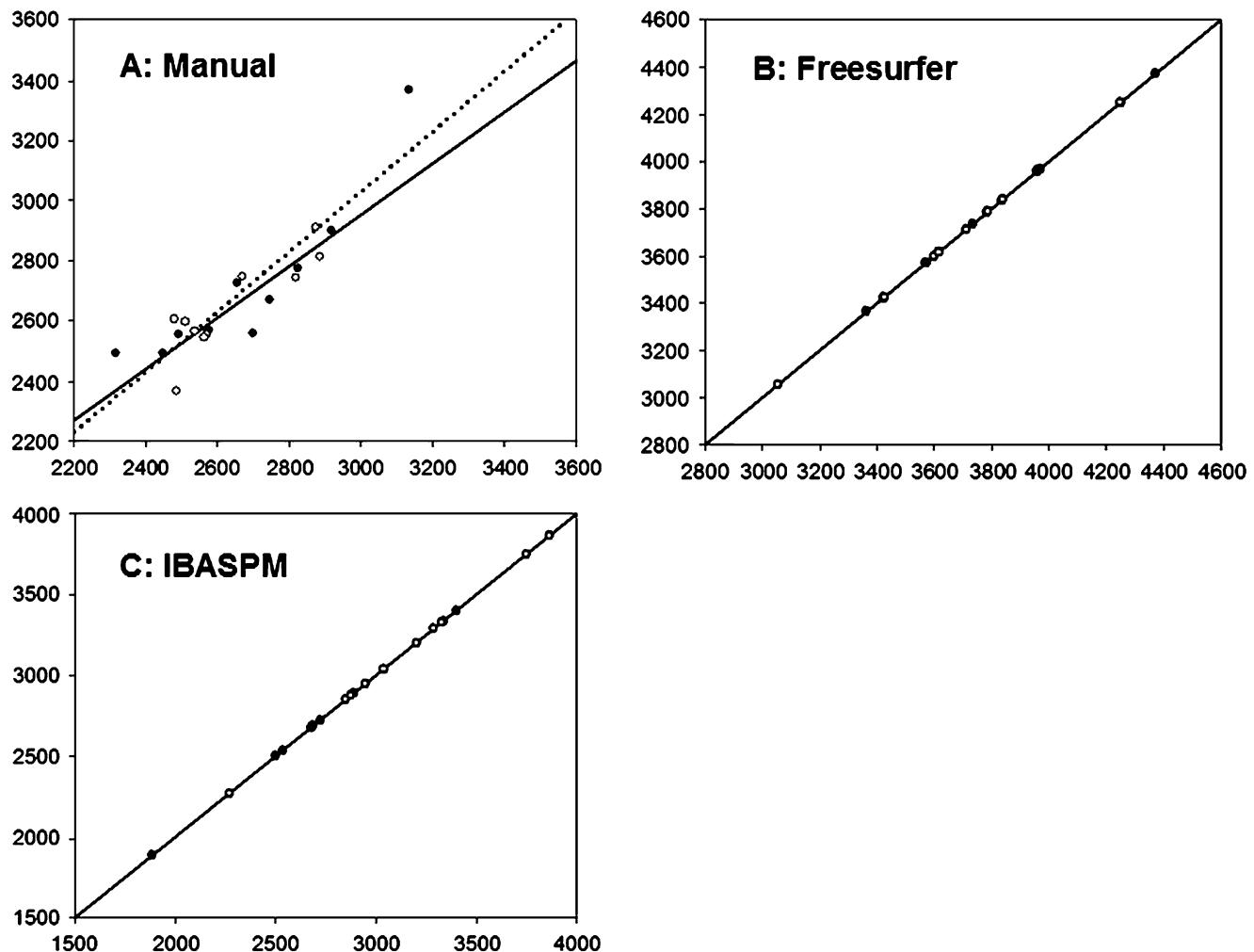


Fig. 3 Plots of repeated measurements from the three volumetric methods. The ICCs were calculated for the manual (a), FreeSurfer (b), and IBASPM (c) volumetric methods. The ICCs for manual volumetry were 0.947 (right) and 0.934 (left), for FreeSurfer were 1.000 (right) and 1.000 (left), and for IBASPM were 1.000 (right) and 1.000 (left).

The *closed symbols* and *solid regression lines* represent the right hippocampus measurements, and the *open symbols* and *dotted regression lines* represent the left hippocampus measurements. The volumes are in cubic millimeters

IBASPM, $P=0.0008$), while the variance of the FreeSurfer and IBASPM methods did not differ significantly ($P=0.54$).

Discussion

Reducing the laboriousness of the work is an important merit of automated volumetric methods, and perfect reproducibility is a major benefit because it guarantees consistent findings regardless of rater. However, accuracy is more important for volumetry studies.

In this study, we obtained excellent intrarater reliability with manual hippocampal volumetry, and the two automated methods showed perfect repeatability (Fig. 3). The hippocampal volumes measured using FreeSurfer showed good

agreement with the manual volumetric method, but both mean hippocampal volumes measured using FreeSurfer were significantly larger than the volumes determined using the manual volumetric method by about $1,000 \text{ mm}^3$ (Fig. 5). IBASPM showed questionable agreement with the manual volumetric method (Fig. 4).

Clinical considerations

The ICV in the normal controls was similar to that in the patients with MDD. As measured by manual hippocampal volumetry, the left hippocampal volume of the patients was significantly smaller than that of the controls, and, although the statistical significance was borderline, the right hippocampal volume of the patients was smaller than that of the

Fig. 4 Correlation plots for hippocampal volumes between manual and automated volumetric methods. The *closed symbols* and *solid regression lines* compare manual volumetry and the Free Surfer (*FS*) measurements, and the *open symbols* and *dotted regression lines* compare manual volumetry and the IBASPM measurements. The manual measurement volumes are shown on the horizontal axis, and the automated measurements are shown on the vertical axis. Volumes are in cubic millimeters

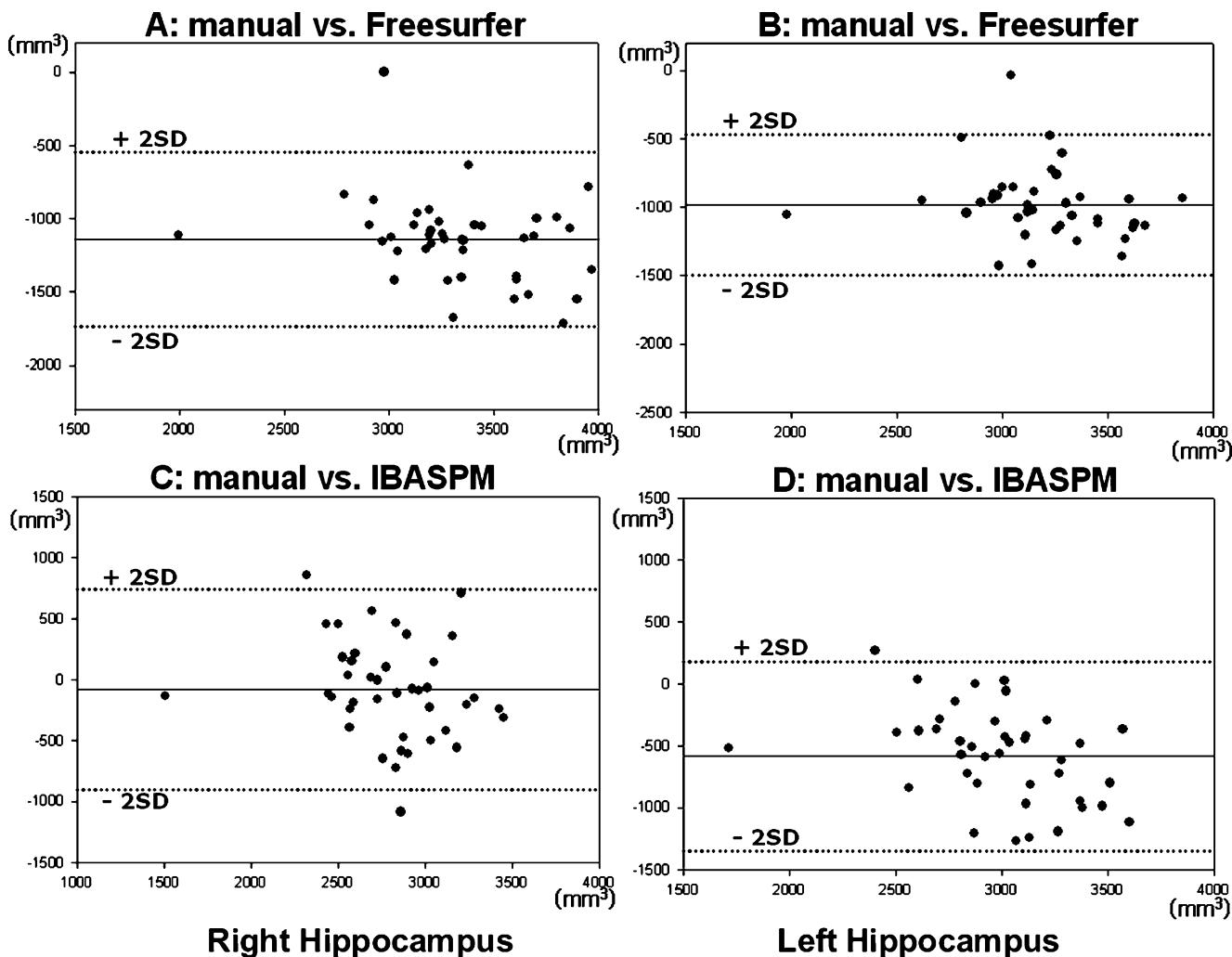
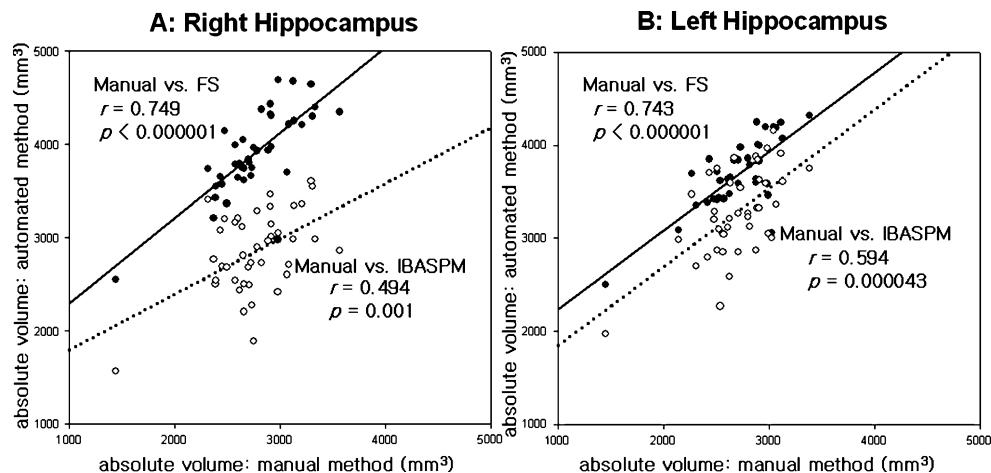


Fig. 5 Bland-Altman mean difference plots for hippocampal volumes between the manual and automated volumetric methods: **a, b** manual minus FreeSurfer measurements for the right (**a**) and left (**b**) hippocampus; **c, d** manual minus IBASPM measurements for the

right (**c**) and left (**d**) hippocampus. The average of the manual and automated volume measurements are shown on the horizontal axes the absolute volume differences between the two measurements are shown on the vertical axis. Volumes are in cubic millimeters

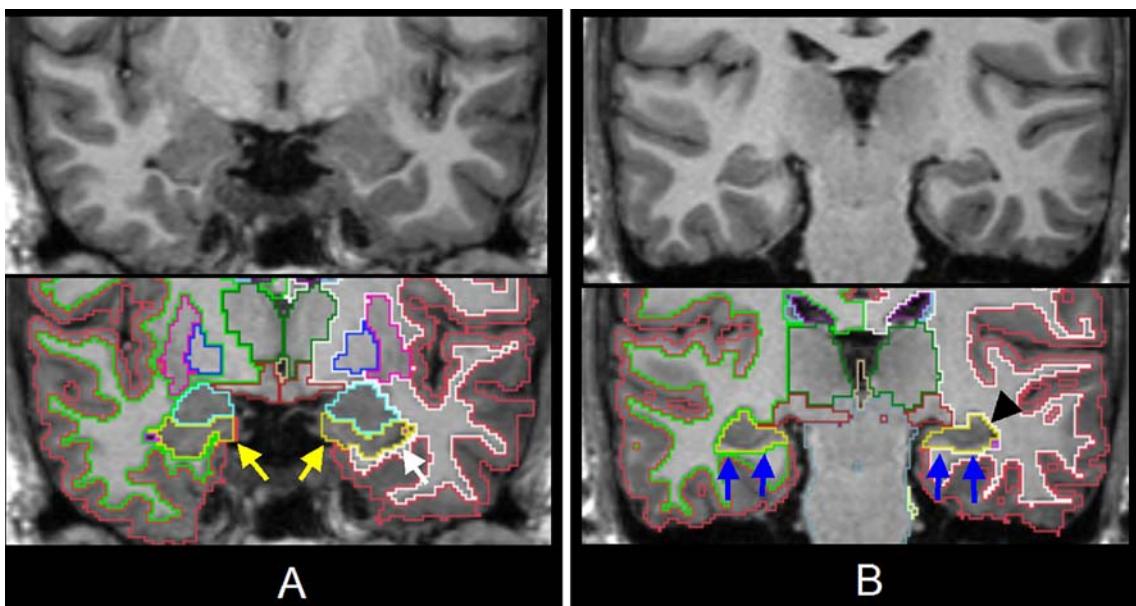


Fig. 6 Hippocampal parcellation with FreeSurfer. **a** The hippocampal tracing line (yellow line) included the anterior head of the hippocampus, entorhinal cortices (yellow arrows), and inferior cornu of the

lateral ventricle (white arrow). **b** The hippocampal ROI expanded to the adjacent white matter area of the inferior hippocampus (blue arrows)

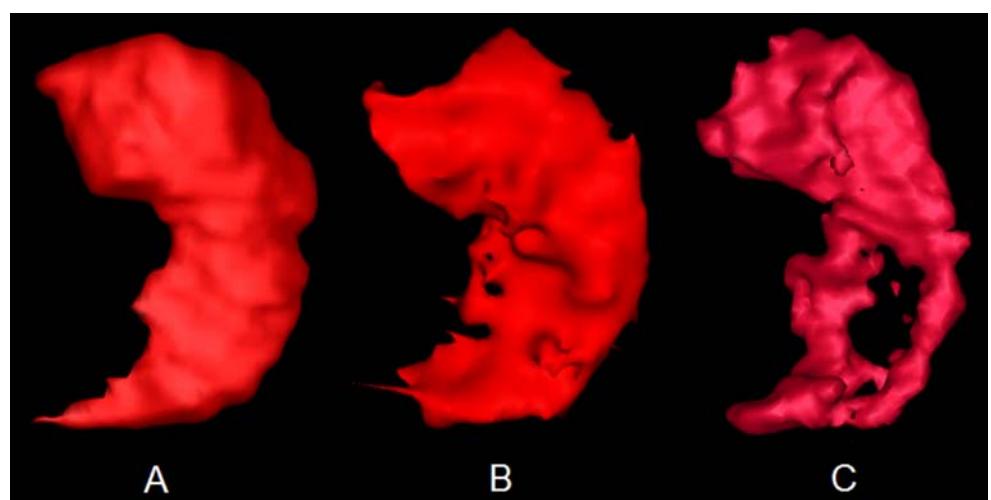
normal controls. After controlling for the ICV, both hippocampal volumes of the patients with MDD were significantly smaller than the hippocampal volumes of the normal controls using manual hippocampal volumetry (Table 1).

A reduction in hippocampal volume in depression has been reported repeatedly [40–44]. A reduction in the total and posterior hippocampal volumes without a reduction in the anterior hippocampal volume has been reported in unmedicated patients with MDD and those in remission [42]. The right hippocampus of older patients with depression showed a progressive volume reduction in a longitudinal study of hippocampal volume, and the hippo-

campal volume reduction was correlated with memory deficits at 6 months [43]. Another study showed that the lower hippocampal volumes of patients with MDD were correlated with poorer performance in the Wisconsin card sorting test [40]. A recent meta-analysis of MRI studies concluded that the hippocampal volume is reduced in patients with MDD, but not in patients with bipolar depressive disorder [44]. When we consider previous studies, our results regarding hippocampal volumes support their findings.

Right-left hippocampal asymmetry was first reported in 52 normal young Caucasian adults [16] and in 34 Asian adults [45]. In our study, the manual volumetric method did

Fig. 7 The hippocampus surface rendering of manual and automated segmentation. The 3-D surface of a normal right hippocampus was visualized using manual (**a**) and automated segmentation with FreeSurfer (**b**) and IBASPM (**c**). The 3-D surface of the manually segmented hippocampus shows a realistic hippocampal shape. The anterior head of the hippocampus segmented using FreeSurfer is not included (**b**). A large hole is seen, indicating the loss of hippocampal tissue in the posterior half of the hippocampus segmented using IBASPM (**c**)



not reveal significant hippocampal asymmetry ($P=0.109$), but this could have been attributable to the relatively small sample size of 20 subjects (Table 2); we therefore need to investigate a larger normal group. As measured using FreeSurfer the right hippocampal volume was significantly larger than the left, supporting the findings of previous studies. Conversely, as measured by the IBASPM method, the left hippocampal volume was larger than the right. This unacceptable finding might have arisen from methodological problems with IBASPM.

Manual hippocampal volumetry

The factors affecting the accuracy of hippocampal volumetry are slice thickness, MRI noise level, MRI orientation, tracing line inclusion, anatomic boundary criteria, separation between the amygdala and hippocampus, the software used for volume measurement, and the experience of the rater [1, 12–15, 17]. The manual hippocampal volumetric method used here was defined precisely and has been validated with high accuracy before [14]. The measured volume of a 2.70-cm³,

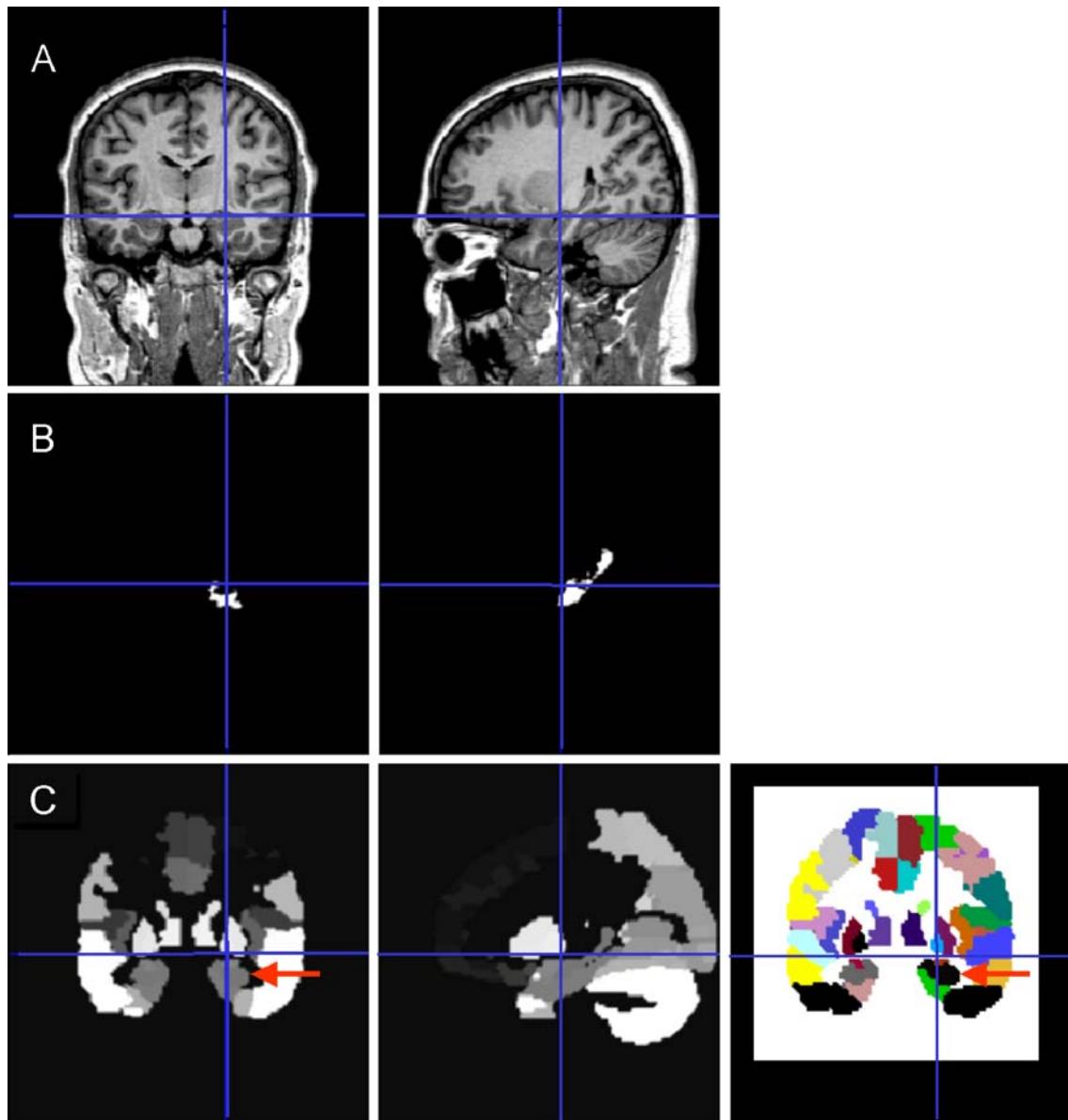


Fig. 8 Inaccurate registration between a normal MR image and the IBASPM atlas. The MR image of a normal subject (a), the left hippocampus segmented using IBASPM (b), and the IBASPM atlas (c) are shown. The blue crosshairs indicate the same location, and the red arrows indicate the left hippocampus in the IBASPM atlas. Since

inaccurate registration occurred between the subject's MR image and atlas (a and c), the upper and lateral parts of the hippocampal tissue are lost (b). The crosshairs were redrawn with a thick line for review purposes only

2.8-cm-long cylindrical phantom filled with CuSO₄ was 2.74 cm³ (difference 1.48%) using essentially the same method as our manual volumetric protocol (software, MRI acquisition, in-plane anatomic boundary criteria, ROI tracing method). For precise volumetry, this protocol originally scans or reformats a coronal T1-W MR image to oblique coronal to the perpendicular long axis of the hippocampus, and magnifies it four times to provide high anatomical resolution [12]. The tracing line was included in the volume because the tracing line contributes over 26% to the total hippocampal volume [15]. However, the MRI reformatting to oblique coronal by the software could produce a partial volume artifact, which could make it difficult to determine the hippocampus boundary. Therefore, MRI scanning perpendicular to the long axis of the hippocampus or the AC–PC line without software reconstruction is preferable. The delineation between the amygdala and the head of the hippocampus is very important [15], and inaccurate delineation of the hippocampal head from the amygdala seriously affects the volume of the entire hippocampus, which could give a misleading result [46]. The inclusion of the posterior tail of the hippocampus is important because tail-specific atrophy in patients with MDD has been reported [42, 47].

Automated hippocampal volumetry

Various automated methods for hippocampal volumetry have been developed and validated in the past 10 years [3, 4, 6–11]. However, only two packages (FreeSurfer and IBASPM) are publicly available to researchers, and a third is commercial [11].

The hippocampal ROI criteria of FreeSurfer were not clearly presented in the program's technical manual [8, 22]. The anterior-most head and posterior tail of the hippocampus are included in the in-plane boundaries, and sectors CA-1 through CA-4, the dentate gyrus, and subiculum appear to be included. Although the automated hippocampal volumetry using FreeSurfer has been validated [8], the hippocampal volumes measured using FreeSurfer were larger than the volumes measured with the manual method in this study (Table 1).

The major cause of the increased hippocampal volume with the FreeSurfer volumetric method arose from the expansion of the hippocampal ROI to the adjacent white matter area at the bottom of the hippocampus. Moreover, the inclusion of part of the entorhinal cortex and inferior cornu of the lateral ventricle also increased the hippocampal volume (Fig. 6). Although the developers of FreeSurfer have reported that this problem has been corrected [24], all of our subjects showed a hippocampal ROI that was expanded to include the adjacent white matter. In addition, the inaccurate differentiation of the anterior-most hippocampal

campus from the amygdala could affect the measured total hippocampal volume (Fig. 7b).

The hippocampal volumes measured using IBASPM showed questionable agreement with the manual volumetric method. The errors in hippocampal volume with the IBASPM volumetric method arose from inaccurate spatial normalization of the mesial temporal area (Fig. 8) and the use of the MNI single-subject MRI for the manually predefined ROI of the hippocampus [23]. We reviewed the accuracy of spatial normalization in the mesial temporal area, and found frequent inaccurate registrations in the IBASPM process. This problem resulted in inaccurate labeling of the hippocampus, and we confirmed the inclusion or exclusion of adjacent brain structures in the hippocampal labels defined by IBASPM visually. The inclusion of the voxel of the gray matter image in the volume calculation when the voxel values of the gray matter image are higher than the voxel values of the white matter and CSF images could have been a source of error.

Conclusions

We successfully showed hippocampal atrophy in patients with chronic MDD using a manual volumetric method. The accuracy of the automated methods, however, did not match that of the manual method. Compared to manual hippocampal volumetry, automated hippocampal volumetry using FreeSurfer showed good agreement, although the measured volume was 35% larger, while the IBASPM volumetric method showed questionable agreement. Although the automated methods were successful in detecting hippocampal atrophy in the patients with MDD, our results indicate that manual hippocampal volumetry is still the gold standard, while automated volumetric methods need to be improved.

Acknowledgement This study was supported by a research grant from Kangwon National University Hospital.

Conflict of interest statement We declare that we have no conflict of interest.

References

1. Geuze E, Vermetten E, Bremner JD (2005) MR-based in vivo hippocampal volumetrics: 2. Findings in neuropsychiatric disorders. *Mol Psychiatry* 10:160–184
2. Geuze E, Vermetten E, Bremner JD (2005) MR-based in vivo hippocampal volumetrics: 1. Review of methodologies currently employed. *Mol Psychiatry* 10:147–159
3. Alemán-Gómez Y, Melie-García L, Valdés-Hernandez P (2006) IBASPM: toolbox for automatic parcellation of brain structures. Presented at the 12th Annual Meeting of the Organization for Human Brain Mapping, 11–15 June 2006, Florence, Italy. Available on CD-Rom in NeuroImage, vol. 27, no. 1

4. 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
5. Chupin M, Mukuna-Bantumbakulu AR, Hasboun D, Bardinet E, Baillet S, Kinkingnehu 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
6. Csernansky JG, Joshi S, Wang L, Haller JW, Gado M, Miller JP, Grenander U, Miller MI (1998) Hippocampal morphometry in schizophrenia by high dimensional brain mapping. *Proc Natl Acad Sci U S A* 95:11406–11411
7. Duchesne S, Pruessner JC, Collins DL (2002) Appearance-based segmentation of medial temporal lobe structures. *Neuroimage* 17:515–531
8. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrave 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 structures in the human brain. *Neuron* 33:341–355
9. Ghanei A, Soltanian-Zadeh H, Windham JP (1998) Segmentation of the hippocampus from brain MRI using deformable contours. *Comput Med Imaging Graph* 22:203–216
10. Haller JW, Banerjee A, Christensen GE, Gado M, Joshi SC, Miller MI, Sheline YI, Vannier MW, Csernansky JG (1997) Three-dimensional hippocampal MR morphometry with high-dimensional transformation of a neuroanatomical atlas. *Radiology* 202:504–510
11. Hsu YY, Schuff N, Du AT, Mark K, Zhu X, Hardin D, Weiner MW (2002) Comparison of automated and manual MRI volumetry of hippocampus in normal aging and dementia. *J Magn Reson Imaging* 16:305–310
12. Jack CR Jr (1994) MRI-based hippocampal volume measurements in epilepsy. *Epilepsia* 35 [Suppl 6]:S21–S29
13. Jack CR Jr, Bentley MD, Twomey CK, Zinsmeister AR (1990) MR imaging-based volume measurements of the hippocampal formation and anterior temporal lobe: validation studies. *Radiology* 176:205–209
14. Jack CR Jr, Sharbrough FW, Twomey CK, Cascino GD, Hirschorn KA, Marsh WR, Zinsmeister AR, Scheithauer B (1990) Temporal lobe seizures: lateralization with MR volume measurements of the hippocampal formation. *Radiology* 175:423–429
15. 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
16. Jack CR Jr, Twomey CK, Zinsmeister AR, Sharbrough FW, Petersen RC, Cascino GD (1989) Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. *Radiology* 172:549–554
17. Luft AR, Skale JM, Welte D, Kolb R, Klose U (1996) Reliability and exactness of MRI-based volumetry: a phantom study. *J Magn Reson Imaging* 6:700–704
18. Devanand DP, Pradhaban G, Liu X, Khandji A, De Santi S, Segal S, Rusinek H, Pelton GH, Honig LS, Mayeux R, Stern Y, Tabert MH, de Leon MJ (2007) Hippocampal and entorhinal atrophy in mild cognitive impairment: prediction of Alzheimer disease. *Neurology* 68:828–836
19. Grundman M, Sencakova D, Jack CR Jr, Petersen RC, Kim HT, Schultz A, Weiner MF, DeCarli C, DeKosky ST, van Dyck C, Thomas RG, Thal LJ (2002) Alzheimer's Disease Cooperative Study: brain MRI hippocampal volume and prediction of clinical status in a mild cognitive impairment trial. *J Mol Neurosci* 19:23–27
20. Jack CR Jr, Petersen RC, Xu YC, O'Brien PC, Smith GE, Ivnik RJ, Boeve BF, Waring SC, Tangalos EG, Kokmen E (1999) Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology* 52:1397–1403
21. Pantel J, O'Leary DS, Cretsinger K, Bockholt HJ, Keefe H, Magnotta VA, Andreasen NC (2000) A new method for the in vivo volumetric measurement of the human hippocampus with high neuroanatomical accuracy. *Hippocampus* 10:752–758
22. Walhovd KB, Fjell AM, Reinvang I, Lundervold A, Dale AM, Eilertsen DE, Quinn BT, Salat D, Makris N, Fischl B (2005) Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiol Aging* 26:1261–1270
23. Tzourio-Mazoyer N, Landau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289
24. Han X, Fischl B (2007) Atlas renormalization for improved brain MR image segmentation across scanner platforms. *IEEE Trans Med Imaging* 26:479–486
25. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961) An inventory for measuring depression. *Arch Gen Psychiatry* 4:561–571
26. Hamilton M (1960) A rating scale for depression. *J Neurol Neurosurg Psychiatry* 23:56–62
27. Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113
28. Kruggel F (2006) MRI-based volumetry of head compartments: normative values of healthy adults. *Neuroimage* 30:1–11
29. Lemieux L, Hammers A, Mackinnon T, Liu RS (2003) Automatic segmentation of the brain and intracranial cerebrospinal fluid in T1-weighted volume MRI scans of the head, and its application to serial cerebral and intracranial volumetry. *Magn Reson Med* 49:872–884
30. MacLullich AM, Ferguson KJ, Deary IJ, Seckl JR, Starr JM, Wardlaw JM (2002) Intracranial capacity and brain volumes are associated with cognition in healthy elderly men. *Neurology* 59:169–174
31. Wolf H, Kruggel F, Hensel A, Wahlund LO, Arendt T, Gertz HJ (2003) The relationship between head size and intracranial volume in elderly subjects. *Brain Res* 973:74–80
32. Eritaia J, Wood SJ, Stuart GW, Bridle N, Dudgeon P, Maruff P, Velakoulis D, Pantelis C (2000) An optimized method for estimating intracranial volume from magnetic resonance images. *Magn Reson Med* 44:973–977
33. Chey J, Na DG, Tae WS, Ryoo JW, Hong SB (2006) Medial temporal lobe volume of nondemented elderly individuals with poor cognitive functions. *Neurobiol Aging* 27:1269–1279
34. Duvernoy HM (1991) The human brain: surface, three-dimensional sectional anatomy, and MRI. Springer-Verlag, New York, NY
35. Hasboun D, Chantome M, Zouaoui A, Sahel M, Deladoeuille M, Sourour N, Duyne M, Baulac M, Marsault C, Dormont D (1996) MR determination of hippocampal volume: comparison of three methods. *AJR Am J Neuroradiol* 17:1091–1098
36. Arndt S, Swayze V, Cizadlo T, O'Leary D, Cohen G, Yuh WT, Ehrhardt JC, Andreasen NC (1994) Evaluating and validating two methods for estimating brain structure volumes: tessellation and simple pixel counting. *Neuroimage* 1:191–198
37. Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* 16:297–334
38. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310
39. Pitman EJG (1939) A note on normal correlation. *Biometrika* 31:9–12
40. Frodl T, Schaub A, Banac S, Charypar M, Jager M, Kummler P, Bottlender R, Zetsche T, Born C, Leinsinger G, Reiser M, Moller

- HJ, Meisenzahl EM (2006) Reduced hippocampal volume correlates with executive dysfunctioning in major depression. *J Psychiatry Neurosci* 31:316–323
41. Hickie I, Naismith S, Ward PB, Turner K, Scott E, Mitchell P, Wilhelm K, Parker G (2005) Reduced hippocampal volumes and memory loss in patients with early- and late-onset depression. *Br J Psychiatry* 186:197–202
42. Neumeister A, Wood S, Bonne O, Nugent AC, Luckenbaugh DA, Young T, Bain EE, Charney DS, Drevets WC (2005) Reduced hippocampal volume in unmedicated, remitted patients with major depression versus control subjects. *Biol Psychiatry* 57:935–937
43. O'Brien JT, Lloyd A, McKeith I, Ghokar A, Ferrier N (2004) A longitudinal study of hippocampal volume, cortisol levels, and cognition in older depressed subjects. *Am J Psychiatry* 161:2081–2090
44. Videbech P, Ravnkilde B (2004) Hippocampal volume and depression: a meta-analysis of MRI studies. *Am J Psychiatry* 161:1957–1966
45. Li YJ, Ga SN, Huo Y, Li SY, Gao XG (2007) Characteristics of hippocampal volumes in healthy Chinese from MRI. *Neurol Res* 29:803–806
46. Tae WS, Hong SB (2001) Boundary of amygdala and hippocampus. *Am J Psychiatry* 158:820–821
47. Maller JJ, Daskalakis ZJ, Fitzgerald PB (2007) Hippocampal volumetrics in depression: the importance of the posterior tail. *Hippocampus* 17:1023–1027