ABSTRACT: Studies of memory-impaired patients will be most useful when quantitative neuroanatomical information is available about the patients being studied. Toward that end, in the case of medial temporal lobe amnesia, protocols have been developed from histological material that identify the boundaries of relevant structures on magnetic resonance images. Because the size of these structures varies considerably in the normal population, some correction for overall brain size is usually employed when calculating volume measurements. Although different correction procedures have been used to normalize for brain size, there has been little study of how well different methods reduce variability and which methods might be most useful. We measured the volume of the hippocampal region (hippocampus proper, dentate gyrus, and subiculum complex) and the volumes of the temporopolar, entorhinal, perirhinal, and parahippocampal cortices in five memory-impaired patients and 30 controls. We then compared three different methods for normalizing the volume measurements: normalization by intracranial volume, normalization by aligning the brain to a standard atlas, and normalization by brain area at the level of the anterior commissure. Normalization by intracranial volume reduced variability in the volume measurements of nearly all brain regions to a greater extent than did normalization by other methods. When normalized by intracranial volume, the patients exhibited a mean reduction in hippocampal volume of about 40% and negligible reductions in the volumes of other medial temporal lobe structures. On the basis of earlier histological analysis of two other patients (L.M. and W.H.), who also had reductions in hippocampal size of about 40%, we suggest that a volume reduction in this range likely indicates a nearly complete loss of hippocampal neurons. Published 2004 Wiley-Liss, Inc.*

KEY WORDS: hippocampus; MRI; amnesia; parahippocampal gyrus

INTRODUCTION

Beginning with the earliest case descriptions (Winslow, 1861; Ribot, 1881), the study of memory impairment has provided useful information about the structure and organization of human memory (Scoville and Milner, 1957; Talland, 1965; Baddeley, 1982; Gabrieli, 1998; Squire et al., 2004). In contrast, neuropathological information has only occasionally been available about the patients who have been studied. Yet neuroanatomical information is critical in order to classify patients and to address questions about how specific brain structures might contribute differently to memory functions (e.g., hippocampus and adjacent medial temporal cortex).

Beginning in the late 1980s, with the development of improved neuroimaging methods, it became possible to relate memory impairment to specific neuropathological change in living patients (Press et al., 1989; Squire et al., 1990; Corkin et al., 1997; Cipolotti et al., 2001; Kopelman et al., 2003; Levy et al., 2003; Vargha-Khadem et al., 2003). These techniques have been especially useful in the case of medial temporal lobe pathology. In most applications, magnetic resonance images (MRIs) are acquired for each patient, anatomic landmarks are identified, and the volume of each region of interest is measured (for another method based on local gray matter density, see Ashburner and Friston, 2000). The hippocampus itself is straightforward to identify and measure (Squire et al., 1990), but the adjacent cortical areas do not have readily identifiable borders. However, it has proved possible to establish anatomical landmarks that are visible in MRI, based on histological analysis of healthy brains, and to develop protocols for identifying the temporopolar, entorhinal, perirhinal, and parahippocampal cortices that lie adjacent to the hippocampus (Insausti et al., 1998a,b, 2003).

A further difficulty is that measurements of the volume of medial temporal lobe structures can vary substantially among individuals. For example, in one group of 20 healthy controls, the volume of the left temporopolar cortex ranged from 1,793 mm³ to 5,016 mm³ (Insausti et al., 1998a). Such variability makes it difficult to detect small amounts of volume loss in patients.

Following the intuition that variation in the volume of a particular brain structure may be related to variation in brain volume, a common approach to the problem of variability has been to employ some correction for overall brain size. Although a number of different normalization procedures have been employed, there has been little study of how well different methods reduce variability and which methods might be most useful. One study of patients with temporal lobe epilepsy (Free et al., 1995) considered six kinds of corrections and identified three that reduced variability in estimates of hippocampal volume (normalization by cranial area, cranial volume, and intracranial volume). Cranial area refers to the area of the cranial cavity as measured on a single midsagittal slice. Cranial volume refers to the volume of the cranial cavity.
plus the temporal bones and the convexity of the skull. Intracranial volume refers to intradural volume.

Similar comparisons have not been carried out in memory-impaired patients, and no studies have been done at all to compare methods for normalizing estimates of volumes of other medial temporal lobe structures. Drawing on MRI data from five memory-impaired patients and 30 controls, we here evaluate three different methods for normalizing volume measurements of medial temporal lobe structures.

### Materials and Methods

#### Participants

MP-RAGE MRIs were collected for 5 memory-impaired patients (4 male and 1 female; Table 1) and 30 matched controls (19 male and 11 female). Three patient scans were done on a 1.5-tesla (T) Siemens magnet at Thornton Hospital, UCSD, and 2 were done on a 1.5-T GE magnet at LDS Hospital in Salt Lake City. Nine control scans (7 male, 2 female) were done on the UCSD scanner, and the remaining scans (12 male, 9 female) were performed on a 1.5-T GE magnet at the San Diego VA hospital.

Patient J.R.W. became amnesic after an ischemic episode associated with cardiac arrest. G.W. and R.S. became amnesic after respiratory failure associated with drug overdoses. J.S. became amnesic during a 6-month period with known precipitating event. The MRI scans for patients R.S., J.R.W., and J.S. (Fig. 1) were then reevaluated in coronal view with attention paid to the separation between the hippocampus and the posterior aspect of the pulvinar, the separation between the subicular complex and entorhinal cortex, and white matter/gray matter segmentation.

Segmentation of the parahippocampal gyrus (here including temporopolar, perirhinal, entorhinal, and parahippocampal cortices) proceeded according to the guidelines defined histologically by Insausti et al. (1998a), whereby the authors related histological boundaries to landmarks readily visible on MRI. The temporopolar cortex included the entire temporal pole rostrally and, at the appearance of the inferior temporal sulcus, was limited to the medial portion of the temporal lobe between the medial bank of the inferior temporal sulcus and the fundus of the most lateral temporopolar sulcus. The perirhinal cortex extended from the appearance of the collateral sulcus, at its rostral border, to its caudal border at the coronal slice 4 mm posterior to the disappearance of the gyrus intralimbicus of the hippocampus. Through its rostrocaudal extent, perirhinal cortex included the cortex between the lateral bank of the collateral sulcus and the midpoint of the medial bank of the collateral sulcus (with a few qualifications as noted by Insausti et al., 1998a). The entorhinal cortex extended from the sulcus semiannularis to the perirhinal cortex, when the sulcus semiannularis could be visualized; otherwise, the medial border of the entorhinal cortex was the medial border of the subiculum. The parahippocampal cortex was defined rostrally by the coronal section 4 mm posterior to the disappearance of the gyrus intralimbicus (caudal to Insausti slice 24 in Fig. 5 of Insausti et al., 1998a) and caudally by the splenium of the corpus callosum (Insausti et al., 1998b). The lateral border of the parahippocampal cortex was the lateral bank of the collateral sulcus, and the medial border was the medial border of the subiculum. Specific guidelines are presented by Insausti et al. (1998a) for altering these boundaries to account for normal variation in the length, depth, and number of branches.

### Regions of Interest

Regions of interest (ROIs) for the left and right hippocampal regions (hippocampus proper, dentate gyrus, and subiculum complex) were drawn in the sagittal view, beginning laterally at the appearance of hippocampal tissue within the lateral ventricle. The drawing continued medially, observing the separation between the hippocampal region and the amygdala. The ROIs were limited to the medial portion of the temporal lobe between the medial bank of the inferior temporal sulcus and the fundus of the most lateral temporopolar sulcus. The perirhinal cortex extended from the appearance of the collateral sulcus, at its rostral border, to its caudal border at the coronal slice 4 mm posterior to the disappearance of the gyrus intralimbicus of the hippocampus. Through its rostrocaudal extent, perirhinal cortex included the cortex between the lateral bank of the collateral sulcus and the midpoint of the medial bank of the collateral sulcus (with a few qualifications as noted by Insausti et al., 1998a). The entorhinal cortex extended from the sulcus semiannularis to the perirhinal cortex, when the sulcus semiannularis could be visualized; otherwise, the medial border of the entorhinal cortex was the medial border of the subiculum. The parahippocampal cortex was defined rostrally by the coronal section 4 mm posterior to the disappearance of the gyrus intralimbicus (caudal to Insausti slice 24 in Fig. 5 of Insausti et al., 1998a) and caudally by the splenium of the corpus callosum (Insausti et al., 1998b). The lateral border of the parahippocampal cortex was the lateral bank of the collateral sulcus, and the medial border was the medial border of the subiculum. Specific guidelines are presented by Insausti et al. (1998a) for altering these boundaries to account for normal variation in the length, depth, and number of branches.
of the collateral sulcus. For a complete description of the segmentation procedure, see Insausti et al., (1998a,b, 2003).

These procedures resulted in five ROIs (the hippocampal region and temporopolar, entorhinal, perirhinal, and parahippocampal cortices) for each hemisphere. When brains were analyzed by a second scorer, the volumes for all ROIs were within 10% of the volumes reported here. Next, three different methods were employed to normalize the volumetric data.

Normalization by Intracranial Volume

An ROI was drawn in the sagittal view around all brain tissue (gray and white matter, including ventricular space) in every fifth section on average, including ventricular space, and excluding the brainstem below the level of the pons. AFNI then filled in the intermediate sections, and the area within each section was summed to yield the intracranial volume (ICV) measurement. The raw volumes for each of the five ROIs in each hemisphere were then divided by ICV to obtain the normalized measurement, which is equivalent to expressing each volume as a percentage of ICV.

Normalization by Area at the Anterior Commissure

The anterior commissure (AC) was identified visually on each scan, and an ROI was drawn around all brain tissue in the coronal section at that level. Raw volumes for each of the five ROIs in each hemisphere were then divided by this area to obtain the normalized measurement, which is equivalent to expressing each volume as a percentage of brain area at the level of the AC.
Normalization by Conversion Into Talairach Space

Standard landmarks were defined manually on the anatomical scans as described by Talairach and Tournoux (1998). The anatomical scans and the raw volumes for each of the five ROIs in each hemisphere were then resampled into Talairach space by AFNI using nearest-neighbor interpolation. The volume of each area after resampling was taken as the normalized measurement.

Comparison of Methods for Representing Volumetric Data

We began by calculating coefficients of variation (CoV) for measurements of each of the five ROIs, first when measured as raw volumes and then after each of the three normalization procedures was applied to the data. The CoV is the standard deviation of a sample divided by the sample mean, which is equivalent to expressing standard deviation as a percentage of the sample mean. The CoV was used as a measure of variability because it is independent of the magnitude of the measurement. By comparing CoVs, it is possible to assess to what extent the variability in estimates of regional brain volume can be reduced by applying different corrections (normalization procedures) for differences in overall brain size. A sample that has a smaller coefficient of variation is more homogeneous than a sample with a large coefficient of variation. The CoVs were compared using Miller’s test for the equivalence of coefficients of variation (Zar, 1999).

RESULTS

Volumes in Patients vs. Volumes in Controls

Figures 2 and 3 compare the volumes of the hippocampal region and parahippocampal gyrus in memory-impaired patients and matched controls (normalized by ICV). Measurements for patients J.R.W., J.S., R.S., and G.W. were compared to the measurements for 19 male controls (mean = 49 years old, SEM = 2.4), and measurements for patient L.J., were compared with the measurements for 11 female controls (mean = 67 years old, SEM = 1.0). The male patients as a group showed significant reduction in hippocampal volume (t[21] = 4.4, P < 0.05) but not in the volume of the parahippocampal gyrus (t[21] = 1.1, P > 0.10). Patients J.R.W., R.S., and G.W. all had hippocampal volumes more than 2 SD below the mean of the controls (z = −4.5, −3.3, and −4.9, respectively). Patient J.S. had a hippocampal volume within the normal range (z = 0.3), but focal lesions were present (see Fig. 1). No male patient had a parahippocampal gyrus volume of <1.3 SD below the mean control volume. Female patient L.J. also had a markedly reduced hippocampal volume (z = −4.5) and no reduction in the volume of the parahippocampal gyrus (z = 0.8). Lastly, the mean ICV of the patients was similar to the mean ICV of the controls (t[21] = 0.9, P > 0.1 for the male patients and t[10] = 1.0, P > 0.1 for patient L.J.).

The mean volume reductions were similar whether we normalized the measurements by ICV or employed the non-normalized (raw) volumes (for the hippocampus, 34% volume reduction vs. 39% volume reduction; for the parahippocampal gyrus, 3% volume reduction vs. 10% volume reduction). The mean reductions were also similar when the other methods of normalization were
The coefficient of variation (CoV) is the standard deviation of a sample divided by its mean. A smaller CoV reflects more homogeneous, less variable data. CoVs were calculated from (1) volumes normalized by intracranial volume, whereby each raw volume was divided by intracranial volume; (2) volumes normalized by converting raw volumes into Talairach space; (3) volumes normalized by the area of a single coronal section at the level of the anterior commissure; and (4) raw (unnormalized) volumes. Z-scores are shown from Miller’s test for equivalence of CoVs (Zar, 1991). Positive Z-scores indicate that the data in the first column are less variable than the data in the second column. Values above 1.96 in bold indicate a significant reduction in variability for the two methods being compared. In this group of 19 healthy males, the volumetric data were almost always numerically less variable than data obtained by the other methods.

**Comparison of Methods for Representing Volumetric Data**

Tables 2 and 3 compare coefficients of variation (CoV) in the male and female control population, respectively, for measurements of the hippocampal region and the cortical regions that lie along the parahippocampal gyrus. The CoV provides a measure of variability in each sample. A smaller CoV reflects more homogeneous, less variable data. Measurements were based on raw (unnormalized) volumes and three different normalization procedures, as described above. The Tables show z-scores computed using Miller’s test for the equivalence of coefficients of variation, which allows the CoVs of two different samples to be compared. For the tables, a positive z-score indicates that the CoV of the data acquired according to the normalization procedure in the first column is smaller than the CoV of the data acquired according to the normalization procedure in the second column.

Among healthy males (Table 2), the variability of hippocampal volume was significantly smaller for the measurements normalized by ICV than for the raw (un-normalized) measurements of the hippocampal region (z = 1.99, P < 0.05). The findings were similar for the parahippocampal gyrus (z = 2.37, P < 0.05). The other normalization procedures also reduced the CoV for the hippocampal and parahippocampal gyrus volumes, but the reduction was not significant. Lastly, normalization by ICV reduced variability in nearly all brain regions studied (9 of 10 regions) to a greater extent than did normalization by the other methods, although these differences did not reach significance.

Among healthy females (Table 3), variability was not significantly reduced by any of the normalization methods. Nevertheless, normalization by ICV reduced variability more than the other normalization methods in most brain regions studied (8 of 10 regions).

### TABLE 2.

**Comparison of Methods for Representing the Volume of Medial Temporal Lobe Structures in Healthy Males***

<table>
<thead>
<tr>
<th>Methods being compared</th>
<th>Hippocampus</th>
<th>Total</th>
<th>TPC</th>
<th>PRC</th>
<th>ERC</th>
<th>PHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Normalized by intracranial volume vs. 4. Raw volume (unnormalized)</td>
<td>1.99</td>
<td>2.37</td>
<td>−0.16</td>
<td>1.06</td>
<td>0.29</td>
<td>1.65</td>
</tr>
<tr>
<td>2. Normalized by Talairach conversion vs. 4. Raw volume (unnormalized)</td>
<td>1.86</td>
<td>1.14</td>
<td>−0.38</td>
<td>−0.51</td>
<td>−0.19</td>
<td>1.14</td>
</tr>
<tr>
<td>3. Normalized by area at anterior commissure vs. 4. Raw volume (unnormalized)</td>
<td>1.59</td>
<td>0.84</td>
<td>0.47</td>
<td>0.42</td>
<td>−0.16</td>
<td>0.66</td>
</tr>
<tr>
<td>1. Normalized by intracranial volume vs. 2. Normalized by Talairach conversion</td>
<td>0.14</td>
<td>1.27</td>
<td>0.22</td>
<td>1.56</td>
<td>0.47</td>
<td>0.53</td>
</tr>
<tr>
<td>1. Normalized by intracranial volume vs. 3. Normalized by area at anterior commissure</td>
<td>0.43</td>
<td>1.57</td>
<td>−0.63</td>
<td>0.64</td>
<td>0.45</td>
<td>1.01</td>
</tr>
</tbody>
</table>

TPC, temporopolar cortex; PRC, perirhinal cortex; ERC, entorhinal cortex; PHC, parahippocampal cortex; total, TPC + PRC + ERC + PHC.

*Z-scores from Miller’s test for equivalence of coefficients of variation (Zar, 1991). In this group of 11 healthy females, the data normalized by intracranial volume were almost always numerically less variable than the data obtained by other methods.

### TABLE 3.

**Comparison of Methods for Representing the Volume of Medial Temporal Lobe Structures in Healthy Females***

<table>
<thead>
<tr>
<th>Methods being compared</th>
<th>Hippocampus</th>
<th>Total</th>
<th>TPC</th>
<th>PRC</th>
<th>ERC</th>
<th>PHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Normalized by intracranial volume vs. 4. Raw volume (unnormalized)</td>
<td>0.48</td>
<td>0.56</td>
<td>−0.39</td>
<td>1.27</td>
<td>0.49</td>
<td>−0.05</td>
</tr>
<tr>
<td>2. Normalized by Talairach conversion vs. 4. Raw volume (unnormalized)</td>
<td>−0.19</td>
<td>−0.23</td>
<td>−1.02</td>
<td>0.26</td>
<td>0.15</td>
<td>0.91</td>
</tr>
<tr>
<td>3. Normalized by area at anterior commissure vs. 4. Raw volume (unnormalized)</td>
<td>0.22</td>
<td>0.45</td>
<td>0.04</td>
<td>0.38</td>
<td>−0.37</td>
<td>−0.14</td>
</tr>
<tr>
<td>1. Normalized by intracranial volume vs. 2. Normalized by Talairach conversion</td>
<td>0.67</td>
<td>0.79</td>
<td>0.63</td>
<td>1.02</td>
<td>0.34</td>
<td>−0.96</td>
</tr>
<tr>
<td>1. Normalized by intracranial volume vs. 3. Normalized by area at anterior commissure</td>
<td>0.26</td>
<td>0.11</td>
<td>−0.43</td>
<td>0.90</td>
<td>0.12</td>
<td>0.09</td>
</tr>
</tbody>
</table>

TPC, temporopolar cortex; PRC, perirhinal cortex; ERC, entorhinal cortex; PHC, parahippocampal cortex; total = TPC + PRC + ERC + PHC.

*Z-scores from Miller’s test for equivalence of coefficients of variation (Zar, 1991). In this group of 11 healthy females, the data normalized by intracranial volume were almost always numerically less variable than the data obtained by other methods.
We measured the volume of the hippocampus and parahippocampal gyrus bilaterally in five memory-impaired patients and 30 controls. Four of the five patients exhibited significant reduction in hippocampal volume, and none of the patients exhibited significant reduction in the volume of other medial temporal lobe structures. We then compared three different methods for reducing variability in the measurements, all of which involve corrections based on brain size: normalization by ICV, normalization by aligning the brain to the atlas of Talairach and Torneaux (1988), and normalization by brain area at the level of the anterior commissure.

Normalization by ICV reduced variability in volume measurements in nearly all brain regions to a greater extent than did normalization by other methods. ICV normalization also has the advantage that it produces an intuitively meaningful number (e.g., the percentage of the total brain volume that is hippocampus). Corrections based on ICV have been used previously when volume measurements are presented, although a standard method for defining ICV is not in regular use. For example, ICV has been defined as the volume of the supratentorial skull cavity (Kaye et al., 1997) or by an automated segmentation procedure that estimates the volume of white matter, gray matter, and cerebrospinal fluid (Callen et al., 2001). Commonly, a correction based on ICV is used without describing the method (cf. Cipolotti et al., 2001; Mayes et al., 2002). Inasmuch as our method of measuring ICV and the two methods considered by Free et al. (1995) all effectively reduced the variability of volume measurements within the medial temporal lobe, the specific method used to calculate ICV may be less important than ensuring that the method is applied identically and reproducibly across brains. An alternative method for reducing variability is to measure a small area or volume as a proxy for total brain size (Cendes et al., 1993; Free et al., 1995; Insausti et al., 1998a).

The area of a coronal section at the level of the anterior commissure is an example of such a proxy measure. This method is attractive because it can be done quickly and is readily explained. However, it has a number of drawbacks. First, abnormally shaped but normal-size brains may be incorrectly normalized. Second, because the region used as a proxy is typically small, the size of the region can be influenced by volume loss due to pathology, resulting in inaccurate normalization. Third, while the proxy method does provide a numerical benefit over uncorrected measures of volume, it is not as effective at reducing variability as normalization by ICV. Indeed, in agreement with our findings, Insauti et al. (1998a) reported that the proxy method reduced variability in only some regions of the medial temporal lobe but not in others.

Difficulties can also arise when normalizing volume measurements by aligning brains to a standard atlas (e.g., Talairach space; for recent applications of this method, see Bernasconi et al., 2003; Pruessnner et al., 2002). Our findings suggest that Talairach normalization is useful for normalizing measurements of hippocampal volume but is less useful for other regions of the medial temporal lobe. This method can be viewed as a regional brain size correction. It calculates, for example, the distance between the anterior commissure and the most anterior point of the brain, compares this distance to the standard brain, and then scales the tissue in that region accordingly. Thus, a normal brain with a slightly large frontal lobe and a slightly small occipital lobe will have the anterior portion of the medial temporal lobe (temporopolar cortex) shrunk and the posterior portion (parahippocampal cortex) stretched. Most likely, it is because of this variation in regional brain volume in the normal population that conversion of brains to Talairach space did not provide as great a reduction in variability as ICV normalization.

An additional technique for normalizing volume measurements deserves mention. Pruessnner et al. (2002) used the surface area of the collateral sulcus as a basis for normalizing the volumes of medial temporal lobe structures in a large population of healthy individuals. These authors noted, as did Insauti et al. (1998a), that the collateral sulcus presents with a different length, depth, and number of branches in every brain. On the basis of histological observations, Insauti et al. (1998a) defined the boundaries of medial temporal lobe cortices for all common collateral sulcus lengths, depths, and number of branches. Pruessnner et al. (2002) additionally sought to take into account the variable appearance of the collateral sulcus by measuring its surface area and then dividing the volumes of entorhinal, perirhinal, and parahippocampal cortices by this area.

This method of normalization, while useful for many applications, is less useful for estimating volumes in patients with medial temporal lobe damage. First, corrections based on the collateral sulcus are not useful for patients with damage that includes the sulcus itself. Second, as the authors reported, this normalization procedure improved the variability of measurements of perirhinal cortex, but the variability of measurements of entorhinal and parahippocampal cortices increased for 11 of the 12 groups in the study, sometimes by as much as a factor of three (Table 3 in Pruessnner et al., 2002).

A final observation about the interpretation of hippocampal volume loss is of interest. In the present study, the mean loss of hippocampal volume in the five patients was 34% (if J.S. is excluded, because he did not exhibit significant volume loss, the mean was 43%). Cipolotti et al. (2001) described a patient (V.C.) with approximately 45% loss of hippocampal volume. Isaacs et al. (2003) described six patients with developmental amnesia who had a mean volume loss in the hippocampus of 40%. Lastly, Mayes et al. (2002) described a patient (Y.R.) with a mean volume loss in the hippocampus of 46%. Interestingly, two patients studied previously (L.M. and W.H.) also had an estimated mean reduction in hippocampal size of 41% (based on MRI scans and corrected for temporal lobe size) (Squire et al., 1990). On subsequent histological examination (Rempel-Clower et al., 1996), this degree of reduction in hippocampal size was found to correspond to a loss of nearly all cells in the CA fields of the hippocampus. There was also extensive cell loss in the dentate gyrus, some subicular damage, and some cell loss in entorhinal cortex. These observations suggest that a reduction in hippocampal volume of approximately 40%, as estimated from MRI scans, likely indicates the nearly complete loss of hippocampal neurons. The tissue collapses with the result that
the hippocampus is markedly reduced in volume, but the tissue does not disappear entirely. Thus, a loss of approximately 40% of hippocampal volume as measured from MRI scans should not be taken to mean that 60% of the hippocampus remains functional.

Acknowledgments

The authors thank Peter Bayley, Jennifer Frascino, Joseph Manns, Gary Press, and Craig Stark for assistance. We also thank Terry Jernigan for providing MRI scans for many of the control brains.

REFERENCES


