Smaller hippocampal volumes in patients with bipolar disorder are masked by exposure to lithium: a meta-analysis

Tomas Hajek, MD, PhD; Miloslav Kopecek, MD, PhD; Cyril Höschl, MD; Martin Alda, MD

Hajek, Alda — Department of Psychiatry, Dalhousie University, Halifax, NS; Hajek, Kopecek, Höschl, Alda — Prague Psychiatric Centre, Department of Psychiatry and Medical Psychology, 3rd School of Medicine, Charles University, Prague, Czech Republic

Background: Smaller hippocampal volumes relative to controls are among the most replicated neuroimaging findings in individuals with unipolar but not bipolar depression. Preserved hippocampal volumes in most studies of participants with bipolar disorder may reflect potential neuroprotective effects of lithium (Li). Methods: To investigate hippocampal volumes in patients with bipolar disorder while controlling for Li exposure, we performed a meta-analysis of neuroimaging studies that subdivided patients based on the presence or absence of current Li treatment. To achieve the best coverage of literature, we categorized studies based on whether all or a majority, or whether no or a minority of patients were treated with Li. Hippocampal volumes were compared by combining standardized differences between means (Cohen d) from individual studies using random-effects models. Results: Overall, we analyzed data from 101 patients with bipolar disorder in the Li group, 245 patients in the non-Li group and 456 control participants from 16 studies. Both the left and right hippocampal volumes were significantly larger in the Li group than in controls (Cohen d = 0.53, 95% confidence interval [CI] 0.18 to 0.88; Cohen d = 0.51, 95% CI 0.21 to 0.81, respectively) or the non-Li group (Cohen d = 0.93, 95% CI 0.56 to 1.31; Cohen d = 1.07, 95% CI 0.70 to 1.45, respectively), which had smaller left and right hippocampal volumes than the control group (Cohen d = −0.36, 95% CI −0.55 to −0.17; Cohen d = −0.38, 95% CI −0.63 to −0.13, respectively). There was no evidence of publication bias. Limitations: Missing information about the illness burden or lifetime exposure to Li and polypharmacy in some studies may have contributed to statistical heterogeneity in some analyses. Conclusion: When exposure to Li was minimized, patients with bipolar disorder showed smaller hippocampal volumes than controls or Li-treated patients. Our findings provide indirect support for the negative effects of bipolar disorder on hippocampal volumes and are consistent with the putative neuroprotective effects of Li. The preserved hippocampal volumes among patients with bipolar disorder in most individual studies and all previous meta-analyses may have been related to the inclusion of Li-treated participants.

Introduction

Neuroprotective effects of lithium (Li) have been well documented in tissue cultures and animal models. Some neuroimaging studies have reported findings suggestive of similar effects of Li in humans, although the data continue to be limited and inconsistent.

Smaller hippocampal volumes are among the most replicated neuroimaging findings in patients with major depressive disorder compared with healthy controls. These structural changes are thought to reflect excitotoxicity that may be related to hypothalamic–pituitary–adrenal (HPA) axis dysregulation during illness recurrences. In contrast, studies of patients with bipolar disorder have been contradictory, with findings of mostly comparable, but also smaller or even larger hippocampal volumes in patients with bipolar disorder relative to controls. As a result, all 8 published meta-analyses found preserved hippocampal volumes in patients with bipolar disorder.

The absence of hippocampal volume abnormalities among patients with bipolar disorder in most studies is unexpected. The clinical features, as well as HPA changes, overlap between unipolar and bipolar depression. In addition, bipolar depression is typically the main manifestation of the illness, and may start earlier than unipolar depression. Patients with unipolar and bipolar depression, however, differ broadly in terms of medication exposure.

Although Li shows antidepressant properties even in patients...
with unipolar depression, it is predominantly used as a mood stabilizer in those with bipolar disorder. Previous studies and meta-analyses have typically investigated hippocampal volumes in participants with bipolar disorder without controlling for exposure to Li. Lithium treatment leads to increased hippocampal volumes reported in prospective studies. Lithium-exposed patients in retrospective studies have typically shown larger hippocampal volumes than unmedicated participants. Thus, it is possible that hippocampal volume changes in patients with bipolar disorder are masked by exposure to putative neuroprotective effects of Li.

In this study, we investigated whether there is an association between hippocampal volumes and Li treatment and whether the absence of hippocampal volume differences between patients with bipolar disorder and controls in most available studies was related to the inclusion of Li-treated participants. To this end, we separately analyzed hippocampal volumes in samples of patients with bipolar disorder with or without current exposure to Li. The mostly post hoc results for such subgroups have not yet been investigated in their entirety. The use of a meta-analysis allowed us to achieve a greater statistical power and a more precise estimate of the effect size relative to individual comparisons. Our a priori hypotheses were that hippocampal volumes would be smaller among patients with bipolar disorder with no current exposure to Li relative to Li-treated patients, who would have comparable hippocampal volumes to controls.

Methods

Data sources

We carried out a systematic search of MEDLINE, EMBASE and SCOPUS databases for articles published from Jan. 1, 1980, to June 15, 2011, using the following Medical Subject Headings: “magnetic resonance imaging,” “hippocampus,” “lithium” and “bipolar disorder.” Review articles relating to neuroimaging in bipolar disorder and reference lists of the included studies were also searched for eligible reports.

Study selection

Studies were considered for inclusion if they were indexed in MEDLINE, EMBASE or SCOPUS as published or in press in peer-reviewed journals by June 15, 2011; investigated hippocampal volumes in patients with bipolar disorder; and provided information about the proportion of patients treated with Li or provided separate volumetric results for patients who were treated with Li or those who were not currently exposed to Li. When a study reported mean hippocampal volumes and standard deviations (SD) adjusted for confounders, these were used in the meta-analysis in place of the raw means. To achieve the best coverage of the available literature, the studies in which the majority of patients was treated with Li were analyzed together with the studies where all patients were treated with Li. Similarly, the studies in which the minority of patients was treated with Li were analyzed together with the studies where no patient was treated with Li.

We decided to analyze the most complete set of studies. Rather than setting stringent exclusion criteria, which would yield only a small, nonrepresentative fraction of the literature, we attempted to identify sources of heterogeneity in subgroup and sensitivity analyses. This approach not only reduces the risk of publication bias, but also increases the generalizability of results and helps identify potential confounders.

Studies were excluded if information about Li treatment at the time of scanning was not provided, voxel-based morphometry was used or noncontiguous slices were collected for the measurements. We also extracted variables that are important for interpretation of the results, including treatment-related factors such as the duration, dose and compliance with the Li treatment and previous exposure to Li, as well as information about illness burden such as duration of illness and number of episodes. This is important as, regardless of medication exposure, hippocampal volume abnormalities are unlikely to be present in the early stages of illness and tend to appear only after a substantial illness burden. Thus, differences in the duration of illness or the numbers of episodes could confound the results.

Two reviewers (T.H., M.K.) assessed each study for eligibility and checked that all data were transcribed correctly.

Statistical analyses

Meta-analyses were performed using Comprehensive Meta-Analysis software, version 2 (Biostat). As a measure of effect size, we used standardized difference between means (Cohen d). Since we could not expect a constant population effect size across studies (fixed effects), we decided to use the random-effects model, with study as the random effect. This assumes that the “population” of studies has variable true effects that are normally distributed. We performed 3 meta-analyses to investigate the differences in hippocampal volumes between 1) patients with bipolar disorder, most or all of whom were treated with Li (Li group), and controls; 2) patients with bipolar disorder, most or all of whom were not currently treated with Li (non-Li group), and controls; and 3) Li-treated patients with bipolar disorder and patients who were not treated with Li at the time of scanning.

We used the volumes of the left and right hippocampus as our primary outcome measure, as most of the available studies did not provide mean or variance estimates for the total hippocampal volume. Whenever there were at least 3 studies reporting only total hippocampal volumes for a particular comparison, we analyzed these separately to provide a maximum coverage of the available studies.

We calculated the I² to provide an easily interpretable measure of consistency among the studies. The I² is an estimate of the percentage of the total variation across studies due to true heterogeneity rather than chance. The I² is placed between 0% and 100%, with a value of 0% indicating no heterogeneity and larger values showing increasing heterogeneity. The I² values of 25%, 50% and 75% indicate low, moderate and high levels of heterogeneity, respectively.

We performed a jackknife sensitivity analysis, omitting one study at a time, to assess whether the results would change
after the exclusion of any study and whether they would remain significant in replication studies (i.e., after the exclusion of the first positive published study).

We used the Egger regression test of funnel plot asymmetry20 and the classic fail-safe N to examine the risk of publication bias. We adopted a significance level of \( p = 0.05 \), 2-tailed, for all these analyses.

We also carried out a power analysis for an independent samples \( t \) test using Statistica 6.0 software (StatSoft) to calculate the numbers of participants needed to detect effect sizes identified in this meta-analysis as statistically significant (\( p = 0.05, 80\% \) power, 2-tailed).

Results

Results of the systematic search

Out of 144 articles initially found in our systematic search, 47 were reviews, 30 did not investigate patients with bipolar disorder, 11 used voxel-based morphometry or other morphometry, 6 used functional magnetic resonance imaging, 8 used magnetic resonance spectroscopy and 8 did not measure hippocampal volume. In all, 34 studies measured hippocampal volumes in patients with bipolar disorder. Among these, 17 nonoverlapping studies provided details about exposure to Li.19,22,23,30–32,47,48,54–56 All but 1 of these studies were included in the meta-analysis. The study by Altshuler and colleagues47 could not be included as it was the only study of Li-treated patients versus controls that provided results for total hippocampal volumes only.

Lithium versus control groups

We analyzed results from 7 of 8 studies that compared hippocampal volumes in patients with bipolar disorder, most or all of whom were not currently treated with Li (non-Li group), and controls.22,23,26–31,32,48 Four studies20,27,29,55 provided results for the total hippocampal volumes only and were thus analyzed separately. In 8 of the analyzed studies (4 studies of bilateral hippocampal volumes20,27,29,55 and 4 studies of total hippocampal volume20,27,29,55), the minority of patients was treated with Li at the time of scanning (range 9% to 35% of patients), whereas in the remaining 6 studies none of the patients was treated with Li at the time of scanning. All of the analyzed studies recruited both men and women and reported mean ages ranging from 10.4 (SD 3.0) to 69.0 (SD 5.9) years. With the exception of 2 studies that used the 3-T scanner20 and the 1-T magnet,55 all of the other studies used 1.5-T scanners with a slice thickness ranging from 0.9 to 3.0 mm. In all of the analyzed studies, the tracings were blind to the diagnostic or medication status of the participants.

Six studies recruited patients with bipolar I disorder only,20,27,31,32,47,56 and 5 recruited patients with bipolar I and II disorders.22,23,26,29,48 The remaining 3 studies did not provide information about the proportion of patients with bipolar I or II disorders.20,29,35

A single study recruited medication-naive patients,7 and another study required a 1-month Li-free period before scanning.20 Two studies provided information about the lifetime history of Li exposure.20,26 The patients in these studies were treated with anticonvulsants, antidepressants, antipsychotics, psychostimulants and benzodiazepines. The average duration of illness ranged from 2.4 (SD 3.1) to 29.7 (SD 15.7) years, but this information was missing in 4 studies.20,27,29,48 The average number of episodes ranged from 6.4 (SD 3.6) to 20.9 (SD 18.6), but this information was missing in 7 studies.20,22,23,26,30,32,48 A single study provided separate estimates for the number of manic and depressive episodes7 (Table 1). No study used the duration of illness or the number of episodes as inclusion criteria.

Lithium versus non-Li groups

We analyzed all 7 studies that compared hippocampal volumes in Li-treated patients with bipolar disorder and patients who were not treated with Li at the time of scanning.22,23,30–32,47,48 With the exception of the study by Foland and colleagues,47 the other studies were described in the previous comparisons.

In all of these studies, all patients in the Li group were...
Table 1: Neuroimaging studies of hippocampal volumes in patients with bipolar disorder who were grouped according to the presence or absence of lithium treatment; values reported are mean (standard deviation) [and range] for the lithium (Li), non-Li and control groups, unless otherwise indicated (part 1 of 2)

<table>
<thead>
<tr>
<th>Study</th>
<th>No. (Li, Non-Li, Control groups)</th>
<th>Age, yr (Li, Non-Li, Control groups)</th>
<th>Groups matched by age?</th>
<th>Duration of current Li exposure, wk</th>
<th>Li dose, mg</th>
<th>Other med used</th>
<th>Mediated criteria for lifetime Li exposure to Li?</th>
<th>Controlled</th>
<th>Illness burden</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hajek et al.</td>
<td>21 (Li, Non-Li, Control)</td>
<td>32.9 (11.4) 35.6 (10.4) 32.7 (10.1)</td>
<td>Yes</td>
<td>2 wk Li</td>
<td>123 (203)</td>
<td>[2–1000]</td>
<td>Avg Li level at scan 0.79 (0.29) [1.40–1.49]</td>
<td>17 Li mono</td>
<td>128.5 (34.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.15 (0.75)</td>
<td>[675–2100]</td>
<td>8 ADYS, 1 AP, 1 BZD</td>
<td>3 ABD</td>
<td>14.9 (7.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS, not separated by polarity</td>
</tr>
<tr>
<td>Bertolino et al.</td>
<td>17 (Li, Non-Li, Controls)</td>
<td>40.1 (12.9) 37.6 (10.3) 31.9 (11.3)</td>
<td>Yes</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Li at scan</td>
<td>NR</td>
<td>Overall 6 pts &lt; 5,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NR</td>
<td>NR</td>
<td>Non-Li = controls</td>
<td>NR</td>
<td>Overall 9 pts &lt; 5,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>not separated for med subgroup</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BD &gt; controls,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Li &gt; non-Li</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brambilla et al.</td>
<td>24 (Li, Non-Li, Controls)</td>
<td>35 (10) 37.1 (10.3)</td>
<td>Yes</td>
<td>NR</td>
<td>91 (116)</td>
<td>[median 29]</td>
<td>Li mono</td>
<td>WR</td>
<td>Overall 3.9 (2.4),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>not separated for med subgroups</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chepenik et al.</td>
<td>20 (7 Li, Non-Li, Controls)</td>
<td>14.6 (2.8) 14.1 (2.8)</td>
<td>Yes</td>
<td>NR</td>
<td>16.0 (26)</td>
<td>[14.6 (26)]</td>
<td>Li at scan</td>
<td>WR</td>
<td>Overall 3.9 (2.4),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>not separated for med subgroups</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen et al.</td>
<td>10 (Li, Non-Li, Controls)</td>
<td>16.0 (36) 14.6 (26)</td>
<td>Yes</td>
<td>NR</td>
<td>4 ACS, 3 ABD, 1 BZD, 1 AP, 1 PS</td>
<td>26 P S, 12 ACS, 10 VPA, 7 AP, 7 Li</td>
<td>6 ACS, 8 ADYS, 1 AP, 1 BZD</td>
<td>NR</td>
<td>Overall 3.9 (2.4),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>not separated for med subgroups</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frazier et al.</td>
<td>43 (Li, Non-Li, Controls)</td>
<td>11.3 (11.3) 10.5 (10.5)</td>
<td>Yes</td>
<td>NR</td>
<td>13% of pts on Li at scan 19 (5 on Li)</td>
<td>10 P S, 9 A D S, 9 PS, 3 A P</td>
<td>5.5 (8.6), 17.8 (11.4)</td>
<td>NR</td>
<td>Overall 3.9 (2.4),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>not separated for med subgroups</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hartberg et al.</td>
<td>121 (Li, Non-Li, Controls)</td>
<td>34.7 (11.4) 36.1 (11.6)</td>
<td>Yes</td>
<td>NR</td>
<td>13% of pts on Li at scan 19 (5 on Li)</td>
<td>10 P S, 9 A D S, 9 PS, 3 A P</td>
<td>5.5 (8.6), 17.8 (11.4)</td>
<td>NR</td>
<td>Overall 3.9 (2.4),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>not separated for med subgroups</td>
</tr>
</tbody>
</table>

ACS = anticonvulsants; ADS = antidepressant; AP = antipsychotics; avg = average; BD = bipolar disorder; BZD = benzodiazepines; eps = episodes; HC = hippocampus; Li = lithium; med = medication; mono = monotherapy; MS = mood stabilizers; NA = not applicable; NOS = not otherwise specified; NR = not reported; NS = nonsignificant; PS = psychostimulants; pts = patients; SD = standard deviation; VPA = valproate.
### Table 1: Neuroimaging studies of hippocampal volumes in patients with bipolar disorder who were grouped according to the presence or absence of lithium treatment; values reported are mean (standard deviation) [and range] for the lithium (Li), non-Li and control groups, unless otherwise indicated (part 2 of 2)

<table>
<thead>
<tr>
<th>Study</th>
<th>No. (Li, Non-Li, Control groups)</th>
<th>Age, yr</th>
<th>Groups matched by age?</th>
<th>Medication exposure</th>
<th>Duration of illness, yr (Li, Non-Li groups)</th>
<th>Illness burden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Javadapour et al.</td>
<td>12/12/24</td>
<td>Combined Li and non-Li</td>
<td>NR</td>
<td>Li at scan</td>
<td>NA</td>
<td>Li = controls; Li &gt; non-Li (trend)</td>
</tr>
<tr>
<td>van Erp et al.</td>
<td>10/8/32</td>
<td>14.2 (11.0)</td>
<td>32</td>
<td>50.2 (11.1)</td>
<td>NA</td>
<td>Li = controls; Li &gt; non-Li (trend)</td>
</tr>
<tr>
<td>Altshuler et al.</td>
<td>24 (17 on Li)</td>
<td>50.2 (11.7)</td>
<td>18</td>
<td>53.4 (11.1)</td>
<td>NA</td>
<td>Li = controls; Li &gt; non-Li (trend)</td>
</tr>
</tbody>
</table>

ACS = anticonvulsants; ADS = antidepressants; AP = antipsychotics; avg = average; BD = bipolar disorder; BZD = benzodiazepines; eps = episodes; HC = hippocampus; Li = lithium; med = medication; mono = monotherapy; MS = mood stabilizers; NA = not applicable; NOS = not otherwise specified; NR = not reported; NS = nonsignificant; PS = psychotherapists; pts = patients; SD = standard deviation; VPA = valproate.
included 245 patients with bipolar disorder (214 of whom were not currently treated with Li) and 456 controls for the left hippocampus and 9 studies that included 233 patients with bipolar disorder (202 of whom were not currently treated with Li) and 432 controls for the right hippocampus, since 1 study\(^3\) provided results for only the left hippocampus. Both the left and right hippocampi were significantly smaller in the non-Li than in the control group (Cohen d = –0.36, 95% CI –0.55 to –0.17; Cohen d = –0.38, 95% CI –0.63 to –0.13, respectively; Fig. 2). The statistical heterogeneity among these studies was low to moderate for the left and right hippocampal volumes (\(I^2 = 12.19\%, p = 0.33\) and \(I^2 = 36.29\%, p = 0.13\), respectively).

Four additional studies of 116 patients with bipolar disorder, 84 of whom were not currently treated with Li, and 111 controls provided results only for the total hippocampal volume. The non-Li group had significantly smaller total hippocampal volume than the control group (Cohen d = –0.58, 95% CI –1.04 to –0.12), with moderate statistical heterogeneity (\(I^2 = 61.15\%, p = 0.05\)).

Sensitivity analysis: The results remained significant after exclusion of each individual study for both the left and right hippocampus and in one-quarter of studies investigating the total hippocampal volumes. In a separate analysis of studies in which no patient was exposed to Li at the time of scanning, the hippocampi remained significantly smaller in the non-Li patients than controls on the right (5 studies, 40 non-Li patients, 174 controls, Cohen d = –0.40, 95% CI –0.77 to –0.03), but not the left side (6 studies, 52 non-Li patients, 198 controls, Cohen d = –0.22, 95% CI –0.58 to 0.13). The statistical heterogeneity for the left and right hippocampus was low (\(I^2 = 8.64\%, p = 0.36\); \(I^2 = 21.2\%, p = 0.27\), respectively). When we excluded the 3 reports that studied pediatric patients,\(^4\)–\(^6\) the hippocampi remained significantly smaller in the non-Li group than in the control group on both the left and right sides (7 studies, 184 non-Li patients, 386 controls, Cohen d = –0.34, 95% CI –0.59 to –0.09; 6 studies, 172 non-Li patients, 362 controls, Cohen d = –0.46, 95% CI –0.65 to –0.28) with low statistical heterogeneity (\(I^2 = 21.63\%, p = 0.26\) and \(I^2 = 0.0\%, p = 0.64\), respectively).

Publication bias: There was no evidence of a publication bias, as tested by visual inspection of the funnel plots and by using Egger regression intercept (\(p = 0.51–0.87\)). The number of additional negative studies required to bring the \(p\) value above 0.05 was 26 for the left hippocampus, 29 for the right hippocampus and 13 for the total hippocampus.

Sample size calculations: A single study would need to recruit 123, 110 and 48 participants per group to detect effect sizes of –0.36, –0.38 and –0.58, respectively, as seen in this meta-analysis for the differences in the left, right and total hippocampal volumes between the non-Li and control groups (\(p = 0.05\), 80% power, 2-tailed).

Lithium versus non-Li groups
Overall, we analyzed data from 7 studies that included 89 Li-treated patients with bipolar disorder and 89 patients who were not currently treated with Li for the left hippocampus and 6 studies that included 77 Li-treated patients with bipolar disorder and 77 patients who were not

<table>
<thead>
<tr>
<th>Side; study</th>
<th>Left hippocampus</th>
<th>Right hippocampus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bearden et al.(^2)</td>
<td>Bearden et al.(^2)</td>
</tr>
<tr>
<td></td>
<td>Beyer et al.(^3)</td>
<td>Beyer et al.(^3)</td>
</tr>
<tr>
<td></td>
<td>Brambilla et al.(^2)</td>
<td>Brambilla et al.(^2)</td>
</tr>
<tr>
<td></td>
<td>Chen et al.(^2)</td>
<td>Chen et al.(^2)</td>
</tr>
<tr>
<td></td>
<td>Javadapour et al.(^2)</td>
<td>Javadapour et al.(^2)</td>
</tr>
<tr>
<td></td>
<td>Yucel et al.(^4)</td>
<td>Yucel et al.(^4)</td>
</tr>
<tr>
<td></td>
<td>van Erp et al.(^3)</td>
<td>van Erp et al.(^3)</td>
</tr>
</tbody>
</table>

\(\text{Cohen d (95\% confidence interval)}\)

\[\text{Li} < \text{controls} \quad \text{Li} > \text{controls} \]

Fig. 1: Comparison of hippocampal volumes in the lithium (Li) and control groups.
currently treated with Li for the right hippocampus, since 1 study\textsuperscript{31} provided results for only the left hippocampus. Both the left and right hippocampi were significantly larger in the Li-treated patients with bipolar disorder than in the non-Li group (Cohen d = 0.93, 95% CI 0.56–1.31; Cohen d = 1.07, 95% CI 0.70–1.45, respectively; Fig. 3). The statistical heterogeneity among these studies was low for both the left and right hippocampal volumes ($I^2 = 19.84\%$, $p = 0.28$ and $I^2 = 6.6$, $p = 0.37$, respectively).

Sensitivity analysis: The results for both the left and the right hippocampus remained significant after the exclusion of any individual study, including the single study that enrolled only pediatric patients.\textsuperscript{22}

Publication bias: There was no evidence of a publication bias, as tested by visual inspection of the funnel plots and by using Egger regression intercept ($p = 0.38–0.69$). The number of additional negative studies required to bring the $p$ value above 0.05 was 23 for both the left and right hippocampi.

Sample size calculations: A single study would need to recruit 20 and 15 participants per group to detect effect sizes of 0.93 and 1.07, respectively, as seen in this meta-analysis for the differences in the left and right hippocampal volumes between the non-Li and Li-treated patients ($p = 0.05$, 80% power, 2-tailed).

Discussion

We found significantly smaller bilateral hippocampal volumes in patients with bipolar disorder who were currently not treated with Li relative to healthy controls or Li-treated patients with bipolar disorder, who had significantly larger hippocampal volumes than the controls. This is in contrast to the majority of previous studies,\textsuperscript{19,21–25} as well as to 8 previous meta-analyses,\textsuperscript{33–40} all of which reported comparable hippocampal volumes between patients with bipolar disorder and controls.

Unlike any of the previous meta-analyses, we subdivided participants based on the presence or absence of Li treatment. The effect size for the larger hippocampal volume in the Li versus the control group (0.51–0.53) marginally exceeded the effect size for the smaller hippocampal volume in the non-Li versus the control group (–0.36 to –0.38). In those studies and in meta-analyses where patients with and without current Li treatment were analyzed jointly, the larger hippocampal volumes in the Li-treated patients could have masked the smaller hippocampal volumes in the non-Li subgroups, thus yielding hippocampal volumes comparable to those of controls.

Perhaps the most interesting and novel finding from the present study is the cumulative evidence for smaller hippocampal volume in patients with bipolar disorder relative to controls once exposure to Li was minimized. It challenges
the prevailing view of the field that patients with bipolar disorder have preserved hippocampal volumes.\textsuperscript{19,21–25,33–40} The previously reported differences between unipolar depression and bipolar disorder in hippocampal volumes\textsuperscript{14} were perhaps not related to any unique genetic or pathophysiological underpinning of each of these mood disorders, but rather to a different exposure to Li. Our results suggest that once current exposure to Li is accounted for, smaller hippocampal volumes are present in patients with bipolar disorder to a similar extent as in those with unipolar depression.\textsuperscript{14}

In keeping with our study, Kempton and colleagues,\textsuperscript{37} using meta-regression, reported that grey matter volume increased with the proportion of patients using Li. In addition, a recent mega-analysis summarized individual data from 5 previous studies and found significantly larger hippocampal volumes in Li-treated patients with bipolar disorder than in those who were not treated with Li or healthy controls.\textsuperscript{5} The methods (mega- v. meta-analysis) and, more importantly, the numbers of participants and the numbers of studies included markedly differed between these 2 investigations. In particular, only 3 of 7 studies included in our Li versus control comparison, 2 of 10 studies in our non-Li versus control comparison and 2 of 7 studies from the Li versus non-Li comparison were also included in the mega-analysis. Although the method of meta-analysis has lower statistical power than that of mega-analysis, the number of participants included was greater in this report than in the mega-analysis for both the Li (101 v. 94) and the non-Li (245 v. 68) groups. The findings of this study thus provide complementary results from different methods in a series of mostly different studies.

The larger hippocampal volumes in Li-treated patients with bipolar disorder may reflect patient heterogeneity or the effects of Li. Perhaps, contrary to the Li responders, the Li non-responders had a neurodegenerative type of illness or an illness that stemmed from pre-existing hippocampal volume changes. This explanation is unlikely, as none of the studies specifically selected unambiguous Li responders. Furthermore, smaller hippocampal volumes were not reported among participants at genetic risk for bipolar disorder or early in the course of illness in any of the previous studies.\textsuperscript{24,32,49,50}

Alternatively, perhaps Li affects the structure of the brain, either directly through neurochemical pathways involved in neurogenesis, or indirectly through the prevention of further episodes. It is clear that only a proportion of patients with diagnosed bipolar disorder will respond to Li.\textsuperscript{46,57} Thus, response to Li likely varied within the available studies. Furthermore, recent reports have shown an association between Li levels in the brain and N-acetylaspartate, a marker of neuronal density,\textsuperscript{58} as well as increases in the brain-derived neurotrophic factor levels in Li-treated patients.\textsuperscript{59} These studies argue for a direct effect of Li. In keeping with this, Li slowed the decline in cognitive performance and decreased the levels of phosphorylated tau protein in patients with amnestic mild cognitive impairment.\textsuperscript{60} Arguing for the indirect, clinically mediated effect are studies showing that grey matter volume increases after treatment with Li were only detected among responders, but not non-responders, to treatment\textsuperscript{61} and were associated with clinical improvement.\textsuperscript{62} Thus, both the direct neurochemical effects and the indirect effects through the treatment/prevention of episodes of illness likely underlie the larger grey matter volumes in Li-treated patients.

It should be noted that subdividing the patients according to the presence or absence of current Li treatment increased the signal, but the statistical heterogeneity remained moderate.\textsuperscript{340}

![Fig. 3: Comparison of hippocampal volumes in lithium (Li) and non-Li groups.](image)
in some analyses. This is not surprising considering the methodological issues still present among the evaluated studies. For example, a short duration of treatment, low dose, subtherapeutic levels of Li, poor compliance or a history of Li intoxications could decrease the effect size in comparisons involving Li-treated patients. Even a short duration of Li exposure (1–8 wk) is associated with hippocampal volume increases. We do not know how long this effect persists after the discontinuation of Li treatment. Thus, the inclusion of participants with previous Li exposure or those who stopped Li treatment only shortly before scanning could have decreased the effect size and increased the heterogeneity in studies of participants who were not currently treated with Li. In addition, smaller hippocampal volumes are unlikely to be present in the early stages of illness and tend to appear only after a substantial illness burden. Thus, differences between studies in duration of illness or number of episodes would also contribute to statistical heterogeneity.

The cumulative results from this meta-analysis have practical implications for future studies investigating hippocampal volumes in patients with bipolar disorder. A study would need to recruit 110–123 patients with bipolar disorder who were not treated with Li to detect significantly smaller hippocampal volumes relative to controls. Most individual studies investigated less than 30 individuals per group. Increasing the sample size may be one strategy for future studies. Focusing on contrasts, which showed larger effects (Li v. non-Li groups) or designing studies that decrease the heterogeneity might be a better strategy. The previously mentioned methodological issues could preferentially skew the results toward false negatives. The real positive effect of Li and the negative effect of bipolar disorder on hippocampal volumes could thus be even greater. A methodologically refined study would likely require fewer participants to detect these effects.

Limitations

Our study had several limitations. Each meta-analysis depends on the quality of the primary data, comparability of the methods and control for known confounding variables. The comparability with regard to magnetic resonance imaging methods was good. All studies used at least 1.0-T scanners, 3-mm slice thickness (at most) and 3-dimensional acquisitions. These methods are sufficient for determining hippocampal volumes. The exclusion of studies that used a 1-T scanner or a 3-mm slice thickness in the sensitivity analyses did not change the results. Furthermore, it has been shown that slice thickness in the commonly used range does not affect hippocampal volume. In keeping with this, we, and others, previously combined data obtained from a range of slice thicknesses for meta-analytical purposes. All of the analyzed studies were blind to the diagnostic and treatment status of the participants. When reported, the inter-rater reliability estimates for measuring hippocampal volumes were sufficiently high (> 0.9).

In contrast to the neuroimaging methods, several potential clinical confounders were not well controlled for in the analyzed studies and could have contributed to the residual statistical heterogeneity in some of the analyses. The absence of these clinical factors makes the interpretation of results as an effect of Li difficult, since it was unclear whether or not the patients were compliant, were treated with a sufficient dose of Li for a sufficient duration of time or whether the clinical groups were comparable in illness burden. To achieve the best coverage of the available literature, we categorized studies based on whether all or a majority, or whether no or a minority, of patients were treated with Li. The results remained mostly unchanged in subgroups of patients where all or no participants were treated with Li. Previous meta-analyses combined studies of patients with and without exposure to Li. By categorizing research participants based on the presence or absence of treatment with Li, we accounted for this important confounder as best as the available literature allowed us. It is of note that including the studies where only the majority or only the minority of patients was exposed to Li makes this conservative, as such studies would dilute the overall effect size. Indeed, when we excluded these studies, the effect size increased in most cases.

To make the results generalizable and representative of the full body of literature, the present as well as other meta-analyses included studies of pediatric patients. Excluding these studies did not change the results. Only 2 studies investigated patients with bipolar disorder treated with Li monotherapy. This is not surprising, as patients stabilized on Li monotherapy represent a minority of patients with bipolar disorder. Thus, most studies included patients treated with polypharmacy. Whereas we systematically varied the exposure to Li, we did not set any criteria for use of any other medications. Thus, it is unlikely that the results were related to effects of other compounds, which were nonsystematically used in both the Li and non-Li groups.

Meta-analytical techniques could be misled by preferential publication of positive findings. There was no evidence of a publication bias in the reviewed studies and, in fact, most of the analyzed studies reported nominally nonsignificant results. Only a single study could not be included in this meta-analysis. In keeping with the overall results, this study showed numerically larger hippocampal volumes in patients with bipolar disorder, most of whom (17 of 24) were treated with Li, than in controls.

Conclusion

When exposure to Li was minimized, patients with bipolar disorder showed smaller hippocampal volumes than controls or Li-treated patients, who had larger hippocampal volumes than controls. The preserved hippocampal volumes among patients with bipolar disorder in the majority of individual studies and all previous meta-analyses may have been related to the inclusion of Li-treated participants. To allow for a better interpretation of these effects, future studies should attempt to maximize the illness burden, as well as use the duration, doses and monitoring of Li treatment as inclusion criteria. Our findings provide indirect support for the negative effects of bipolar disorder on hippocampal volumes and are consistent with the hypothesis that Li may have neuroprotective effects.
Acknowledgements: This study was supported by funding from the Canadian Institutes of Health Research, the Nova Scotia Health Research Foundation, the Ministry of Health of the Czech Republic (MZO10CP2005) and a Dalhousie Clinical Research Scholarship to T. Hajek. The sponsors of the study had no role in the design or conduct of this study; in the collection, management, analysis and interpretation of the data; or in the preparation, review or approval of the manuscript.

Competing interests: None declared.

Contributors: All authors designed the study, analyzed the data, reviewed the article and approved its publication. T. Hajek and M. Kopecek acquired the data. T. Hajek wrote the article.

References


