Last draft of Otten & Meeter (2014 or 2015). Journal of Affective Disorders

Hippocampal Structure and Function in Individuals with Bipolar Disorder: A Systematic Review

Mara Otten*, and Martijn Meeter

Department of Cognitive Psychology, VU University Amsterdam, The Netherlands

Van der Boechorststraat 1, 1081BT, Amsterdam, The Netherlands,

*Corresponding author.

Email address: mara.otten@vu.nl

Abstract

Introduction: Bipolar disorder (BD) is a psychiatric disorder accompanied by deficits in declarative memory. Given the importance of the hippocampus in declarative memory, it is not surprising that BD patients have been reported to show hippocampal abnormalities. **Objectives**: Review evidence about structural and functional hippocampal abnormalities in BD. **Methods**: Systematic review of studies comparing BD patients and healthy controls with respect to hippocampal structure or function. **Results**: Twenty-five studies were included, together involving 1043 patients, 21 of which compared patients to controls. Decrease in hippocampal volume was found in four of eighteen studies using adult samples, and two of three samples using adolescents. Four studies revealed localized hippocampal deficits. Meta-analysis revealed a significant but small effect with lower hippocampal volumes when comparing all BD patients with controls. Lithium treatment was associated with larger hippocampal volumes across studies. The three functional studies yielded contradictory evidence. **Limitations:** Studies were only cross-sectional in nature and all used (f)MRI to investigate hippocampal volume. Heterogeneous patients groups and different methodologies for hippocampal segmentation, may have contributed to difficulties when comparing the different studies. **Conclusions:** There seems to be a small reduction in hippocampal volume in BD, which perhaps is more pronounced in early-onset BD and is counteracted by a neuroprotective effectof lithium treatment, However, how these structural abnormalities relate to functional deficits is largely unclear. .

Given the few functional neuroimaging studies and given the lack of congruence in these results, further investigation of especially hippocampal function in BD is recommended.

Bipolar disorder (BD), also called manic depressive disorder, is a psychiatric affliction in which manic and depressive episodes alternate (Goodwin & Jamison, 2007). During mania, individuals with BD experience an elevated mood, which can either be extremely happy and optimistic, or highly irritated. Individuals who experience a manic episode are very energetic, active, restless, easily distracted and impulsive, experience racing thoughts, barely need sleep and often lose sight of what is real. On the other hand, during depression, individuals with BD experience feelings such as sadness, hopelessness and worthlessness. They lose interest in activities, cannot concentrate and feel extremely tired (Goodwin & Jamison, 2007; Marohn, 2011). BD affects around 1% of the worldwide population (Belmaker, 2004; Weissman, Bland, Canino, Faravelli, Greenwald, Hwu et al., 1996) and is ranked by the World Health Organisation as the sixth leading cause of disability worldwide (Lopez & Murray, 1998). To develop and innovate both therapies as well as medication, understanding the underlying pathophysiology of this disorder is of great importance.

BD is characterized by global cognitive impairment (Bora, Yucel & Pantelis, 2008; Green, 2006; Martínez-Arán, Vieta, Reinares, Colom, Torrent, Sánchez-Moreno et al., 2004; Robinson & Ferrier, 2006). For example, deficits are frequently manifest in executive functions, verbal learning, memory, sustained attention, response inhibition and psychomotor speed (Green, 2006; Robinson, Thompson, Gallagher, Goswami, Young, Ferrier et al., 2006). These deficits are not only present during acute phases of illness, but also during remission and in the so called euthymic state, in which patients experience a normal mood which is neither depressive nor highly elevated (Bora et al., 2008; Robinson et al., 2006). Greater neuropsychological impairment is associated with a greater number of manic episodes, greater number of hospitalizations and a long-term duration of the illness (Robinson & Ferrier, 2006), and moreover contributes to poor overall daily functioning (Martínez-Arán et al., 2004).

A cognitive domain that is consistently identified as being severely impaired in BD, is declarative memory (Altshuler, Ventura, Van Gorp, Green, Theberge & Mintz, 2004; Bearden, Glahn, Monkul, Barrett, Najt, Kaur et al., 2006; Robinson et al., 2006; van Gorp, Altshuler, Theberge & Mintz, 1999). Given these large deficits in declarative memory in patients with BD (Robinson et al., 2006), and given that even first-degree unaffected relatives of patients with BD show small impairments in this domain, declarative memory is a candidate endophenotype for BD (Arts, Jabben, Krabbendam & van Os, 2008). Nevertheless, the neural basis of these impairments is still a matter of debate.

The hippocampus is an essential structure for the acquisition, consolidation and retrieval of declarative memory (Eichenbaum, 2000). Declarative memory deficits in BD might therefore be associated with abnormalitiess in this brain structure. Postmortem studies suggest that there might be abnormalities in the hippocampal structure and function (Frey, Andreazza, Nery, Martins, Quevedo, Soares et al., 2007), but structural and functional imaging studies in living individuals show less consistent findings (Brambilla, Hatch & Soares, 2008; Hall, Whalley, Marwick, McKirdy, Sussmann, Romaniuk et al., 2010; Matsuo, Walss-Bass, Nery, Nicoletti, Hatch, Frey et al., 2009). Therefore, to increase the understanding of the pathophysiology in BD, the aim of this review is to look at the structure and function of the hippocampus in BD in a systematic way. Specifically, the focus will be on differences between individuals with BD and healthy controls as regards hippocampal structure and function, while examining structural and functional magnetic resonance imaging studies.

Study Selection

To identify relevant papers, we searched through PubMed with the following search combination: “(bipolar disorder) or (manic depress*) or (bipolar) AND hippocamp* AND fMRI or MRI or (magnetic resonance) AND morphometry or memory”. We thus restricted the search to BD patient studies looking at the hippocampus either at a structural level, or with functional neuroimaging of memory-related tasks. The first and second search terms were applied to titles and abstracts while the

last two search terms were applied to the whole article. There was no limit with regard to the date of publication.

This primary search netted 48 studies. Studies were included when the following criteria were met: (i). concerned a research report (so no review papers); (ii). patients with BD were compared with healthy controls based on structural characteristics of the hippocampus or hippocampal function; (iii). when research was conducted while using more than one patient group, it had to be clear that BD patients were directly compared with a healthy control group; (iv). as regards functional imaging, the task had to measure some form of memory performances.

Fifteen studies met the inclusion criteria. In addition, articles were scanned for additional relevant studies, which resulted in the inclusion of ten additional studies. Studies were categorized based on examining the hippocampal structure or function. Nineteen studies investigated hippocampal volume in BD patients, two investigated hippocampal activity during memory tasks (Glahn et al., 2010; Whalley et al., 2009) and one study addressed both topics (Avery et al., 2013). When relevant data was available, the standardized effect size '*d*' (Cohen, 1988) was calculated. All statistical analyses were performed using IBM SPSS statistics for Windows (Version 20) and MS Excel. The funnel plot was created using MS Excel.

Results

Participant characteristics of all studies are summarized in Table 1. Participant number, sex, age, illness duration, education, disease state, medication, the percentage of lithium-taking patients, and IQ are listed when available, as well as bipolar subtype. There are different subtypes of BD, of which BD type 1 and 2 are the most common. While severe depression occurs in both subtypes, the mania symptoms are less severe in BD type 2 (hypomania) than in BD type 1 (Marohn, 2011). The 25 studies together involved 1043 BD patients (41% male, 57% female, 2% not specified) and 1513 healthy controls (51% male, 48% female, 1% not specified), with samples of BD patients ranging from 11 to 192 individuals. In three studies, samples consisted of adolescents (Bearden et al., 2008a.; Dickstein et al., 2005; Gao et al., 2013), while one study focused on older BD patients (Delaloye et al., 2009). For the 19 studies that reported it, illness duration was on average 11.8 years.

In all studies the healthy controls were matched based on demographics, except for seven studies (Altshuler et al., 2000; Haukvik et al., 2013; Haukvik et al., 2014; Killgore et al., 2009; Mathew et al., 2014; Radonic et al., 2011; Rimol et al., 2010; exact differences between groups are listed in Table 1). One study only compared the two groups statistically on the variable of age (Killgore et al., 2009), and another on no variable at all (Radonic et al., 2011). Three studies addressed the difference between patients who did or did not receive lithium treatment (Bearden et al., 2008b; Hajek et al., 2012; Van Erp et al., 2012), and one study looked at differences between psychotic and non-psychotic BD individuals (Haukvik et al., 2013).

Structural results

The results of the 23 studies addressing hippocampal volume are summarized in Table 2. In 21 studies the hippocampal volume of BD patients and healthy controls was directly compared, the two other studies (Hajek et al., 2012; Van Erp et al., 2012) differentiated within the BD group between patients treated with or without lithium and did not directly compare all BD patients with the control group. Of these 21 studies, six found significantly smaller hippocampal volume in patients with BD compared to healthy controls (Bearden et al., 2008a; Chepenik et al., 2012; Gao et al., 2013; Mathew et al., 2014; Rimol et al., 2010; Wijeratne et al., 2013), with effect sizes (Cohen, 1988) ranging from small to large, and one study (Haukvik et al., 2014) found 9 out of 14 hippocampal subfields to be smaller in BD patients compared to controls. On the other hand, one study (Javadapour et al., 2010) found significantly larger hippocampal volumes in the left (but not in the right) hemisphere of patients with

BD compared to controls, with a medium effect size. The other thirteen studies did not find any differences as regards hippocampal volume (one study additionally differentiated between psychotic and non-psychotic BD patients, but this distinction was not related to hippocampal volume; Haukvik et al., 2013). Of the thirteen studies that did not find significant effects, four showed numerically smaller hippocampal volumes in BD patients than in healthy controls (Avery et al., 2013; Brambilla et al., 2003; Delaloye et al., 2009; Haukvik et al., 2013), with small to medium effect sizes, and five (Altshuler et al., 2000; Bearden et al., 2008b; Brown et al., 2011; McDonald et al., 2006; Strakowski et al., 1999) numerically larger hippocampal volume in BD patients, with small effect sizes. The last four studies did not provide the mean values of both groups.

These results are summarized in a funnel plot in Figure 1. In this funnel plot the effect size of each study (x-axis) is plotted against the sample size (y-axis). The weighted average effect size of all studies (shown by the thick line) was $\bar{d}$ = 0.22, which differs significantly different from 0, *95%-CI* [0.134;0.309], z = 4.98, p < .001 This suggests that BD is indeed characterized by lower hippocampal volumes. To test whether there was a large between-study variance we performed a homogeneity test (Hedges & Olkin, 1985). This test was not significant, $\chi^2(14)$ = 5.41, p = .979, which means that differences between studies are probably due only to sampling errors and not to between-studies variability. Moreover, the funnel plot displayed symmetry, as shown by the linear regression test proposed by Egger et al. (1997), $t(11)$ = 0.56, p = .587, which shows there was no evidence for publication bias (Egger, Smith, Schneider & Minder, 1997).

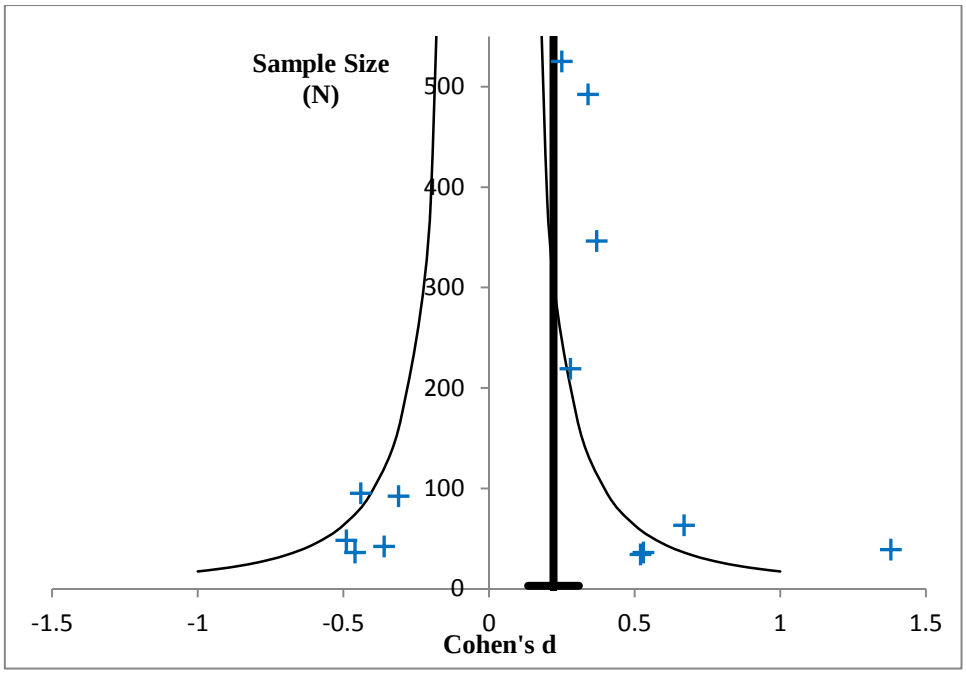

Figure 1. *Funnel plot of the observed hippocampal volume differences between patients with BD and healthy controls (effect size) and the corresponding sample sizes. A positive Cohen's d indicates larger hippocampal volume in controls. The curved vertical lines represent the border between results that differ significantly from 0 and those that do not. The thick line indicates the weighted average effect size of all studies, with the line at the x axis giving the confidence interval around this average.*

Additionally, four studies (Bearden et al., 2008a; Bearden et al., 2008b; Haukvik et al., 2014; Mathew et al., 2014) used statistical 3-D maps to investigated hippocampal subfields. Bearden et al. (2008a) found localized deficits in the head and tail of the left hippocampus in patients with BD compared with controls, which were most pronounced in the subiculum. Moreover, two recent studies

revealed a number of localized deficits in the hippocampi of patients with BD compared with controls. The first study (Mathew et al., 2014) showed reductions of the CA2/3, CA4/DG and subiculum in the right hippocampus and reduction of the CA2/3 and presubiculum in the left hippocampus. The second one (Haukvik et al., 2014) revealed reductions of the CA2/3, subiculum, CA4/DG and hippocampal formation in the left hippocampus in patients compared with controls, while reductions of the CA1, CA2/3, subiculum, CA4/DG and hippocampal formation were present in the right one. Interestingly, when comparing the volumes of the total hippocampi, both studies revealed smaller hippocampi in patients, but with very small effect sizes. This could imply that patients with BD show very localized deficits of the hippocampus, while the effect on total hippocampal volume is very modest. We will come back to this possibility in the Discussion.

Age groups

Of six studies that found a significant decrease in hippocampal volume, two sampled BD adolescents (Bearden et al., 2008a; Gao et al., 2013). Additionally, Gao et al. (2013) found a negative correlation between the hippocampal volume and the score on Young Mania Rating Scale, r = -.60, even after controlling for age, gender, illness duration and age of onset of the BD. However, a third study that also addressed BD in adolescents (Dickstein et al., 2005), did not find significant differences between patients and controls. There is little to explain these inconsistent results. All three studies included 16 to 20 patients with mostly with type 1 BD who were in a euthymic state at the moment of testing. All patients used mood stabilizers, anti-epileptics and in two out of three studies antipsychotics. However, the fact remains that two of three studies using adolescents with BD found smaller hippocampal volume compared with controls, while only four out of eighteen studies found smaller hippocampal volume in adults.

Lithium treatment

Three studies (Bearden et al., 2008b; Hajek et al., 2012; Van Erp et al., 2012) specifically looked at the effects of lithium on hippocampal volume. The first study (Bearden et al., 2008b) compared patients who were treated with lithium at least two weeks before scanning, with patients who were at least one month deprived from lithium. After controlling for total brain volume, hippocampal volume was larger in patients treated with lithium than in both patients off lithium and in healthy controls, with very large effect sizes (comparison with patients off lithium, Cohen's d = 1.1; with controls, d = 0.84). Patients without lithium treatment did not differ from healthy controls. Interestingly, additional analysis using statistical 3-D maps revealed localized deficits in the right hippocampus of patients who did not receive lithium, compared with healthy controls and patients treated with lithium. These very specific deficits were found in the lateral CA1 subfields and subiculum. The second study (Hajek et al., 2012) compared patients with a current lithium treatment lasting a minimum of 24 months with patients who had less than three months of lithium life time exposure, at least 24 months before scanning. In contrast with Bearden et al. (2008b), they found smaller left hippocampal volume in patients without lithium treatment, compared with controls. However, patients who received lithium did not differ from healthy controls, and showed larger hippocampal volume than those of patients off lithium, only on the left side and at statistical trend level. Finally, Van Erp and colleagues (2012) compared BD patients with an average dose of 900 mg lithium per day, with BD patients who were at least one year deprived from lithium. In accordance with Bearden et al. (2008b), they found larger hippocampal volume in patients treated with lithium than in healthy controls. Moreover, when patients treated with lithium were compared with their non-bipolar co-twins, again larger hippocampal volume was found in patients. Additionally, in accordance with Hajek et al. (2012), Van Erp et al. (2012) showed larger hippocampal volumes for BD patients treated with lithium compared with patients off lithium at statistical trend level.

Aside from different inclusion criteria for lithium-treated and lithium-untreated groups, there were some differences between these studies. Bearden et al. (2008b) used a combination of depressed

(n = 11) and euthymic (n = 22) patients, with an average illness duration of 14 years, and with both BD type 1 and 2 within the sample of patients. Hajek et al. (2012) only used euthymic patients (n = 29), also with both BD type 1 and 2, with a much longer average illness duration of 26 years who were also older (mean age 47 vs 34 years). Finally Van Erp et al. (2012) did not report the current disease state of the patients nor the illness duration, and they only used BD type 1 patients in their study.

Despite these differences, all studies show a positive effect of lithium treatment on hippocampal volume: two studies show larger hippocampal volumes for patients treated with lithium, compared with healthy controls, and moreover two studies show a trend towards lager hippocampal volumes for patients treated with lithium in comparison with patients off lithium. Therefore lithium seems to have a protective effect on the hippocampus in patients with BD. We additionally tested this hypothesis by looking at the correlation between the effect sizes of the different studies, and the percentage of patients taking lithium[1] within each study. We found a negative correlation, $r(16) = -.62$, $p = .011$, showing smaller effects when the percentage of patients taking lithium was greater[2]. In other words, the more patients treated with lithium, the larger the hippocampal volume of the total BD group compared with controls. Since this correlation is only based on fifteen studies, it should be interpreted with caution. However, it provides additional evidence for the protective effect of lithium on the hippocampus in patients with BD.

Relation hippocampal volume and memory performance

In three studies in which the hippocampal volume was researched, the participants additionally performed a memory test (Avery et al., 2013; Chepenik et al., 2012; Delaloye et al., 2009). Two of these related hippocampal volume to the memory performance: either by a relational memory task (Avery et al., 2013) or by the California Verbal Learning Test (Chepenik et al., 2012), a test to assess verbal memory abilities (Delis, Freeland, Kramer & Kaplan, 1988). When group (BD vs. controls) was included as covariate, Avery et al. (2013) found a positive correlation (r = .34) between memory accuracy and hippocampal volume, which explained 12% of the variance of the accuracy scores. As group was added as covariate, this positive correlation was present in both BD patients and controls. Chepenik et al. (2012) found the same positive relationships between hippocampal volume and memory performance for total immediate recall (r = .52), short delay cued recall (r = .39), long delay free recall (r = .40) and long delay cued recall (r = .44) in BD patients, though not in controls.

Both studies thus show a positive correlation between hippocampal volume and memory performances in BD patients. This suggests that memory performance in this disorder are related to the volume of the hippocampus, with smaller volumes being associated with an increase in memory difficulties.

Hemispheric differences

Seven studies researched the possibility of hemispheric differences concerning hippocampal volume. Two studies did not find differences between left and right hippocampal volumes, neither in BD patients nor in healthy controls (Chepenik et al., 2012; Killgore et al., 2009), three studies found a smaller left hippocampus than the right one in both patients and controls (Delaloye et al., 2009; Javadapour et al., 2010; Van Erp et al., 2012), one study found the left hippocampus being smaller only in controls (Avery et al., 2013), while the last study found the left hippocampus being smaller only in BD patients (Radonic et al., 2011). Taken together, most studies show evidence for a smaller

[1] When the number of patients with lithium treatment was known, we used this to calculate the correlation. If only patients using mood stabilizers was mentioned, we used the number of patients taking these in the calculation, since lithium is the mood stabilizer most often prescribed (Geddes & Miklowitz, 2013).
[2] The study of Wijeratne et al. (2013) is somewhat of an outlier. After excluding this study, the correlation coefficient was very high, $r(15) = -.80$, $p < .001$.

left hippocampus compared with the right one, and this seems to be the case in both patients and controls.

Functional results

Only three studies addressed the topic of functional activity in the hippocampus of BD patients during memory tasks (Avery et al., 2013; Glahn et al., 2010; Whalley et al., 2009). Avery et al. (2013) used a transitive inference paradigm to look at differences in hippocampal activation between BD patients and healthy controls during a relational memory task. No differences were detected. Glahn et al. (2010) let participants perform a face-name paired associate task. No differences in hippocampal activation between controls and BD patients were found during encoding. However, despite adequate task performances, BD patients showed less hippocampal activation than controls during recognition. Moreover controls showed a positive correlation between task performance and hippocampal activation during both encoding (r = .46) and recognition (r = .45), which was missing in patients with BD. In patients the relationship between hippocampal activity during recognition and task performance was even negative, (r = -.62, p < .05). However, these results must be interpreted with strong caution, as both correlations did not differ significantly between the patient and the control group. This suggests that any difference in correlation could also be due to chance. In the last study (Whalley et al., 2009), participants performed an emotional memory task in which images were shown that depicted either emotionally positive or neutral scenes. Afterwards participants had to indicate whether or not they recognized the images. Despite equal recognition performances, BD patients demonstrated greater activation for emotional versus neutral memory in the left hippocampus during recognition, compared with healthy controls. Moreover, for both groups there was a positive correlation between recognition accuracy and bilateral hippocampal activation.

It is difficult to derive unequivocal conclusions from only three functional studies. Since one study demonstrated less hippocampal activation in BD patients during recognition compared to controls, one demonstrated the opposite pattern during recognition of emotionally arousing stimuli, and the third did not find any differences in hippocampal function between both groups, no consistent pattern emerges. Conducting more functional research is necessary to see whether the functional activity of the hippocampus during memory tasks really differs between individuals with BD and controls.

4. Discussion

Declarative memory deficits are common in BD (Altshuler et al., 2004; Bearden et al., 2006; Robinson et al,. 2006; van Gorp et al., 1999). Given the important function of the hippocampus in declarative memory (Eichenbaum, 2000), hippocampal abnormalities in BD, either structural or functional, might be one cause of these memory deficits. The purpose of this study was to review evidence about hippocampal deficits in BD. A systematic review of MRI evidence indeed yielded evidence in favour of decreased hippocampal volume in BD compared with healthy controls. Only four of eighteen studies using adult samples showed significantly smaller hippocampal volumes in patients on their own, but the meta-analysis nevertheless revealed a significant but small effect with lower hippocampal volumes in BD. There was no evidence for publication bias.

Of the three studies that were carried out with adolescents, two found a decrease in hippocampal volume (Bearden et al., 2008a, Gao et al., 2013). It may thus be that hippocampal abnormalities are particularly salient in individuals with early-onset BD. Indeed, early-onset BD is often associated with worse outcomes. Perlis and colleagues (2004) for example, examined 983 BD patients of whom the age of onset could be determined, and divided them into two groups: very early-onset and early-onset. They showed that earlier onset of BD was associated with greater comorbidity for most axis I disorders, more recurrences, shorter euthymic periods, greater risk of suicide attempts and moreover they showed worse functioning and quality of life at the time of research (for

corresponding evidence see e.g. Birmaher, Axelson, Goldstein, Strober, Gill, Hunt et al., 2009). Moreover, Frazier and colleagues (2005) showed that early-onset BD is associated with overall structural abnormalities as well as global cortical gray matter deficits of greater magnitude compared with late-onset BD (Frazier, Ahn, DeJong, Bent, Breeze & Giulioano, 2005). However, early-onset is also associated with longer duration of the disorder when patients on one age group are tested, and there has been a debate whether BD can be more accurately described as a neurodevelopmental disorder, in which the brain develops on an abnormal path, or a neuroprogressive disorder, in which abnormalities increase with the duration of the disorder (Schneider, DelBello, McNamara, Strakowski, & Adler, 2012). There is evidence for both sides. For example, Lim et al. (2013) recently reviewed longitudinal neuroimaging studies, and showed that an earlier onset of BD may incate a more malignant course of illness, and is related to more brain abnormalities (although not strongly to abnormalities in the medial temporal lobe). On the other hand, Javadapour et al. (2010) showed a negative association between illness duration and hippocampal volume. Other studies do not find such a strong relation. Recently, Gildengers et al. (2014) found in a sample of older BD patients that while duration of illness was related to overall gray matter volume, this was not the case for hippocampal volume specifically. Instead, lower hippocampal volume was related to more extensive exposure to antipsychotic drugs. The findings reviewed here do not firmly support either the neurodevelopmental or the neuroprogressive view.[3] More research is necessary to clarify this ongoing debate.

Three studies found that lithium treatment was associated with increases in hippocampal volume relative to patients treated otherwise, to non-BD co-twins (Van Erp et al., 2012), or to healthy controls. Though the exact observed effects differed by study, all showed a positive effect of lithium on hippocampal volumes. In a recent meta-analysis Hajek and colleagues (2012) also found increased hippocampal volumes in lithium-treated patients compared with both lithium naïve patients and healthy controls. Lithium thus seems to have a neuroprotective effect, which is also supported by research in mice (Chen, Rajkowska, Du, Seraji-Bozorgzad & Manji, 2000), by longitudinal MRI research in BD patients (Yucel, McKinnon, Taylor, MacDonald, Alda, Young, et al., 2007) and by research in which BD patients who used different types of medication were compared (Yucel, Taylor, McKinnon, MacDonald, Alda, Young et al., 2008). This effect of lithium on hippocampal volume could provide an explanation for findings of preserved hippocampal volume in BD patients in many individual studies reviewed here. Indeed, we observed a negative association between reported effect size in each study and the percentage of patients treated with lithium.

However, lithium does not seem to have the same effects for each individual with BD. For example, Van Erp et al.(2012) showed that familial inheritance factors were related to the effect of lithium on regionally thickening of the hippocampi in BD, which suggests that this thickening may be partly due to familial factors and partly due to the effects of lithium treatment. Moreover, a study of Selek and colleagues (2013) seemed to show that there are only positive effects of lithium on hippocampal volumes in patients that respond to lithium as a treatment for BD.

While we thus found a small effect for lower hippocampal volumes in BD, this effect was absent in most individual studies. As discussed, variations in the percentage of patients taking lithium might be one factor in null effects reported in some studies. There might be other reasons as well. First, we identified four studies that found localized deficits in hippocampal subfields in BD compared with controls, accompanied by inconclusive evidence for lower total hippocampal volume of patients with BD. It is thus possible that only specific hippocampal subfields are affected in BD, though as of yet there is little consistency in reports of what subfields that would be. Moreover, only some subtypes of BD may show volumetric hippocampal deviations (Frey, Andreazza, Nery, Martins, Quevedo, Soares et al., 2007). This would imply that assessing the whole BD spectrum in one study could obscure true effects.

[3] We computed the correlation across the studies that reported both variables, of mean illness duration and mean effect size. This correlation was not significant, $r(12) = .24$, $p = .449$. Although we thus did not find evidence in favor of such a correlation between studies, our analysis does not preclude the possibility that duration and hippocampal volume are correlated within studies.

While there is thus evidence that BD is associated with a somewhat smaller hippocampal volume, the link from hippocampal abnormalities to memory dysfunction has been demonstrated less convincingly. Given the large number of studies into anatomical abnormalities, there is a surprising lack of studies linking anatomy to function. Only two studies reported positive relationships between hippocampal volume and measures of memory performances in patients with BD, and these relationships were obtained using very different measures (Avery et al., 2013; Chepenik et al., 2012). Just three studies have looked with functional imaging at how patients with BD perform memory tasks. Since one study demonstrated less hippocampal activation in BD during recognition in a face-name paired association task compared with controls, one demonstrated the opposite pattern during recognition of emotional arousing stimuli, and the third did not find difference between BD patients and controls, it is difficult to derive unequivocal conclusions based on only these three studies. Another complication factor is that the hippocampus is certainly not the only brain structure involved in memory performance. One structure that is clearly important in declarative memory is the prefrontal cortex (e.g., Blumenfeld & Ranganath, 2007), and there is evidence for prefrontal dysfunction in BD (e.g. Drevets, Price, Simpson, Todd, Reich, Vannier et al., 1997; Hall et al., 2010; López-Larson, Delbello, Zimmerman, Schwiers & Strakowski, 2002). Moreover, brain regions are part of neural circuits that could dysfunction at the circuit level. Phillips and Swartz (2014) recently investigated the functional connectivity of the hippocampus. They showed that the hippocampus contributes to a PFC-hippocampus-amygdala circuit, which is highly involved in emotion processing and regulation. This sheds light on the findings of Whalley and colleagues (2009) who demonstrated greater activation in BD compared with controls in the left hippocampus during recognition for emotional versus neutral memory. Future research will thus have to focus, more than is currently the case, on brain function and functional connectivity to uncover how structural abnormalities contribute to symptomatology in BD.

The current study has some limitations. First of all, we restricted ourselves to functional and structural magnetic resonance imaging data, leaving aside, e.g., PET imaging data. Secondly, out of the twenty-three studies addressing hippocampal volumes, only fifteen studies gave effect sizes or sufficient data to calculate them, restricting statistical power of our analyses. Thirdly, all studies were cross-sectional in nature. This makes it difficult to really look at the hippocampal structure and function as the disorder develops and more episodes of illness are passed through. Moreover, the high heterogeneity of the available studies makes it difficult to compare their results. Finally, the comparability of findings from different studies is hampered by the use of different segmentation protocols to delineate the hippocampus. Reviews of methodologies for this delineation show a wide variety of landmarks used to detect the borders of the hippocampus in MR images, which results in large volumetric differences (Geuze, Vermetten & Bremner, 2005; Konrad, Ukas, Nebel, Arolt, Toga & Narr, 2009). In the case of our review, for example, one study (Brambilla et al., 2003) used the corona radiate, the ambient cistern, the inferior horn of the lateral ventricle and the mammillary bodies as borders to delineate the hippocampus, while another (Wijteratne et al., 2013) used the hippocampal head, the crus of the fornix, the medial wall of the temporal horn, the ambient cistern, the transverse fissure, the parahippocampal gyrus, the amygdala and the choroid plexus as borders for the delineation. These methodological differences can lead to variability in structural hippocampal measures, which makes it very difficult to derive unequivocal conclusions.

Summarizing, twenty-five studies were reviewed that addressed hippocampal structure and function in patients with BD. Studies in adult samples seem consistent with a small reduction in hippocampal volume in patients relative to controls. This reduction is perhaps more pronounced in early-onset patients, although more research, especially longitudinal, is needed to determine whether BD is better seen as a neurodevelopmental or a neuroprogressive disorder. Moreover, lithium treatment seems to be related to hippocampal growth, which may be one reason why many individual studies using adult samples did not show a decreased hippocampal volume.

While the presence of structural hippocampal deviations thus seems to be established, this does not imply that they are causally related to, for example, memory dysfunction. There is a marked lack of studies into how structural abnormalities may relate to function, as well as how structural

abnormalities may affect broader functional networks of brain regions. Especially needed is thus more research into how the hippocampus functions in BD.

References

Altshuler, L. L., Bartzokis, G., Grieder, T., Curran, J., Jimenez, T., Leight, K., et al. (2000). An MRI study of temporal lobe structures in men with bipolar disorder or schizophrenia. *Biological Psychiatry, 48(2),* 148-162.

Altshuler, L. L., Ventura, J., Van Gorp, W. G., Green, M. F., Theberge D. C., & Mintz, J. (2004). Neurocognitive function in clinically stable men with bipolar I disorder or schizophrenia and normal control subjects. *Biological Psychiatry, 56,* 560-569.

Arts, B., Jabben, N., Krabbendam, L., & van Os, J. (2007). Meta-analyses of cognitive functioning in euthymic bipolar patients and their first-degree relatives. *Psychological Medicine, 38,* 771- 785.

Avery, S. N., Williams, L. E., Woolard, A. A., & Heckers, S. (2013). Relational memory and hippocampal function in psychotic bipolar disorder. *European Archives of Psychiatry and Clinical Neurosciences, 264(3),* 199-211.

Bearden, C. E., Soares, J. C., Klunder, A. D., Nicoletti, M., Dierschke, N., Hayashi, K. M., et al. (2008a). Three-dimensional mapping of hippocampal anatomy in adolescents with bipolar disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 47(5),* 515-525.

Bearden, C. E., Glahn, D. C., Monkul, E. S., Barrett, J., Najt, P., Kaur, M., et al. (2006). Sources of declarative memory impairment in bipolar disorder: mnemonic processes and clinical features. *Journal of Psychiatric Research, 40,* 47-58.

Bearden, C. E., Thompson, P. M., Dutton, R. A., Frey, B. N., Peluso, M A., Nicoletti, M., et al. (2008b). Three-dimensional mapping of hippocampal anatomy in unmedicated and lithium-treated patients with bipolar disorder. *Neuropsychopharmacology, 33(6),* 1229-1238.

Belmaker, R. H. (2004). Bipolar Disorder. *The New England Journal of Medicine, 351,* 476-486.

Blumenfeld, R. S., & Ranganath, C. (2007). Prefrontal cortex and long-term memory encoding: an integrative review of findings from neuropsychology and neuroimaging. *Neuroscientist, 13(3),* 280-291.

Birmaher, B., Axelson, D., Goldstein, B., Strober, M., Gill, M., Hunt, J., et al. (2009). Four-year longitudinal course of children and adolescents with bipolar spectrum disorders: the Course and Outcome of Bipolar Youth (COBY) study. *American Journal of Psychiatry, 166(7),* 795- 804.

Bora, E., Yucel, M., & Pantelis, C. (2008). Cognitive endophenotypes of bipolar disorder: A meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. *Journal of Affective Disorders, 13(1),* 1-20.

Brambilla, P., Harenski, K., Nicoletti, M., Sassi, R. B., Mallinger, A. G., Frank, E., et al. (2003). MRI investigation of temporal lobe structures in bipolar patients. *Journal of Psychiatric Research, 37,* 287-295.

Brambilla, P., Hatch, J. P., & Soares, J. C. (2008). Limbic changes identified by imaging in bipolar patients. *Current Psychiatry Reports, 10,* 505-509.

Brown, G. G., Lee, L. S., Strigo, I. A., Caligiuri, M. P., Meloy, M. J., & Lohr, J. (2011). Voxel-based morphometry of patients with schizophrenia or bipolar I disorder: a matched control study. *Psychiatry Research, 194(2),* 149-156.

Chen, G., Rajkowska, G., Du, F., Seraji-Bozorgzad, N., & Manji, H. K. (2000). Enhancement of hippocampal neurogenesis by lithium. *Journal of Neurochemistry, 75(4),* 1729-1734.

Chepenik, L. G., Wang, F., Spencer, L., Spann, M., Kalmar, J. H., Womer, F., et al. (2012). Structure-function associations in hippocampus in bipolar disorder. *Biological Psychology, 90(1),* 18- 22.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Delaloye, C., De Bilbao, F., Moy, G., Baudios, S., Weber, K., Campos, L., et al. (2009). Neuroanatomical and neuropsychological features of euthymic patients with bipolar disorder. *American Journal of Geriatric Psychiatry, 17(12),* 1012-1021.

Delis, D. C., Freeland, J., Kramer, J. H., & Kaplan, E. (1988). Integration clinical assessment with cognitive neuroscience: construct validation of the California Verbal Learning Test. *Journal of Consulting and Clinical Psychology, 56(1),* 123-130.

Dickstein, D. P., Milham, M. P., Nugent, A. C., Drevets, W. C., Charney, D. S., Pine, D. S., et al. (2005). Frontotemporal alterations in pediatric bipolar disorder: results of a voxel-based morphometry study. *Archives of General Psychiatry, 62(7),* 734-741.

Drevets, W. C., Price, J. L., Simpson, J. R., Todd, R. D., Reich, T., Vannier, M., & Raichle, M. E. (1997). Subgenual prefrontal cortex abnormalities in mood disorders. *Nature, 386,* 824–827.

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews Neuroscience, 1(1),* 41-50.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315,* 629–634.

Frey, B. N., Andreazza ,A. C., Nery, F. G., Martins, M. R., Quevedo, J., Soares, J. C., et al. (2007). The role of hippocampus in the pathophysiology of bipolar disorder. *Behavioral Pharmacology, 18,* 419-430.

Frazier, J. A., Ahn, M. S., DeJong, S., Bent, E. K., Breeze, J. L., & Giuliano, A. J. (2005). Magnetic resonance imaging studies is early-onset bipolar disorder: a critical review. *Harvard Review of Psychiatry, 13(3),* 125-140.

Gao, W., Jiao, Q., Qi, R., Zhong, Y., Lu, D., Xiao, Q., et al. (2013). Combined analyses of gray matter voxel-based morphometry and white matter tract-based spatial statistics in pediatric bipolar mania. *Journal of Affective Disorders, 150(1),* 70-76.

Geddes, J. R., & Miklowitz, D. J. (2013). Treatment of bipolar disorder. *The Lancet, 381(9878),* 1672-1682.

Geuze, E., Vermetten, E., & Bremner, J. D. (2004). MR-based in vivo hippocampal volumetrics: 1. Review of methodologies currently employed. *Molecular Psychiatry, 10(2),* 147-159.

Gildengers, A. G., Chung, K. H., Huang, S. H., Begley, A., Aizenstein, H. J., & Tsai, S. Y. (2014). Neuroprogressive effects of lifetime illness duration in older adults with bipolar disorder. *Bipolar Disorders, 16,* 617-623.

Glahn, D. C., Robinson, J. L., Tordesillas-Gutierrez, D., Monkul, E. S., Holmes, M. K. Green, M. J., et al. (2010). Fronto-temporal dysregulation in asymptomatic bipolar I patients: a paired associate functional MRI study. *Human Brain Mapping, 31(7),* 1041-1051.

Goodwin, F. K., & Jamison, K. R. (2007). *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression,* 2nd edition. New York: Oxford University Press.

Green, M. F. (2006). Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. *The Journal of Clinical Psychiatry, 67(10),* 3-8.

Hajek, T., Cullis, J., Novak, T., Kopecek, M., Höschl, C., Blagdon, R., et al. (2012). Hippocampal volumes in bipolar disorders: opposing effects of illness burden and lithium treatment. *Bipolar Disorders, 14(3),* 261-270.

Hajek, T., Kopecek, M., Höschl, C., & Alda, M. (2012). Smaller hippocampal volumes in patients with bipolar disorder are masked by exposure to lithium: a meta-analysis. *Journal of Psychiatry and Neuroscience, 37(5),* 333-343.

Hall, J., Whalley, H. C., Marwick, K., McKirdy, J., Sussmann, J., Romaniuk, L., et al. (2010). Hippocampal function in schizophrenia and bipolar disorder. *Psychological Medicine, 40,* 761-770.

Haukvik, U. K., McNeil, T., Lange, E. H., Melle, I., Dale, A. M., Andreassen, O. A. et al. (2013). Pre- and perinatal hypoxia associated with hippocampus/amygdala volume in bipolar disorder. *Psychological Medicine, 27,* 1-11.

Haukvik, U. K., Westlye, L. T., Mørch-Johnsen, L., Jørgensen, K. N., Lange, E. H., Dale, A. M., et al. (2014). In vivo hippocampal subfield volumes in schizophrenia and bipolar disorder. *Biological Psychiatry*. http://dx.doi.org/10.1016/j.biopsych.2014.06.020.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academi Press.

Javadapour, A., Malhi, G., Ivanovski, B., Chen, X., Wen, W., & Sachdev, P. (2010). Hippocampal volumes in adults with bipolar disorder. *The Journal of Neuropsychiatry and Clinical Neurosciences, 22(1),* 55-62.

Killgore, W. D., Rosso, I. M., Gruber, S. A., & Yurgelun-Todd, D. A. (2009). Amygdala volume and verbal memory performance in schizophrenia and bipolar disorder. *Cognitive and Behavioral Neurology, 22(1),* 28-37.

Konrad, C., Ukas, T., Nebel, C., Arolt, V., Toga, A. W., & Narr, K. L. (2009). Defining the human hippocampus in cerebral magnetic resonance images—an overview of current segmentation protocols. *Neuroimage, 47(4),* 1185-1195.

Lim, C. S., Baldessarini, R. J., Vieta, E., Yucel, M., Bora, E., & Sim, K. (2013). Longitudinal neuroimaging and neuropsychological changes in bipolar disorder patients: review of the evidence. *Neuroscience & Biobehavioral Reviews, 37(3),* 418-435.

Lopez, A. D., & Murray, C. C. (1998). The global burden of disease, 1990-2020. *Nature Medicine, 4(11),* 1241-1243.

López-Larson, M. P., Delbello, M. P., Zimmerman, M. E., Schwiers, M. L., & Strakowski, S. M. (2002). Regional prefrontal gray and white matter abnormalities in bipolar disorder. *Biological Psychiatry, 52(2),* 93-100.

Marohn, S. (2011). *Natural Medicine Guide to Bipolar Disorder*. Charlottesville, Virginia: Hampton Roads Publishing Company.

Martínez-Arán, A., Vieta, E., Reinares, M., Colom, F., Torrent, C., Sánchez-Moreno, J. et al. (2004). Cognitive function across manic or hypomanic, depressed, and euthymic states in bipolar disorder. *American Journal of Psychiatry, 161,* 262-270.

Mathew, I., Gardin, T. M., Tandon, N., Eack, S., Francis, A. N., Seidman, L. J., et al. (2014). Medial Temporal Lobe Structures and Hippocampal Subfields in Psychotic Disorders: Findings From the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) Study. *JAMA Psychiatry, 71(7),* 769-777.

Matsuo, K., Walss,-Bass, C., Nery, F. G., Nicoletti, M. A., Hatch, J. P., Frey, B. N., et al. (2009). Neuronal correlated of brain-derived neurotrophic factor Val66Met polymorphism and morphometric abnormalities in bipolar disorder. *Neuropsychopharmacology, 34,* 1904-1913.

McDonald, C., Marshall, N., Sham, P. C., Bullmore, E. T., Schulze, K., Chapple, B., et al. (2006). Regional brain morphometry in patients with schizophrenia or bipolar disorder and their unaffected relatives. *American Journal of Psychiatry, 163,* 478-487.

Perlis, R. H., Miyahara, S., Marangell, L. B., Wisniewski, S. R., Ostacher, M., DelBello, M. P., et al. (2004). Long-term implications of early onset in bipolar disorder: data from the first 1000 participants in the systematic treatment enhancement program for bipolar disorder (STEP- BD). *Biological Psychiatry, 55(9),* 875-881.

Phillips, M. L., & Swartz, H. A. (2014). A critical appraisal of neuroimaging studies of bipolar disorder: toward a new conceptualization of underlying neural circuitry and a road map for future research. *American Journal of Psychiatry,171,* 829-843.

Radonić, E., Radoš, M., Kalember, P., Bajs-Janović, M., Folnegović-Šmac, V., & Henigsberg, N. (2011). Comparison of hippocampal volumes in schizophrenia, schizoaffective and bipolar disorder. *Collegium Antropologicum, 35,* 249-252.

Rimol, L. M., Hartberg, C. B., Nesvåg, R., Fennema-Notestine, C., Hagler Jr, D. J., Pung, C. J., et al. (2010). Cortical thickness and subcortical volumes in schizophrenia and bipolar disorder. *Biological Psychiatry, 68(1),* 41-50.

Robinson, L. J., & Ferrier, I. N. (2006). Evolution of cognitive impairment in bipolar disorder: a systematic review of cross-sectional evidence. *Bipolar Disorders, 8,* 103-116.

Robinson, L. J., Thompson, J. M., Gallagher, P., Goswami, U., Young, A. H., Ferrier, I. N., et al. (2006). A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *Journal of Affective Disorders, 93,* 105-115.

Selek, S., Nicoletti, M., Zunta-Soares, G. B., Hatch, J. P., Nery, F. G., Matsuo, K., et al. (2013). A longitudinal study of fronto-limbic brain structures in patients with bipolar I disorder during lithium treatment. *Journal of Affective Disorders, 150,* 629–633.

Schneider, M. R., DelBello, M. P., McNamara, R. K., Strakowski, S. M., & Adler, C. M. (2012). Neuroprogression in bipolar disorder. *Bipolar Disorders, 14(4),* 356-374.

Strakowski, S. M., DelBello, M. P., Sax, K. W., Zimmerman, M. E., Shear, P. K., Hawkins, J. M., et al. (1999). Brain magnetic resonance imaging of structural abnormalities in bipolar disorder. *Archives of General Psychiatry, 56(3),* 254-260.

Van Erp, T. G., Thompson, P. M., Kieseppä, T., Bearden, C. E., Marino, A. C., Hoftman, G. D., et al. (2012). Hippocampal morphology in lithium and non-lithium-treated bipolar I disorder patients, non-bipolar co-twins, and control twins. *Human Brain Mapping, 33(3),* 501-510.

Van Gorp, W. G., Altshuler, L., Theberge, D. C., & Mintz, J. (1999). Declarative and procedural memory in bipolar disorder. *Biological Psychiatry, 46,* 525-531.

Watson. D. R., Anderson, J. M. E., Bai, F., Barrett, S. L., McGinnity, T. M., Mulholland, C. C., et al. (2012). A voxel based morphometry study investigating brain structural changes in first episode psychosis. *Behavioral Brain Research, 227,* 91-99.

Weissman, M. M., Bland, R. C., Canino, G. J., Faravelli, C., Greenwald, S., Hwu, H., et al. (1996). Cross-national epidemiology of major depression and bipolar disorder. *The Journal of the American Medical Association, 276(4),* 293-299.

Whalley, H. C., McKirdy, J., Romaniuk, L., Sussmann, J., Johnstone, E. C., Wan, H. I., et al. (2009). Functional imaging of emotional memory in bipolar disorder and schizophrenia. *Bipolar Disorders, 11(8),* 840-856.

Wijeratne, C., Sachdev, S., Wen, W., Piguet, O., Lipnicki, D. M., Malhi, G. S., et al. (2013). Hippocampal and amygdala volumes in an older bipolar disorder sample. *International Psychogeriatrics, 25(1),* 54-60.

Yucel, K., McKinnon, M. C., Taylor, V. H., MacDonald, K., Alda., Young, L. T., et al. (2007). Bilateral hippocampal volume increases after long-term lithium treatment in patients with bipolar disorder: a longitudinal MRI study. *Psychopharmacology, 195,* 357-367.

Yucel, K., Taylor, V. H., McKinnon, M. C., MacDonald, K., Alda, M., Young, L. T., et al. (2008). Bilateral hippocampal volume increase in patients with bipolar disorder and short-term lithium treatment. *Neuropsychopharmacology, 33,* 361-367.

Table 1

Participant characteristics

Study	Controls	BD patients									
	N (M/F); Matched?	N (M/F)	Age M ± SD (years)	Duration illness M ± SD (years)	Education M ± SD (years)	BD type 1/2 (N)	Medication (N)	% BD taking lithium (li)	IQ M ± SD	State (N)	Structural (S) or functional (F) research
Altshuler et al. (2000)	18 (18/0) Not Matched	24 (24/0)	50.2±12.7	23.6±11.4	15.5±2.5	NS	Antipsychotics (6), mood stabilizers (17), anticonvulsants (8), antidepressants (7), benzodiazepines (3)	70.8%	NS	Euthymic (24)	S
Avery et al. (2013)	22 (11/11) Matched	17 (10/7)	37.41± 10.37	15.90± 8.69	13.41± 2.18	Type 1 (17)	Antipsychotic + antidepressant + mood stabilizers (10, li = 3), antipsychotics (4), antidepressant (3)	58.8%	Verbal IQ 111.35±6.75	Mild depressive(5), euthymic (12)	S + F
Bearden et al. (2008a)	20 (11/9) Matched	16 (8/8)	15.5±3.4	3.92±2.4	NS	Type 1 (12), type 2 (3), NOS (1)	Mood stabilizers + antiepileptics (4), mood stabilizers (6), antiepileptics (4), thyroid hormone (3), additional medication (1)	37.5%	NS	Depressed (2), euthymic (14)	S
Bearden et al. (2008b)**	62 (33/29) Matched	Li+: 21(12/9) Li-: 12(4/8)	Li+: 32.9±11.4 Li-: 36.5±10.4	Li+: 13.7±8.8 Li-: 14.9±7.8	Li+: 14.6±2.6 Li-: 15.4±3.4	Li+: Type 1 (19), type 2 (2) Li-: Type 1 (7), type 2 (5)	Li+: all lithium for at least 2 weeks; antidepressants (2), thyroid hormone (1), antidepressants + thyroid hormone (1) Li-: no additional medication, at least one month off of lithium	63.6%	NS	Li+: depressed (4), euthymic (17) Li-: depressed (7), euthymic (5)	S
Brambilla et al. (2003)	36 (22/14) Matched	24 (15/9)	35±10	15±9	NS	Type 1 (18), type 2 (6)	Mood stabilizers (15); other 9 patients off psychotropic drugs for at least 2 weeks	62.5%	NS	Euthymic (13), depressed (10), hypomanic (1)	S

Brown et al. (2011)	21 (10/11) Matched	15(7/8)	46.2±10.6	18.9±6.9	15.5±2.0	Type 1 (15)	Antipsychotics (3), mood stabilizers (11), antidepressants (9), benzodiazepines (3)	73.3%	NS	Acute mood episode, manic or depressed (15)	S
Chepenik et al. (2012)	32 (10/22) Matched	31 (13/18)	33.2±11.7	19±11	14.8±1.7	NS	Mood stabilizers (8), antiepileptics (15), atypical antipsychotics (12), antidepressant (8), benzodiazepines (6), stimulant (2)	25.8%	WASI IQ 114.1±13.9	Depressed (7), manic/hypomanic/ mixed (7), euthymic (17)	S
Delaloye et al. (2009)	17 (NS) Matched	17 (NS)	69.00±5.85	29.65±15.7	13.12± 3.72	Type 1 (9), type 2 (8)	Mood stabilizers (14, li = 4), atypical antipsychotics (3), antidepressants (2), benzodiazepines (2)	23.5%	NS	Euthymic (17)	S

* N = number, M = male, F = female, NS = not specified, NOS = not otherwise specified, li = lithium, WASI = Wechsler Abbreviated Scale of Intelligence.

** Li+ = patients treated with lithium at least two weeks before scanning (mean duration treatment 123±226 weeks), Li- = patients at least one month deprived from lithium.

Table 1 (continuation)

Participant characteristics

	Controls	**BD patients**									
Study	**N (M/F); Matched?**	**N (M/F)**	**Age M ± SD (years)**	**Duration illness M ± SD (years)**	**Education M ± SD (years)**	**BD type 1/2 (N)**	**Medication (N)**	**% BD taking lithium (li)**	**IQ M ± SD**	**State (N)**	**Structural (S) or functional (F) research**
Dickstein et al. (2005)	20 (13/7) Matched	20(13/7)	13.4±2.5	2.3 ± NS	NS	Type 1(15), type 2(5)	Mood stabilizers (10), antiepileptics (19), antidepressant (4), atypical antipsychotics (13), psychostimulants(4), benzodiazepines (3), thyroid hormone (5), other(4)	50.0%	109±13.6	"On average euthymic" (20)	S

Gao et al. (2013)	18 (6/12) Matched	18(6/12)	15.1±1.81	1.3±1.1	8.17±1.69	Type 1 (14), type 2 (4)	Mood stabilizers (7), antiepileptics (8), antipsychotics (13), antidepressant (3)	38.9%	98.5±13.5	Psychotic (9), non-psychotic (9)	S
Glahn et al. (2010)	24 (10/14) Matched	15 (5/10)	38.00±13.1	7.64±7.4	14.57±2.3	Type 1 (15)	Mood stabilizers (11, li = 1), antidepressants (5), atypical antipsychotics (7)	6.7%	108.21±7.8	Remitted/ euthymic (15)	F
Hajek et al. (2012)**	11 (3/8) Matched	Li+: 17(5/12) Li-: 12(6/6)	Li+: 47.8±10.1 Li-: 45.6±8.9	Li+: 27.1±8.2 Li-: 25.6±9.8	Li+: 14.47±4.27 Li-: 15.67±2.53	Li+: Type 1 (11), type 2 (6) Li-: Type 1 (9), type 2 (3)	Li+: antiepileptics (4), antidepressants (7), antipsychotics (2), mood stabilizers (minimum 24 months) Li-: antiepileptics (9), antidepressants (6), antipsychotics (4)	58.6%	NS	Euthymic (29)	S
Haukvik et al. (2013) ***	140 (74/66) Not matched, difference in age	pBD: 48(23/25) npBD: 31(8/23)	pBD: 29.9±6.7 npBD: 28.7±4.8	NS	pBD: 13.6±2.3 npBD: 13.6±2.2	Type 1 (47), type 2 (28), NOS (4)	pBD: antipsychotics (30), antiepileptics (21), antidepressants (20), mood stabilizers (8); npBD: antipsychotics (4), antiepileptics (17), antidepressants (13), mood stabilizers (4)	15.2%	pBD: 106.2±4.1 npBD: 106.6±3.2	Psychotic (48), non-psychotic (31)	S
Haukvik et al. (2014)	300 (158/142) Not matched, differences in sex, years education, IQ	192 (77/115)	35.1±11.5	Age onset 27.3±10.1	13.5±2.3	Type 1 (117), type 2 (66), NOS (9)	Mood stabilizers (33), antipsychotics (103), antiepileptics (87)	17.2%	WASI: 109.1±12.0	NS	S
Javadapour et al. (2010)	24 (6/18) Matched	24(6/18)	38.2±11.0	14.17 ±10.26	14.54±2.75	Type 1 (24)	Mood stabilizers (8), mood stabilizers and antiepileptics (4), antiepileptics (8)	50.0%	NS	Psychotic (14), non-psychotic (10)	S

* N = number, M = male, F = female, NS = not specified, NOS = not otherwise specified, li = lithium, WASI = Wechsler Abbreviated Scale of Intelligence.

** Li+ = patients treated with lithium for minimum of 24 months, Li- = patients at least 24 months deprived from lithium, maximum three months of lifetime exposure.

*** pBD = psychotic BD; npBD = non-psychotic BD.

Table 1 (continuation)

Participant characteristics

Study	Controls	BD patients									
	N (M/F); Matched?	**N (M/F)**	**Age M ± SD (years)**	**Duration illness M ± SD (years)**	**Education M ± SD (years)**	**BD type 1/2 (N)**	**Medication (N)**	**% BD taking lithium (li)**	**IQ M ± SD**	**State (N)**	**Structural (S) or functional (F) research**
Killgore et al. (2009)	20 (18/2) Not matched	11 (9/2)	23.7±3.6	3.0±3.0	NS	NS	Stabilized on medication for at least one week before study	DNA	NS	Acute mood episode, manic or depressed (11)	S
Mathew et al. (2014)	337 (152/185) Not matched, differences in sex, race, site	188 (57/131)	36.1±NS	NS	NS	NS	Mood stabilizers (53), antipsychotics (66)	28.2%	NS	Psychotic (188)	S
McDonald et al. (2006)	54 (25/29) Matched	38 (15/23)	41.0±11.7	NS	14.1±3.3	Type 1 (38)	Mood stabilizer (33), antipsychotics (11)	86.8%	NS	NS	S
Radonic et al. (2011)	15 (6/9) Not matched	15 (5/10)	54±4.8	9±4.9	NS	NS	NS	DNA	NS	Manic (2), depressed (3), euthymic (10)	S

Rimol et al. (2010)	207 (108/99) Not matched, differences in sex, ethnicity, IQ	139(54/85)	35.4±11.3	6.5±6.5	14.2±3.0	Type 1 (87), type 2 (52)	Antipsychotics (60), mood stabilizers (19), antiepileptics (51), antidepressants (48), sedatives (13)\	13.7%	WASI: 109.3±11.8	NS	S
Strakowski et al. (1999)	22 (13/9) Matched	24 (17/7)	27±6	6±6	13±2	"Manic" (14), "mixed" (10	Antipsychotics (17), mood stabilizers (24)	100%	NS	Psychotic (20), non-psychotic (4)	S
Van Erp et al. (2012)**	32 (17/15) Matched	Li+: 10(3/7) Li-: 8(5/3)	Li+: 45.60±8.69 Li-: 42.5±6.76	NS	Li+: 4.30±2.36 Li-: 3.63±2.97	Type 1 (18)	Li+: mood stabilizers (10), antidepressants (9) Li-: antidepressants (3)	55.6%	NS	NS	S
Watson et al. (2012)	24 (8/16) Matched	24(8/16)	36.0±10.0	First episode BD	NS	NS	Atypical antipsychotics (19), mood stabilizers (2)	8,3%	WAIS: 106±9 NART: 104±16	Manic/hypomanic (24)	S

* N = number, M = male, F = female, NS = not specified, NOS = not otherwise specified, li = lithium, WASI = Wechsler Abbreviated Scale of Intelligence, WAIS = Wechsler Adult Intelligence Scale.

** Li+ : patients treated with lithium, dose (mg/day) : 900 (600-1200), Li- = patients at least one year deprived from lithium.

Table 1 (continuation)

Participant characteristics

Study	Controls	BD patients									Structural (S) or functional (F) research
	N (M/F); Matched?	N (M/F)	Age M ± SD (years)	Duration illness M ± SD (years)	Education M ± SD (years)	BD type 1/2 (N)	Medication (N)	% BD taking lithium (li)	IQ M ± SD	State (N)	
Whalley et al.(2009)	14 (10/4) Matched	14(9/5)	41.5±10.0	NS	NS	Type 1 (14)	Antipsychotics (8), mood stabilizers (4), antidepressants (3)	28.6%	NART: 110.9±9.0	Manic (3), depressed (4)	F
Wijeratne et al. (2013)	21 (13/8) Matched	18(6/12)	57.0±9.9	27.8 ±12.5	13.1±3.9	Type 1 (18)	Mood stabilizers (15, li = 9)	50.0%	NS	Euthymic (18)	S

* N = number, M = male, F = female, NS = not specified, NOS = not otherwise specified, li = lithium, NART = National Adult Reading Test.

Table 2

Hippocampal volume (HV) measured by MRI. Positive effect sizes indicate a larger hippocampus in HC vs BD or Li- patients and in Li+ vs HC or Li- patients

Study	Memory test?**	HV – M(SD)	Results/measure	Effect size	Main conclusion	Extra conclusions
Altshuler et al. (2000)		BD: 4577.7 (782) mm^3 HC: 4320.3 (640) mm^3	H volume BD = HC (p = .59)	d = -0.36	No differences in hippocampal volume between BD and HC.	
Avery et al. (2013)	Transitive inference (TI) paradigm	BD(LH/RH): 1664/1677 ml HC (LH/RH): 1665/1761 ml	H volume BD = HC (p > .05), for both LH and RH	DNA	No differences in hippocampal volume between BD and HC (LH and RH).	- LH < RH in HC but not in BD. - H volume positively correlated with accuracy on TI paradigm (r^2 = .12).
Bearden et al. (2008a)		DNA	When controlling for total brain volume: - H adolescents BD < HC (p = .03); 9.2% smaller. - LH: BD < HC (p = .03) - RH: BD = HC (p = .06)	total H: d = 0.53	Total H volume smaller in adolescents with BD than HC. LH significantly smaller, but not the RH.	- Localized deficits in the head and tail of the LH in BD compared with HC (most pronounced in the subiculum).
Bearden et al. (2008b)***		DNA	When controlling for total brain volume, H volume: - BD > HC (p = .002) - Li+ > HC (both RH, p = .002 and LH, p = .005) - Li+ >Li- (both RH, p= .005 and LH, p = .01)	BD vs HC: d = -0.44 Li+ vs HC: d = 0.84 Li+ vs Li-: d = 1.1	H volume significantly larger in BD patients than in HC. Patients with lithium: significantly larger H volume compared to HC and unmedicated patients (both LH and RH). Unmedicated patients did not differ from HC.	- Statistical 3-D maps revealed very localized deficits in the RH of Li- patients as compared to HC (p = .01) and Li+ (p = .03; most prominent in lateral CA1 subfields and subiculum). - Same results after excluding patients with additional medication or BD type 2 patients. - No correlations between H volume and age onset, episodes or duration illness.

		- Li- = HC (p = .42)	Li- vs HC: d = 0.26	
Brambilla et al. (2003)	BD(LH/RH): 3.93(0.68) / 3.91 (0.71) ml HC(LH/RH): 4.04(0.60) / 3.85(0.54) ml	When controlling for age, gender and ICV, H volume: - LH: BD = HC (p = .07) - RH: BD = HC (p = .63)	LH/RH/total d = 0.17 d = -0.10 d = 0.04	No differences in hippocampal volume between BD and HC (LH and RH).
Brown et al. (2011)	BD: 250(30) mm^3 HC: 230(70) mm^3	H volume BD = HC (p = .244)	d = -0.46	No differences in hippocampal volume between BD and HC.

* HV = hippocampal volume; DNA = data not available; HC = healthy controls; H = hippocampus, LH = left hippocampus, RH = right hippocampus.
** Transitive inference (TI) paradigm: make inferential judgements on novel stimulus pairings based on previously learned relationships.

*** Li+ = patients treated with lithium at least two weeks before scanning (mean duration treatment 123±226 weeks), Li- = patients at least one month deprived from lithium.

Table 2 (continuation)

Hippocampal volume (HV) measured by MRI. Positive effect sizes indicate a larger hippocampus in HC vs BD or Li- patients and in Li+ vs HC or Li- patients

Study	**Memory test?****	**HV – M(SD)**	**Results/measure**	**Effect size**	**Main conclusion**	**Extra conclusions**

Chepenik et al. (2012)	CVLT	BD: ls mean 2588 (SE 43) HC: ls mean 2748 (SE 42)	H volumes BD < HC (p = .010)	d = 0.67	H volume significantly smaller in BD patients than in HC.	- Lower CVLT performance in BD. - Positive correlation between CLVT performances and H volume in BD. - Rapid cycling BD larger H volume. No other factor associated with H volume (duration, hospitalizations, mood state, medication etc). - No differences volume LH and RH in BD.
Delaloye et al. (2009)	CR48 + CERAD Word List Memory Test	BD: 3.42(0.55) HC: 3.69(0.49) (mean normalized volumes: [volume H (mm^3) / intracranial volume (mm^3)] x 1000)	H volume BD = HC (p = .13)	d = 0.52	No differences in hippocampal volume between BD and HC in elderly BD patients.	- Older euthymic BD patients scored lower on CR48 and lower on delayed recall on CERAD Word List Memory Test. - Volume RH > LH, for both BD patients and HC.
Dickstein et al. (2005)		DNA	When controlling for total gray matter volume, H volume BD = HC (p > .05)	DNA	No differences in hippocampal volume between BD adolescents and HC.	
Gao et al. (2013)		DNA	LH volume BD < HC (p = .014)	DNA	LH volume significantly smaller in BD adolescents than in HC.	- Negative correlation between hippocampal volume and YMRS (Young Mania Rating Scale) score, r = -.60, p = .029, even when controlled for age, gender, illness duration, and the onset age in adolescents with BD.
Hajek et al. (2012)***		DNA	LH volume Li- < HC (p = .015) LH volume Li - = Li+ (p = .087) H volume Li + = HC (p > .05)	DNA	BD without lithium: LH volume significantly smaller than HC. No differences between BD patients treated with lithium and HC.	

Haukvik et al. (2013)****	BD (LH/RH): 4038(426) / 4144 (403) mm^2 HC (LH/RH): 4150(391) / 4259 (419) mm^2	When controlled for age, sex, intracranial volume and lithium use: H volume BD(pBD + npBD) = HC, p = .221; for LH (p = .222) and RH (p = .388)	d = 0.28	No differences in hippocampal volume between BD and HC, also not when differentiated between pBD and npBD.

* HV = hippocampal volume; DNA = data not available; HC = healthy controls; H = hippocampus, LH = left hippocampus, RH = right hippocampus.
**CR48: Cued Recall 48 items test, memory test; CERAD Word List Memory Test = memory test assessing recall and recognition; CLVT: California Verbal Learning Test to assess verbal learning and memory.
*** Li+ = patients treated with lithium for minimum of 24 months, Li- = patients at least 24 months deprived from lithium, maximum three months of lifetime exposure.
**** pBD = psychotic BD; npBD = Non-psychotic BD.

Table 2 (continuation)

Hippocampal volume(HV) measured by MRI. Positive effect sizes indicate a larger hippocampus in HC vs BD or Li- patients and in Li+ vs HC or Li- patients

Study	HV – M(SD)	Results/measure	Effect size	Main conclusion	Extra conclusions
Haukvik et al. (2014)	BD (LH/RH)**: 3028(67.3) / 3076(65.8) ml HC (LH/RH): 3154(68.0) / 3218 (70.5)ml	BD < HC: LH: CA2/3 (p = .004), subiculum (p = .034), CA4/DG (p = .003), hippocampal formation (p = .001) RH: CA1 (p = .009), CA2/3 p = .001), subiculum (p = .003), CA4/DG (p = .002), hippocampal formation (p < .001) BD = HC: LH: Presubiculum, CA1, Fimbria (p > .05) RH: Presubiculum, Fimbria (p > .05)	d = 0.29	LH: volume of CA2/3, subiculum, CA4/DG and hippocampal formation smaller in BD than in HC. RH: volume of CA1, CA2/3, subiculum, CA4/DG and hippocampal formation smaller in BD than in HC.	
Javadapour et al. (2010)	BD (LH/RH): 2.52(0.26) / 2.65(0.28)	After controlling for intracranial volume, LH volume BD > HC (p = .015)	LH/RH/total d = - 0.68	LH volume significantly larger in BD adolescents than in HC. No differences for RH volume.	- RH significantly larger than LH, for both groups. - No differences in HV between psychotic/non-

	HC (LH/RH): 2.34(0.27) / 2.56(0.31)	RH volume BD = HC (p = .346)	d = - 0.30 d = - 0.49		psychotic patients or between lithium treated and untreated patients. - Episodes and duration are negatively associated with hippocampal volume.
Killgore et al. (2009)	DNA	H volume BD = HC (p > .05)	DNA	No differences in hippocampal volume between BD and HC.	- No significant difference between volume LH and RH, for both BD and HC.
Mathew et al. (2014)	BD (LH/RH): 3986.1(431.9) / 3995.6(430.5) mm^3 HC (LH/RH): 4070.4(327.7) / 4124.6(326.3) mm^3	When controlling for sex, race and site: LH volume BD < HC (p = .01) RH volume BD < HC (p <.001)	LH/RH/total d = 0.19 d = 0.30 d = 0.25	H volume significantly smaller in BD patients than in HC, both for LH and RH.	- When looking at subfields (CA1, CA2/3, CA4/DG, presubiculum, subiculum): smaller volumes for BD compared with controls for CA2/3 (bilateral), presubiculum (LH) and CA4/DG and subiculum (RH).
McDonald et al. (2006)	BD (LH/RH): 2.64(0.30) / 2.69(0.29) ml HC (LH/RH): 2.54(0.28) / 2.61(0.28) ml	When controlled for age, gender, height, handedness and lifetime alcohol or substance abuse: H volume BD = HC, for both LH (p = .140) and RH (p = .240)	d = -0.31	No differences in hippocampal volume between BD and HC, for both LH and RH.	
Radonic et al. (2011)	DNA	H volume BD = HC, for both LH (p = .123) and RH (p = .881)	DNA	No differences in hippocampal volume between BD and HC, for both LH and RH.	- LH volume significantly smaller than RH volume in BD, but not in HC.

* HV = hippocampal volume; DNA = data not available; HC = healthy controls; H = hippocampus, LH = left hippocampus, RH = right hippocampus.
** Data about HV was only available for six subdivisions of the LH and RH. Total volume was computed by summation of these volumes.

Table 2 (continuation)

Hippocampal volume(HV) measured by MRI. Positive effect sizes indicate a larger hippocampus in HC vs BD or Li- patients and in Li+ vs HC or Li- patients

Study	HV – M(SD)	Results/measure	Effect size	Main conclusion	Extra conclusions
Rimol et al. (2010)	DNA	H volumes BD < HC (p = .002)	d = 0.37	H volume significantly smaller in BD patients than in HC.	
Strakowski et al. (1999)	BD(LH/RH/total): 4.3(0.6) / 4.3(0.6) / 8.6(1.2) cm^3 HC(LH/RH/total): 4.2(0.4) / 4.2(0.4) / 8.4(0.8) cm^3	H volume BD = HC (p > .05)	LH/RH/total d = -0.20 d = -0.20 d = -0.20	No differences in hippocampal volume between BD and HC.	
Van Erp et al. (2012)**	Li+ (LH/RH): 3828(221) / 3877(272) ml Li- (LH/RH): 3436(196) / 3570(293) ml HC(LH/RH): 3534(396) / 3619(347) ml	H volume Li+ > HC (p = .005) H volume Li+ = Li- (p = .06; trend) H volume Li+ > non-BD co-twin (p = .003)	LH/RH Li+ vs Li-: d = 1.88 / d = 1.09 Li+ vs HC: d = 0.92 / d = 0.83 Li- vs HC: d = 0.15/ d = 0.31	H volume of lithium treated BD patients larger than HC and non-BD co-twins of BD patients; trend towards larger H volume than lithium naive patients.	- Across groups: RH > LH. - Compared with HC, non-BD co-twins showed right regional hippocampal thickening, which was partially overlapping with the lithium-associated regional thickening in BD probands. Suggests that familial factors could play a role in regionally thickened hippocampi in BD.
Watson et al. (2012)	DNA	H volume BD = HC (p > .05)	DNA	No differences in hippocampal volume between BD and HC.	- No difference in total brain volume between BD and HC.
Wijeratne et al. (2013)	BD(LH/RH/total): 2411(269) / 2445(356) / 4856(588)	H volume BD < HC (p = .001), for both LH (p = .001) and RH (p = .002)	d = 1.38	H volume significantly smaller in BD patients than	After controlling for intracranial volume, sex, age and years of education, left and total hippocampal volumes were negatively associated with total duration of

	mm^3	in HC, both for LH and RH.	manic, depressive and all mood episodes.
	HC(LH/RH/total)		Patients with or without current lithium treatment did not differ as regards hippocampal volume.
	2856(353) / 2911(406) / 5767 (727) mm^3		

* HV = hippocampal volume; DNA = data not available; HC = healthy controls; H = hippocampus, LH = left hippocampus, RH = right hippocampus.
** Li+ : patients treated with lithium, dose (mg/day) : 900 (600-1200), Li- = patients at least one year deprived from lithium.