Structural MRI volumetric analysis in patients with organic amnesia, I: methods and comparative findings across diagnostic groups

A Colchester, D Kingsley, D Lasserson, B Kendall, F Bello, C Rush, T G Stevens, G Goodman, G Heilpern, N Stanhope, M D Kopelman

Abstract

Background—If they are to be replicable, MRI volume measurements require explicit definitions of structures and of criteria for delineating these structures on MRI. Previously published volumes in healthy subjects show considerable differences in measurements across different studies, including a fourfold variation in estimates of hippocampal volume. Previous neuroimaging reports in patients with Korsakoff syndrome have generally found widespread or non-specific change, whereas in patients with herpes encephalitis the extent of pathological involvement reported beyond the temporal lobes has varied.

Method—In the present study, a clear set of anatomical criteria and detailed MRI segmentation procedures were applied to measure whole brain, frontal and temporal lobe, and anterolateral and medial temporal volumes, as well as thalamic areas in patients with organic amnesia (from Korsakoff’s syndrome, herpes encephalitis, and focal frontal lesions) as well as healthy controls.

Results—Patients with Korsakoff’s syndrome showed decreased thalamic measurements but no significant changes in the medial temporal lobes, whereas patients with herpes encephalitis showed severe medial temporal but not thalamic atrophy. In the patients with known frontal lobe lesions, quantitative analysis on MRI showed reduced frontal lobe volume but no significant temporal lobe or thalamic atrophy.

Conclusion—Quantified MRI can be a useful technique with which to examine brain-cognitive relations, provided that detailed techniques are explicitly described. In particular, specific patterns of volume change can be found in vivo in patients with Korsakoff’s syndrome and those with herpes encephalitis.

Keywords: magnetic resonance imaging; volumetrics; amnesia; memory

Quantified radiological techniques can now be employed to measure crucial brain structures or lesions in neurological patients. Most published reports have focused on only a few disorders—for example, Alzheimer’s dementia and epilepsy—and have measured a restricted range of anatomical structures. There have been relatively few reports in organic amnesia.1–4 Accurate quantification of areas or volumes of structures on MRI is difficult. There are no satisfactory and fully automated methods, and some degree of user interaction is required. Fully manual planimetric methods require the user to identify a structure and to trace its entire boundary but, generally, boundary placement in low resolution images has poor reproducibility. An alternative approach is stereology,5 in which quantification is reduced to counting points within a structure, but this is still dependent on the user’s perception of where boundaries lie. Atlas based methods have also been developed,6 but these have difficulties where brain structures are grossly atrophied or distorted.

In practice, most published studies have employed planimetry or a variant: this requires that anatomical boundaries are clearly defined, using criteria identifiable on MRI. Variation in the criteria employed complicates comparisons across studies and is noticeable, for example, in studies reporting medial temporal volumes.7–10 Bronen and Cheung11 described particular difficulties in measuring the hippocampi.

Reports have also varied widely in the explicitness of their description of what was actually done. Given the neuroanatomical variation seen in normal subjects,11 clear rules are important and an element of pragmatism (“rules of thumb”) is required in assessing boundaries when the normal anatomy is unclear or absent. Only a few reports7–9 give such details.

The accuracy of measuring a three dimensional neuroscientific structure by a series of cross sectional areas, as occurs in MRI volumetric analysis, is also limited by the slice thickness and the gap between sequential scan slices. The variation in slice thickness has been substantial in measurements of hippocampal volume, ranging between 1.5 mm and 8.5 mm.7 8 10 12 Large interslice distances reduce accuracy and provide a further obstacle to meaningful comparisons.

The repeatability or reliability of measurements varies within the same raters and between different raters, but intrarater reliability has often not been reported. The validation of a segmentation technique (for delineating structures) cannot rely entirely on within rater measurements, especially when only a few slices,13 only one slice,7 or few subjects14 are examined.
In this paper, we have carried out a volumetric analysis of MR scans from patients with organic amnesia. We employed an interactive computer analysis package, which required the user to indicate a few points inside and outside a structure to be measured. From this, the computer selected the most appropriate boundaries from a set of automatically precomputed candidate boundaries. The user then edited these boundaries at a level of detail which he or she controlled. We measured a range of neurological structures, applying an explicit set of anatomical criteria and segmentation (demarcation) rules: these have been documented in sufficient operational detail that our measurements could be repeated by others. We used all the available slices in our data acquisition to improve the accuracy of our volume measurements. Structural volume findings are analysed in terms of (1) measurement repeatability in healthy subjects; and (2) comparisons between findings in Korsakoff, herpes encephalitis, and focal frontal lesion groups. Correlations with neuropsychological impairments are reported in the companion paper.

**Methods**

**MR ACQUISITION PROTOCOL**

The first 17 subjects were scanned on a Siemens Magnetom SP 63 1.5 Tesla scanner and the next 19 subjects on a General Electric 1.5 Tesla Signa scanner (because of delays in installation of the second scanner). All subjects had a conventional proton density and T2 weighted sequence for basic visual analysis, using 5 mm thick slices with an interslice gap of 2.5 mm. The field of view (FOV) was 250 mm (Magnetom) and 240 mm (Signa). In addition, separate T1 and T2* weighted gradient echo three dimensional data sets were obtained in the sagittal plane using 64 partitions with the Magnetom (effective slice thickness 2.34 mm) and 128 partitions with the Signa scanner (effective slice thickness 1.20 mm). The matrix size was 256x256 on both scanners with FOVs as above. This gave a voxel volume of 2.23 mm³ on the Magnetom (0.98x0.98x2.34 mm) and 1.05 mm³ on the Signa scanner (0.94x0.94x1.20 mm). Patients from each diagnostic group were fairly evenly distributed across sequential MR slices, using locally developed software. An application specific graphical user interface (GUI) incorporated all the necessary visualisation, manipulation, and storage or retrieval functions. The segmentation tools included a multislice two dimensional hierarchical segmentation program, a standard two dimensional polylime tool for drawing a sequence of connected straight lines, and a three dimensional plane cutting tool.

The hierarchical segmentation tool was based on an agglomerative clustering algorithm which worked by grouping pixels for which the edge strength was low locally, according to local grey level differences and edge stability in scale space. Boundaries which corresponded to ridges of locally high edge strength were established between pixels that belonged to different clusters. The resulting regions were then subjected to the same clustering process to form larger regions and this process was repeated until the regions were grouped hierarchically into the whole image. Computation was fully automatic and was done at the beginning of a session. The results were not shown on the monitor at this stage. Once the hierarchy had been computed, the user marked points inside and outside the required anatomical structure by using the mouse buttons. The computer then searched the hierarchy from the top down to find the largest structure which contained all the inside points and none of the outside points. If there was no such structure, the hierarchy was amended to create a valid structure complying with the constraints specified by the user.

A polyline was defined by the user providing a set of control points that could be edited interactively. It was not necessary for a polyline to form a closed structure, but a facility was available to obtain closed structures from an open ended polyline definition. The plane cutting tool allowed the user to generate a two or three dimensional plane from a series of interactively specified points that could be edited simultaneously. Once a plane had been defined, it could be used to define part of the boundary of structures which were then processed further.

Integration of the hierarchical segmentation, polyline, and plane cutting tools was controlled interactively by the user so that the structures of interest could be segmented according to specific criteria, as described below.

**DEFINITION OF BOUNDARIES OF NEUROANATOMICAL STRUCTURES**

*Intracranial vault volume*

The intracranial space above the tentorium was segmented along the interface between CSF and dura using the T2* three dimensional data set. Laterally, the tentorium is easy to identify, separating the cerebellum from the cerebral hemisphere. Medially, the cerebral peduncles were cut off by drawing a line from the estimated position of the anterior free margin of the tentorium to the posterior clinoid processes. Where the clinoids were not visible, the posterior margin of the floor of the anterior cranial fossa (tuberculum sellae) was used as the inferomedial marker of the tentorial attachment to the skull base.

*Whole brain (supratentorial)*

T2 weighted scans in the sagittal plane were used to give good contrast between the cortical surface and CSF which was excluded from measurements of brain volume. As far as possible, all sulci were outlined, and vessels, the fa, and infratentorial structures were excluded from segmentation.
Frontal lobe

The superior, medial, and lateral surfaces are bounded by CSF, and were defined in the whole brain segmentation. Additional boundaries were defined on serial coronal sections starting at the front and working backwards. In the anterior sections, the inferior surface is bounded by subarachnoid space. In progressively more posterior sections, deep structures including the anterior basal ganglia start to appear as forward invaginations into the frontal lobe, and a little further back the temporal stem becomes visible. Before the temporal stem was reached, the deep structures above the optic chiasm were not excluded from the frontal lobe as the volume contributed by these structures is small, and it is difficult otherwise to define a plane clearly. Once the temporal stem was present (fig 1A), the segmentation involved drawing a polyline from the inferolateral edge of the frontal operculum (1) into the sylvian fissure to the insula, then upwards into the fundus of the circular sulcus (2). From there a straight line was drawn to the superolateral tip of the lateral ventricle (3). The roof of the lateral ventricle was followed to the midline. This procedure was continued to the posterior limits of the frontal lobe.

The posterior boundary was formed by a plane defined by three points:

Point 1: the superomedial limit of the central sulcus—The central sulcus was recognised on axial slices as the most prominent sulcus running obliquely backwards and reaching the interhemispheric fissure. The point landmark of its superomedial limit was defined by extrapolation from the gyral convexities on either side by viewing the MR slices axially to establish the anteroposterior coordinates of the point, and by viewing a coronal section to judge the superoinferior and the mediolateral position (x, y, z coordinates).

Point 2: the posterior commissure—The posterior commissure (PC) in the midline, lying just above the superior colliculus, was the second point landmark. If not clearly seen on T2*, its position could be located in relation to the surrounding structures.

Point 3: the inferolateral limit of the central sulcus—The central sulcus was followed downwards through its full extent. At its tip, a line was drawn perpendicular from its inferior limit to the sylvian fissure, and their junction was the third landmark.

Temporal lobe

The anterior, lateral, and inferior boundaries were defined in the whole brain segmentation. The medial boundary was followed up into the suprasellar cistern and into the choroidal fissure.

The superior boundary of the temporal lobe was segmented by viewing coronal slices sequentially from front to back. Anteriorly, the temporal pole was fully isolated by CSF. However, once the temporal stem was encountered, the superomedial margin was defined by a line from the inferior fundus of the circular sulcus of the insula (5) to the most lateral invagination of the choroidal fissure (6 in fig 1A). Posteriorly, the temporal stem disappears, although the sylvian fissure is still present. At this point, the superomedial boundary was drawn from the medial limit of the sylvian fissure (laterally) to the tip of the trigone of the lateral ventricle, and then to the choroidal fissure (medially).

The posterior boundary was formed by a plane defined by three points (the superomedial limit of the central sulcus and the posterior commissure in the midline) were as above. The third point for the temporal lobe was an approximation of the preoccipital notch. On the appropriate axial slice (fig 2) the line joining the AC-PC line was established by identifying the anterior commissure (AC) and the point on the axial slice closest to the position of the PC. The AC-PC line was extended anteriorly and posteriorly to the surface of the brain (a), and a point was identified three quarters of the way back along this line, from which a line was drawn perpendicular to the AC-PC line. Its intersection with the
was clear. Because only a limited number of
density (PD) images where the lateral boundary
measurements were made on axial proton den-
T2 weighted images. Consequently, thalamic
delineated from the internal capsule on T1 or
The lateral thalamic boundary was not clearly
Thalamus

sulcus until it met the inferolateral point of the
gyrus, and then deep into the collateral (rhinal)
the cortical surface of the parahippocampal

tions were taken from the subiculum around
margins but, in the coronal plane, segmenta-
tions. The “medial temporal” measurement

arachnoid surface of the brain defined the third
point landmark (b).

Medial temporal lobe structures
Separate segmentations were carried out on
coronal sections to measure hippocampal
volume and “medial temporal” (combined
hippocampal and parahippocampal) volume
(fig 1 B). Our boundary definitions for the hippocampus were closely similar to those described by Mori et al,12 except that we included the subiculum as part of the hippocampus. Anteriorly, the alveolar covering of the hippocampus provided a border with the amygdala. The posterior limit was the coronal slice in which the fornix clearly emerged from the fim-
bria of the hippocampus, just anterior to the
spleenium of the corpus callosum. These margins were checked in sagittal and axial sec-
tions. The “medial temporal” measurement
employed the same anterior and posterior
margins but, in the coronal plane, segmenta-
tions were taken from the subiculum around the
cortical surface of the parahippocampal
gyrus, and then deep into the collateral (rhinal)
sulcus until it met the inferolateral point of the
hippocampus.

Thalamus
The lateral thalamic boundary was not clearly
delineated from the internal capsule on T1 or
T2 weighted images. Consequently, thalamic
measurements were made on axial proton den-
sity (PD) images where the lateral boundary
was clear. Because only a limited number of

PD axial slices were available, we made a
thalamic area measurement at the slice giving the
maximum thalamic area. The volume of the
third ventricle was obtained by segmenting its
outline in the sagittal slices where it was visible.
Its posterior margin was drawn from the supe-
rior colliculus along the inferior margin of the internal cerebral veins. We calculated, thereby,
a “thalamic index” (the ratio of thalamic area: third ventricle volume) as an alternative
method of estimating thalamic atrophy.

Lesion volume
Where there was a discrete lesion (or lesions),
usually in the temporal or frontal lobes, this
was segmented in sagittal views, and its volume
was subtracted from that of the respective lobe.

Mammillary bodies
We considered that valid volumetric measure-
ments of these could not be determined,
because of the problem of partial volume arte-
fact across the few sections transecting this very
small structure. Consequently, ratings on MRI
films (checked against three dimensional im-
ages on the GUI viewing screen) were made by
a consultant neuroradiologist (DK), blind to
each subject’s clinical diagnosis. Ratings were
classified along a three point scale: normal, vis-
ible but small, and not visible.

Subjects
From 20 healthy subjects, who had taken part
in cognitive studies18 and who matched our
total group of patients in terms of mean age
(F(6, 57)=1.35 NS), 10 healthy volunteers
participated in the MR studies. These 10
volunteers were somewhat younger (mean
39.1(SD 13.5) years than the larger group of
controls (mean 45.9 (SD 17.3 years)), but oth-
erwise they were matched in terms of current
IQ, memory indexes, and scores on frontal/
executive tests.

Comparisons were made between findings of
different lesion groups. Patients with focal
frontal lesions were included to establish that
such pathology could indeed be quantified in
vivo with a view to investigating the relation
between frontal volume and cognitive function.
There were six patients with focal frontal
pathology, selected on CT evidence of lesions
which sappeared confined to the frontal lobes,
and which resulted from a haematoma, inf-
arcts, craniotomy, or (in one case) leucotomy
scars from a bilateral operation 18 years earlier.
Three of these patients had right sided lesions,
one a left sided lesion, and two had bilateral
lesions. Secondly, patients with herpes en-
cephalitis were included (1) to quantify the
temporal lobe changes, and (2) to examine the
extent of pathological change elsewhere, which
has varied in previous MR reports on such
patients.19 There were nine patients selected
on the basis of a history of viral encephalitis,
consistent with herpes simplex, serologically
proved in most cases, and with CT evidence of
severe temporal lobe pathology. Thirdly, a

group of patients with Korsakoff’s syndrome
was examined: on the basis of pathological
findings, diencephalic changes would be pre-
Structural MRI volumetric analysis in patients with organic amnesia: methods and findings

17

Table 1 Background cognitive scores: means (SD) (tests as described in Kopelman et al 

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>Healthy controls (n=10)</th>
<th>Korsakoff patients (n=11)</th>
<th>Herpes encephalitis patients (n=9)</th>
<th>Focal frontal lesion patients (n=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background intelligence tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NART-R IQ</td>
<td>106.1 (14.9)</td>
<td>108.9 (11.9)</td>
<td>103.3 (11.6)</td>
<td>107.5 (14.0)</td>
</tr>
<tr>
<td>WAB-R PS IQ</td>
<td>104.0 (18.3)</td>
<td>99.5 (14.8)</td>
<td>96.9 (12.9)</td>
<td>102.3 (21.7)</td>
</tr>
<tr>
<td>Memory tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WMS-R GMQ†</td>
<td>103.3 (16.1)</td>
<td>69.5 (20.4)</td>
<td>66.6 (14.7)</td>
<td>80.2 (19.6)</td>
</tr>
<tr>
<td>WMS-R DMQ†</td>
<td>103.3 (20.3)</td>
<td>61.0 (16.0)</td>
<td>65.6 (13.3)</td>
<td>76.5 (16.6)</td>
</tr>
<tr>
<td>Executive tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAS verbal fluency</td>
<td>44.2 (15.0)</td>
<td>31.0 (14.3)</td>
<td>36.3 (13.7)</td>
<td>29.7 (15.0)</td>
</tr>
<tr>
<td>Card sorting categories</td>
<td>5.1 (3.4)</td>
<td>3.8 (1.9)</td>
<td>4.7 (3.6)</td>
<td>4.8 (1.7)</td>
</tr>
<tr>
<td>Card sorting % perseverations</td>
<td>9.7 (9.7)</td>
<td>31.7 (27.6)</td>
<td>26.6 (29.2)</td>
<td>23.3 (14.3)</td>
</tr>
<tr>
<td>Cognitive estimates</td>
<td>4.1 (3.5)</td>
<td>6.5 (4.4)</td>
<td>5.4 (3.4)</td>
<td>9.2 (1.9)</td>
</tr>
</tbody>
</table>

*General memory index (or quotient) from Wechsler memory scale-revised.
†Delayed memory index (or quotient) from Wechsler memory scale-revised.


dicted but, to date, it has proved very difficult to demonstrate these in vivo with any precision or consistency on neuroimaging.13–27 There were 11 patients with a clinical diagnosis of Korsakoff’s syndrome, of whom 10 had a clearly documented history of Wernicke’s features. None of these patients showed focal cortical pathology on naked eye inspection of their scans.

Table 1 gives background cognitive data for these subject groups, demonstrating severe impairment in the patient groups on the general (F(1,34)=8.70, p<0.0002) and delayed (F(1,34)=12.97, p<0.0001) memory indexes, and a more moderate impairment on executive tests.19

ASSESSMENT OF MEASUREMENT REPEATABILITY

Comparisons were made between two assessors for control whole brain and hippocampal measurements in terms of interrater reliability coefficients (the intraclass correlation11 is a Pearson interrater correlation corrected for differences in the within rater (intraclass) variance)25 and mean error measurements. Intraclass reliability correlations were also examined for hippocampal values in a subgroup of 12 patients and in a combined group of 22 patients and controls.

HYPOTHESES

It was hypothesised that the patient groups would show regional volume loss consistent with their clinical diagnoses—medial temporal lobe volume in the patients with herpes encephalitis, thalamic volume in the patients with Korsakoff’s syndrome, and frontal volume in the frontal group. The extent of concomitant involvement of other structures—thalamic and frontal in the herpes group, frontal and medial temporal in the Korsakoff group, and thalamic/medial temporal in memory disordered frontal patients—remains controversial from previous investigations, and was also a matter of interest: the null hypothesis was that these structures would be unaffected.

Results

ASSESSMENT OF MEASUREMENT REPEATABILITY

Whole brain volume

Comparison was made between segmentations in the 10 healthy controls by an assessor at Stonybrook Hospital, NY (under the direction of Dr Lynn De Lisi) and by DL. The mean error between them was 2.4% (range 0.1 to 4.7%). The intraclass reliability correlation was 0.96 (p<0.0001).

Hippocampal volumes

Volumetric measures by two United Kingdom assessors were compared for the left and right hippocampus across 10 healthy control subjects. The mean difference between assessors was 8.0%, and their intraclass reliability correlation across the 20 measurements (10 left; 10 right) was 0.79 (p<0.001). In a subgroup of 12 patients of various aetiologies, the intraclass correlation was 0.81 (p<0.001), and in the combined subset of 22 controls and patients, interrater reliability was 0.85 (p<0.001).

Interscanner comparisons

Because the subjects had been scanned in two different machines, we carried out a two way analysis of variance (ANOVA), examining the effects of scanner and diagnostic group, across each of the outcome measures. None of the effects of scanner or scanner by diagnostic group was significant. In the total subject group, mean intracranial volume was 1.251 l (SD 0.014) on scanner 1 versus 1.257 l (SD 0.022) on scanner 2, and among the patients mean brain volume was 0.941 l (SD 0.111) versus 0.890 l (SD 0.104). Findings within the individual patient groups were consistent with this—for example, mean brain volume was 0.874 l (SD 0.106) in five patients with herpes encephalitis on scanner 1 versus 0.872 l (SD 0.112) in four patients with herpes encephalitis on scanner 2.

STRUCTURAL VOLUME CHANGES AND DIAGNOSTIC CATEGORY

Brain volumes were evaluated for both actual observed volumes and volumes normalised for intracranial volume using the formula of Jack et al.24 This formula corrects an observed regional brain volume according to the difference between that subject’s intracranial volume and the control mean, and the regression of regional brain volumes against intracranial volumes in controls. However, we found that it was much more legitimate in correcting the volumes of large structures than of small ones, in terms of the percentage of variance accounted for by the regression equation and its statistical significance. For example, regressing left frontal volume against intracranial volume...
accounted for 63.7% of left frontal variance in controls (p<0.01), but the equivalent regression accounted for only 0.01% of left medial temporal variance (non-significant). We report the findings for actual volumes, but statistical analyses were performed on both actual and corrected volumes in all instances, and only minor differences were obtained.

Frontal lobe volumes in vivo across the four subject groups showed significant differences (left: F(3, 30)=11.02, p<0.0001; right: F=11.03, p<0.0001). Student-Newman-Keuls post hoc tests showed significant differences between each of the patient groups and healthy controls. Figure 3 A shows the left and right mean frontal volumes in healthy controls and the focal frontal group.

Temporal lobe volume also differed significantly across the diagnostic groups (F(3,30)=12.58, p<0.0001). On the post hoc tests, the herpes group differed significantly from the healthy controls, Korsakoff group, and focal frontal group (left: p<0.01; right: p<0.05). There were no other statistically significant differences between the groups. Figure 3 C shows a comparison of mean left and right medial temporal volumes in the herpes and control groups.

Parahippocampal volume was determined by subtracting hippocampal from medial temporal volume. Figure 4 A shows the mean total (left and right) parahippocampal volume across the four groups, which differed significantly (F(3, 16)=6.19, p<0.0005), and on the post hoc tests, the herpes group differed significantly (p<0.05) from the other three groups. For left parahippocampal volume, the groups also differed significantly (F(3, 16)=3.58, p<0.05) with the herpes group showing a significantly (p<0.05) smaller volume than the controls; but for right parahippocampal volume, the groups did not differ.

![Figure 3](image-url)
Table 2 Mammillary bodies

<table>
<thead>
<tr>
<th>Group</th>
<th>Normal</th>
<th>Small</th>
<th>Not visible</th>
<th>% Abnormal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Focal frontal</td>
<td>4</td>
<td>1 (bilateral vascular)</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Herpes encephalitis</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>Korsakoff</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>64</td>
</tr>
</tbody>
</table>

Atrophy e no atrophy $\chi^2=10.12; p<0.02$.

Figure 4 Comparison of mean (SEM) medial temporal and thalamic measurements across all four groups (controls, herpes, Korsakoff and focal frontal): (A) total (left and right) parahippocampal volume; (B) total (left + right) hippocampal volume; (C) total (left + right) thalamic area. The figure shows significant t test findings compared with healthy controls. For ANOVAs and post hoc tests see text.

Discussion

Sources of error and measurement repeatability

Table 3 shows that published volumes of brain structures in healthy subjects vary widely. Data are presented as actual volumes, normalised, or as a percentage of intracranial volumes, as in the original publications, except for our own (KCL/UKC) results, which are presented in all three ways. Hippocampal volumes vary particularly widely: total (left plus right) hippocampal volume ranged from 2519 to 2614 mm$^3$ to 10 166 and 10 360 mm$^3$. This fourfold variation is not accounted for simply by age or sex differences. Given the lack of published details about definitions of boundaries and their application during segmentation in many studies, this wide variability is likely to reflect differences in measurement technique. It also demonstrates the limitations of simply publishing reliability coefficients in the absence of these explicit details.

For whole brain volume, our interrater error (mean 2.4%) and reliability coefficient (intraclass correlation=0.96) were very satisfactory. Interrater reliability was less accurate for the control, Korsakoff, and focal frontal groups. The groups also differed significantly for left (F=3.16; p<0.05) and right (F=5.67; p<0.01) hippocampal volume. On post hoc tests, the herpes group differed significantly from the other three groups for right (p<0.05) but not left sided volume. However, left hippocampal volume was significantly smaller in the herpes group than controls on a $t$ test ($t=2.99, p<0.02$).

Figure 4 C compares total thalamic area across the four groups, and figure 3 D compares patients with Korsakoff's syndrome and controls for the left and right thalami. The groups differed significantly for total area ($F(3, 16)=5.10, p<0.005$). The Korsakoff group differed significantly from the focal frontal group (p<0.05), and from healthy controls on $t$ tests (total area: $t=2.61$; left: $t=2.08$; right: $t=2.56$; p<0.05 in each case) and from the herpes group for total and right thalamic area ($t=2.40$ and 2.44, p<0.05) and at near significance for the left thalamus ($t=2.01, p<0.06$). Table 2 shows the findings for ratings of mammillary body atrophy. Sixty four per cent of the patients with Korsakoff's syndrome showed atrophy versus 33% or less in the other groups ($\chi^2=10.12, p<0.02$).

In summary, the patients with herpes encephalitis showed significant atrophy of the medial temporal lobe structures, in the presence of normal thalamic areas and a much lower incidence of mammillary body atrophy than the Korsakoff group. By contrast, the patients with Korsakoff's syndrome showed significant atrophy of the thalami and mammillary bodies, but not of medial temporal structures (hippocampi, parahippocampal gyri, or the two measures combined). All three patient groups showed evidence of frontal lobe loss of volume, but the focal frontal lesion group did not show any atrophy in the temporal lobes or thalami.

www.jmp.com
Table 3  MR measurements of structural volume in healthy subjects: present study (KCL/UKC) and previous investigations

<table>
<thead>
<tr>
<th>Actual:</th>
<th>Whole brain (litres)</th>
<th>LF (cm³)</th>
<th>RF (cm³)</th>
<th>LF (cm³)</th>
<th>RF (cm³)</th>
<th>Left hipp (mm³)</th>
<th>Right hipp (mm³)</th>
<th>Total hipp (mm³)</th>
<th>LP/hipp (mm³)</th>
<th>RP/hipp (mm³)</th>
<th>Total P/hipp (mm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KCL/UKC</td>
<td>1.14 (0.13)</td>
<td>250.2 (39.1)</td>
<td>249.0 (36.8)</td>
<td>83.9 (11.0)</td>
<td>91.2 (12.6)</td>
<td>3642 (172)</td>
<td>4130 (680)</td>
<td>7772 (1008)</td>
<td>3402 (548)</td>
<td>3150 (497)</td>
<td>6553 (960)</td>
</tr>
<tr>
<td>Cook et al (^5)</td>
<td>3292 (397)</td>
<td>3185 (393)</td>
<td>6414 (786)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watson et al (^5)</td>
<td>4903 (684)</td>
<td>5263 (652)</td>
<td>10166 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breier et al (^5)</td>
<td>3800 (600)</td>
<td>5700 (400)</td>
<td>7500 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.19 (0.16)</td>
<td>64.4 (7.8)</td>
<td>68.6 (8.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.07 (0.13)</td>
<td>79.9 (7.9)</td>
<td>82.3 (8.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatia et al (^5)</td>
<td>3.32 (0.18)</td>
<td>3800 (600)</td>
<td>3700 (400)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aylward et al (^5)</td>
<td>35</td>
<td>2229 (397)</td>
<td>2185 (393)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watson et al (^5)</td>
<td>4903 (684)</td>
<td>5263 (652)</td>
<td>10166 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breier et al (^5)</td>
<td>3800 (600)</td>
<td>5700 (400)</td>
<td>7500 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.19 (0.16)</td>
<td>64.4 (7.8)</td>
<td>68.6 (8.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.07 (0.13)</td>
<td>79.9 (7.9)</td>
<td>82.3 (8.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatia et al (^5)</td>
<td>3.32 (0.18)</td>
<td>3800 (600)</td>
<td>3700 (400)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aylward et al (^5)</td>
<td>35</td>
<td>2229 (397)</td>
<td>2185 (393)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watson et al (^5)</td>
<td>4903 (684)</td>
<td>5263 (652)</td>
<td>10166 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breier et al (^5)</td>
<td>3800 (600)</td>
<td>5700 (400)</td>
<td>7500 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.19 (0.16)</td>
<td>64.4 (7.8)</td>
<td>68.6 (8.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.07 (0.13)</td>
<td>79.9 (7.9)</td>
<td>82.3 (8.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatia et al (^5)</td>
<td>3.32 (0.18)</td>
<td>3800 (600)</td>
<td>3700 (400)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aylward et al (^5)</td>
<td>35</td>
<td>2229 (397)</td>
<td>2185 (393)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watson et al (^5)</td>
<td>4903 (684)</td>
<td>5263 (652)</td>
<td>10166 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breier et al (^5)</td>
<td>3800 (600)</td>
<td>5700 (400)</td>
<td>7500 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.19 (0.16)</td>
<td>64.4 (7.8)</td>
<td>68.6 (8.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.07 (0.13)</td>
<td>79.9 (7.9)</td>
<td>82.3 (8.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatia et al (^5)</td>
<td>3.32 (0.18)</td>
<td>3800 (600)</td>
<td>3700 (400)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aylward et al (^5)</td>
<td>35</td>
<td>2229 (397)</td>
<td>2185 (393)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watson et al (^5)</td>
<td>4903 (684)</td>
<td>5263 (652)</td>
<td>10166 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breier et al (^5)</td>
<td>3800 (600)</td>
<td>5700 (400)</td>
<td>7500 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.19 (0.16)</td>
<td>64.4 (7.8)</td>
<td>68.6 (8.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.07 (0.13)</td>
<td>79.9 (7.9)</td>
<td>82.3 (8.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatia et al (^5)</td>
<td>3.32 (0.18)</td>
<td>3800 (600)</td>
<td>3700 (400)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aylward et al (^5)</td>
<td>35</td>
<td>2229 (397)</td>
<td>2185 (393)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watson et al (^5)</td>
<td>4903 (684)</td>
<td>5263 (652)</td>
<td>10166 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breier et al (^5)</td>
<td>3800 (600)</td>
<td>5700 (400)</td>
<td>7500 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.19 (0.16)</td>
<td>64.4 (7.8)</td>
<td>68.6 (8.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.07 (0.13)</td>
<td>79.9 (7.9)</td>
<td>82.3 (8.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatia et al (^5)</td>
<td>3.32 (0.18)</td>
<td>3800 (600)</td>
<td>3700 (400)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aylward et al (^5)</td>
<td>35</td>
<td>2229 (397)</td>
<td>2185 (393)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watson et al (^5)</td>
<td>4903 (684)</td>
<td>5263 (652)</td>
<td>10166 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breier et al (^5)</td>
<td>3800 (600)</td>
<td>5700 (400)</td>
<td>7500 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.19 (0.16)</td>
<td>64.4 (7.8)</td>
<td>68.6 (8.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.07 (0.13)</td>
<td>79.9 (7.9)</td>
<td>82.3 (8.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatia et al (^5)</td>
<td>3.32 (0.18)</td>
<td>3800 (600)</td>
<td>3700 (400)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aylward et al (^5)</td>
<td>35</td>
<td>2229 (397)</td>
<td>2185 (393)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watson et al (^5)</td>
<td>4903 (684)</td>
<td>5263 (652)</td>
<td>10166 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breier et al (^5)</td>
<td>3800 (600)</td>
<td>5700 (400)</td>
<td>7500 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.19 (0.16)</td>
<td>64.4 (7.8)</td>
<td>68.6 (8.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al (^5)</td>
<td>1.07 (0.13)</td>
<td>79.9 (7.9)</td>
<td>82.3 (8.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatia et al (^5)</td>
<td>3.32 (0.18)</td>
<td>3800 (600)</td>
<td>3700 (400)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aylward et al (^5)</td>
<td>35</td>
<td>2229 (397)</td>
<td>2185 (393)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watson et al (^5)</td>
<td>4903 (684)</td>
<td>5263 (652)</td>
<td>10166 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breier et al (^5)</td>
<td>3800 (600)</td>
<td>5700 (400)</td>
<td>7500 (NA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data presented as in the original published studies—actual or normalised volumes or % intracranial volume. Normalised volumes by formula of Jack et al \(^{24}\) except * where own formula used. NA=not available.
hippocampus. Our mean hippocampal error in controls (8.0%) compares with 7%27 5.1%,28 and 4.1%,29 and an intrarater error of 6.7%.30 However, our hippocampal intraclass correlation (r=0.79) is in line with other studies—for example, 0.7813 and 0.82.23 Interrater reliability in patients has seldom been reported before: in the present study, the intraclass correlation for hippocampi was 0.81, and in combined patients and controls it was 0.85.

STRUCTURAL VOLUME CHANGES AND DIAGNOSTIC CATEGORY

The focal frontal group showed statistically significant losses of volume across left, right, and total frontal lobe measurements, but did not show any evidence of loss of volume in the thalami or on our measures of medial temporal, anterolateral temporal, and total temporal lobe volume. In previous neuropsychological studies,31 similar patients have been reported but without any MR quantification: such measurements permit in vivo correlation of loss of frontal tissue with cognitive performance.18 The patients with Korsakoff’s syndrome showed significant atrophy of the left and right thalami, mammillary bodies, and frontal lobes. All the patients with Korsakoff’s syndrome showed evidence of thalamic atrophy on at least one of our measures. Sixty four per cent of the Korsakoff group in the study also showed mammillary body atrophy. By contrast, this group did not differ significantly from healthy controls for anterolateral, medial temporal, parahippocampal, or hippocampal volume. Although this in vivo finding of thalamic and mammillary body atrophy is consistent with results obtained from neuropathological studies, conflict remains about which thalamic sites and pathways are critical for memory,32 33 and previous neuroimaging studies have generally reported either widespread or non-specific changes only.20-22 Jernigan et al found significant atrophy in eight patients with Korsakoff’s syndrome within a block of tissue labelled “diencephalic grey matter”, which encompassed the thalami, mammillary bodies, other hypothalamic grey matter, and the septal nuclei. The present findings indicate that focal changes can be found in the thalami (and the mammillary bodies) taken in isolation.

Our patients with herpes encephalitis showed a highly significant anterolateral and medial temporal loss of volume. Consistent with this, they also showed significantly reduced parahippocampal and hippocampal volumes. By contrast, their thalamic readings did not differ significantly from (and tended to be slightly larger than) those of the control group. Unlike the Korsakoff group, the mammillary bodies were visible in all our patients with herpes encephalitis, although three out of nine (33%) showed a moderate degree of atrophy. The herpes group also showed evidence of frontal lobe atrophy. To our knowledge, the only comparable study of patients with herpes encephalitis is that by Yonez et al who examined five patients and reported significant atrophy in the hippocampal formation and parahippocampal gyri. There was a lesser degree of atrophy in the amygdaloid body and anterior temporal neocortex, but these findings were not statistically significant in this small group of patients. These authors did not make measurements elsewhere. In a non-quantitative study, Kapur et al reported damage in the anterior and inferior temporal lobe gyri, the fornix, the mammillary bodies, the thalamic nuclei (in 50% of cases), the frontal lobes, and the striatum, as well as “limited involvement” of the cingulate gyrus, the parietal and occipital lobes. This study was non-quantitative, and consequently comparisons are difficult. However, neither visual inspection nor quantification suggested any damage to the thalamic nuclei in our patients, and the mammillary bodies were affected in only a minority of cases, and then to only a minor degree.

Conclusions

We have emphasised the importance of clear definitions of brain structures and of operational procedures in carrying out MR volumetric measurements, arguing that reliability and validity measurements are only useful in this context and pointing to the wide variation of findings in healthy subjects in previous studies. We obtained mean error and interrater reliability coefficients for each structure which were very satisfactory, and hippocampal values which were comparable with other studies. We found that patients with Korsakoff’s syndrome showed thalamic but not medial temporal atrophy on these in vivo measurements, whereas patients with herpes encephalitis manifested medial temporal but not thalamic atrophy (a double dissociation). Patients with known frontal lobe pathology showed in vivo frontal volume reduction on quantification, but did not show atrophy elsewhere (a single dissociation). Our criteria and techniques can be developed further to examine other, specific brain regions.

This research was funded by a Wellcome Trust project grant to MDR and BK and by European Union grant 950845 (Biomorph) grant to AC and colleagues.