Hippocampal Volume and Depression: A Meta-Analysis of MRI Studies

Poul Videbech, M.D.
Barbara Ravnkilde, Ph.D.

Objective: Several studies have found reduced hippocampal volume in patients with unipolar depression, but discrepancies exist. The authors performed a systematic review and meta-analysis of volumetric studies of the hippocampus in patients with mood disorders.

Method: Studies of hippocampal volume in unipolar and bipolar patients were identified. A meta-analysis of the 12 studies of unipolar depression fulfilling specific criteria was performed. The sample comprised 351 patients and 279 healthy subjects.

Results: The studies were highly heterogeneous regarding age and gender distribution, age at onset of the disorder, average number of episodes, and responsiveness to treatment, but the pooled effect size of depression was significant in both hemispheres for the unipolar patients. The weighted average showed a reduction of hippocampal volume of 8% on the left side and 10% on the right side. The causes of the heterogeneity were analyzed, and a meta-regression showed that the total number of depressive episodes was significantly correlated to right but not left hippocampal volume reduction.

Conclusions: Hippocampal volume is reduced in patients with unipolar depression, maybe as a consequence of repeated periods of major depressive disorder. Bipolar patients did not seem to show a reduction in hippocampal volume, but this has been much less investigated.

Increasing evidence has shown structural cerebral abnormalities in patients with unipolar and bipolar depression. Several studies have thus indicated an increased ventricle/brain ratio and other signs of both generalized and localized cerebral atrophy of the prefrontal cortex, cingulate gyrus, caudate nucleus, cerebellum, and hippocampus (for reviews see references 1–4). Often this atrophy is found to correlate with poor treatment response and shorter time to recurrence of the disease. Functional neuroimaging has also pointed to widespread abnormalities in the brain during depression (5).

The hippocampus is one of the areas in the brain that has been extensively studied in patients with mood disorders. This interest rests on a large body of neuropsychological and neuroimaging studies. The hippocampus is involved in episodic, declarative, contextual, and spatial learning and memory (6, 7), deficits which often accompany depression (8, 9). Furthermore, extensive rodent and human research has shown that its mnemonic functions and its neuroplasticity are highly sensitive to stress, i.e., increased cortisol levels (see reference 10 for an excellent review), which is found in a large proportion of patients with major depression (11).

In a Danish positron emission tomography (PET) study that included 42 acutely depressed patients and 47 matched healthy volunteers, one of the main findings was increased blood flow to the right hippocampus (12, 13). Accordingly, several other PET studies have found abnormalities in this structure in depression under various scanning conditions (14–19).

The hippocampus of patients with unipolar depression has been studied since 1993 using magnetic resonance imaging (MRI) techniques to reveal changes in volume, density, and water contents. Some volumetric studies have found significant bilateral volume deficits in depression (20–22). Others have found significantly lower volume in the right hemisphere (23, 24) or in the left hemisphere (25–27), but several studies have failed to find any differences (28–34).

Likewise, the picture is inconsistent regarding the correlation between measurements of hippocampal volume and clinical characteristics of the patient groups. One study found a correlation between age at onset of depression and hippocampal volume, namely that patients with late onset tended to have smaller hippocampi, especially in the right hemisphere (24). Other studies found the opposite to be the case (23) or could not confirm any relationship at all (33). Several authors have also tried to correlate the accumulated duration of episodes of depression to the volume of hippocampus and found that longer total duration of the disease or more episodes was correlated to smaller volumes (20–22). Important discrepancies do, however, still exist (26, 28, 33, 34). Finally, responsiveness to treatment has been correlated to volume reduction, which is often most pronounced in the right hippocampus (27, 32, 35).
Because of these discrepancies, which make it very difficult to reach a conclusion by simple summation of previous results, we decided to perform a meta-analysis of the effect of depression on hippocampal volume, hypothesizing that at least some of the discrepancies can be explained by between-group differences in number of recurrences.

Method

The MEDLINE and EMBASE electronic databases were searched using the following medical subject heading (MeSH): “Mood disorders” and “Magnetic Resonance Imaging” and “Hippocampus.” To make sure no study was missed, a free-text search was performed on the words “depression” and “MRI” and “hippocampus.” The search covered the years from 1966 through 2003. Furthermore, all reference lists of the obtained papers were scrutinized for studies not indexed in the electronic databases.

If not otherwise stated, all the studies reviewed herein fulfill the following criteria: 1) thorough clinical characterization of the patients with DSM-IV, ICD-10, or an equivalent system used as a diagnostic tool; 2) a comparison group of nearly the same average age; 3) exclusion of patients and comparison subjects with neurological disorders or medical diseases that could affect brain function; 4) exclusion of subjects with alcohol or drug dependency/abuse; and 5) comparison groups screened for psychiatric disorders. Since the first studies used scanners with low resolution unable to distinguish between the hippocampus and the adjacent amygdala, these studies were not considered.

In the present report all relevant studies were scrutinized, but only studies stating the mean and standard deviation of the hippocampal volume in each hemisphere separately were included in the meta-analysis. We converted all volumes to mm³ before entering them into the meta-analysis. Furthermore, this analysis was only carried out for studies of patients with unipolar depression and not for bipolar disorder patients, since these studies were few and very heterogeneous regarding the scanning techniques and actual measurements of hippocampal volume compared with the studies of unipolar depression patients.

The calculations were performed by using STATA, version 8 (Stata Corp., College Station, Tex.) by means of the Metan, Metareg, Metainf, and Metabias programs. The meta-analyses were performed by using a random effects model weighting the studies by the inverse variance and calculating the Dersimonian-Laird effect size. The random effects model was chosen because it is considered more conservative than a fixed effects model, since it takes into account the variability between studies leading to wider confidence intervals (CIs). Furthermore, the analyses were repeated excluding one study at a time to ensure that the results were not skewed by a single outlier. Heterogeneity, i.e., whether the differences between studies were greater than would be expected by chance alone, was assessed by the Q test and further analyzed by so-called meta-regression, a linear regression of the effect sizes against selected covariates. Meta-regression using the Metareg program was conducted to evaluate factors that could affect results between studies, such as differences in gender distribution or average age. A variable called RECUR was defined for each study and assigned a value of 1 if the study comprised patients with first-episode depression only and a value of 3 if all the patients participating in the study had recurrent depression. In three studies a value of 2 was assigned because the patient group was considered to consist of both

<table>
<thead>
<tr>
<th>Study and Year</th>
<th>Magnet Strength (T)</th>
<th>Slice Thickness (mm)</th>
<th>N</th>
<th>Age (years) Mean SD</th>
<th>Male (%)</th>
<th>Left Volume (mm³) Mean SD</th>
<th>Right Volume (mm³) Mean SD</th>
<th>RECUR¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posner et al. (28), 2003</td>
<td>1.5</td>
<td>1.0</td>
<td>27</td>
<td>33.0 ± 10.7</td>
<td>44</td>
<td>2546.0 ± 392.7</td>
<td>2948.4 ± 446.7</td>
<td>3</td>
</tr>
<tr>
<td>MacQueen et al. (20)², 2003</td>
<td>1.5</td>
<td>1.2</td>
<td>20</td>
<td>28.4 ± 11.8</td>
<td>35</td>
<td>2738.0 ± 301.1</td>
<td>2793.0 ± 303.8</td>
<td>1</td>
</tr>
<tr>
<td>MacQueen et al. (20)², 2003</td>
<td>1.5</td>
<td>1.2</td>
<td>17</td>
<td>35.9 ± 11.1</td>
<td>35</td>
<td>2381.0 ± 273.5</td>
<td>2392.0 ± 256.7</td>
<td>3</td>
</tr>
<tr>
<td>Frodl et al. (29), 2002</td>
<td>1.5</td>
<td>1.0</td>
<td>30</td>
<td>40.3 ± 12.6</td>
<td>43</td>
<td>3681.0 ± 393.0</td>
<td>3847.0 ± 400.0</td>
<td>1</td>
</tr>
<tr>
<td>Rusch et al. (30), 2001</td>
<td>1.5</td>
<td>1.2</td>
<td>25</td>
<td>33.2 ± 9.5</td>
<td>44</td>
<td>1770.0 ± 230.0</td>
<td>1870.0 ± 240.0</td>
<td>2</td>
</tr>
<tr>
<td>Mervaala et al. (25), 2000</td>
<td>1.5</td>
<td>3.0</td>
<td>34</td>
<td>42.2 ± 12.2</td>
<td>47</td>
<td>3104.0 ± 391.0</td>
<td>3462.0 ± 405.0</td>
<td>3</td>
</tr>
<tr>
<td>Vakili et al. (32), 2000</td>
<td>1.5</td>
<td>3.0</td>
<td>38</td>
<td>38.5 ± 10.0</td>
<td>45</td>
<td>2640.0 ± 550.0</td>
<td>2610.0 ± 580.0</td>
<td>2</td>
</tr>
<tr>
<td>von Gunten et al. (31), 2000</td>
<td>1.5</td>
<td>1.5</td>
<td>14</td>
<td>57.6 ± 42.6</td>
<td>42</td>
<td>2498.6 ± 294.8</td>
<td>2597.9 ± 244.0</td>
<td>1</td>
</tr>
<tr>
<td>Steffens et al. (24), 2000</td>
<td>1.5</td>
<td>3.0</td>
<td>66</td>
<td>71.7 ± 8.4</td>
<td>23</td>
<td>2920.0 ± 360.0</td>
<td>2980.0 ± 390.0</td>
<td>3</td>
</tr>
<tr>
<td>Bremner et al. (26), 2000</td>
<td>1.5</td>
<td>——³</td>
<td>16</td>
<td>43.0 ± 8.0</td>
<td>63</td>
<td>940.0 ± 208.0</td>
<td>982.0 ± 269.0</td>
<td>3</td>
</tr>
<tr>
<td>Ashtari et al. (33), 1999</td>
<td>1.0</td>
<td>3.1</td>
<td>40</td>
<td>74.3 ± 6.0</td>
<td>29</td>
<td>1745.0 ± 380.0</td>
<td>1742.0 ± 345.0</td>
<td>3</td>
</tr>
<tr>
<td>Sheline et al. (50), 2003</td>
<td>1.5</td>
<td>1.3</td>
<td>24</td>
<td>50.8 ± 17.1</td>
<td>10</td>
<td>2171.0 ± 316.0</td>
<td>2203.0 ± 315.0</td>
<td>3</td>
</tr>
</tbody>
</table>

¹ 1=first-episode patients, 2=mixed group, 3=patients with recurrent depression.

² Data on first time depression and recurrent depression presented separately in this table.

³ Data not given.

http://ajp.psychiatryonline.org

Am J Psychiatry 161:11, November 2004
types of patients (Table 1). This variable was also used in a meta-regression.

Begg's and Egger's tests were used to test for publication bias, i.e., the phenomenon in which for instance studies with negative results are not published.

Results

Unipolar Depression

Twelve studies comprising 351 patients and 279 healthy subjects fulfilled the aforementioned criteria and were entered into the meta-analysis (Table 1). The studies deviate markedly on several demographic characteristics of the study groups: the mean age varies from 28 to 74 years and the percentage of male subjects in each group varies from 0 to 63. Clinically, some of the studies comprise patients with first-episode depression (20, 29, 31) or treatment-resistant depression (25). Furthermore, the average volumes measured varied somewhat, with one study deviating especially noticeably (26).

The Q test of heterogeneity (df=11) was highly significant as expected (left side: p<0.003; right side: p<0.01). For this reason the effect size was calculated under the assumption of a random effects model. The Derimoni-Laird pooled effect size revealed bilateral statistical significance: \( -0.38 \) (95% CI=\(-0.65\) to \(-0.11\)) for the left hippocampus (Figure 1) and \( -0.32 \) (95% CI=\(-0.56\) to \(-0.08\)) for the right hippocampus (Figure 2). The average volume reduction weighted by sample size was 8% in the left hemisphere and 10% in the right.

Begg's and Egger's tests for publication bias were both far from significant (smallest p=0.135), confirmed graphically by a funnel plot (Figure 3). The meta-analysis was repeated omitting one study at a time to ensure that the result was not skewed by a single study. This procedure did not change the random-effect estimate notably, as it in all cases continued to be statistically significant.

The significant heterogeneity was then analyzed using meta-regression. A priori we assumed that interstudy differences in age and gender distribution could explain some of the variation. Analyzed separately and together these variables were, however, not significantly correlated with the random effect estimate in either hemisphere (data available upon request). Meta-regression using the variable RECUR (Table 1) as covariate showed a significant negative correlation with the random-effect estimate in the right hemisphere (r=\(-0.30\); z=\(-2.36\), p<0.02) and nonsignificant correlation in the left (r=\(-0.20\); z=\(-1.18\), p=0.24). This means that the higher the proportion of patients with recurrent depression, the smaller the volume of the right hippocampus.

Bipolar Patients

Studies of the hippocampus in bipolar disorder patients are shown in Table 2. Most of the studies except one (36) showed no significant differences between patients and
comparison subjects (37–42). Several of the studies did, however, use a very crude slice thickness to calculate the volumes, which increased the variation. Furthermore, one of the studies reports very deviating volumes of the hippocampus. Despite such methodological shortcomings the results are rather uniform, indicating that the volume of hippocampus is not changed in bipolar disorder. However, in an uncontrolled study Ali et al. (43, 44) found that

\[ \frac{\text{Overall difference represents the Derimonian-Laird pooled effect size, calculated under the assumption of a random effects model. Studies are grouped by their RECUR variable, a value assigned on the basis of patient group type (1=first-episode patients, 2=mixed group, 3=patients with recurrent depression).} } {\text{Overall difference represents the Derimonian-Laird pooled effect size, calculated under the assumption of a random effects model. Studies are grouped by their RECUR variable, a value assigned on the basis of patient group type (1=first-episode patients, 2=mixed group, 3=patients with recurrent depression).}} \]
larger right hippocampal volume was associated with longer duration of the illness and poorer neuropsychological functioning.

Discussion

The question whether depression is associated with shrinkage of the hippocampus is indeed important for our understanding of the disease. From Table 1 it is seen that two studies found significant bilateral volume deficits, one found significantly reduced volume in the right hemisphere, and two found reduction in the left, whereas seven studies failed to find any differences. In contrast, the meta-analysis of the 12 studies included showed a significant effect size of depression on the volume of hippocampus in both hemispheres, most pronounced in the right. Furthermore, the volume reduction in the right hippocampus was significantly correlated to the number of episodes. Tests for publication bias both fell out negatively.

**Heterogeneity of Studies**

The marked differences among patient groups regarding age and gender distribution, age at first depression, average number of episodes, and responsiveness to treatment were a priori expected to increase the variation of hippocampal volume. Further increase in variation was expected considering differences in scanning protocols and delineation of the structures in question. Meta-analysis plays an important role precisely because of these possible confounders, since it is most likely that some of the confounding effects are diluted or even cancel each other out in the large number of patients analyzed. This increases the extendibility of the results to the general population of depressed patients (45). The risk, of course, is that the results of the studies point in so many directions that they too cancel each other out and obscure important links between volume and depression in certain subpopulations of patients.

It is therefore important to analyze the causes of the significant heterogeneity found among the studies included. The studies of unipolar depression patients are comparable regarding the MRI scanner used and spatial resolution applied in contrast to the studies of bipolar disorder patients. None of the studies showed any statistical differences in total intracranial volume between patients and healthy subjects, but all authors corrected for this, either by using relative measurements or by using total intracranial volume as a covariate in the statistical analysis.

Differences in scanning protocols and the delineation of the hippocampal boundaries on the MRI scans are thus important sources of variation between the measurements (46, 47). The meta-analysis is, however, relatively robust against this, since the effect of such moderators in the individual studies were the same in both patients and comparison subjects, and studies using protocols that cause large variance have less influence on the summarized effect size in the meta-analysis because of the weighting of studies.

One study especially stood out with very deviating measurements (26) but did not skew the analysis, since stepwise exclusion of one study at a time did not change the effect sizes significantly. The results of the meta-analysis thus cannot be attributed to any single study with extreme results. Clinical and demographical variables can, however, play an important role and were therefore controlled.

**Age and Gender**

The hippocampus is generally larger in men than in women, a fact accounted for in the selection of comparison subjects or, in a few of the included studies, by statistical means. Furthermore, decreased hippocampal volume has been reported with increasing age in male but not female healthy volunteers (48). Moreover, significant interaction between hippocampal size, depression, and gender...
was observed in at least one study of patients with first-episode depression (29). In this study, the volume of the left hippocampus was smaller in male patients, whereas the right was larger in female patients. If correct, this could confound the results of the studies mentioned in Table 1 as the female-to-male ratio varies considerably (from 0% to 63% male). Generally it is also problematic to draw any conclusions from a study of predominantly male participants to the predominantly female population of depressed patients. Using linear meta-regression we were, however, unable to demonstrate any significant confounding of age and gender on the summarized effect size.

**Treatment Response**

The ratio between treatment response and treatment resistance in the study populations may also influence the results. In three studies smaller volume in right hippocampus (32, 35) or reduced density in the left (27) was linked to poor response to antidepressant medication. It is difficult to account for the importance of this fact, since the frequency of refractory depression is practically never stated to account for the importance of this fact, since the frequency of refractory depression is practically never stated.

**Cumulative Time Being Depressed**

Several studies (20, 23) have found a negative correlation between total lifetime duration of depression and volume of the hippocampus since Sheline et al. (21, 22) reported this in women. However, others did not find any relationship between duration of depression or number of episodes and hippocampal volume (26, 28, 33). One study in fact even found nearly the opposite to be the case: smaller hippocampal volumes in late-onset depression (24), supporting the notion that late-onset depression has a different etiology and pathophysiology compared with early-onset depression (1, 49). Omitting this study from the present analysis does not, however, change the results significantly.

Two studies of first-episode patients found no differences in hippocampal volume (20, 29). Accordingly, a meta-regression with the variable RECUR designating the proportion of first-episode patients versus patients with recurrent depression showed a highly significant correlation with the estimate of effect size in the right hemisphere. This means that the number of depressed episodes was correlated with lower volume of right but not left hippocampus, and that some of the heterogeneity can be explained by this variable. The RECUR variable is indeed a very crude measurement of recurrences and probably only loosely correlated to the accumulated time of depression. This crudeness will increase the risk of type II error, thus making our conclusion even stronger. Sheline et al. extended their original data using a much more precise estimate, namely the number of days of untreated depression, and correlated it with hippocampal volume. Their results ($R^2=0.28, p=0.0006$) revealed that 28% of the variation in volume can be explained by this variable (50).

Other studies have used statistical parametric mapping to estimate hippocampal size and found significantly smaller right hippocampi in depressed patients, particularly in patients with a longer course of illness (23). Others found that subjects with chronic depression showed reduced gray matter density in the left temporal cortex, including the hippocampus, and a tendency toward reduction in the right hippocampus (27).

**Limitations of the Study**

In principle, cross-sectional studies such as those included in the present analyses cannot conclude about causality. Does the depression cause shrinkage of the hippocampus or are subjects with small hippocampi susceptible to depression? It is tempting to conclude the former on the basis of our meta-regression and the data of Sheline et al. (50), but longitudinal follow-up studies are necessary. We have therefore initiated a study along these lines at our department.

A confounding effect of posttraumatic stress disorder and early lifetime stress, which both are often followed by...
depression, cannot be completely excluded. In some studies, but not all, these conditions have been associated with reduced hippocampal size (51, 52). Hence, it is important in future studies to account for such factors.

Other factors can also act as confounders, such as adolescent-onset alcohol abuse, which has been connected to smaller hippocampi (53), but this has been accounted for in the studies.

We abstained from performing a meta-analysis of the data on bipolar disorder patients because of the small number of studies and because some of the studies used scanning techniques that today must be considered suboptimal. The conclusions on this topic are therefore tentative.

**Depression, Hippocampal Shrinkage, Cognitive Deficits, Dementia?**

Volume reduction of the hippocampus offers an explanation of recent epidemiological and clinical findings of depression being a risk factor for dementia. A large register study showing that affective patients had an increased risk of developing dementia compared with the general Danish population (54) has recently been confirmed by meta-analyses (55, 56). Moreover, cognitive impairment has been demonstrated even in the euthymic phase in patients with unipolar depression and bipolar disorder (57, 58), and severity of the deficits has been shown to correlate with the number of affective episodes (59).

A few MRI studies have supported a connection between hippocampal abnormalities in depressed patients and cognitive deficits. In a study of patients with chronic depression, reduced gray matter density was found in the left temporal cortex, including the hippocampus, as well as a tendency toward reduction in the right hippocampus. Left hippocampal gray matter density was correlated with verbal recognition memory: the higher the density, the better the performance (27). Relative to matched comparison subjects, euthymic women with recurrent depression showed smaller bilateral hippocampal volumes and a lower score in verbal memory, which is a neuropsychological measure of hippocampal function. In contrast, no difference in overall brain size or general intellectual performance was found (22). Concurrently, another study found impairments on hippocampus-dependent verbal memory tests in both patients with first-episode depression and those with multiple episodes. However, only the latter group had hippocampal volume reductions, which suggests that dysfunctions of the hippocampus predate detectable structural changes (20).

Two studies of geriatric depression found correlations between the brief assessment of memory and attention from the Mini-Mental State Examination and volume deficits in the left (24) and bilateral (33) hippocampus, although one study did not find any associations (23). It is of interest that having a small left hippocampus has been found to predict dementia at 5-year follow-up in a group of 115 older nondemented depressed individuals (60).

**What Is the Mechanism Behind the Decreased Hippocampal Volume?**

The nature of the volume reduction of hippocampus is not known. The elevated glucocorticoid levels often seen in severely depressed patients (11) along with the decreased hippocampal volume suggest a mechanism for putative neuronal loss seen within depressive patients either by apoptosis (programmed cell death) or inhibition of neurogenesis (61–63). Other mechanisms are, however, also possible, such as reduction of the volume of individual neurons or reduction of glia tissue (64, 65). Numerous animal studies have shown that glucocorticoids are toxic to the hippocampus, analogous to what is seen in Cushing’s syndrome, in which the patients exhibit cognitive dysfunction, depression, and reduced hippocampal volume in addition to the other symptoms characteristic for this disease (66). It is thus well established that approximately half of depressed patients have hypothalamic-pituitary-adrenal (HPA) axis hyperactivity (11, 67, 68). This
abnormality could implicate hippocampal dysfunction because of its inhibitory influence on the HPA axis (69–71). In a PET study of relatively acutely depressed patients, we found markedly increased blood flow to the hippocampus (12), whereas others have found decreased activity in the parahippocampal area in a study of patients with treatment-resistant depression with a very long depression history (72). It is therefore tempting to hypothesize that in some types of depression, stressful life events may initiate a vicious circle in which increased cortisol levels gradually overstimulate the hippocampal cells, leading to their death and further decreasing the inhibitory regulation of the HPA axis (73–76). However, only one volumetric study of depression has been performed where the authors also measured cortisol after a dexamethasone suppression test, but unfortunately their MRI technique did not allow separation of amygdala from the hippocampus (77). In future research the combination of measuring HPA activity together with hippocampal volume in longitudinal studies is important.

It is not known whether the reduction in volume is reversible. Several studies have, however, suggested that treatment of depression can stop hippocampal atrophy or even reduce it (78–80), and a recent study has suggested that the behavioral effects of chronic antidepressant treatment may be mediated by the stimulation of neurogenesis in the hippocampus (81). Furthermore, neuropathological evidence from postmortem studies of patients with major depressive disorder or bipolar disorder suggests that depression is a disorder of neuroplasticity and cellular resilience and not a neurodegenerative disorder (65, 82). The aforementioned connection between depression, stress, cortisol, and reduced hippocampal volume is intriguing. It is tempting to speculate that the hippocampus in patients with bipolar disorder being of normal size points to differences in pathogenesis between unipolar depression and bipolar disorder, but this requires further research because of the small number of volumetric studies of bipolar disorder patients.

Conclusion

In the present meta-analysis we found an average reduction of hippocampal volume of 8% in the left hemisphere and 10% in the right hemisphere in depressed patients relative to comparison subjects. It is interesting that a recent PET study of acutely depressed patients also found abnormalities in the right hippocampus (12). Reduced hippocampal volume is, however, not specific for depression, since it is also seen to a much larger degree in Alzheimer’s disease (83).

The present findings of reduced hippocampal volume in unipolar depression and a correlation with the number of episodes are clinically interesting and in accordance with the predictions of the so-called glucocorticoid cascade hypothesis, although other explanations are also possible. If hippocampal volume reduction is a consequence of untreated depression, secondary prophylaxis to prevent the damage to the hippocampus becomes extremely important, especially since several studies suggest that treatment can stop the shrinkage or even reduce it. To test these hypotheses, longitudinal studies are necessary.

References


http://ajp.psychiatryonline.org
Am J Psychiatry 161:11, November 2004