Brain structure in bipolar disorder

A longitudinal neuroimaging study in twins

Florian Bootsman

Funding for this study was provided by the Netherlands Organization for Scientific Research (NWO, ZonMW TOP grant no. 9120818).

ISBN	978-94-028-0183-5
Cover image and design	Florian Bootsman
Layout	Ilse Stronks, persoonlijkproefschrift.nl
Printing	Ipskamp Printing, Enschede

Brain structure in bipolar disorder
A longitudinal neuroimaging study in twins

Hersenstructuur bij de bipolaire stoornis
Een longitudinaal neurobeeldvormend onderzoek bij tweelingen
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op dinsdag 31 mei 2016 des ochtends te 10.30 uur

door

Florian Bootsman
geboren op 1 juli 1982
te Zeist

Promotoren:	Prof.dr. R.S. Kahn
	Prof.dr. W.A. Nolen
Copromotoren:	Dr. N.E.M. van Haren
	Dr. R.M. Brouwer

Voor mijn ouders

Voor Frodo

"Unfathomable mind, now beacon, now sea"

Samuel Beckett (Molloy, 1951)

TABLE OF CONTENTS

Chapter 1

Introduction

"Numberless as the sands of the sea are the passions of man, and none is like any other, and all of them, the base and the beautiful, in the beginning are submissive to man, and only later do they establish a terrible tyranny over him"

Nikolai Gogol (Dead Souls, 1842)

1.1 BIPOLAR DISORDER, A SEVERE MENTAL ILLNESS

1.1.1 A short history of nosology

The first documented diagnosis of the disease that nowadays is referred to as bipolar disorder (BD) dates back to 1851, when the French psychiatrist Jean-Pierre Falret described what he termed *la folie circulaire* (circular insanity), a mental condition in which the patient experiences alternating periods of manic excitement and depression. Contrary to popular belief at the time that psychiatric symptoms may be different manifestations of one single disease entity ('mental alienation'), Falret asserted that such manifestations should be viewed as separate entities or syndromes (Falret, 1851; Pichot, 2004). Around 1900, Emil Kraepelin, a German psychiatrist, suggested a clinical dichotomy separating two diseases that both featured psychotic manifestations, *dementia praecox* (redefined as schizophrenia by Eugene Bleuler in 1908) and *manic-depressive madness,* the main difference between the two being that the former supposedly progresses to a final state of mental deterioration whereas the latter does not. *La folie circulaire* was incorporated in the latter disease (Kraepelin, 1899; Pichot, 2004; Akiskal, 2006). In 1903 Carl Jung was the first to distinguish between manic-depressive states involving psychosis and those not involving psychosis (Jung, 1903). The term 'bipolar disorder' first appeared in the third edition of the Diagnostic and Statistical Manual (DSM) for mental disorders in 1980, replacing the term 'manic-depressive disorder'. Here, the distinction between unipolar and bipolar disorder was also made (American Psychiatric Association, 1980). In the fourth edition of the DSM, subtypes of BD were first introduced, such as BD type I and type II (American Psychiatric Association, 1994, 2000). In 2013, the current, fifth, edition of the DSM was published, in which further refinements of BD subtypes were made (American Psychiatric Association, 2013)

1.1.2 Phenomenology and clinical manifestation

BD is a serious mental disorder that is defined by episodic elevations in mood – referred to as mania or hypomania - and episodes of depression, which alternate with periods of

euthymia (normal mood) or subsyndromal symptoms. During episodes of (hypo)mania, the individual may experience, for example, (extremely) elevated mood (euphoria), increased activity, high energy levels, lack of the need for sleep, impulsivity/reckless behaviour, flight of ideas, talkativeness and distraction. In contrast, during episodes of depression, the individual may experience depressed mood, inability to experience pleasure, low activity, low energy, difficulty concentrating/making decisions, lethargy, changes in appetite, sleep problems (insomnia/hypersomnia), feelings of worthlessness and guilt, thoughts of death or suicidal ideation. During mood episodes, symptoms need to be present for at least 4 days (hypomania), 1 week (mania) or 2 weeks (depression), and elicit suffering and cause interference with daily functioning to meet criteria for diagnosis (American Psychiatric Association, 1994, 2000). Please refer to **box 1** for diagnostic criteria of mania, hypomania, depression and mixed episode.

The fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) distinguishes several BD subtypes, based on specific symptoms and/or aetiological criteria. For the diagnosis of **BD type I** at least one *manic* episode needs to have been present. *Depressive* episodes are not required for the diagnosis but do occur in almost all patients. **BD type II** is diagnosed when the individual has experienced one or more *hypomanic* episodes as well as one or more *major depressive* episodes. **BD not otherwise specified** is a category that includes individuals who do experience mood episodes that clearly mark a departure from the individual's normal functioning but do not fulfill the criteria associated with BD subtypes I or II. A diagnosis of **cyclothymia** is made when the individual experiences hypomanic episodes and depressive episodes that do not meet criteria for major depressive episodes. Individuals with BD may experience *rapid cycling,* i.e. having four or more episodes of mania, hypomania, major depression or mixed state, within one year (American Psychiatric Association, 1994, 2000; Goodwin et al., 2007). BD often manifests itself in late adolescence or early adulthood with at least half of all cases starting before age 25 (Kessler et al., 2005). It has an estimated lifetime prevalence of 2.4% (Merikangas et al., 2011). The World Health Organization lists this disorder as the sixth leading cause of disability worldwide (Goodwin et al., 2007) and it has been estimated that 25%-50% of patients with BD attempt suicide at least once (Jamison, 2000).

Although the pathogenesis of BD is poorly understood, there is compelling evidence of genetic involvement in the aetiology. Studies in monozygotic (identical, MZ) and dizygotic (fraternal, DZ) twins have demonstrated higher concordance rates (both twins having the same trait, in this case a disease) for BD in MZ twin pairs relative to DZ twins pairs. For example, Kieseppä et al. (2004) found probandwise concordance rates for BD of 43% in MZ twins and 6% in DZ twins in a Finnish nationwide population-based twin sample. Another study reported similar concordance rates for BD in MZ and DZ twins (McGuffin et al., 2003).

Box 1. Diagnostic criteria for mood episodes (American Psychiatric Association, 1994)

Manic episode

A. A distinct period of abnormally and persistently elevated, expansive, or irritable mood for at least a week (or any duration if hospitalization is necessary)

B. During this period of mood disturbance, three (or more) of the following symptoms have persisted (four if the mood is only irritable) and have been present to a significant degree:

1. Inflated self-esteem or grandiosity
2. Decreased need for sleep
3. Pressure to keep talking
4. Flight of ideas and/or racing thoughts
5. Distractibility
6. Increased goal-directed activity or psychomotor agitation
7. Engaging in activities with potentially painful consequences

C. The symptoms do not meet criteria for a *Mixed episode*

D. The mood disturbance causes marked impairment in occupational functioning or social activities or relationships with others, or necessitates hospitalization to prevent harm to self or other, or there are psychotic features

E. The symptoms are not due to the direct physiological effects of a substance or a general medical condition

Hypomanic episode

Same as the above where episodes are not severe enough to cause marked social or occupational impairment (or necessitate hospitalization) and have no psychotic features but do represent an unequivocal change in functioning that is uncharacteristic of the person and persists for a minimum duration of 4 days.

Depressive episode

A. Five (or more) of the following symptoms have been present during the same two-week period and represent a change from previous functioning; at least one of the symptoms is either (1) depressed mood or (2) loss of interest or pleasure

1. Depressed mood
2. Loss of interest or pleasure in activities
3. Significant change in appetite or weight
4. Insomnia or hypersomnia
5. Psychomotor agitation or retardation
6. Fatigue or loss of energy
7. Feelings of worthlessness or inappropriate guilt
8. Diminished ability to think, concentrate or make decisions
9. Recurrent thoughts of death, suicidal ideation or attempt

B. The symptoms do not meet criteria for a *Mixed episode*

C. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning

D. The symptoms are not due to the direct physiological effects of a substance or a general medical condition

E. The symptoms are not better accounted for by bereavement, i.e., after the loss of a loved one, the symptoms persist for longer than two months or are characterized by marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation

Mixed episode

A. The criteria are met both for a *Manic episode* and for a *Major depressive episode* (except for duration) nearly every day during at least a 1-week period

B. The mood disturbance is sufficiently severe to cause marked impairment in occupational functioning or in usual social activities or relationships with others, or to necessitate hospitalization to prevent harm to self or others, or there are psychotic features

C. The symptoms are not due to the direct physiological effects of a substance or a general medical condition

A review by Craddock and Jones (1999) revealed an MZ concordance for BD ranging from 40% to 70%. Moreover, the approximate lifetime risk of BD in first-degree relatives of a BD patient appears to be between 5% to 10% (Craddock and Jones, 1999) (**Table 1**). Furthermore, the heritability of BD has been found to be high, around 90% (McGuffin et al., 2003; Kieseppä et al., 2004). With the technological advances allowing researchers to investigate the human genome in more detail, a fairly large number of molecular genetic studies have been carried out to determine which genes or gene variants contribute to BD. Here, several studies have implicated specific gene variants (including variants within the genes *CACNA1C*, *ODZ4*, and *NCAN*) in the aetiology of BD but also note that polygenic contributions to the risk of developing BD is likely, probably with many more genes exerting small effects (Craddock and Sklar, 2013). Although the above discussed findings indicate strong genetic involvement in the disease, BD is not entirely explained by genes, as MZ concordance is not 100%, which suggests considerable environmental contributions to the disease as well.

Table 1. Approximate lifetime rates of bipolar disorder (adapted from Craddock and Jones, 1999)

	Risk of bipolar disorder
Degree of relationship to bipolar proband	**%**
Monozygotic co-twin	40–70
First degree relative	5–10
General population (ie, unrelated)	0.5–1.5

1.1.3 Treatment

In the treatment of individuals with BD, three phases can be discriminated: acute management of mood episodes, continuation treatment aimed at preventing relapse (i.e. return of the index episode) and maintenance treatment aimed at preventing recurrences (i.e. new mood episodes). Treatment for patients with BD often includes psychotherapeutic and pharmacological interventions, with the primary aims of reducing the frequency and severity of mood episodes, preventing psychosocial and relational complications during episodes (including suicide) and promoting optimal recovery between episodes. Pharmacotherapy, psychoeducation and promoting self-management are the cornerstones of treatment in every phase of the illness, aside from support and coaching, and psychotherapy, when necessary (Kupka et al., 2015).

1.2 BRAIN IMAGING IN BIPOLAR DISORDER

1.2.1 Brain imaging studies in bipolar disorder

As the pathophysiology of BD is poorly understood, several lines of investigation attempting to identify the biological markers associated with the disease have been set up in the previous century, not in the least of which those concerning the fields of neurology and neuroimaging. A great number of studies have been carried out to identify the neural correlates of BD. A large share of those studies investigating brain function and structure in BD have employed functional and structural magnetic resonance imaging (fMRI and sMRI, respectively) and Diffusion Tensor Imaging (DTI), but other imaging modalities such as Computed Tomography (CT), Positron Emission Tomography (PET) and Magnetic Resonance Spectroscopy (MRS) have also been used (Strakowski et al., 2005; Sullivan et al., 2009; Fusar-Poli et al., 2012).

Functional MRI (fMRI) studies assessing changes in cerebral blood flow in the brains of BD patients during task performance have particularly focused on brain regions assumed to be involved in emotion processing, such as dorsal and ventral portions of the prefrontal cortex, and subcortical brain areas, such as the amygdala, hippocampus, striatum, thalamus and nucleus accumbens (Phillips et al., 2003; Chen et al., 2011; Phillips and Swartz, 2014). For example, a meta-analysis of 65 fMRI studies with adult BD patients and healthy control subjects that was carried out by Chen et al. (2011) showed hyperactivation in multiple subcortical regions in BD patients, including the amygdala, hippocampus, caudate nucleus, pallidum and thalamus, particularly when performing tasks requiring emotion processing. In contrast, hypoactivation of the inferior frontal gyrus during both cognitive and emotion processing was also apparent, particularly when BD patients were in a manic state (Chen et al., 2011). Here, one of the most replicated and robust findings in BD is that of hyperactivation of the amygdala (Strakowski et al., 2005; Chen et al., 2011). Furthermore, an increasing number of studies employing resting-state fMRI have measured functional connectivity between brain regions during rest. These have shown, for example, decreased positive or negative resting state connectivity among frontal, temporal and subcortical brain regions, as well as abnormally increased resting connectivity in meso/paralimbic and fronto-temporal/paralimbic networks, encompassing the mesotemporal cortex, amygdala, parahippocampus and hippocampus, and orbitofrontal, ventrolateral prefrontal and subgenual cingulate cortices (Meda et al., 2012; Phillips and Swartz, 2014). Regarding structural connectivity between brain regions, a number of studies using diffusion imaging techniques have assessed white matter structure and indicated impairments in fronto-limbic and fronto-parieto-temporal circuits, as well as in the corpus callosum, compromising interhemispheric structural connectivity (Brambilla et al., 2009).

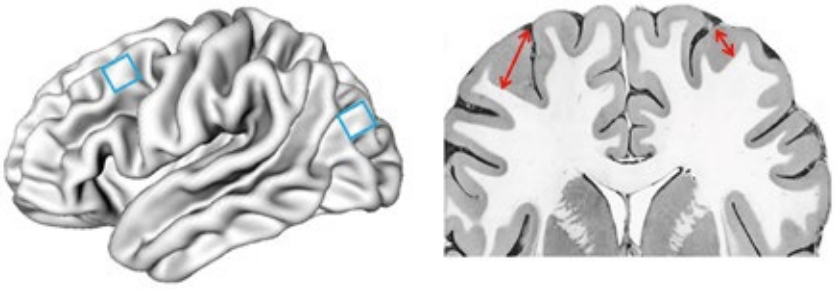
Apart from assessment of brain activation during task performance, and functional and structural connectivity in BD, many MRI studies have been devoted to determining whether

BD is associated with deficits in global and regional measures of brain volume, surface and thickness in BD. The focus of the studies described in this thesis was exclusively on brain structure and not function.

1.2.2 Cross-sectional structural brain imaging studies in bipolar disorder

A number of studies have compared structural brain measures between BD patients and healthy control subjects. However, results are notoriously inconsistent across studies, where clinical heterogeneity of included samples, medication use, age, illness duration and differences in imaging methodology may all contribute to differences in findings between studies (Emsell and McDonald, 2009; Savitz and Drevets, 2009; Hajek et al., 2012). Nevertheless, in BD patients compared to healthy controls, findings of smaller volumes of the prefrontal grey matter, white matter and corpus callosum, as well as larger ventricular volumes have been demonstrated in cross-sectional studies. Furthermore, both smaller and larger volumes of the amygdala, hippocampus and striatum, brain regions assumed to be involved in emotion processing, have been reported previously. In addition, inconsistencies regarding the volume of the thalamus, an important hub in the cortico-striato-thalamic loop linked to emotion processing, have been noted, as there are reports of BD patients showing a smaller volume as well as no differences in volume relative to healthy controls (Phillips et al., 2003; Emsell and McDonald, 2009; Savitz and Drevets, 2009; Rimol et al., 2010; Blond et al., 2012; Hajek et al., 2012; Phillips and Swartz, 2014).

Although cortical surface area and cortical thickness comprise cortical volume (Rakic, 1988, 1995; Panizzon et al., 2009; Winkler et al., 2010), it appears there has been only limited interest in investigating these structural brain elements in BD (see **Figure 1**). Regarding cortical surface area, two earlier studies did not find cortical surface area abnormalities in BD (Fornito et al., 2009; Rimol et al., 2012). In contrast, a few studies have found cortical thickness abnormalities in BD, reporting thinner (pre)frontal, anterior cingulate, temporal,



(a) (b)

Figure 1a and b. Cortical surface area (cyan squares) and cortical thickness (red arrows) (adapted from the CIVET template, described in chapters 4 and 5 (a) and Roberts et al., 1987 (b)).

parietal and occipital cortices in patients relative to healthy controls (Lyoo et al., 2006; Rimol et al., 2010; Foland-Ross et al., 2011).

1.2.3 Longitudinal structural brain imaging studies in bipolar disorder

Studies assessing structural brain changes over time in BD are rare and those that have been carried out show great variability in results. For example, in a review by Lim et al. (2013), preservation, increases and decreases over time in volumes of the prefrontal and temporal cortex, amygdala and hippocampus were all described. Furthermore, both increases and decreases in cingulate volumes have been shown, whereas total brain volume appears to be relatively stable over time in BD. In addition, there is some evidence of volume increases in the thalamus and caudate nucleus but this has to be confirmed in future studies (Lisy et al., 2011; Lim et al., 2013).

1.2.4 Lithium

For some part the inconsistency across studies in findings on brain structure in BD may be attributable to the influence of medication on the brain, where the influence of lithium on the brain has been particularly noted (Hafeman et al., 2012). Here, several studies have indicated larger volumes of the total grey matter, amygdala, hippocampus and anterior and subgenual cingulate cortex in BD patients who had taken lithium compared to BD patients that had not (Kempton et al., 2008; Hallahan et al., 2011; Hafeman et al., 2012; Hajek et al., 2012). Therefore, accounting for lithium use by BD patients when assessing brain structure is an important undertaking, albeit a complicated one. In the studies described in this thesis, correction for lithium use was applied where possible.

1.3 THE TWIN DESIGN

In research areas concerned with studying human phenotypes, a twin design is sometimes used. It allows for estimation of the *relative* contribution of genes and environment to observable traits without specifying which specific gene variants or environments are involved. When using both monozygotic (MZ) and dizygotic (DZ) twin pairs, it is possible to disentangle the relative contribution of genes and (shared and unique) environmental factors to observable traits by comparing the variance and covariance within and between MZ and DZ twins. MZ twins are assumed to be genetically identical whereas dizygotic twins only share on average 50% of their segregating genes. Because of this difference in genetic relatedness within twins pairs, the magnitude of genetic and environmental involvement in a particular trait can be estimated (Boomsma et al., 2002). For example, if total brain volume is highly

correlated between twins of MZ pairs but not between twins of DZ pairs, this could suggest that additive genetic factors (A) are primarily involved in the trait, since both types of twin pairs (MZ and DZ) are assumed to have been exposed to equal environments (Rijsdijk and Sham, 2002). If, however, the correlations within MZ and DZ pairs are similar, then the involvement of shared environmental factors (C) in the trait is more likely. If there is no strong correlation within (either MZ or DZ) twin pairs, then unique environmental factors (E) likely influence the trait. It is also possible to estimate non-additive or dominant genetic influences (D) on traits (Boomsma et al., 2002; Rijsdijk and Sham, 2002). The magnitude of the relative influence of these factors is expressed in terms of heritability and 'environmentability', although the latter term is not as widely used as the former. The heritability (h^2) of a trait denotes the variance in a phenotype that is attributable to the variance in genetic factors in a particular population (Wray and Visscher, 2008). The percentage of phenotypic variance that is not attributable to genetic variance is automatically defined as that part of the phenotypic variance that is due to non-genetic, environmental variance (1-h^2), which is a composite percentage indicating the total influence of environmental factors on a trait that can be further decomposed in shared and unique environmental factors (c^2, e^2). It is important to note that heritability does explicitly not refer to how much of a trait is explained by genes in any specific individual, it is a population-based statistic that only indicates the variance in a phenotypic trait that is due to genetic variance in a population (Wray and Visscher, 2008). Therefore, changes in population characteristics may influence the magnitude of the heritability, it is not a static parameter. In addition to estimating the influence of genetic and/or environmental factors to single traits, the twin pair correlations can also serve as a basis to estimate the influences of A, D, C, and E on the *association* between two or more traits (Rijsdijk and Sham, 2002). For example, genes influencing BD could, to some degree, overlap with genes influencing brain volume and this overlap can be calculated. In the classical twin design, several types of correlations are used to determine influences genetic and environmental factors:

- **Within-trait/cross-twin correlations** are the correlations in one trait between members in a MZ or DZ twin pair and serve as the basis for estimation of heritability and environmental influence on that trait.
- **Within-twin/cross-trait correlations** pertain to the correlations between traits in single individuals and serve as the basis for estimating the magnitude of association between those traits.
- **Cross-twin/cross-trait correlations** are the correlations between trait 1 of twin 1 and trait 2 of twin 2 (and vice versa) within a twin pair and can be compared between MZ and DZ twin pairs. These correlations allow for estimation of the degree to which genes and environmental factors influence both traits.

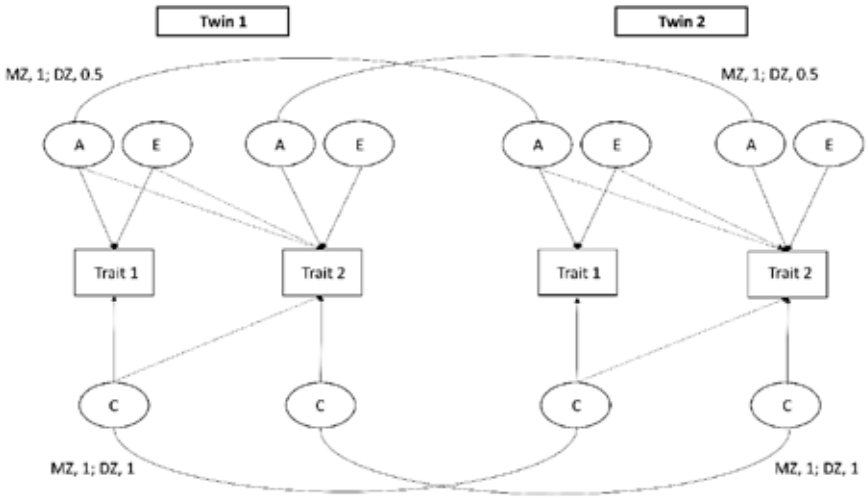


Figure 2. Example of a path diagram, depicting a bivariate twin model. Latent additive genetic (A), shared (common) environmental (C) and unique environmental (E) factors influence two traits in each twin, as indicated by arrows. The additive genetic factors (A) of monozygotic (MZ) twins are perfectly correlated (1.0), whereas those of dizygotic (DZ) twins are correlated at 0.5; shared environmental factors are perfectly correlated for both MZ and DZ twins; unique environmental influences (E) are always uncorrelated between twins. Dotted arrows indicate where A, C and E factors may influence both traits.

In order to estimate the genetic and environmental influences and correlations, a structural equation model (SEM) is usually applied. Structural equation modeling involves estimating regression coefficients between observed (e.g. brain volume) and not observed (latent) variables (e.g. genes). The SEM allows for comparison of 'goodness of fit' of various models (using maximum likelihood estimation) that can be nested within one another. Usually, a saturated model is first constructed, estimating all the means and variances without restrictions. Then, the more restricted ACE-model is fitted and its goodness of fit compared to the saturated model. Subsequently, even more restricted models (AE-model, CE-model or E-model) can be fitted to the data to assess which one describes the data best. For example, if the data indicate larger MZ twin correlations than DZ twin correlations, the C-factor may possibly be dropped from the model without seriously compromising the goodness of fit. Because an AE-model is more parsimonious than an ACE-model, the former is preferred over the latter to describe the data with. Note however, that the E-factor can never be dropped from the model as this component also includes measurement error (Neale and Cardon, 1992; Sham, 1998; Rijsdijk and Sham, 2002).

In the studies described in this thesis, bivariate twin models were implemented in structural equation modeling software estimating genetic and environmental variance components, and phenotypic, genetic and environmental correlations between two traits (**Figure 2**).

1.4 STRUCTURAL BRAIN IMAGING STUDIES IN PATIENTS WITH BIPOLAR DISORDER AND THEIR RELATIVES

Family and twin studies assessing brain structure in BD are extremely rare. Only a small number of cross-sectional structural imaging studies in BD patients and their family members have been carried out. For example, Kieseppä et al. (2003) found smaller white matter volume to be genetically linked to the risk for BD. A previous study from our group yielded similar findings, where smaller volume of the white matter was influenced by genes contributing to BD. Furthermore, in that study, smaller cortical grey matter volume was found to be associated with unique environmental factors related to BD (van der Schot et al., 2009) whereas regional indices of grey and white matter density were subsequently found to be both genetically and uniquely environmentally associated with BD (van der Schot et al., 2010). In another study assessing cortical brain structure, BD was genetically linked to thicker supramarginal, fusiform and postcentral cortices, and thinner orbitofrontal and parahippocampal cortices, whereas unique environmental factors overlapped between BD and thicker precentral and thinner calcarine cortices (Hulshoff Pol et al., 2012). In addition, volumetric deficits of the ventral striatum and anterior cingulate gyrus were associated with the genetic risk for BD (McDonald et al., 2004). One study found larger caudate nuclei in MZ twin pairs discordant for BD compared to MZ control pairs (Noga et al., 2001). However, an absence of genetic or environmental links between BD and brain structure has also been reported (McDonald et al., 2006; McIntosh et al., 2006).

Although there have been a few twin and family studies assessing brain structure in BD cross-sectionally, there have been no longitudinal studies assessing brain structure change over time in BD in genetically informative samples. The study described in this thesis is the first longitudinal structural brain imaging twin study with BD patients and their co-twins in the literature to date.

1.5 SETUP AND OUTLINE THESIS

1.5.1 Rationale of this study

The studies described in this thesis include baseline and longitudinal examinations of brain structure in twins concordant and discordant for BD. From the outset, the goal was to assess whether BD is associated with structural brain abnormalities at baseline and over time, and, if so, to what extent genes and environment contribute to these associations. It has been documented that schizophrenia, a disease showing overlap in genes (Craddock and Owen, 2010) and clinical manifestation with BD (Lin and Mitchell, 2008), is associated with

progressive volume loss in frontal, temporal and parietal grey matter, and widespread cortical thinning, particularly in bilateral temporal and left frontal cortices, during the course of the illness (van Haren et al., 2011; Vita et al., 2012). Although the studies described in this thesis did not directly compare schizophrenia patients and BD patients, findings may contribute to the ongoing debate as to whether or not the clinical dichotomy once suggested holds up (Kraepelin, 1899).

However, longitudinal studies assessing brain structure in BD are rare, particularly those including cohorts with twins concordant and discordant for the disease such as the one described in this thesis. Therefore, in this study, baseline and change measures of the entire brain, as well as those of subcortical and cortical brain regions were investigated in BD patients, their co-twins and healthy controls. Here, a number of different structural measures were obtained, such as the volume, surface and thickness of brain regions. Moreover, applying bivariate genetic modeling, the relative contributions of genes and environment to the association between the liability to BD and cross-sectional and longitudinal brain measures were estimated. However, in one study we assessed the contribution of genes and environment to the association between stressful life events and hippocampal volume in healthy twins only.

1.5.2 Subjects, materials and methods

The twin cohort

In this study, data were collected in a twin sample comprising monozygotic (MZ) and dizygotic (DZ) twin pairs that were concordant and discordant for BD, and healthy MZ and DZ twin pairs with no psychiatric diagnosis that were matched to the patient group for zygosity, gender, age and parental education. All twin pairs were recruited between 1997 and 2006, except for one DZ patient pair that was first included in 2011. Patient twin pairs were mainly recruited via newspaper ads and through referral by mental health institutions and practitioners (see van der Schot et al., 2009, 2010) whereas many of the healthy twin pairs were recruited from the healthy twin sample of the department of Psychiatry of the University Medical Center Utrecht and the Netherlands Twin Registry (Boomsma, 1998; Baaré et al., 2001). The total baseline sample consisted of 51 patient twin pairs and 67 healthy control twin pairs. Between 2010 and 2013, the majority of these subjects were approached for the second measurement. Zygosity was determined by DNA fingerprinting using high polymorphic microsatellite markers 9-11 in the laboratory of the Division Biomedical Genetics, University Medical Center Utrecht.

Clinical assessment

All study subjects underwent elaborate psychiatric assessment at both time points. Clinical

diagnosis of patients (and co-twins) was confirmed with the Structured Clinical Interview for DSM-IV (SCID) and the Structured Interview for DSM-IV Personality (SIDP and SCID-II)(First et al., 1997; Pfohl et al., 1997), as well as through available medical records. Moreover, current mood state (i.e., (hypo)manic, depressive or mixed episode) was assessed with the Young Mania Rating Scale (YMRS) (Young et al., 1978), the Inventory for Depressive Symptomatology (IDS) (Beck et al., 1961) and the Hamilton Depression Rating Scale (HDRS, at follow-up only) (Hamilton, 1960). Medical histories were registered with a medical checklist. Family histories were obtained with the Family Interview Genetic Studies (FIGS) (Nurnberger, jr. et al., 1994). In addition, a large part of the subject sample was interviewed on life events with the Bedford College Life Events and Difficulties Schedule (LEDS), an elaborate semi-structured interview collecting information about life events and long-term difficulties from the age of 5 and onwards (Brown and Harris, 1978,1989).

Brain imaging

MRI procedures were the same in all subjects that participated in the study. MRI was carried out using 1.5 Tesla scanners (Philips, The Netherlands) with identical imaging and processing parameters across time points. Segmentation of global cortical brain volumes was usually carried out with an in-house developed semi-automated processing pipeline (Schnack et al., 2001; Brouwer et al., 2010), whereas subcortical brain volumes were segmented with the automated FreeSurfer structural imaging pipeline (http://surfer.nmr.mgh.harvard.edu/, Fischl et al., 2002, 2004). Regional measures of cortical surface area, cortical thickness and cortical volume were obtained with the CLASP algorithm (Constrained Laplacian Anatomic Segmentation Using Proximity) in a custom implementation of CIVET that was developed at the McConnell Brain Imaging Centre of the Montreal Neurological Institute (MacDonald et al., 2000; Kabani et al., 2001; Kim et al., 2005).

Statistical modeling

Imaging data were analyzed with Statistical Package for the Social Sciences (IBM SPSS, release 21.0.0.0), R statistical programming software (R Development Core Team, 2008) and structural equation modelling software OpenMx (Kenny et al., 2009). Within and between group comparisons were carried out, as well as genetic model fitting procedures, allowing for estimation of genetic and environmental contributions to phenotypes and associations between phenotypes.

1.5.3 Outline thesis

In **chapter 2**, findings of a cross-sectional study on the association between hippocampal volume and life events in healthy twins are presented. The hippocampus is an important

brain structure involved in emotion processing (Phillips et al., 2003, 2008) that may be compromised in multiple stress-related psychiatric disorders such as BD (Hajek et al., 2012), major depressive disorder (MDD) (Kempton et al., 2011), post-traumatic stress-disorder (PTSD) (Karl et al., 2006; Kuhn and Gallinat, 2013) and schizophrenia (Adriano et al., 2012; Shepherd et al., 2012). In this study, the extent of genetic and environmental involvement in the association between life events and hippocampal volume was evaluated. In **chapter 3**, the results of cross-sectional and longitudinal analysis of subcortical brain regions implicated in emotion processing are described in twins concordant and discordant for BD. This study focused particularly on the degree of genetic and environmental overlap between cross-sectional and longitudinal measures of subcortical volumes and BD. In **chapter 4**, the results of a cross-sectional study on global and regional measures of the volume, surface and thickness of the cortex in twins concordant and discordant for BD are described. Here too, the degree of genetic and environmental influences on the association between BD and these measures was investigated. In **chapter 5**, longitudinal findings of global brain volume and global and regional measures of cortical surface, cortical thickness and cortical volume in twins concordant and discordant for BD are presented. The extent to which genes and environment influence brain changes over time in BD was evaluated. Finally, in **chapter 6**, the results and implications of all the individual studies that are included in this thesis are discussed jointly.

Chapter 2

The association between hippocampal volume and life events in healthy twins

Florian Bootsman*, Sanne M. Kemner*, Manon H.J. Hillegers, Rachel M. Brouwer, Ronald Vonk, Astrid C. van der Schot, Hilleke E. Hulshoff Pol, Willem A. Nolen, René S. Kahn, Neeltje E.M. van Haren

* Authors contributed equally to this manuscript and share first authorship

In press

2.1 ABSTRACT

Hippocampal volume deficits have been linked to life stress. However, the degree to which genes and environment influence the association between hippocampal volume and life events is largely unknown. In total, 123 healthy twins from monozygotic and dizygotic twin pairs underwent magnetic resonance imaging (MRI), and 57 healthy twins were interviewed with the Life Events and Difficulties Schedule (LEDS), with an overlap of 54 twins undergoing both MRI and the life events interview. Hippocampal volumes were segmented with FreeSurfer software. Data were analyzed with OpenMx software. Smaller hippocampal volume was associated with higher severe life event load (r_{ph}= -0.39), where shared environmental factors influencing both measures fully explained the association. Hippocampal volume was not associated with total or mild life event load. Hippocampal volume showed high heritability (range, h^2: 57%-81%) whereas life event measures were influenced by shared (c^2) and unique (e^2) environmental factors only (range, c^2:40%-64%, e^2: 36%-60%). The results suggested that shared environmental factors influenced the relationship between smaller hippocampal volume and severe (but not mild) stress. This indicated that particularly severe life events that were shared between twins were associated with smaller hippocampal volume. Furthermore, it is suggested to distinguish between mild and severe life events in life event research.

2.2 INTRODUCTION

The hippocampus is a highly plastic brain structure (Gu et al., 2013; Spalding et al., 2013), implicated in memory and emotional processing (Milner et al., 1998; Bird and Burgess, 2008; Phillips et al., 2008; Aldhafeeri et al., 2012). Importantly, it provides negative feedback signaling to the hypothalamus during the hypothalamic-pituitary-adrenal (HPA) activated response to environmental stress, a process which results in secretion of glucocorticoids that are vital to short-term survival (Jankord and Herman, 2008). However, chronic stress has been linked to disruption of hippocampal negative feedback signaling (Conrad, 2006) and to glucocorticoid toxicity in the hippocampus itself, which, as a result, may show structural atrophy and diminished neurogenesis (Sapolsky et al., 1990; Brown et al., 2004; Mirescu and Gould, 2006; Gianaros et al., 2007; McEwen, 2007).

Only a few studies have examined hippocampal structural deficits in relation to environmental stress by specifically investigating the association between stressful life events and hippocampal volume in healthy subjects, suggesting evidence for a negative association between them (Gianaros et al., 2007; Papagni et al., 2011; Rabl et al., 2014; Shepherd et al., 2012). Remarkably, there is a large variability in stress-assessment strategies when determining the number and impact of life events, as well as differences in demographics (e.g. age) of studied groups. For example, one recent study in adults reported that smaller hippocampal volume was predicted by total number of self-reported stressful life events, but only in interaction with specific genetic variants (COMTVal158Met, BDNFVal66Met and 5-HTTLPR) (Rabl et al., 2014). In contrast, a study in postmenopausal women showed that smaller hippocampal grey matter volume was predicted by higher stress scores, where subjects were required only to indicate the perceived degree of control, predictability and overload of life events on a 4-item self-report questionnaire without describing the event itself (Gianaros et al., 2007). Yet another, longitudinal study reported that hippocampal volume loss over time was associated with higher number of stressful life events in a sample of young adults (Papagni et al., 2011).

Importantly, little is known about the genetic and environmental contributions to the association between hippocampal volume and life events. In this respect, twin studies provide a suitable approach to determine the relative influence of genes, and shared and unique environmental factors on measured traits as well as on the association between traits. Therefore, in this healthy twin study, we examined (1) whether hippocampal volume was associated with life events as assessed with the elaborate semi-structured Bedford College Life Events and Difficulties Schedule (LEDS, Brown and Harris, 1978, 1989), and (2) the relative contribution of genes and (shared and unique) environmental factors to hippocampal volume, life events, and the association between them.

2.3 MATERIALS AND METHODS

2.3.1 Subjects

A total of 123 healthy twins from monozygotic (MZ) and dizygotic (DZ) twin pairs (MZ: 35 twin pairs and 4 subjects from incomplete pairs; DZ: 22 twin pairs and 5 subjects from incomplete pairs) underwent magnetic resonance imaging. Furthermore, a group of 57 healthy twins (MZ: 15 twin pairs; DZ: 13 twin pairs and 1 twin from an incomplete pair) were interviewed on life events with the semi-structured Life Events and Difficulties Schedule (LEDS). Here, 54 twins underwent both MRI scanning and the LEDS interview. Due to the limited availability and motivation of some of the subjects in the MRI sample, the LEDS interview could not be conducted in all subjects that underwent MRI. However, as there were no differences in inclusion or exclusion criteria between the LEDS sample and the MRI sample, we feel confident that the twins whose life events were assessed are representative of the larger sample.

Healthy twins were taken from previously described twin samples (Baaré et al., 2001; Brans et al., 2008; van der Schot et al., 2009). Of the total group of 126 twins that either underwent MRI or were interviewed on life events, 47 healthy control twins were originally recruited by van der Schot et al. (2009). Of the remainder of 79 twins, 18 twins were taken from the cohort that was described by Baaré et al. (2001) and 61 twins were taken from the cohort that was included by Brans et al. (2008). These twins were originally recruited from the (healthy) twin sample of the department of Psychiatry of the University Medical Center Utrecht and the Netherlands Twin Registry (Boomsma, 1998; Baaré et al., 2001). Subjects were between 18 and 60 years of age (see **Table 1** for demographic information). MZ and DZ twin pairs were matched for age and parental education.

As many of these twins served as control subjects in studies with patients with schizophrenia and bipolar disorder, they had no history of Axis I psychiatric disorder or Axis II personality disorder according to *DSM-IV* criteria. This was confirmed with the Structured Clinical Interview for DSM-IV (SCID) (First et al., 1997), and the Structured Interviews for DSM-IV Personality (SIDP and SCID-II) (First et al., 1997; Pfohl et al., 1997). Moreover, they had no history of severe medical illness. None of the subjects had drug or alcohol dependency in the last six months prior to inclusion. Furthermore, the twins had no first-degree relatives with a history of a major Axis I psychiatric disorder (DSM-IV) such as bipolar disorder, schizophrenia, psychotic disorder, major mood disorder, anxiety disorder or substance-related disorder. Family histories of the twins were obtained via the Family Interview Genetic Studies (FIGS, Nurnberger et al., 1994), performed with both twins. Zygosity was determined in the laboratory of the Division Biomedical Genetics, University Medical Center Utrecht (UMCU) with DNA fingerprinting using high polymorphic microsatellite markers 9 to 11. The medical

ethics review board of the UMCU approved the study and all participants gave written informed consent after full explanation of the study aims and procedures.

Table 1. Demographic characteristics and descriptive statistics of hippocampal volume and life event measures for all subjects.

	Total group[a] (n=126)	**MZ** (n=75)	**DZ** (n=51)	**F**	**df**	**p**
Gender, m/f	54/72	30/45	24/27			
Age, yr, mean (sd)	38.44 (8.96)	38.64 (9.92)	38.14 (7.41)	0.10	1, *124*	0.76
Parental educ., yr, mean (sd)	11.49 (3.36)	11.41 (3.33)	11.61 (3.44)	0.10	1, *124*	0.75
Education, yr, mean (sd)	13.56 (2.52)	13.73 (2.66)	13.29 (2.3)	0.92	1, *124*	0.34
Hippocampal volume (in ml), mean (sd)[b]	8.52 (0.74)	8.59 (0.75)	8.41 (0.72)	1.61	1, *121*	0.21
Total life event load, mean (sd)[c]	80.74 (21.55)	73.4 (21.23)	88.89 (19.15)	8.30	1, *55*	0.01
Mild life events, load, mean (sd)[c]	66.09 (16.84)	60.97 (16.64)	71.78 (15.43)	6.42	1, *55*	0.01
Severe life events, load, mean (sd)[c]	14.65 (8.42)	12.43 (7.8)	17.11 (8.54)	4.67	1, *55*	0.04

Abbreviations: MZ, Monozygotic; DZ, Dizygotic
n= number of individuals
[a] Total number of subjects who were assessed for at least one of the two measures (i.e. hippocampus and/or life events); a total of 54 subjects were assessed for both types of measures (i.e. hippocampus AND life events).
[b] Sample size of hippocampal volume only: Total group= 123; MZ=74 , DZ=49.
[c] Sample size of life events measures only: Total group= 57; MZ=30 , DZ=27.

2.3.2 Brain imaging

Magnetic resonance images were acquired on a Philips Intera 1.5 Tesla scanner (Philips, the Netherlands), with the following imaging parameters: T1-weighted 3D fast field echo scans with 160–180 contiguous coronal slices (echo time=4.6 ms, repetition time=30 ms, flip angle=30°, 1x1x1.2 mm^3 voxels) (van der Schot et al., 2009). MZ and DZ twins were randomly assigned to MRI slots, eliminating possible between-group biases due to scanner drifts.
Processing of brain images and hippocampal volumetric segmentation was performed with the FreeSurfer 5.1.0 structural imaging pipeline (http://surfer.nmr.mgh.harvard.edu/). Anatomic volume of the bilateral hippocampus was delineated using information on image intensity, probabilistic atlas location and spatial relationships between subcortical structures (Fischl et al., 2002; Fischl et al., 2004). All hippocampal segmentations were visually inspected to assure high quality data.

2.3.3 Life Events and Difficulties Schedule (LEDS)

Life events were assessed with the investigator-based Bedford College Life Events and Difficulties Schedule (LEDS), a semi-structured interview assessing life events and long-term difficulties (Brown and Harris, 1978, 1989). The present study focused exclusively

on life events. The LEDS collects detailed information about the event itself, the timing of its occurrence (date) and relevant contextual information for each event. Based on the contextual information, the threat for each event is rated via standardized rating procedures. The threat score represents the severity of the event, ranging from mild (1) to severe (4). The contextual threat is conceptualized as: "What most people would be expected to feel about an event in a particular set of circumstances and biography, taking no account of what the respondent says either about his or her reaction or about any psychiatric or physical symptoms that followed it" (Brown and Harris, 1989). Severe events could be negative as well as positive, for example: moving to another country can be a very positive, but at the same time a stressful, life event. The interview covered life events from early childhood (after age 5) up to assessment. Life events were rated on a yearly basis. All interviewers and raters were trained by MH, who herself was trained by Brown and Harris who developed the LEDS. The events were rated from written transcripts by two independent raters who had not been involved in the interview. A panel consisting of five raters, including FB, SK and MH, reached consensus on the events that raised rating problems.

The cumulative life event load was calculated as the sum of the threat scores of all life events in the year of MRI and all preceding years. We calculated three types of life event load. The total life event load; representing the sum of the threat scores of all life events (threat score 1 to 4), mild event load; representing the sum of the threat scores of all mild life events (threat scores 1 and 2) and severe event load; representing the sum of the threat scores of all severe life events (threat scores 3 and 4). Please refer to **Supplementary Table S1** for some examples of life event load calculations.

2.3.4 Statistical analyses

Calculation of descriptive statistics and rendering of standardized residuals (obtained after regression on age, gender and intracranial volume (ICV) for hippocampal volume and on age only for life event measures) suitable for genetic model fitting were performed with the Statistical Package for the Social Sciences (IBM SPSS, release 21.0.0.0).

Genetic model fitting

To estimate relative genetic and environmental contributions to the association between hippocampal volume and life event measures, a bivariate continuous model was chosen and implemented in structural equation modeling software OpenMx (Kenny et al., 2009), running under the statistical programming environment R (R Development Core Team, 2008). A bivariate Cholesky decomposition was fitted to the standardized residuals of our hippocampal and life event measures (hippocampal volume was corrected for age, gender and ICV, and life event measures were corrected for age) to estimate additive

genetic (A), and shared/common (C) and unique environmental (E) variance components of hippocampal volume and life event measures, and phenotypic, genetic and environmental overlap between these measures. The phenotypic correlation (r_{ph}), an index of association between phenotypes (e.g. hippocampal volume and severe life event load), was based on calculations of within-twin/between-trait correlations. Heritability (h^2) and influence of shared (c^2) and unique environment (e^2), as well as disentanglement of the observed correlation between measures into genetic and environmental components, was based on polychoric cross-twin/within-trait and cross-twin/cross-trait correlations within MZ and DZ groups (Neale and Miller, 1997). The heritability of hippocampal volume and life event measures was determined within the bivariate model. A larger correlation between traits in MZ twins than in DZ twins suggests higher genetic contribution due to MZ twins being genetically identical, whereas DZ twins only share on average 50% of their segregating genes. If there is no difference between MZ and DZ correlations then a larger influence of shared environmental factors is more likely (Boomsma et al., 2002). The genetic (r_g) and shared (r_c) and unique environmental (r_e) correlations respectively indicate the degree of overlap in genes or shared and unique environment influencing phenotypes. The phenotypic correlation can be written as the sum of the genetic correlation weighted by the square root of the heritabilities of the two traits ($r_g * \sqrt{(h^2_{\text{hippocampus}})} * \sqrt{(h^2_{\text{life event measure}})}$) and the environmental correlations weighted by the square root of environmental variances associated with the two traits ($r_c * \sqrt{(c^2_{\text{hippocampus}})} * \sqrt{(c^2_{\text{life event measure}})}$; $r_e * \sqrt{(e^2_{\text{hippocampus}})} * \sqrt{(e^2_{\text{life event measure}})}$). These quantities are written as $r_{ph\text{-}g}$, $r_{ph\text{-}c}$ and $r_{ph\text{-}e}$ (Toulopoulou et al., 2007). In order to determine the model that best explained our data, we fitted different nested models and compared their goodness of fit using Akaike's Information Criterion (AIC). A saturated model in which means, variances and correlations are estimated freely served as a baseline model to which more restrictive models were compared. We performed an exhaustive analysis and determined the best fitting model (i.e. the model with the lowest AIC-value) of all the possible bivariate models (variable 1: hippocampal volume, variable 2: life event measure). This was done for each of the three bivariate relationships that we assessed (i.e. 'hippocampal volume and total life event load', 'hippocampal volume and mild life event load' and 'hippocampal volume and severe life event load'). Significance of parameter estimates and correlations within the best fitting model was determined based on 95% confidence intervals (CI) (Neale and Miller, 1997).

Correction for multiple comparisons

Correction for multiple testing was applied by dividing the alpha of 0.05 by the number of life event loads we assessed, which was 3. This resulted in a Bonferroni threshold for significance of $\alpha=0.05/3=0.017$. Significance at $\alpha=0.05$ and $\alpha=0.017$ is indicated in all tables and figures.

2.4 RESULTS

2.4.1 Demographic information and descriptive statistics

Table 1 shows the demographic information and mean values of hippocampal volume and life events measures for all groups. There were no differences in age, parental education, or own education between MZ and DZ twins. MZ twins had a significantly lower total life event load ($F_{1, 55}$=8.3, p=0.01), mild life event load ($F_{1, 55}$=6.42, p=0.01), and severe life event load ($F_{1, 55}$=4.67, p=0.04) than DZ twins.

2.4.2 Genetic model fitting

The best fitting models for the three bivariate comparisons are depicted in **Supplementary Figure S1**. Here, one model fitted best for the association of hippocampal volume with both total and severe life event loads. This model consisted of an A variance component for the hippocampus, a C variance component that was shared between the hippocampus and both total and severe life event loads, and two E variance components for the hippocampus and the life event load (total/severe) separately. For the association between hippocampal volume and mild life events, a model estimating A and E variance components for the hippocampus and C and E variance components for the mild life events but no shared factors between measures had the best fit. For the hippocampus, the variance components were obtained from all three bivariate models (i.e. 'hippocampus' vs 'total life event load', 'hippocampus' vs 'mild life event load', 'hippocampus' vs 'severe life event load').

Table 2 shows the MZ and DZ within-trait/cross-twin correlations, and variance components of genetic and environmental influences on all measures. Here, hippocampal volume showed high heritability (h^2: 57% to 81%) and low environmental influence in all three bivariate analyses whereas life event measures were not influenced by genes. The variance in total, mild and severe life event measures was attributable to shared and unique environmental factors (total, c^2: 52%, e^2: 48%; mild, c^2: 40%, e^2:60%; severe, c^2: 64%, e^2: 36%). **Table 3** shows the MZ and DZ cross-trait/cross-twin correlations, as well as the phenotypic, genetic and environmental correlations between hippocampal volume and life events measures (see also **Figure 1**). Hippocampal volume was significantly phenotypically associated with severe life event load (r_{ph}: -0.39), after Bonferroni correction for multiple comparisons. In the best-fitting model, this association was entirely explained by shared environmental factors. This indicates that shared events that influence twins similarly determined the association between smaller hippocampal volume and higher load of severe life events . In contrast, in the best fitting models, total and mild life event loads were not significantly phenotypically, genetically or environmentally associated with hippocampal volume.

Table 2. Within-trait/cross-twin correlations in MZ and DZ pairs measures (with 95% confidence intervals), and parameter estimates of genetic and environmental influences on all measures (with 95% confidence intervals).

	Within-trait/cross-twin correlations		Variance components					
	MZ	DZ		h^2		c^2		e^2
Measure[a]				%				
Hippocampal volume (in ml)[b,c]	**0.82**†(0.67 to 0.91)	0.3 (-0.2 to 0.65)	(T)	**70**† (39 to 87)		11 (0 to 41)		**19**† (11 to 33)
			(M)	**81**† (68 to 89)		-		**19**† (11 to 32)
			(S)	**57**† (27 to 81)		**24** (2 to 52)		**19**† (11 to 34)
Total life events, load[d]	0.34 (-0.43 to 0.79)	**0.63**† (0.17 to 0.87)		0		**52**† (20 to 74)		**48**† (26 to 80)
Mild life events, load[d]	0.17 (-0.51 to 0.77)	0.37 (-0.17 to 0.75)		0		**40** (5 to 67)		**60**† (33 to 95)
Severe life events, load[d]	0.37 (-0.28 to 0.73)	**0.86**†(0.64 to 0.95)		0		**64**† (37 to 81)		**36**† (19 to 63)

†Significant at α=0.017 (Bonferroni threshold), estimates in bold face are significant at α=0.05.
[a] Life event measures are corrected for age, and hippocampal volume is corrected for age, gender and intracranial volume.
[b] Sample size of hippocampal volume only: Total group= 123; MZ=74 , DZ=49.
[c] Heritability of the hippocampus was calculated for the 3 respective bivariate comparisons between hippocampal volume and life event loads. T: from total life events load, M: from mild life events load, S: from severe life events load.
[d] Sample size of life events measures only: Total group= 57; MZ=30 , DZ=27.

Table 3. Cross-trait/cross-twin correlations in MZ and DZ twin pairs (with 95% confidence intervals), and phenotypic, genetic and environmental correlations with hippocampal volume (with 95% confidence intervals).

	Cross-trait/cross-twin correlations		Parameter estimates						
Category (load)[a]	MZ (n=29)	DZ (n=25)	r_{ph}	r_g	r_c	r_e	$r_{ph\text{-}g}$	$r_{ph\text{-}c}$	$r_{ph\text{-}e}$
Total life events	-0.31	-0.41	-0.24	-	-1	-	-	-0.24	-
	(-0.66 to 0.21)	(-0.7 to 0.04)	(-0.49 to 0.07)		(-1 to -1)			(-0.49 to 0.07)	
Mild life events	-0.23	-0.26	-	-	-	-	-	-	-
	(-0.59 to 0.31)	(-0.59 to 0.19)							
Severe life events	-0.41	**-0.47**	**-0.39**†	-	-1	-	-	**-0.39**†	-
	(-0.7 to 0.02)	(-0.72 to -0.06)	(-0.61 to -0.11)		(-1 to -1)			(-0.61 to -0.11)	

†Significant at α=0.017 (Bonferroni threshold), estimates in bold face are significant at α=0.05.
[a] Life event measures are corrected for age, and hippocampal volume is corrected for age, gender and intracranial volume.

2.5 DISCUSSION

In this study, we assessed the influence of genes and environment on the relation between hippocampal volume and life events in healthy twins. Our main finding is that severe stressful life events are strongly associated with a smaller hippocampal volume and that this

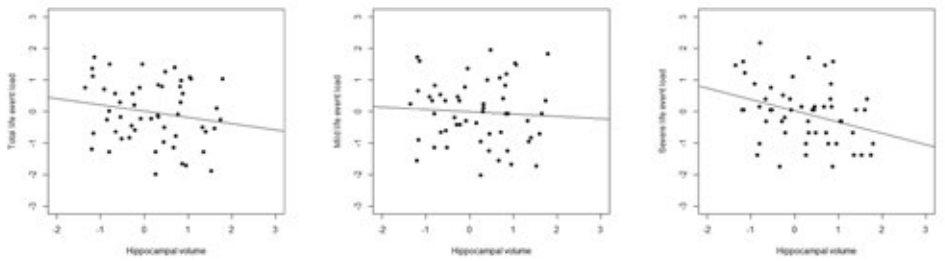


Figure 1. Uncorrected hippocampal volumes and life event loads (total, mild and severe).

association is fully explained by shared environmental factors that influence twins similarly within the twin pairs. This suggests that severe shared events (e.g. divorce of parents, learning that a family member is seriously ill) may lead to a decrease in hippocampal volume. Despite study differences in MRI processing and assessment of life events, our findings are in line with earlier studies showing suggestive evidence for higher levels of stress to be associated with hippocampal volume loss (Gianaros et al., 2007; Papagni et al., 2011). We extend these findings by showing that stressful life events are associated with (smaller) hippocampal volume, especially when they are severe. In a previous twin study with combat exposed veterans who developed PTSD and their combat unexposed co-twins, Gilbertson et al. (2002) suggest that a smaller hippocampus may constitute a risk factor for the development of stress-related pathology. Particularly since both the combat exposed veterans with PTSD and their combat unexposed co-twins showed smaller hippocampi than combat exposed veterans without PTSD and their co-twins. However, in that study, the contribution of shared environmental factors that influence twins similarly to hippocampal volume and life stress was not assessed, so their involvement may have been overlooked.

The high heritability of hippocampal volume is in line with a previous study reporting a heritability of 69% (van Erp et al., 2004), although moderate genetic influence on the hippocampus has also been noted (Schmitt et al., 2007; Blokland et al., 2012). In contrast, life event measures showed large influences of shared and unique environmental factors with no influence of genetic factors. This contradicts previous studies that suggest that life events are influenced by genetic factors (see review by Kendler and Baker, 2007 and the study by Vinkhuyzen et al., 2010). However, in those studies, the observed genetic influences were modest and mostly based on self-report measures of life events, which may have inflated heritability estimates (Vinkhuyzen et al., 2010). Furthermore, a strong influence of shared environmental factors on life events is to be expected as twins often experience the same life events and may be affected similarly by them. Nevertheless, significant unique environmental influence on hippocampal and life event measures should be interpreted with caution as, in a twin model, this variance component also includes measurement error.

Our findings may be particularly relevant in psychiatry as hippocampal abnormalities have been shown in a number of stress-related psychiatric disorders, including bipolar disorder (Hajek et al., 2012), major depressive disorder (MDD) (Kempton et al., 2011), post-traumatic stress-disorder (PTSD) (Karl et al., 2006; Kuhn and Gallinat, 2013) and schizophrenia (Adriano et al., 2012; Shepherd et al., 2012). Therefore, addressing the link between stress and hippocampal structure is important in understanding the mechanisms involved in serious mental illness.

Our findings show that it may be relevant to differentiate between life events with different levels of threat when investigating the relationship between stress and the brain. For future research, we also recommend to include functional measures of stress responsiveness, which allow for assessing potential mediatory effects of HPA axis reactivity (see Rabl et al., 2014)

There are several limitations that need to be taken into account when interpreting our findings. First, methodological limitations are a major issue when interpreting life events measures (Johnson, 2005). Regardless of the number of queries in an interview, people gradually forget life events (Paykel, 1997; Brown and Harris, 1982; Harris, 2001). The average participant in our sample had to report life events over a time span of 35 years. One could question the reliability of the LEDS when it is used retrospectively to collect lifetime life event data. However, the LEDS is probably more reliable compared to (retrospective) checklist inventories (Ormel et al., 2001; Hillegers et al., 2004), as the LEDS minimizes recall bias by actively obtaining information in a very structured interview with detailed questions in ten domains. Second, using a relatively small sample size to estimate genetic and environmental sources of variance in bivariate designs reduces statistical power (Posthuma and Boomsma, 2000; Visscher, 2004). Furthermore, the confidence intervals that we reported were wide, diminishing precision of the extent of the estimates. The relatively small size of the sample may also have played a role in the finding that our MZ and DZ twins had a different number of life events. There is no a priori reason we can think of why this would be the case. However, we applied Bonferroni correction for multiple testing to ensure our estimates were robust. Third, females were overrepresented. Although a recent study showed that statistically controlling for ICV eliminates gender differences in hippocampal volume (among others) (Jancke et al., 2015), other studies reported both larger (Inano et al., 2013) and smaller (Fjell et al., 2009) hippocampal volumes in females compared to males, after controlling for ICV. Therefore, in addition to including ICV as a covariate, we also controlled for gender to eliminate any influence it may have on hippocampal volume. Fourth, the twins have been selected previously as control subjects in studies with psychiatric patients, where stringent exclusion criteria with respect to the presence of Axis I or II psychopathology were applied. Therefore, these subjects may not be representative of the general population. Fifth, it was not possible to ascertain the association between hippocampal volume and life events in

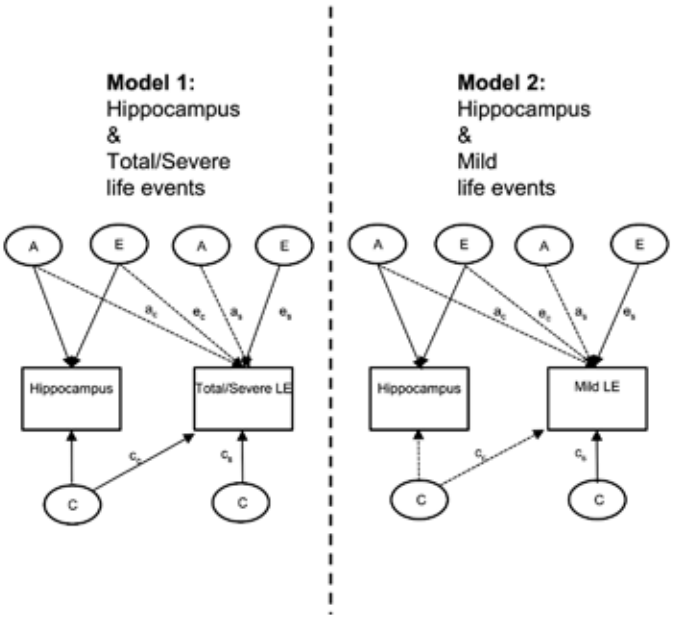
different age groups in our sample, as doing so would have seriously affected statistical power. Nevertheless, it may be conceivable that shared environmental factors have the largest influence during childhood and early adulthood, when the brain (and hippocampus) is also at its most plastic. Therefore, future studies could benefit from analyzing the influences of early versus late life events on hippocampal volume. Last, unfortunately our study does not allow for any statements regarding causality. Therefore, it remains unclear whether a smaller hippocampus is cause or consequence of severe stress. For example, a smaller hippocampal volume could also increase vulnerability to severe life events (but perhaps primarily those that are nonshared among twins), inflating the strength of the association (Gilbertson et al., 2002).

In conclusion, an association was found between smaller hippocampal volume and severe stressful life events, which was attributable to a complete overlap of shared environmental factors influencing both phenotypes. This indicates that severe events that are shared between twins predominantly contribute to both a smaller hippocampal volume and severe life events. In contrast, we did not observe a relationship between hippocampal volume and mild or total life event load. In this respect, our study highlights the importance of addressing the influence of shared environmental factors on the relation between stress and hippocampal volume, and to distinguish between mild and severe stressful life events in life event research.

SUPPLEMENTARY INFORMATION

Supplementary Table S1. Example LEDS report and life event load calculation.

Description event	Threat score	Severity
Buying a house with partner	2	Mild
Daughter moves out of house	2	Mild
Borrowing 4000 euros	2	Mild
Winning 10.000 euros in lottery	3	Severe
House was broken into whilst on vacation, no expensive items missing	2	Mild
Broken arm, in cast for 6 weeks	3	Severe
Son & daughter in law announce pregnancy	1	Mild
Operation for hernia, unable to work for 3 months	4	Severe
Total life event load	19	
Mild life event load	9	
Severe life event load	10	



Supplementary Figure S1. Best fitting models for all three bivariate comparisons. Latent additive genetic (A), shared (C) and unique environmental (E) factors influence the hippocampus and life events, as indicated by arrows. Path coefficients (a_c and a_s) quantify the effects of genetic influences on life events, where a_c represents genetic influences that also influence the hippocampus and a_s represents genetic influences that are unique for life events. Similarly, path coefficients c_c, c_s, e_c and e_s quantify the effect of shared (C) and unique environmental (E) influences on phenotypes, respectively. Dotted arrows indicate genetic and environmental factors that would normally influence the hippocampus and life events under the full ACE-model but did not in the best fitting models at present.

Chapter 3

Contribution of genes and unique environment to cross-sectional and longitudinal measures of subcortical volumes in bipolar disorder

Florian Bootsman, Rachel M. Brouwer, Sanne M. Kemner, Hugo G. Schnack, Astrid C. van der Schot, Ronald Vonk, Manon H.J. Hillegers, Dorret I. Boomsma, Hilleke E. Hulshoff Pol, Willem A. Nolen, René S. Kahn, Neeltje E.M. van Haren

European Neuropsychopharmacology 25, 2197-2209

3.1 ABSTRACT

The influence of genes and environment on the association between bipolar disorder (BD) and volumes of subcortical brain regions involved in emotion processing has rarely been studied. Furthermore, as far as we know, longitudinal twin studies of subcortical brain volume change in BD have not been carried out at all. In this study, we focused on the genetic and environmental contributions to cross-sectional and longitudinal measures of subcortical brain volumes in BD.

A total of 99 twins from monozygotic and dizygotic pairs concordant or discordant for BD and 129 twins from monozygotic and dizygotic healthy control pairs underwent magnetic resonance imaging at baseline. Longitudinal assessment was carried out in 48 twins from monozygotic and dizygotic patient pairs and 52 twins from monozygotic and dizygotic control pairs. Subcortical volume measures were obtained with FreeSurfer software and analyzed with structural equation modeling software OpenMx.

At baseline, BD was phenotypically and genetically associated with smaller volumes of the thalamus, putamen and nucleus accumbens. BD was not associated with subcortical brain volume change over time in any of the examined regions. Heritability of subcortical volumes at baseline was high, whereas subcortical volume change had low heritability.

Genes contributing to BD showed overlap with those associated with smaller volumes of the thalamus, putamen and nucleus accumbens at baseline. Further evaluation of genetic contributions to abnormalities in subcortical brain regions assumed to be involved in emotion processing is recommended.

Keywords: longitudinal twin study, bipolar disorder, magnetic resonance imaging, subcortical brain volumes, heritability

3.2 INTRODUCTION

Abnormal processing of emotion is considered a key feature of bipolar disorder (BD) (Goodwin et al., 2007). In BD, particular attention has been given to abnormalities in subcortical brain regions that are part of or associated with the cortico-striato-thalamic and limbic networks involved in emotion processing, including the amygdala, hippocampus, striatum and thalamus (Phillips et al., 2003, 2008; Emsell and McDonald, 2009; Marchand and Yurgelun-Todd, 2010; Aldhafeeri et al., 2012; Blond et al., 2012; Phillips and Swartz, 2014). However, neuroimaging studies disagree on the extent and variety of morphological abnormalities in BD in some of these regions. For example, both smaller and larger volumes of the amygdala, hippocampus and striatum have been reported (Emsell and McDonald, 2009; Savitz and Drevets, 2009; Rimol et al., 2010; Hajek et al., 2012; Phillips and Swartz, 2014). Furthermore, smaller volumes of the thalamus and nucleus accumbens have been shown in BD (Rimol et al., 2010) although the majority of studies investigating the thalamus report no differences in volume between BD patients and healthy controls (Emsell and McDonald, 2009).

In contrast to the relatively large number of cross-sectional neuroimaging studies investigating subcortical volume, there have only been a few longitudinal studies assessing subcortical volume change over time in BD. Here too, findings are inconclusive. For example, volume preservation, increases and decreases have all been demonstrated in the amygdala and hippocampus (see review by Lim et al., 2013). There is some evidence of volume increase over time in the caudate nucleus and thalamus (Lisy et al., 2011), although there is limited data available. Moreover, measures of brain volume and brain volume change in BD are influenced by lithium use (often resulting in larger volumes in patients), age, familial load, mood status and variability in imaging methodology (Emsell and McDonald, 2009; Savitz and Drevets, 2009; Hallahan et al., 2011; Hajek et al., 2012), which complicates reliable assessment of subcortical abnormalities associated with the disease.

The heritability of BD has been estimated to be up to 85% (McGuffin et al., 2003). The degree to which genes and environmental factors contributing to BD are associated with subcortical brain volumes has not been studied extensively, and, as far as we know, volume change over time in subcortical structures has not been investigated at all. In one cross-sectional study, the genetic risk for BD was associated with smaller volume of the ventral striatum (McDonald et al., 2004). A different study found larger caudate nuclei in discordant monozygotic BD twin pairs compared to monozygotic healthy control twin pairs (Noga et al., 2001), which suggests that genes associated with BD influence volume of this region. Another group investigating subcortical brain regions in a genetically informative cohort found no association between liability to BD and grey matter of the amygdala-hippocampal complex or thalamus (McIntosh et al., 2006).

As far as we know, it has not been studied whether cross-sectional subcortical volume deficits in twins with BD show progressive change over time and to what extent subcortical volume change is related to genetic and environmental factors associated with the disease. Therefore, in this twin study, we set out to assess whether baseline volume and volume change over time in subcortical brain regions is associated with BD in monozygotic (MZ) and dizygotic (DZ) twin pairs concordant and discordant for the disease. In addition, the degree to which genes and environment influence the association between BD and the subcortical measures at baseline and over time is estimated.

3.3 EXPERIMENTAL PROCEDURES

3.3.1 Subjects

A total of 99 twins from pairs concordant or discordant for BD (MZ: 15 discordant and 9 concordant pairs; DZ: 20 discordant and 4 concordant pairs, and 1 patient and 2 co-twins from incomplete pairs) and 129 twins from healthy control pairs (MZ: 37 pairs and 2 twins from incomplete pairs; DZ: 25 pairs and 3 twins from incomplete pairs) were included at baseline. BD patients, their co-twins, and 49 healthy control twins were originally recruited by van der Schot et al. (2009), except for 1 DZ bipolar twin pair that was presently included for the first time. Of the remainder of 80 healthy control twins, 18 twins were taken from the cohort that was included by Baaré et al. (2001) and 62 twins were taken from the cohort that was included by Brans et al. (2008). These healthy twins were originally recruited from the (healthy) twin sample of the department of Psychiatry of the University Medical Center Utrecht and the Netherlands Twin Registry (Boomsma, 1998). Between 2010 and 2013 we conducted follow-up measurements with subjects from this total sample. Ultimately, longitudinal assessment of subcortical brain volume change was carried out in 48 twins from patient pairs (MZ: 10 discordant and 2 concordant pairs, and 1 patient and 2 co-twins from incomplete pairs; DZ: 6 discordant and 2 concordant pairs, and 5 co-twins) and 52 twins from control pairs (MZ: 13 pairs and 6 twins from incomplete pairs; DZ: 8 pairs and 4 twins from incomplete pairs). All twins were raised together, except for one control pair who were separated at 12 years of age when both parents died. The subjects were between 18 and 60 years of age at baseline. Cross-sectional and longitudinal demographic information is presented in **Tables 1a and 1b**. Clinical diagnosis of Axis I psychiatric disorders was confirmed using the Structured Clinical Interview for DSM-IV (SCID) (First et al., 1997), for Axis II personality disorders using the Structured Interview For DSM-IV Personality (SIDP and SCID-II) (Pfohl et al., 1997; First et al., 1997), and for both through available medical records. At both measurements, current mood state was assessed using the Young Mania Rating Scale (YMRS) (Young et al., 1978), Inventory for

Depressive Symptomatology (IDS) (Beck et al., 1961) and the Hamilton Depression Rating Scale (HDRS, at follow-up only) (Hamilton, 1960). At baseline, all patients were euthymic (ie, were not in a depressive, manic, or hypomanic episode) or were in an episode in partial remission with a YMRS score of 4 or less and an IDS score of 12 or less, except for two patients who were (hypo) manic (YMRS scores of 5 and 17, respectively) and eight patients who were mildly to severely depressed (IDS scores of 14, 14, 15, 17, 20, 22, 29, 38, respectively). At follow-up, the majority of patients were euthymic except for one patient who was mildly hypomanic (YMRS score of 14) and two other patients who were mildly depressed (HDRS scores of 12 and 18 respectively).

At baseline, the twin pairs had no history of drug or alcohol dependency for the last six months. At follow-up, three patients met diagnostic criteria for alcohol abuse and/or dependence (one of them also for abuse of cannabis, sedatives and morphine), one patient had a cocaine dependency and two co-twins of patients met diagnostic criteria for alcohol abuse in the six months prior follow-up measurement. None had severe medical illness, verified with a medical history inventory.

The healthy control pairs were matched to the bipolar pairs for zygosity, sex, age, parental education, and birth order. At baseline, healthy control pairs had no history of Axis I psychiatric disorder or Axis II personality disorder according to DSM-IV criteria (SCID and SIDP, respectively) and no history of severe medical illness. Furthermore, they had no first-degree relative with a history of a major Axis I psychiatric disorder (DSM-IV) such as schizophrenia, psychotic disorder, mood disorder, anxiety disorder, or substance-related disorder. However, at follow-up, two control subjects met diagnostic criteria for a major depressive episode, two control subjects were diagnosed with an adjustment disorder (one of which in full remission), one control subject was diagnosed with a specific phobia and three control subjects met diagnostic criteria for alcohol abuse. All clinical ratings were carried out by trained and experienced clinicians (psychologists and psychiatrists) and discussed in weekly meetings until consensus was obtained.

The family histories of both the affected and control twins were obtained via the Family Interview Genetic Studies (FIGS) (Nurnberger jr. et al., 1994) performed with both the proband and co-twin. Zygosity was determined by DNA fingerprinting using high polymorphic microsatellite markers 9 to 11 in the laboratory of the Division Biomedical Genetics, University Medical Center Utrecht. The study was approved by the medical ethics review board of the University Medical Center Utrecht and all participants gave written informed consent after full explanation of the study aims and procedures.

3.3.2 Brain imaging

Magnetic resonance images were acquired on Philips 1.5 Tesla scanners (at baseline Intera, at follow-up Achieva, Philips, the Netherlands). Both scanners were simultaneously used in a

Table 1a. Demographic and clinical characteristics of the bipolar and matched healthy control twin pairs at baseline.

	Bipolar patients and their co-twins (n=99)		**Matched control twins** (n=129)	
	MZ[a] (n=48)	**DZ** (n=51)	**MZ**[b] (n=76)	**DZ** (n=53)
Gender, m/f	14/34	18/33	31/45	24/29
Age, y[c]	37.9 (10.5)	43.5 (8.1)	39.0 (9.8)	39.4 (8.0)
Parental educ., y[d]	10.9 (3.5)	11.4 (3.8)	11.4 (3.3)	11.3 (3.5)
Education,y[e]	12.0 (2.1)	13.3 (2.8)	13.7 (2.6)	13.0 (2.7)
Handedness (right/left or both)	34/14	44/7	62/14	44/9

	Patient	**Co**	**Patient**	**Co**
	(MZ)		(DZ)	
Onset, age	26.5 (8.9)	-	30.8 (9.8)	-
Birth order, $_{1st/2nd}$	15/18	9/6	13/16	13/9
Lithium, *on/off on day MRI, No.*	26/7	0/15	18/11	0/22
Antipsychotics, *on/off on day of MRI*	3/30	0/15	6/23	1/21
Antidepressants, *on/off on day of MRI*	9/24	0/15	8/21	1/21
YMRS score[f]	1.1 (1.5)	0.5 (0.8)	1.0 (3.3)	0.1 (0.7)
IDS score[g]	6.1 (6.6)	2.9 (4.2)	6.0 (8.6)	2.4 (3.7)
Psychotic sympt.	14	1	18	1

Abbreviations: MZ, Monozygotic; DZ, Dizygotic
n= number of individuals
[a] 15 MZ discordant pairs, 9 MZ concordant pairs; 20 DZ discordant pairs, 4 DZ concordant pairs; 1 DZ patient and 2 DZ co-twins from incomplete pairs
[b] 37 MZ healthy control pairs, 2 MZ healthy controls from incomplete pairs; 25 DZ healthy control pairs, 3 DZ healthy controls from incomplete pairs
[c] Significant effect of group [$F_{3,224}$=3.7, p=0.013]. Post-hoc test: DZ patient pairs were significantly older than MZ patient pairs [p=0.016] and MZ control pairs [p=0.045].
[d] Years of parental education could not be determined for 2 bipolar twin pairs (1 MZ, 1 DZ).
[e] Significant effect of group [$F_{3,224}$=4.8, p=0.03]. Post-hoc test: MZ control pairs had significantly more years of education than MZ patient pairs [p=0.001].
[f] YMRS score was not assessed for 2 MZ patients, 2 MZ co-twins, 2 DZ patients and 1 DZ co-twin.
[g] IDS score was not assessed for 2 MZ patients, 2 MZ co-twins, 2 DZ patients and 1 DZ co-twin.

large multicenter collaboration, attesting to their cross-scanner reliability (Hibar et al., 2015). Furthermore, in a test set of 6 subjects that were scanned on both scanners on the same day, we calculated intraclass correlations (ICC) between subcortical measures. Here, ICC's were found to be very high (>0.97), although the small size of the test sample and high variability between the subjects should be considered. The imaging parameters were identical across scanners and measurements: T1-weighted 3D fast field echo scans with 160–180 contiguous coronal slices (echo time=4.6 ms, repetition time=30 ms, flip angle=30°, 1x1x1.2 mm^3 voxels) (van der Schot et al., 2009; Brans et al., 2010). At both assessments, concordant and discordant MZ and DZ patient twins and healthy twins were randomly assigned to MRI slots, eliminating possible between-group biases due to scanner drifts. Processing of brain images and

Table 1b. Demographic and clinical characteristics of the bipolar and matched healthy control twin pairs who were measured at both baseline and follow-up.

	Bipolar patients and their co-twins (n=48)		Matched control twins (n=52)	
	MZ[a] (n=27)	DZ (n=21)	MZ[b] (n=32)	DZ (n=20)
Gender, m/f	7/20	10/11	9/23	8/12
Age at baseline, y	37.9 (12.3)	41.0 (7.2)	39.6 (8.6)	39.6 (6.4)
Interval $_{\text{MRI t0-t1}}$	7.4 (1.4)	7.9 (1.5)	7.4 (1.5)	7.3 (1.1)
Parental educ., y	12.0 (3.1)	12.3 (3.3)	10.6 (3.5)	12.0 (3.9)
Education,y[c]	12.0 (2.3)	13.9 (1.8)	13.7 (2.2)	13.6 (1.7)
Handedness (right/left or both)	19/8	19/2	28/4	19/1

		Patient (MZ)	Co (MZ)	Patient (DZ)	Co (DZ)		
Onset, age		25.4 (10.0)	-	31.3(8.1)	-		
Birth order, $_{\text{1st/2nd}}$		6/9	8/4	4/6	7/4		
Lithium	- both time points	8	0	4	0	0	0
	- never	3	12	3	11	32	20
	- baseline only	3	0	1	0	0	0
	- follow-up only	1	0	2	0	0	0
Antipsychotics	- both time points	2	0	3	1	0	0
	- never	8	12	5	9	32	20
	- baseline only	1	0	0	0	0	0
	- follow-up only	4	0	2	1	0	0
Antidepressants	- both time points	3	0	1	0	0	0
	- never	8	10	7	11	32	20
	- baseline only	3	0	1	0	0	0
	- follow-up only	1	2	1	0	0	0
$YMRS_{base}$[d]		0.7 (0.9)	0.5 (0.8)	0.6 (0.8)	0.3 (0.9)		
IDS_{base}[e]		6.0 (7.3)	3.2 (4.4)	10.4 (12.3)	2.6 (4.0)		
$YMRS_{follow\text{-}up}$[f]		1.5 (1.6)	-	4.1 (4.0)	-		
$HDRS_{follow\text{-}up}$[g]		2.8 (3.5)	2.1 (2.7)	5.0 (5.7)	1.1 (1.9)		
$GAF_{follow\text{-}up}$[h]		71.4 (15.3)	81.3 (9.8)	67.4 (15.7)	84.5 (16.3)		
Psychotic sympt.		5	1	7	1		
Substance abuse $_{\text{follow-up}}$[i]		3	2	1	-	1	2

Abbreviations: MZ, Monozygotic; DZ, Dizygotic
n= number of individuals
[a] 10 MZ discordant pairs, 2 MZ concordant pairs, 1 MZ patient and 2 MZ co-twins; 6 DZ discordant pairs, 2 DZ concordant pairs and 5 DZ co-twins.
[b] 13 MZ healthy control pairs and 6 MZ healthy controls from incomplete pairs; 8 DZ healthy control pairs, 4 DZ healthy controls from incomplete pairs.
[c] Significant effect of group [$F_{3,96}$=4.4, p=0.006]. Post-hoc test: MZ patient pairs had significantly fewer years of education than DZ patient pairs [p=0.015] and MZ control pairs [p=0.018].
[d] At baseline, YMRS score was not assessed for 1 MZ patient and 2 MZ co-twins.
[e] At baseline, IDS score was not assessed for 1 MZ patient and 2 MZ co-twins.
[f] At follow-up, YMRS score was not assessed for 3 MZ patients and none of the co-twins. For 3 MZ patients YMRS score was not assessed on the same day of the MRI scan but approximately 2 months after.
[g] At follow-up, HDRS score was not assessed for 4 MZ patients, 2 MZ co-twins, 2 DZ patients and 1 DZ co-twin. For 3 MZ patients HDRS score was not assessed on the same day of the MRI scan but approximately 2 months after.
[h] At follow-up, GAF score was not assessed for 2 MZ patients and 1 DZ patient.
[i] 6 months prior to follow-up measurement, 3 MZ patients met diagnostic criteria for alcohol abuse and/or dependency (and one of them also for abuse of cannabis, sedatives and morphine), 1 DZ patient had a cocaine dependency, 2 MZ co-twins of patients met criteria for alcohol abuse, and 1 MZ and 2 DZ control twins met diagnostic criteria for alcohol abuse.

subcortical volumetric segmentation was performed with the FreeSurfer structural imaging pipeline (http://surfer.nmr.mgh.harvard.edu/). Anatomic volumes of the bilateral thalamus, caudate nucleus, putamen, pallidum, hippocampus, amygdala and nucleus accumbens were delineated using information on image intensity, probabilistic atlas location and spatial relationships between subcortical structures (Fischl et al., 2002, 2004). Data were processed with the cross-sectional (v5.1.0) and longitudinal (v5.3.0) FreeSurfer pipelines. Subcortical brain volumes were extracted from the FreeSurfer output. Based on visual inspection of subcortical segmentations at baseline and follow-up, brain measures were excluded from the analysis if clear segmentation errors were observed. Left and right volumes were added, rendering total volumes of subcortical brain regions. Change in subcortical brain volume was calculated by subtracting the volumes at baseline from the volumes at follow-up (these volumes were obtained with the longitudinal processing stream of FreeSurfer). The resulting individual measures of volume change for each region were then divided by the number of years between measurements for each subject individually, thus yielding a measure of annual subcortical brain volume change.

3.3.3 Statistical analysis

Quality checking

After quality checking of subcortical brain segments, baseline and change measures were explored for statistical outliers. A number of subjects had outlying values for individual brain measures. If removal of these subjects from analysis did not influence results but did ensure normal distribution of the data, we chose to present results with outliers excluded. However, if outlier removal affected the nature of the association between BD and subcortical volume (change), we presented both the data with and without the outliers removed (see **Tables 2** and **3** for the number of subjects that were analyzed for each brain measure).

Furthermore, we assessed whether intracranial volume (ICV) influenced subcortical brain volumes and found that ICV influenced the baseline but not the change measures and was therefore added as a covariate in the baseline analysis only.

Influence of medication use on subcortical brain measures

As the influence of lithium on the brain in BD has been noted repeatedly (Hafeman et al., 2012), we assessed whether lithium use influenced the subcortical brain measures. At baseline, univariate analysis of variance in BD patients with group ('lithium users' and 'lithium non-users') as between-subject variable, and subcortical brain volume as the dependent variable (after the effects of age, gender and intracranial volume had been regressed out) revealed that lithium users had significantly larger volumes of the thalamus ($p=0.012$) and putamen ($p=0.047$) than patients who did not use lithium. Therefore, at baseline, we chose

to correct for lithium use by adding the difference in means between non-users and users to the values of subjects from the latter group, obtained after regression on age, gender and intracranial volume. This normalized the volumes of the lithium users to the level of the non-users (van der Schot et al., 2009). For completeness, we also assessed the differences in subcortical brain volumes between patients who used antipsychotics or antidepressants and patients who did not. Although patients who used antipsychotics had a smaller thalamus (p=0.036) compared to patients who did not use antipsychotics, correcting for their use did not contribute to volumetric differences between patients and controls (see **Supplementary Table S1** and **Table 2**). Based on this and the fact that the influence of antipsychotics and antidepressants on the brain in BD appears to be limited (Hafeman et al., 2012), we chose not to correct for their use.

For the change analysis, correction for medication use was not applied, given the small sample sizes of the respective medication using groups ('medication use at both measurements', 'started using medication during interval', 'quit using medication during interval' and 'never used medication'), yielding low statistical power to reliably detect group differences.

Univariate analysis of variance

At baseline, univariate analysis of variance was performed with group ('BD patients' versus 'healthy controls' and 'co-twins of patients' versus 'healthy controls') as between-subject variables and subcortical brain volume of each region as the dependent variable, after the effects of age, gender and intracranial volume had been regressed out (and the difference in means between Li- and Li+ patients had been added to the values of the latter group, resulting in volume estimates when no lithium would have been used). For the change measures, we assessed with a linear model whether within-group subcortical volume change in each region as the dependent variable differed significantly from zero, correcting for age at baseline and gender. Subsequently, univariate analysis of variance was performed with group ('BD patients' versus 'healthy controls' and 'co-twins of patients' versus 'healthy controls') as between-subject variable and annual subcortical volume change in each region as the dependent variable, after the effects of age at baseline and gender had been regressed out. Preliminary analysis and rendering of standardized residuals suitable for genetic model fitting was performed with the Statistical Package for the Social Sciences (IBM SPSS, release 21.0.0.0).

Genetic model fitting

To estimate relative genetic and environmental contributions to the association between subcortical brain volume (change) and liability to develop BD, a bivariate liability threshold model was chosen and implemented in structural equation modeling software

OpenMx (Kenny et al., 2009), running under the statistical programming environment R (R Development Core Team, 2008). Here, a bivariate Cholesky decomposition was fitted to the standardized residuals of our brain measures (corrected for age, gender and ICV [ICV correction was applied in baseline measures only], as well as lithium use in patients) to estimate additive genetic (A), and unique environmental (E) variance components of brain volume (changes) and phenotypic, genetic and environmental overlap between BD and the brain change measures. As there is no evidence of shared environment (C) influencing BD (McGuffin et al., 2003), this factor was only estimated for the brain measures and not for BD. Disease status was dichotomous and assumed to represent an underlying continuous liability with a mean (SD) of 0 (1). Patients will have a higher value on the liability scale, thereby crossing a certain threshold (patient status=1). All other individuals will have lower liability scores and not cross the critical threshold (patient status = 0; discordant co-twin of patient or control twin pairs). As we included approximately equal numbers of concordant, discordant, and healthy twin pairs, the critical threshold and heritability (the relative contribution of genetic variance to total variance) for the underlying liability for BD could not be estimated from this sample. Prevalence and heritability of BD were thus fixed to population values. Prevalence was set to 1% (Regeer et al., 2004) and heritability was set to 85% (McGuffin et al., 2003). Varying the prevalence and heritability of BD (e.g. 2% and 75% respectively) did not change the results. To apply the threshold model to the brain measures, the obtained standardized residuals of the subcortical brain measures were rendered into five ordinal categories identical for all subjects – thereby equating them across groups – and put in the model. Thresholding was based on normality plots, with the boundaries of the 'outer' two categories set at -1.5 SD and 1.5 SD respectively, with the other three categories falling in between, being 1 SD wide.

The phenotypic correlation (r_{ph}), an index of association between phenotypes (e.g. liability to develop BD and brain volume (change)), was based on calculations of within-twin/between-trait correlations. Heritability (h^2) and influence of shared and unique environment (c^2, e^2) as well as disentanglement of the observed correlation between liability for BD and subcortical measures into genetic and environmental components was based on polychoric cross-twin/within-trait and cross-twin/cross-trait correlations within MZ and DZ groups (Neale and Miller, 1997). The heritability of brain measures was determined within the bivariate model. A larger correlation between traits in MZ twins than in DZ twins suggests higher genetic contribution due to MZ twins being genetically identical whereas DZ twins only share on average 50% of their segregating genes. If there is no difference between MZ and DZ correlations then a larger influence of shared environmental factors is more likely (Boomsma et al., 2002). The genetic (r_g) and (shared and unique) environmental (r_e) correlations respectively indicate the degree of overlap in genes or (shared or unique) environment influencing phenotypes. The

phenotypic correlation can be written as the sum of the genetic correlation weighted by the square root of the heritabilities of the two traits (r_g*h_{BD}*h_{brain}) and the environmental correlation weighted by the square root of environmental variance associated with the two traits (r_c*c_{BD}*c_{brain}, r_e*e_{BD}*e_{brain}). These quantities are written as $r_{ph\text{-}g}$, $r_{ph\text{-}c}$ and $r_{ph\text{-}e}$ (Toulopoulou et al., 2007). The significance of variance components was tested by fitting different nested models to the data and by comparing their goodness of fit using Akaike's Information Criterion (AIC). A saturated model in which means, variances and correlations are estimated freely served as a baseline model to which more restrictive models were compared. Compared to the saturated model, the AE-model had the best fit in all ROIs and was therefore applied indiscriminately. Significance of parameter estimates and correlations was determined based on 95% confidence intervals (CI) (Neale and Miller, 1997).

Correction for multiple testing

In all tables and figures, significance after Bonferroni correction for multiple testing is indicated. The critical threshold for significance was calculated by dividing the alpha of 0.05 by the number of brain regions assessed, which was 7, yielding an significance threshold of p =0.05/7=0.007.

3.4 RESULTS

3.4.1 Demographic and clinical characteristics

Please refer to **Tables 1a and 1b** for demographic information relevant to the baseline and change analyses, respectively. At baseline, DZ patient pairs were significantly older than MZ patient pairs (p=0.016) and MZ control pairs (p=0.045). Furthermore, at baseline, MZ control pairs had significantly more years of education compared to MZ patient pairs (p=0.001).

3.4.2 Baseline and change analysis of subcortical brain volumes in bipolar disorder

Univariate analysis of variance of baseline and change measures of subcortical volume

Tables 2 and 3 show the raw subcortical brain volumes and volume changes of BD patients, co-twins of patients and healthy controls.

At baseline, univariate analysis of variance revealed smaller volumes of the thalamus ($F_{1,188}$=12.56, p=0.000), putamen ($F_{1,186}$=19.91, p=0.000) and nucleus accumbens ($F_{1,188}$=7.38, p=0.007) in BD patients as compared to healthy controls, when correction for lithium use was applied. Regarding subcortical volume change, BD patients, co-twins and healthy controls showed significant within-group changes in the majority of brain regions, with volume loss in the thalamus, caudate nucleus, putamen and nucleus accumbens (controls

Table 2. Mean uncorrected subcortical volumes for BD patients, co-twins of patients and healthy controls, and between-group comparisons with data uncorrected and corrected for use of lithium and antipsychotics.

	Mean (SD) uncorrected volume			Statistics							
	Patients[+]	Co-twins	Healthy controls[++]	Patients vs HC*				Co-twins vs HC			
Region (volume, in ml)				*F*	*df*	*p*	*Δ*	*F*	*df*	*p*	*Δ*
ICV[a]	1511.26 (193.4)	1547.17 (188.99)	1524.01 (134.55)	0.15	1, *189*	0.699		2.39	1, *164*	0.124	
Thalamus[b]	15.02 (1.82)	15.06 (1.55)	15.17 (1.37)	0.09	1, *188*	0.765		0.58	1, *163*	0.449	
				12.56	**1, *188***	**0.000**[†]	**pt<ctrl**				
				1.33	1, *188*	0.251					
Caudate Nucleus[c]	6.63 (0.74)	6.91 (0.72)	6.80 (0.79)	0.54	1, *181*	0.465		0.83	1, *161*	0.364	
				2.40	1, *181*	0.123					
				0.12	1, *181*	0,726					
Putamen[d]	9.42 (1.30)	9.62 (1.40)	9.82 (1.18)	2.86	1, *186*	0.092		1.11	1, *161*	0.293	
				19.91	**1, *186***	**0.000**[†]	**pt<ctrl**				
				2.57	1, 186	0.111					
Pallidum[e]	3.02 (0.38)	3.07 (0.34)	3.12 (0.30)	3.64	1, *185*	0.058		1.58	1, *161*	0.210	
				0.11	1, *185*	0.741					
				2.96	1, *185*	0.087					
Hippocampus[f]	8.32 (0.95)	8.43 (0.97)	8.52 (0.74)	1.57	1, *179*	0.212		0.98	1, *158*	0.323	
				3.17	1, *179*	0.077					
				0.96	1, *179*	0.329					
Amygdala[g]	3.03 (0.42)	2.99 (0.33)	3.07 (0.35)	0.00	1, *189*	0.952		1.42	1, *163*	0.235	
				2.78	1, *189*	0.097					
				0.02	1, *189*	0.904					
Nucleus Accumbens[h]	0.94 (0.15)	0.96 (0.15)	0.98 (0.15)	1.20	1, *188*	0.275		1.27	1, *164*	0.261	
				7.38	**1, *188***	**0.007**	**pt<ctrl**				
				1.08	1, *188*	0.299					

Note. Table depicts uncorrected subcortical volumes at baseline (in milliliter). Univariate analysis of variance was performed with group (patients versus healthy controls and co-twins versus healthy controls) as independent variable and subcortical brain volume of each region as dependent variable, after the effects of age, gender and intracranial volume had been regressed out (and the difference in means between Li- and Li+ patients and between Ap- and Ap+ patients had been added to the values of the Li+ or Ap+ group, resulting in volume estimates when no lithium or antipsychotics had been used).

Li +/- = Lithium use yes/no; Ap+/- = Antipsychotic use yes/no.

[+] Including concordant pairs, [++] including both twins from complete pairs.

*In the analysis between patients and controls, the values of the subcortical brain volumes of the patients had been either uncorrected for medication use (first row), corrected for lithium use (second row) or corrected for antipsychotic use (third row).

[†] significant at α=0.007 (Bonferroni threshold), p values in bold face are significant at α=0.05.

n= [a] 62 patients, 37 co-twins, 129 healthy controls; [b]62 patients, 37 co-twins, 128 healthy controls (1 subject removed); [c] 57 patients, 37 co-twins and 126 healthy controls (8 subjects removed); [d] 62 patients, 37 co-twins and 126 healthy controls (3 subjects removed); [e] 61 patients, 37 co-twins and 126 healthy controls (4 subjects removed); [f] 58 patients, 37 co-twins and 123 healthy controls (10 subjects removed); [g] 62 patients, 36 co-twins and 129 healthy controls (1 subject removed); [h] 61 patients, 37 co-twins and 129 healthy controls (1 subject removed).

Table 3. Mean uncorrected annual subcortical volume change for BD patients, co-twins of patients and healthy controls.

	Mean (SD) volume change						Statistics							
	Patients[+]	p[±]	Co-twins	p[±]	Healthy controls[++]	p[±]	Patients vs HC				Co-twins vs HC			
Region (volume change, in ml/year x 1000)							*F*	*df*	*p*	*Δ*	*F*	*df*	*p*	*Δ*
Thalamus [a]	-29.34 (50.01)	**0.013**	-31.71 (46.50)	**0.005**[†]	-37.72 (36.54)	**0.000**[†]	0.57	1,*72*	0.455		0.73	1,*70*	0.394	
Caudate Nucleus [b]	-50.80 (33.72)	**0.000**[†]	-38.86 (30.99)	**0.000**[†]	-47.62 (29.79)	**0.000**[†]	0.04	1,*70*	0.834		1.49	1,*71*	0.226	
Putamen [c]	-16.07 (21.85)	**0.002**[†]	-12.85 (23.89)	**0.017**	-18.32 (24.33)	**0.000**[†]	0.19	1,*72*	0.661		0.85	1,*70*	0.360	
Pallidum [d]	7.96 (15.85)	**0.029**	9.11 (12.63)	**0.003**[†]	7.50 (11.41)	**0.000**[†]	0.06	1,*73*	0.808		0.16	1,*72*	0.687	
Hippocampus [e]	43.46 (29.75)	**0.000**[†]	42.81 (26.23)	**0.000**[†]	41.15 (23.02)	**0.000**[†]	0.23	1,*74*	0.634		0.05	1,*72*	0.827	
Amygdala [f]	4.64 (16.76)	0.128	3.12 (12.06)	0.232	5.33 (13.97)	**0.010**	0.02	1,*74*	0.881		0.46	1,*72*	0.501	
Nucleus Accumbens [g]	0.86 (7.63)	0.582	1.98 (9.30)	0.204	-2.58 (7.51)	**0.020**	3.16	1,*74*	0.079		**5.47**	**1,72**	**0.022**	**co>ctrl**
Total brain volume [h]	-1477.47 (2465.39)	**0.004**[†]	-1595.34 (2022.33)	**0.002**[†]	-1922.40 (1754.02)	**0.000**[†]	1	1,*73*	0.321		0.60	1,*71*	0.443	
No outliers removed														
Thalamus	-19.19 (59.43)	0.130	-30.71 (64.44)	**0.011**	-40.66 (41.93)	**0.000**[†]	2.83	1,*75*	0.097		1.36	1,*73*	0.248	

Note. Table depicts uncorrected annual subcortical volume changes in ml/year, multiplied by a factor of 1000 to enhance readability. When removal of outliers changed the nature of the association between BD and subcortical volume change, results from both models are given. Univariate analysis of variance was performed with group (patient, co-twin and healthy control; patients versus healthy controls and co-twins versus healthy controls) as independent variable and annual subcortical brain volume change of each region as independent variable, after the effects of age, and gender had been regressed out.

[+] Including concordant pairs, [++] including both twins from complete pairs.

[±] within-group annual change for each region, corrected for age at baseline and gender. P-value indicates whether within-group mean volume change differs significantly from zero.

[†] significant at α=0.007 (Bonferroni threshold), p values in bold face are significant at α=0.05.

n= [a] 23 patients, 21 co-twins, 51 healthy controls (5 subjects removed); [b] 21 patients, 22 co-twins and 51 healthy controls (6 subjects removed); [c] 23 patients, 21 co-twins and 51 healthy controls (5 subjects removed); [d] 23 patients, 22 co-twins and 52 healthy controls (3 subjects removed); [e] 25 patients, 23 co-twins and 51 healthy controls (1 subject removed); [f] 25 patients, 23 co-twins and 51 healthy controls (1 subject removed); [g] 25 patients, 23 co-twins and 51 healthy controls (1 subject removed); [h] 24 patients, 22 co-twins and 51 healthy controls (3 subjects removed).

Table 4a. Genetic/environmental influences (AE-model, with 95% confidence intervals) on cross-sectional measures of subcortical brain volume, and phenotypic, genetic and environmental correlations with bipolar disorder, *uncorrected* for lithium use in patients.

	h^2	e^2	r_{ph}	r_g	r_e	r_{ph-g}	r_{ph-e}
Region (volume)	%						
Thalamus	60	40	0.02	-0.07	0.31	-0.05	0.08
	(39 to 75)	(25 to 61)	(-0.10 to 0.15)	(-0.29 to 0.13)	(-0.11 to 0.66)	(-0.20 to 0.09)	(-0.03 to 0.17)
Caudate Nucleus	84	16	-0.08	-0.05	-0.25	-0.04	-0.04
	(70 to 91)	(9 to 30)	(-0.21 to 0.06)	(-0.22 to 0.13)	(-0.66 to 0.22)	(-0.19 to 0.11)	(-0.11 to 0.03)
Putamen	74	26	-0.12	**-0.20**	0.17	**-0.16**	0.03
	(58 to 84)	(16 to 42)	(-0.25 to 0.01)	(-0.37 to -0.02)	(-0.20 to 0.52)	(-0.29 to -0.01)	(-0.04 to 0.11)
Pallidum	64	36	-0.11	-0.12	-0.10	-0.09	-0.02
	(43 to 78)	(22 to 57)	(-0.24 to 0.02)	(-0.31 to 0.08)	(-0.48 to 0.29)	(-0.23 to 0.06)	(-0.12 to 0.07)
Hippocampus	84	16	-0.08	-0.13	0.17	-0.11	0.03
	(71 to 92)	(8 to 29)	(-0.21 to 0.05)	(-0.30 to 0.04)	(-0.29 to 0.58)	(-0.25 to 0.03)	(-0.05 to 0.10)
Amygdala	69	31	0	-0.03	0.08	-0.02	0.02
	(52 to 81)	(19 to 48)	(-0.13 to 0.13)	(-0.22 to 0.16)	(-0.33 to 0.46)	(-0.17 to 0.13)	(-0.07 to 0.10)
Nucleus Accumbens	64	36	-0.12	-0.18	0.05	-0.13	0.01
	(45 to 77)	(23 to 55)	(-0.25 to 0.01)	(-0.37 to 0.02)	(-0.34 to 0.45)	(-0.27 to 0.01)	(-0.08 to 0.10)

†Significant at α=0.007 (Bonferroni threshold), estimates in bold face are significant at α=0.05.

Table 4b. Genetic/environmental influences (AE-model, with 95% confidence intervals) on cross-sectional measures of subcortical brain volume, and phenotypic, genetic and environmental correlations with bipolar disorder, *corrected* for lithium use in patients.

	h^2	e^2	r_{ph}	r_g	r_e	r_{ph-g}	r_{ph-e}
Region (volume)	%						
Thalamus	64	36	**-0.20**†	**-0.21**	-0.20	**-0.16**	-0.05
	(45 to 78)	(22 to 55)	(-0.32 to -0.07)	(-0.41 to -0.02)	(-0.59 to 0.23)	(-0.30 to -0.01)	(-0.14 to 0.05)
Caudate Nucleus	83	17	-0.13	-0.07	-0.45	-0.06	-0.07
	(69 to 91)	(9 to 31)	(-0.27 to 0)	(-0.25 to 0.10)	(-0.84 to 0.02)	(-0.21 to 0.09)	(-0.14 to 0)
Putamen	80	20	**-0.28**†	**-0.29**†	-0.23	**-0.24**†	-0.04
	(67 to 89)	(11 to 33)	(-0.40 to -0.15)	(-0.45 to -0.12)	(-0.57 to 0.16)	(-0.37 to -0.10)	(-0.11 to 0.03)
Pallidum	61	39	-0.02	-0.04	0.04	-0.03	0.01
	(40 to 76)	(24 to 60)	(-0.15 to 0.11)	(-0.24 to 0.16)	(-0.33 to 0.41)	(-0.17 to 0.11)	(-0.08 to 0.10)
Hippocampus	85	15	-0.11	-0.14	0.07	-0.12	0.01
	(72 to 93)	(7 to 28)	(-0.24 to 0.02)	(-0.31 to 0.03)	(-0.39 to 0.51)	(-0.26 to 0.02)	(-0.06 to 0.08)
Amygdala	71	29	-0.11	-0.10	-0.13	-0.08	-0.03
	(55 to 82)	(18 to 45)	(-0.23 to 0.03)	(-0.29 to 0.09)	(-0.52 to 0.28)	(-0.22 to 0.07)	(-0.11 to 0.06)
Nucleus Accumbens	66	34	**-0.22**†	**-0.25**	-0.11	**-0.19**	-0.03
	(47 to 79)	(21 to 53)	(-0.34 to -0.08)	(-0.45 to -0.06)	(-0.50 to 0.31)	(-0.33 to -0.04)	(-0.12 to 0.07)

†Significant at α=0.007 (Bonferroni threshold), estimates in bold face are significant at α=0.05.

only), and volume increase in the pallidum, hippocampus and amygdala (controls only). Subcortical volume change was not significantly different between groups, except for a more pronounced increase in volume of the nucleus accumbens over time in co-twins compared to healthy controls ($F_{1, 72}$=5.47, p=0.022).

Association between bipolar disorder and subcortical brain volumes at baseline

Supplementary Tables S2 and S3, Table 4 and Figure 1 show the genetic model estimates for the baseline subcortical volumes, and phenotypic, genetic and unique environmental associations with BD. BD was phenotypically and genetically associated with smaller volumes of the thalamus (r_{ph}= -0.20, r_g= -0.21, r_{ph-g}= -0.16), putamen (r_{ph}= -0.28, r_g= -0.29, r_{ph-g}= -0.24) and nucleus accumbens (r_{ph}= -0.22, r_g= -0.25, r_{ph-g}= -0.19), when correction for lithium use was applied. This indicates that genes contributing to BD also contribute to smaller volumes of these regions. Furthermore, subcortical volumes at baseline showed high heritability in general (range h^2: 61% [Pallidum] to 85% [Hippocampus]).

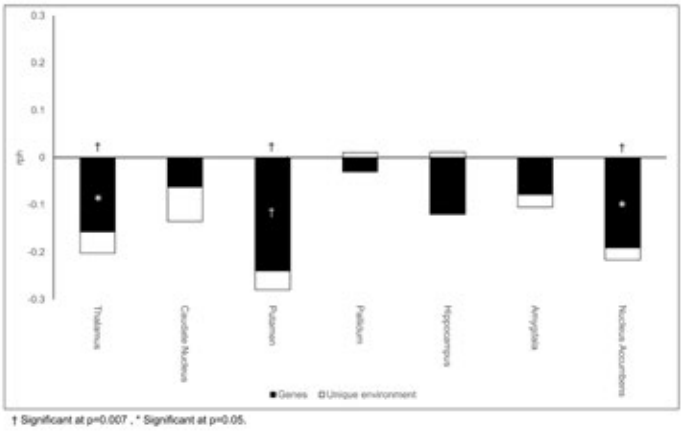


Figure 1. Genetic and unique environmental contributions to the phenotypic correlation between BD and subcortical brain volumes. Level of significance is indicated for the total correlation (outer symbols), and for genetic and unique environmental contributions (inner symbols) separately.

Association between bipolar disorder and subcortical brain volume changes

Please refer to **Supplementary Tables S3 and S4** and **Supplementary Figure S1** for the genetic model estimates for the subcortical volume change measures, and phenotypic, genetic and unique environmental associations with BD. There was a phenotypic and genetic association between BD and volume change of the nucleus accumbens ($r_{ph:}$= 0.2, r_g= 1, r_{ph-g}= 0.3). This effect was in large part due to the difference in volume change between co-twins of patients and healthy controls, whereas patients did not differ significantly from healthy controls but did show a trend level increase in the nucleus accumbens (p=0.079, see **Table 3**). In general, subcortical volume change was predominantly influenced by unique environment, as the heritability of subcortical brain change was very small (range h^2: 0% [hippocampus] to 11% [nucleus accumbens]), except for the caudate nucleus that showed moderate heritability (h^2:43%) when outliers were removed (with no outliers removed h^2 was 0%).

Association of clinical measures with baseline and change measures of subcortical volume in patients

In those regions that were phenotypically correlated to BD (i.e. thalamus, putamen and nucleus accumbens at baseline and change in the nucleus accumbens over time) we found no association with number of hospitalizations, lifetime experience of psychotic symptoms, YMRS score, IDS score (available for the first measurement only), HDRS score (available for the second measurement only) or Global Assessment of Functioning score (GAF, available for the second measurement only) in patients.

3.5 DISCUSSION

In this longitudinal twin study investigating subcortical brain volume in BD, we concentrated specifically on the contribution of genes and environment to the association between BD and baseline and change measures of subcortical brain volumes.

The main finding is that, at baseline, BD was associated with smaller volumes of the thalamus, putamen and nucleus accumbens when we corrected for lithium use. This finding stresses the importance of taking into the account the influence of lithium on the brain as its use may mask brain abnormalities associated with BD. Volumes of these regions were strongly influenced by genes associated with BD. Our findings in twins confirm results from a previous study with BD patients and their first-degree relatives, that found an association between the genetic liability to BD and smaller volume of the ventral striatum, including the anterior putamen (McDonald et al., 2004). Similarly, less grey matter density has been demonstrated in the anterior thalamus in BD patients and unaffected relatives, which the authors interpreted to reflect an association with genetic liability to psychosis in general, as schizophrenia patients and relatives showed the same thalamic grey matter deficit (McIntosh et al., 2004). However, findings across cross-sectional case-control studies are inconsistent. For example, Hibar et al. (2013) and Rimol et al. (2010) did show smaller volume of the thalamus but most studies assessing thalamic volume do not report abnormalities in this structure in BD (Emsell and McDonald, 2009; Hallahan et al., 2011; Womer et al., 2014). Here, the difficulty of isolating thalamic nuclei with current available methods may contribute to the inconsistency of findings (Blond et al., 2012). However, differences across studies could possibly also be attributed to the presence of different disease mechanisms or the inclusion of heterogeneous clinical samples in studies where correction for medication use may or may not have been applied (Emsell and McDonald, 2009; Savitz and Drevets, 2009; Hajek et al., 2012). Furthermore, smaller volumes of the left (Almeida et al., 2009) and right (Haller et al., 2011) putamen have been found previously, as have larger total (DelBello et al., 2004)

and right side (Hallahan et al., 2011) volumes in this region, as well as no differences between BD patients and controls in putamen volume (Womer et al., 2014). Our finding of smaller volume of the nucleus accumbens has also been demonstrated by others (Dickstein et al., 2005; Rimol et al., 2010; Haller et al., 2011) but so has tissue preservation in this structure (Womer et al., 2014).

As the thalamus and striatum are particularly important nodes in the cortico-striato-thalamic loops involved in emotion regulatory processes (Emsell and McDonald, 2009; Marchand and Yurgelun-Todd, 2010; Blond et al., 2012; Strakowski et al., 2012), structural deficiency in these and associated regions could underlie BD (Emsell and McDonald, 2009; Blond et al., 2012), which necessitates closer examination. Moreover, based on this and several other studies, abnormal morphology in emotion processing areas may share some genetic variance with BD (Noga et al., 2001; McDonald et al., 2004; McIntosh et al., 2004), although McIntosh et al (2006) noted an absence of a genetic association between BD and measures of the amygdala-hippocampal complex or thalamus. Therefore, future studies could benefit from evaluating to what extent volume deficiencies in subcortical brain regions are genetically and environmentally mediated, in order to determine by which mechanisms affective dysregulation develops. Here, genome-wide association studies (GWAS) could assist in identifying gene pools associated with subcortical structural abnormalities and BD. For example, specific genetic variants influencing putamen volume were identified recently (Hibar et al., 2015), although these variants do not appear to confer risk for BD (Mühleisen et al., 2014).

In our study, BD was not associated with subcortical volume change over time in any of the examined regions (including those that showed an association with BD at baseline), except for the nucleus accumbens where a significant association was found. However, this association was no longer significant after Bonferroni correction. These findings suggest BD may not be a progressive brain disease. However, due to the relatively small sample size of the group that was assessed longitudinally, this finding should be interpreted with caution. Perhaps with a larger sample a Bonferroni significant increase in volume of the nucleus accumbens would have been found. Previous longitudinal case-control studies showed volumetric preservation, increases and decreases over time in the hippocampus (see review by Lim et al., 2013), as well as thalamic and caudate nucleus increases over time in BD (Lisy et al., 2011). Furthermore, it has been suggested that volumetric abnormalities in the amygdala arise during neurodevelopment but remain stable during adulthood in BD (Lim et al., 2013). However, based on our results, we cannot fully support this hypothesis as we found no difference in amygdala volume between BD patients and healthy controls at baseline.

At baseline, subcortical brain volumes showed high heritability, which is in line with studies by Bohlken et al. (2014) and den Braber et al. (2013) that reported similar heritability estimates

in the same brain regions. In contrast, heritability of subcortical brain volume change was generally low. In part this may reflect measurement error inherent to low statistical power, using a relatively small sample for the change analyses, but could to some extent suggest subcortical brain change to be particularly influenced by factors unique to the individual rather than genes.

A number of limitations are relevant to our study. First, two different MRI scanners were used for the two measurements, therefore we cannot completely rule out the presence of scanner effects on our measures. However, we ensured scanner field strength (1.5 Tesla), imaging parameters and (pre)processing algorithms to have been equal for all subjects across measurements. In addition, all baseline scans were obtained on one scanner while all follow-up scans were obtained on the other scanner. So, within-twin effects are controlled for. Second, females were overrepresented in patient and control pairs. Therefore, subcortical brain measures were corrected for gender. Third, using a relatively small sample size, particularly in the change analyses, reduces the statistical power to estimate genetic and environmental sources of variance in univariate and bivariate designs (Visscher, 2004; Posthuma and Boomsma, 2000). However, this is particularly relevant when heritability is expected to be low, which was not the case for our baseline measures. Moreover, in a few regions the associations between BD and subcortical volume were highly significant, surviving Bonferroni correction. Therefore, we are confident to have had sufficient power for baseline analysis of subcortical volume in BD. Fourth, although we corrected for lithium use in the baseline measures, we did not for the change measures, due to the small sample sizes of the respective lithium using groups.

In summary, BD was associated with smaller volumes of the thalamus, putamen and nucleus accumbens at baseline, with genes contributing to the disease influencing the volumes of these regions. In contrast, BD was not associated with subcortical volume change over time after Bonferroni correction, indicating no differences in change between patients, co-twins and healthy controls. Further evaluation of genetic and environmental contributions to structural brain abnormalities in BD is recommended, in particular regarding subcortical volumes assumed to be involved in emotion processing, to ascertain by which mechanisms affective dysregulation develops.

Supplementary Table S1. Mean uncorrected subcortical volumes and between-group comparisons for BD patients who had used medication and who had not used medication at baseline.

	Lithium use					Antipsychotic use					Antidepressant use				
Region (volume, in ml)	**Li+** (n=44)	**Li-** (n=18)	**F**	**df**	**p**	**Ap+** (n=9)	**Ap-** (n=53)	**F**	**df**	**p**	**Ad+** (n=17)	**Ad-** (n=45)	**F**	**df**	**p**
Thalamus	15.15 (1.72)	14.69 (2.07)	6.73	1, *60*	**0.012**	14.37 (1.72)	15.13 (1.83)	4.61	1, *60*	**0.036**	15.13 (1.85)	14.97 (1.83)	0.85	1, *60*	0.361
Caudate Nucleus[a, b, c]	6.63 (0.70)	6.63 (0.85)	0.45	1, *55*	0.505	6.27(0.51)	6.68 (0.76)	1.86	1, *55*	0.178	6.87 (0.60)	6.54 (0.77)	2.26	1, *55*	0.138
Putamen	9.50 (1.23)	9.23 (1.46)	4.11	1, *60*	**0.047**	9.33 (1.53)	9.44 (1.27)	0.06	1, *60*	0.807	9.57 (1.37)	9.36 (1.28)	1.67	1, *60*	0.202
Pallidum[d, e, f]	2.97 (0.36)	3.15 (0.43)	2.61	1, *59*	0.111	3.0 (0.43)	3.02 (0.38)	0.31	1, *59*	0.582	3.04 (0.30)	3.01 (0.41)	0.32	1, *59*	0.571
Hippocampus[g, h, i]	8.31(0.88)	8.35 (1.15)	0.14	1, *56*	0.706	8.21 (0.66)	8.34 (1.0)	0.57	1, *56*	0.455	8.25 (1.12)	8.35 (0.90)	0.05	1, *56*	0.818
Amygdala	3.04 (0.41)	3.0 (0.47)	1.86	1, *60*	0.178	3.03 (0.47)	3.03 (0.42)	0.03	1, *60*	0.857	3.03 (0.50)	3.03 (0.39)	0.16	1, *60*	0.695
Nucleus Accumbens[j, k, l]	0.95 (0.16)	0.93 (0.13)	1.74	1, *59*	0.192	0.94 (0.15)	0.95 (0.15)	0.03	1, *59*	0.872	0.92 (0.16)	0.95 (0.15)	0.17	1, *59*	0.679

Note. Table depicts uncorrected subcortical volumes at baseline (in milliliter). Univariate analysis of variance was performed to assess differences between patients who had used medication (i.e., lithium, antipsychotics or antidepressants) and patients who had not used medication, with group (Li+/Ap+/Ad+ vs Li-/Ap-/Ad-) as between-group variable and subcortical brain volume for each region as the dependent variable, after the effects of age, gender and intracranial volume had been regressed out.

Li+/- = Lithium use yes/no; Ap+/- = Antipsychotic use yes/no; Ad+/- = Antidepressant use yes/no.

p values in bold face are significant at α=0.05.

[a] 41 BD patients who used lithium and 16 BD patients who did not use lithium;[b] 7 BD patients who used antipsychotics and 50 BD patients who did not use antipsychotics;[c] 16 BD patients who used antidepressants and 41 BD patients who did not use antidepressants; [d] 44 BD patients who used lithium and 17 BD patients who did not use lithium; [e] 8 BD patients who used antipsychotics and 53 BD patients who did not use antipsychotics; [f] 16 BD patients who used antidepressants and 45 BD patients who did not use antidepressants;[g] 42 BD patients who used lithium and 16 BD patients who did not use lithium;[h] 9 BD patients who used antipsychotics and 49 BD patients who did not use antipsychotics; [i] 16 BD patients who used antidepressants and 42 BD patients who did not use antidepressants; [j] 43 BD patients who used lithium and 18 BD patients who did not use lithium; [k] 9 BD patients who used antipsychotics and 52 BD patients who did not use antipsychotics; [l] 16 BD patients who used antidepressants and 45 BD patients who did not use antidepressants

Supplementary Table S2a. Genetic and (common and unique) environmental influences (ACE-model, with 95% confidence intervals) on cross-sectional measures of subcortical brain volume, and phenotypic, genetic and environmental correlations with bipolar disorder, *uncorrected* for lithium use in patients.

	h^2	c^2	e^2	r_{ph}	r_g	r_c	r_e	$r_{ph\text{-}g}$	$r_{ph\text{-}c}$	$r_{ph\text{-}e}$
Region (volume)		%								
Thalamus	60	0	40	0.02	-0.07	0	0.31	-0.05	0	0.08
	(24 to 75)	(0 to 27)	(25 to 61)	(-0.10 to 0.15)	(-0.29 to 0.13)	(0 to 0)	(-0.11 to 0.66)	(-0.20 to 0.09)	(0 to 0)	(-0.03 to 0.17)
Caudate Nucleus	84	0	16	-0.08	-0.05	0	-0.25	-0.04	0	-0.04
	(56 to 91)	(0 to 25)	(9 to 30)	(-0.21 to 0.06)	(-0.22 to 0.13)	(0 to 0)	(-0.66 to 0.22)	(-0.19 to 0.11)	(0 to 0)	(-0.11 to 0.03)
Putamen	74	0	26	-0.12	**-0.20**	0	0.17	**-0.16**	0	0.03
	(39 to 84)	(0 to 31)	(16 to 42)	(-0.25 to 0.09)	(-0.39 to -0.02)	(0 to 0)	(-0.20 to 0.52)	(-0.29 to -0.01)	(0 to 0)	(-0.04 to 0.11)
Pallidum	64	0	36	-0.11	-0.12	0	-0.10	-0.09	0	-0.02
	(19 to 78)	(0 to 0)	(22 to 57)	(-0.24 to 0.02)	(-0.34 to 0.08)	(0 to 0)	(-0.48 to 0.29)	(-0.23 to 0.06)	(0 to 0)	(-0.12 to 0.07)
Hippocampus	84	0	16	-0.08	-0.13	0	0.17	-0.11	0	0.03
	(43 to 92)	(0 to 38)	(8 to 29)	(-0.21 to 0.05)	(-0.31 to 0.04)	(0 to 0)	(-0.29 to 0.58)	(-0.25 to 0.03)	(0 to 0)	(-0.05 to 0.10)
Amygdala	69	0	31	0	-0.03	0	0.08	-0.02	0	0.02
	(15 to 81)	(0 to 47)	(19 to 48)	(-0.13 to 0.13)	(-0.25 to 0.18)	(0 to 0)	(-0.33 to 0.46)	(-0.17 to 0.13)	(0 to 0)	(-0.07 to 0.10)
Nucleus Accumbens	47	15	38	-0.12	-0.20	0	0.04	-0.13	0	0.01
	(0 to 77)	(0 to 59)	(23 to 58)	(-0.24 to 0.01)	(-1 to 0.03)	(0 to 0)	(-0.34 to 0.44)	(-0.27 to 0.02)	(0 to 0)	(-0.09 to 0.10)

†Significant at α=0.007 (Bonferroni threshold), estimates in bold face are significant at α=0.05.

Supplementary Table S2b. Genetic and (common and unique) environmental influences (ACE-model, with 95% confidence intervals) on cross-sectional measures of subcortical brain volume, and phenotypic, genetic and environmental correlations with bipolar disorder, *corrected* for lithium use in patients.

	h^2	c^2	e^2	r_{ph}	r_g	r_c	r_e	$r_{ph\text{-}g}$	$r_{ph\text{-}c}$	$r_{ph\text{-}e}$
Region (volume)		%								
Thalamus	64	0	36	**-0.20**†	**-0.21**	0	-0.20	**-0.16**	0	-0.05
	(25 to 78)	(0 to 31)	(22 to 55)	(-0.32 to -0.07)	(-0.45 to -0.02)	(0 to 0)	(-0.59 to 0.23)	(-0.30 to -0.01)	(0 to 0)	(-0.14 to 0.05)
Caudate Nucleus	83	0	17	-0.13	-0.07	0	-0.45	-0.06	0	-0.07
	(55 to 91)	(0 to 25)	(9 to 31)	(-0.27 to 0)	(-0.25 to 0.10)	(0 to 0)	(-0.84 to 0.02)	(-0.21 to 0.09)	(0 to 0)	(-0.14 to 0)
Putamen	80	0	20	**-0.28**†	**-0.29**†	0	-0.23	**-0.24**†	0	-0.04
	(37 to 89)	(0 to 39)	(11 to 33)	(-0.40 to -0.15)	(-0.52 to -0.12)	(0 to 0)	(-0.57 to 0.16)	(-0.37 to -0.10)	(0 to 0)	(-0.11 to 0.03)
Pallidum	61	0	39	-0.02	-0.04	0	0.04	-0.03	0	0.01
	(13 to 76)	(0 to 38)	(24 to 60)	(-0.15 to 0.11)	(-0.26 to 0.16)	(0 to 0)	(-0.33 to 0.41)	(-0.17 to 0.11)	(0 to 0)	(-0.08 to 0.10)
Hippocampus	85	0	15	-0.11	-0.14	0	0.07	-0.12	0	0.01
	(46 to 93)	(0 to 36)	(7 to 28)	(-0.24 to 0.02)	(-0.32 to 0.03)	(0 to 0)	(-0.39 to 0.51)	(-0.26 to 0.02)	(0 to 0)	(-0.06 to 0.08)
Amygdala	65	6	29	-0.11	-0.11	0	-0.13	-0.08	0	-0.03
	(9 to 82)	(0 to 53)	(18 to 47)	(-0.23 to 0.03)	(-0.41 to 0.10)	(0 to 0)	(-0.52 to 0.28)	(-0.22 to 0.07)	(0 to 0)	(-0.11 to 0.06)
Nucleus Accumbens	58	8	34	**-0.22**†	**-0.27**	0	-0.12	**-0.19**	0	-0.03
	(2 to 79)	(0 to 54)	(21 to 55)	(-0.34 to -0.08)	(-1 to -0.06)	(0 to 0)	(-0.50 to 0.31)	(-0.33 to -0.04)	(0 to 0)	(-0.12 to 0.07)

†Significant at α=0.007 (Bonferroni threshold), estimates in bold face are significant at α=0.05.

Supplementary Table S3. Values of Akaike's Information Criterion of all the genetic models, for baseline and change analyses, with and without lithium correction (baseline only).

	Model			
Region (volume)	**ACE**	**AE**	**CE**	**E**
Baseline measures				
Lithium uncorrected data				
Thalamus	218.455	216.455	289.217	303.228
Caudate Nucleus	199.971	197.971	279.735	311.371
Putamen	210.515	208.515	287.409	313.398
Pallidum	215.741	213.741	284.874	301.62
Hippocampus	176.125	174.125	254.057	291.789
Amygdala	211.861	209.861	279.576	308.481
Nucleus Accumbens	233.471	231.771	299.158	324.498
Lithium corrected data				
Thalamus	215.819	213.819	286.144	303.601
Caudate Nucleus	196.064	194.064	275.286	306.271
Putamen	185.305	183.305	261.531	297.142
Pallidum	225.845	223.845	293.062	308.985
Hippocampus	178.274	176.274	257.252	295.345
Amygdala	211.516	209.56	278.975	311.624
Nucleus Accumbens	215.786	213.862	283.773	307.933
Change measures				
Thalamus	127.626	125.626	140.028	138.028
Caudate Nucleus	139.635	137.635	151.951	151.659
Putamen	141.148	139.148	153.344	151.344
Pallidum	147.349	145.349	159.464	157.464
Hippocampus	141.012	139.012	153.088	151.088
Amygdala	148.873	147.015	161.16	159.336
Nucleus Accumbens	136.331	134.331	155.527	153.527
Total brain	143.786	142.55	157.074	157.779
No outliers removed				
Thalamus	141.348	139.531	154.566	152.913

Note. Table depicts Akaike's Information Criterion (AIC) for each model for each brain region separately, for both lithium corrected and uncorrected data (baseline) and change data.

Supplementary Table S4a. Genetic/environmental influences (ACE-model, with 95% confidence intervals) on change measures of subcortical brain volume, and phenotypic, genetic and environmental correlations with bipolar disorder.

	h^2	c^2	e^2	r_{ph}	r_g	r_c	r_e	r_{ph-g}	r_{ph-c}	r_{ph-e}
Region (volume)		%								
Thalamus	0	0	100	0.06	1	0	0.01	0.05	0	0.01
	(0 to 39)	(0 to 38)	(61 to 100)	(-0.13 to 0.24)	(-1 to 1)	(0 to 0)	(-0.42 to 0.44)	(-0.14 to 0.25)	(0 to 0)	(-0.16 to 0.17)
Caudate Nucleus[a]	43	0	57	-0.02	0.01	0	-0.11	0.01	0	-0.03
	(0 to 76)	(0 to 60)	(24 to 100)	(-0.23 to 0.18)	(-1 to 1)	(0 to 0)	(-0.63 to 0.40)	(-0.20 to 0.21)	(0 to 0)	(-0.18 to 0.13)
Putamen	0	0	100	0	1	0	-0.07	0.03	0	-0.03
	(0 to 19)	(0 to 19)	(80 to 100)	(-0.19 to 0.19)	(-1 to 1)	(0 to 0)	(-0.48 to 0.36)	(-0.16 to 0.21)	(0 to 0)	(-0.18 to 0.14)
Pallidum	0	0	100	-0.03	-1	0	0.01	-0.03	0	0
	(0 to 36)	(0 to 25)	(64 to 100)	(-0.21 to 0.16)	(-1 to 1)	(0 to 0)	(-0.39 to 0.43)	(-0.22 to 0.16)	(0 to 0)	(-0.15 to 0.16)
Hippocampus	0	0	100	0.06	1	0	0.05	0.04	0	0.02
	(0 to 22)	(0 to 20)	(78 to 100)	(-0.12 to 0.23)	(-1 to 1)	(0 to 0)	(-0.34 to 0.42)	(-0.15 to 0.22)	(0 to 0)	(-0.13 to 0.16)
Amygdala	5	0	95	-0.04	-0.93	0	-0.07	-0.02	0	-0.03
	(0 to 41)	(0 to 41)	(58 to 100)	(-0.23 to 0.14)	(-1 to 1)	(0 to 0)	(-0.48 to 0.33)	(-0.21 to 0.17)	(0 to 0)	(-0.18 to 0.13)
Nucleus Accumbens[b]	11	0	89	**0.20**	**1**[†]	0	-0.26	**0.30**[†]	0	-0.09
	(1 to 37)	(0 to 19)	(63 to 99)	(0.01 to 0.39)	(0.37 to 1)	(0 to 0)	(-0.69 to 0.20)	(0.11 to 0.47)	(0 to 0)	(-0.25 to 0.07)
Total brain	31	0	69	0.09	1	0	-0.09	0.12	0	-0.03
	(0 to 60)	(0 to 57)	(39 to 100)	(-0.10 to 0.27)	(-1 to 1)	(0 to 0)	(-0.54 to 0.36)	(-0.08 to 0.31)	(0 to 0)	(-0.17 to 0.12)
No outliers removed										
Thalamus	1	0	99	0.17	1	0	0.23	0.08	0	0.09
	(0 to 38)	(0 to 44)	(56 to 100)	(-0.01 to 0.34)	(-1 to 1)	(0 to 0)	(-0.17 to 0.59)	(-0.10 to 0.26)	(0 to 0)	(-0.06 to 0.23)

Note. Table shows results for the models with outliers removed. When removing outliers changed the nature of the association between BD and subcortical volume change, estimates from both models are presented.

[†]Significant at α=0.007 (Bonferroni threshold), estimates in bold face are significant at α=0.05.

[a] In the model where no outliers were removed, heritability (h^2) of change in the caudate nucleus was 0%. All the other subcortical brain regions showed similar heritabilities between models, whether outliers had been removed or not.

[b] In the model where no outliers were removed, the r_{ph} between BD and change in the nucleus accumbens was no longer significant (r_{ph} : 0.18, CI: -0.01 to 0.36).

Supplementary Table S4b. Genetic/environmental influences (AE-model, with 95% confidence intervals) on change measures of subcortical brain volume, and phenotypic, genetic and environmental correlations with bipolar disorder.

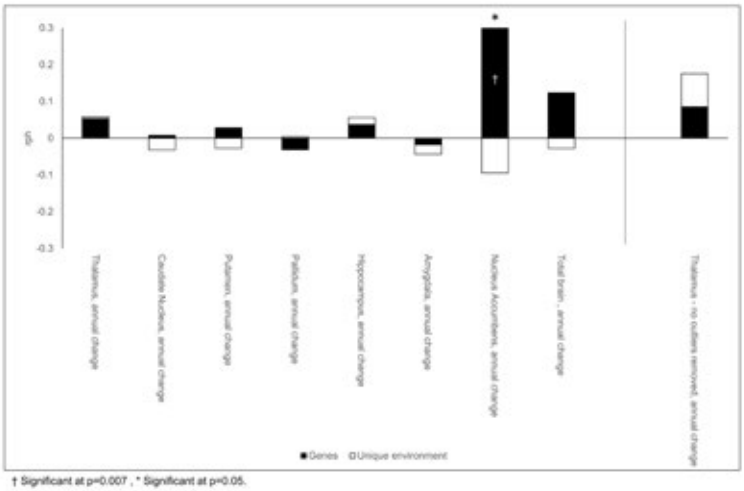
	h^2	e^2	r_{ph}	r_g	r_e	r_{ph-g}	r_{ph-e}
Region (volume)	%						
Thalamus	0	100	0.06	1	0.01	0.05	0.01
	(0 to 39)	(61 to 100)	(-0.13 to 0.24)	(-1 to 1)	(-0.42 to 0.44)	(-0.14 to 0.25)	(-0.16 to 0.17)
Caudate Nucleus[a]	43	57	-0.02	0.01	-0.11	0.01	-0.03
	(0 to 76)	(24 to 100)	(-0.23 to 0.18)	(-1 to 1)	(-0.63 to 0.40)	(-0.20 to 0.21)	(-0.18 to 0.13)
Putamen	0	100	0	1	-0.07	0.03	-0.03
	(0 to 19)	(81 to 100)	(-0.19 to 0.19)	(-1 to 1)	(-0.48 to 0.36)	(-0.16 to 0.21)	(-0.18 to 0.14)
Pallidum	0	100	-0.03	-1	0.01	-0.03	0
	(0 to 36)	(64 to 100)	(-0.21 to 0.16)	(-1 to 1)	(-0.39 to 0.43)	(-0.22 to 0.16)	(-0.15 to 0.16)
Hippocampus	0	100	0.06	1	0.05	0.04	0.02
	(0 to 22)	(78 to 100)	(-0.12 to 0.23)	(-1 to 1)	(-0.34 to 0.42)	(-0.15 to 0.22)	(-0.13 to 0.16)
Amygdala	5	95	-0.04	-0.09	-0.07	-0.02	-0.03
	(0 to 42)	(58 to 100)	(-0.22 to 0.14)	(-1 to 1)	(-0.48 to 0.34)	(-0.21 to 0.17)	(-0.18 to 0.13)
Nucleus Accumbens[b]	11	89	**0.20**	**1**[†]	-0.26	**0.30**[†]	-0.09
	(1 to 37)	(63 to 99)	(0.01 to 0.39)	(0.37 to 1)	(-0.69 to 0.20)	(0.11 to 0.47)	(-0.25 to 0.07)
Total brain	31	69	0.10	0.24	-0.09	0.12	-0.03
	(0 to 62)	(38 to 100)	(-0.09 to 0.28)	(-1 to 1)	(-0.54 to 0.37)	(-0.08 to 0.32)	(-0.17 to 0.13)
No outliers removed							
Thalamus	1	99	0.18	1	0.23	0.09	0.09
	(0 to 39)	(61 to 100)	(0 to 0.34)	(-1 to 1)	(-0.17 to 0.59)	(-0.10 to 0.26)	(-0.06 to 0.23)

Note. Table shows results for the models with outliers removed. When removing outliers changed the nature of the association between BD and subcortical volume change, estimates from both models are presented.

[†]Significant at α=0.007 (Bonferroni threshold), estimates in bold face are significant at α=0.05.

[a] In the model where no outliers were removed, heritability (h^2) of change in the caudate nucleus was 0%. All the other subcortical brain regions showed similar heritabilities between models, whether outliers had been removed or not.

[b] In the model where no outliers were removed, the r_{ph} between BD and change in the nucleus accumbens was no longer significant (r_{ph} : 0.18, CI: -0.01 to 0.36).



Supplementary Figure S1. Genetic and unique environmental contributions to the phenotypic correlation between BD and change in subcortical brain volumes.

Chapter 4

Genetic and environmental influences on cortical surface area and cortical thickness in bipolar disorder

Florian Bootsman, Rachel M. Brouwer, Hugo G. Schnack, G. Caroline M. van Baal, Astrid C. van der Schot, Ronald Vonk, Hilleke E. Hulshoff Pol, Willem A. Nolen, René S. Kahn, Neeltje E.M. van Haren

Psychological Medicine 45, 193-204

4.1 ABSTRACT

The risk of developing bipolar disorder (BD) has been linked to structural brain abnormalities. The degree to which genes and environment influence the association of BD with cortical surface area remains to be elucidated. In this twin study, genetic and environmental contributions to the association between liability to develop BD and surface area, thickness and volume of the cortex were examined.

The study cohort included 44 affected monozygotic (9 concordant, 12 discordant) and dizygotic (4 concordant, 19 discordant) twin pairs, and 7 twins from incomplete discordant monozygotic and dizygotic discordant twin pairs. In addition, 37 monozygotic and 24 dizygotic healthy control twin pairs, and 6 twins from incomplete monozygotic and dizygotic control pairs were included.

Genetic liability to develop BD was associated with larger cortical surface in limbic and parietal regions, and thicker cortex in central and parietal regions. Environmental factors related to BD were associated with larger medial frontal, parietal and limbic, and smaller orbitofrontal surfaces. Furthermore, thinner frontal, limbic and occipital cortex, and larger frontal and parietal, and smaller orbitofrontal volumes were also associated with environmental factors related to BD.

Our results suggest that unique environmental factors play a prominent role in driving the associations between liability to develop BD and cortical measures, particularly those involving cortical thickness. Further evaluation of their influence on the surface and thickness of the cortical mantle is recommended. In addition, cortical volume appeared to be primarily dependent on surface and not thickness.

Keywords: Bipolar disorder, twin study, magnetic resonance imaging, cortical surface area, cortical thickness, unique environmental factors, heritability

4.2 INTRODUCTION

Neuroimaging studies have provided evidence for subtle structural brain abnormalities in patients with bipolar disorder (BD). Most of these studies reported on cerebral volume and density as indices of structural neuropathology. In particular, smaller volumes of prefrontal cortical grey matter, corpus callosum and cerebral white matter, and larger volumes of the lateral ventricles and striatum were found in BD relative to healthy controls (Emsell and McDonald, 2009; Savitz and Drevets, 2009). The volume of the amygdala appears larger in BD adults, but smaller in BD children and adolescents (Savitz and Drevets, 2009). However, an absence of volumetric abnormalities of the total brain, grey and white matter has also been described (Emsell and McDonald, 2009). In addition, a recent meta-analysis (Selvaraj et al., 2012) of voxel-based morphometry studies demonstrated smaller regional grey matter volume in frontal and temporal regions but did not include two studies that noted larger regional volumes in these (Adler et al., 2007; Bearden et al., 2007) and parietal regions (Adler et al., 2007). Measures of brain volume in BD patients are possibly influenced by lithium use, familial load, mood status, illness duration, age, heterogeneity of subject groups and variability in imaging methodology, likely contributing to the inconsistencies across studies (Emsell and McDonald, 2009; Savitz and Drevets, 2009; Hajek et al., 2012).

The heritability of BD is estimated to be around 85% (McGuffin et al., 2003). Twin studies have shown an association between the genetic liability to develop BD and smaller cerebral white matter volume (Kieseppa et al., 2003; van der Schot et al., 2009), implicating that genes involved in the risk to develop BD are related to smaller white matter volume. In contrast, environmental factors related to BD were found to be significantly associated with smaller cortical grey matter volume (van der Schot et al., 2009; Hulshoff Pol et al., 2012). This suggests that environmental factors unique for each individual influence directly (or indirectly through a causal relationship) both BD and grey matter volume.

Cortical volume is the "product" of surface area and thickness (Panizzon et al., 2009). The cortex has a columnar organization (Mountcastle, 1997). The number of cells in ontogenetic columns that run perpendicular to the surface of the brain influence cortical thickness while the size of the surface is suggested to be dependent of the number of columns (Rakic, 1988; Rakic, 1995). Although these measures are inherently biologically correlated, they appear to be genetically distinct but may singly or both drive cortical volume (Rakic, 1988; Rakic, 1995; Pakkenberg and Gundersen, 1997; Jansen and Andermann, 2005; Fornito et al., 2008; Im et al., 2008; Pontious et al., 2008; Panizzon et al., 2009; Rakic et al., 2009; Winkler et al., 2010; Eyler et al., 2011). Consequently, the genetic contribution to brain volume in healthy individuals (Baare et al., 2001; Thompson et al., 2001; Peper et al., 2007) and BD patients (Noga et al., 2001; van der Schot et al., 2009; Hulshoff Pol et al.,

2012) is likely dependent on the genetic influences on cortical surface area and/or cortical thickness.

An association between liability to develop BD and cortical surface area has not yet been established nor have case-control studies found evidence for cortical surface area abnormalities in BD (Fornito et al., 2009; Rimol et al., 2012) whereas cortical thickness in BD has been investigated (Lyoo et al., 2006; Rimol et al., 2010; Foland-Ross et al., 2011; Hulshoff Pol et al., 2012). Quantification of genetic and environmental contributions to cortical surface area and cortical thickness independently in BD may further illuminate the association between grey matter volume deficits and the risk to develop BD (McDonald et al., 2004; van der Schot et al., 2009). Previously, we reported on whether patients with schizophrenia (SZ) and patients with BD showed overlapping abnormalities in cortical thickness (Hulshoff Pol et al., 2012). In the present twin study, using the same BD twin sample as Hulshoff Pol et al. (2012), we set out to quantify the additive genetic, common, and unique environmental contributions to the association between BD liability and cortical measures in 39 cortical regions of interest (ROIs) per hemisphere. In addition, we specifically explored the degree to which associations between regional cortical volume and BD liability reflect the associations of cortical surface area and cortical thickness with BD liability.

4.3 METHODS

4.3.1 Subjects

A total of 118 monozygotic (MZ) and dizygotic (DZ) twin pairs underwent magnetic resonance imaging (MRI) (van der Schot et al., 2009; van der Schot et al., 2010). Of this group, 13 subjects were excluded from analysis due to bad image quality. Consequently, 44 twin pairs affected with BD (9 MZ concordant, 12 MZ discordant; 4 DZ concordant, 19 DZ discordant) and 7 twins from incomplete pairs (3 MZ co-twins from discordant pairs; 2 DZ patients and 2 DZ co-twins from incomplete discordant pairs), and 61 matched control twin pairs (37 MZ, 24 DZ) plus 2 MZ and 4 DZ healthy control twins from incomplete pairs were included in the analyses. All twins were raised together, except for one control pair where twins were separated at 12 years of age when both parents died. Subjects were between 18 and 60 years of age at the time of enrolment in the study. Demographic information is presented in **Table I**. Clinical diagnosis of Axis I psychiatric disorders and Axis II personality disorders was confirmed using the Structured Clinical Interview for DSM-IV (SCID) (First et al., 1997) and the Structured Interview For DSM-IV Personality (SIDP) (Pfohl et al., 1997) respectively, and through available medical records. Patients were also interviewed on their medication history. The twin pairs had no history of drug or alcohol dependency for the last 6 months prior to inclusion in the study, for this was an

exclusion criterion. Moreover, none had severe medical illness, verified with a medical history inventory. The current mood state of BD patients was assessed using the Young Mania Rating Scale (YMRS) (Young et al., 1978) and the Inventory for Depressive Symptomatology (IDS) (Beck et al., 1961). Upon inclusion, all patients were euthymic with a YMRS score of 4 or less and an IDS score of 12 or less, except for 9 BD patients who were mildly to severely depressed (IDS scores of 14, 14, 14, 15, 17, 20, 22 and 38, respectively) or hypomanic (YMRS score of 17). Healthy control pairs were matched to the bipolar pairs for zygosity, gender, age and parental education. Control pairs had no history of Axis I or II disorders according to DSM-IV criteria (confirmed with the SCID and SIDP, respectively) and no history of severe medical illness. Furthermore, they had no first-degree relative with a history of a major Axis I psychiatric disorder (DSM-IV). Family histories of all twins were obtained via the Family Interview Genetic Studies (FIGS) (Nurnberger et al., 1994), performed with both twins of each pair. Zygosity was determined with DNA fingerprinting using high polymorphic microsatellite markers 9 to 11. The medical ethics review board of the University Medical Center Utrecht approved the study and all participants gave written informed consent after full explanation of the study aims and procedures.

4.3.2 Brain imaging

For information on previously described MRI acquisition and pre-processing protocols, please refer to **Supplementary Method**, available online.

Cortical surface and cortical thickness estimates were obtained using the CLASP (Constrained Laplacian Anatomic Segmentation Using Proximity) algorithm in a custom implementation of CIVET, developed at the McConnell Brain Imaging Centre of the Montreal Neurological Institute (MacDonald et al., 2000; Kabani et al., 2001; Kim et al., 2005). The grey/white/CSF segments from our own segmentation algorithm, as described earlier, were used as inputs for the original CLASP algorithm. A 3-dimensional surface consisting of the 81920 polygons and 40962 vertices was fitted to the grey/white matter intersection, thereby creating the inner surface of the cortex. The outer surface of the cortex was then constructed by expanding the inner surface to fit the grey matter/CSF intersection (MacDonald et al., 2000; Kim et al., 2005). The surfaces are modelled as nets of polygons (triangles) with vertices being the points where the polygons meet (i.e. the angular points of the nets). The surfaces of the subjects (Lyttelton et al., 2007) were registered to an average surface, allowing for comparison of cortical thickness locally. Each polygon vertex on the outer surface had its counterpart on the inner surface; cortical thickness was defined as the distance between these inner and outer surface polygons. Subsequently, automated anatomical labeling (AAL) was executed to create parcellations of 39 bilateral brain regions that encompass these vertices (Tzourio-Mazoyer et al., 2002). Cortical surface area was estimated from the mid-surface, positioned halfway between the outer and inner surfaces. Cortical volume was calculated via multiplication of cortical surface area and cortical thickness.

Table 1. Demographic characteristics of the bipolar and matched healthy control twin pairs.

	Bipolar Twins (n=95)		**Matched control twins** (n=128)	
	MZ[a] (n=45)	**DZ** (n=50)	**MZ**[b] (n=76)	**DZ** (n=52)
Gender, m/f	13/32	17/33	31/45	24/28
Age, y[c]	37.2(10.4)	43.9(8.8)	39.1(9.7)	39.5(8.0)
Parental educ., y[d]	10.9(3.6)	11.2(3.9)	11.4(3.3)	11.6(3.5)

	Patient	**Co**	**Patient**	**Co**	**Twin 1**	**Twin 2**	**Twin 1**	**Twin 2**
Onset, age[e]	25.1(8.5)	-	31.5(11.5)	-	-	-	-	-
Education,y	12.0(2.1)	12.1(2.3)	13.6(2.7)	12.2(3.1)	13.8(2.5)	13.7(2.8)	12.4(2.6)	13.6(2.3)
Birth order, $_{1st/2nd}$ [f]	15/15	8/7	11/17	12/8	-	-	-	-
Handedness (right/left or both)	21/9	10/5	24/5	19/2	31/7	31/7	24/1	21/6
Lithium, *on/off on day MRI, No.*	25/5	0/15	18/11	0/21	-	-	-	-
Antipsychotics, *on/off on day of MRI*	2/28	1/14	7/22	1/20	-	-	-	-
Antidepressants, *on/off on day of MRI*	9/21	0/15	8/21	1/20	-	-	-	-
YMRS score[g]	1.1(1.5)	0.5(0.82)	1.2(3.4)	0.1(0.7)	-	-	-	-
IDS score[h]	6.1(5.3)	2(2.6)	6.1(8.6)	2.7(3.7)	-	-	-	-

Abbreviations: MZ, Monozygotic; DZ, Dizygotic
n= number of individuals
[a] 9 MZ concordant pairs; 12 MZ discordant pairs; 3 MZ co-twins from incomplete pairs; 4 DZ concordant pairs; 19 DZ discordant pairs; 2 DZ patients and 2 DZ co-twins from incomplete pairs
[b] 37 MZ healthy control pairs; 2 MZ healthy controls from incomplete pairs; 24 DZ healthy control pairs ; 4 DZ healthy controls from incomplete pairs
[c] DZ patient pairs were significantly older than MZ patient pairs [$F_{1,93}$=11.55, p<0.05)], MZ control pairs [$F_{1, 124}$= 7.95, p < 0.05)] and DZ control pairs [$F_{1, 100}$= 7.30, p <0.05)]
[d] Years of parental education for 2 bipolar twin pairs could not be determined
[e] Age of onset could not be determined for 4 bipolar patients; Significant difference between MZ and DZ [$F_{1, 53}$= 5.47, p < 0.05)]
[f] Birth order was not determined for 1 bipolar twin pair
[g] YMRS score was not ascertained for 2 MZ patients, 2 MZ co-twins and 3 DZ patients.
[h] IDS score was not ascertained for 2 MZ patients, 2 MZ co-twins and 2 DZ patients

4.3.3 Model fitting

Twin designs allow for quantification of the influence of genes and environment on measured traits. Comparison of correlations within MZ and DZ twins provides a basis on which the degree of genetic and environmental influences on such traits can be judged. A larger correlation between traits in MZ twins than in DZ twins suggests that genetic factors play a role, because MZ twins are genetically identical whereas DZ twins only share on

average 50% of their segregating genes. When there are no differences between MZ and DZ correlations, but there is phenotypic resemblance for the trait, then environment shared by twins is more likely influencing the trait (Boomsma et al., 2002).

To estimate the extent of the contribution of genetic and unique environmental factors to the association between cortical measures and BD liability, a bivariate liability threshold model was chosen. This model was implemented in structural equation modeling software OpenMx (Kenny et al., 2009), running under the statistical programming environment R (R Development Core Team, 2008). Prior to model fitting, linear regression on gender, age and handedness (and lithium use for BD patients only, measured as on/off on day of MRI) was performed to eliminate their contribution to the variance in cortical measures. We found that handedness significantly ($\alpha<0.05$) or at trend-level influenced some of the cortical measures. Therefore, correction for handedness was applied, given that doing so would only result in correction of the measures that showed influence of handedness and not of those that did not show such an influence.

A bivariate Cholesky decomposition was fitted to the data to estimate additive genetic (A), and unique environmental (E) variance components of cortical surface area, cortical thickness and cortical volume, and phenotypic, genetic and environmental overlap between BD disease liability and cortical measures (online **Supplementary Figure S1**). No evidence of shared environmental influences (C) on bipolar disorder has been found previously (McGuffin et al., 2003) and we presently found no evidence of such influences on our cortical measures. Shared environmental factors were therefore not included in the model. Disease status was dichotomous and assumed to represent an underlying continuous liability with a mean of 0 (S.D.=1). A patient will have a high value on the liability scale, thereby crossing a certain threshold (patient status=1). All other individuals will have lower liability scores and will not cross the critical threshold (patient status = 0; discordant co-twin of patient or control twin pairs). The critical threshold and heritability for the underlying liability to develop BD was not based on this sample because we included approximately equal numbers of concordant, discordant, and healthy twin pairs. Prevalence and heritability (the relative contribution of genetic variance to total variance) of BD were fixed to population values; prevalence was set to 1% (Regeer et al., 2004) and heritability was set to 85% (McGuffin et al., 2003). Importantly, we have tested whether using alternative BD prevalences of 2.5% and 4% would influence the number and nature of phenotypic associations between BD and cortical measures. It did not. If anything, we may have been too conservative in setting prevalence at 1%, but feel confident in doing so, based on the available literature on BD prevalence in the Netherlands (see Regeer et al., 2004). In order to apply the threshold model to the brain measures, the standardized residuals that were obtained for these measures via linear regression (on gender, age, handedness and lithium use [in patients only]) were rendered

into five categories identical for all subjects – thereby equating them across groups - and put in the model. Thresholding was based on normality plots, with the boundaries of the 'outer' two categories set at -1.5 standard deviation and 1.5 standard deviation, respectively, and the other three categories that fall in between being 1 S.D. wide.

The phenotypic correlation (r_{ph}), an index of association between phenotypes (e.g. liability to develop BD and brain measure), was based on calculations of within-twin/between-trait correlations. Heritability (h^2) and influence of unique environment (e^2) as well as disentanglement of the observed correlation between liability to develop BD and brain measures into genetic (r_g) and environmental (r_e) components was based on polychoric cross-twin/within-trait and cross-twin/cross-trait correlations within MZ and DZ groups (Neale and Miller, 1997). The heritability of brain measures was thus determined within the bivariate model. The genetic correlation indicates the degree of overlap in genes influencing both phenotypes. Unique environmental correlations indicate overlap in unique environmental factors influencing both phenotypes. The phenotypic correlation is a function of the genetic and environmental correlations weighted by the square root of the heritabilities ($r_g * h_{BD} * h_{brain}$) and the square root of e^2($r_e * e_{BD} * e_{brain}$). We refer to these quantities as $r_{ph\text{-}g}$ and $r_{ph\text{-}e}$ (Toulopoulou et al., 2007).

A saturated model in which means, variances and correlations are estimated freely served as a baseline model to which the more restrictive AE-model was compared, using Akaike's Information Criterion (AIC). Compared to the saturated model, the AE-model had the best fit in 95% of ROIs and was therefore applied to all ROIs. For relevant estimates, 95% confidence intervals (CI) were obtained (Neale and Miller, 1997). To account for multiple testing in multiple regions, Bonferroni correction was applied. This was done by dividing the alpha value by the number of regions per hemisphere, which resulted in a critical level of significance of 0.05/39=0.0013 for each parameter (r_{ph}, r_g, r_e, $r_{ph\text{-}g}$, $r_{ph\text{-}e}$). In the tables, significance at $\alpha < 0.05$ and at $\alpha < 0.0013$ is indicated.

4.4 RESULTS

4.4.1 Global cortical measures

Genetic and environmental influences on global cortical measures

Heritability of all cortical measures was estimated irrespective of disease (see **Supplementary Tables S1, S2 and S3**). Total cortical surface area was predominantly explained by genes (h^2: 90% and 91%, left and right hemispheres respectively), as was total cortical grey matter volume (h^2: 88% and 86%, left and right hemispheres respectively). Mean cortical thickness across the entire cortex was moderately influenced by genes (h^2: 51% and 48%, left and right hemispheres respectively).

Association between liability to develop bipolar disorder and global cortical measures
There were no significant phenotypic associations between liability to develop BD and total cortical surface area in either hemisphere, indicating that genetic and environmental factors influencing BD are not associated with global cortical surface area. However, unique environmental factors related to BD were associated with a thinner cortex and smaller total cortical volume of the right hemisphere.

4.4.2 Regional cortical measures

Genetic and environmental influences on regional cortical measures
Regionally, heritability of all cortical measures was estimated irrespective of disease. At this level, cortical surface area, cortical thickness and cortical volume were strongly influenced by unique environmental factors (see online **Supplementary Tables S1, S2** and **S3**).

Association between liability to develop bipolar disorder and regional cortical surface area
In local brain regions, liability to develop BD was associated with both larger and smaller cortices (see online **Supplementary Table S1** and **Figures 1** and **2**). In particular, genetic liability to develop BD was significantly associated with larger cortical surface area in the left inferior parietal gyrus and right posterior cingulate and supramarginal gyri. In contrast, unique environmental factors related to BD were significantly associated with larger cortical surface in the left supplementary motor area, parahippocampal gyrus and right supramarginal gyrus, and smaller cortical surface area in the left gyrus rectus and middle temporal gyrus. This indicates that the genetic liability to develop BD appears to be primarily associated with larger regional cortical surface areas whereas unique environmental factors related to BD are associated with both larger and smaller regional cortical surface areas.

Association between liability to develop bipolar disorder and regional cortical thickness
At the regional level, liability to develop BD was associated with both thicker and thinner cortices in a number of brain regions (see online **Supplementary Table S2** and **Figures 1** and **3**). Particularly, genetic liability to develop BD was significantly associated with thicker cortex in the left rolandic operculum and supramarginal gyrus. In contrast, unique environmental factors related to BD were significantly associated with thinner cortices in the right triangular inferior frontal gyrus, supplementary motor area, posterior cingulate gyrus, calcarine cortex, cuneus, lingual gyrus, superior occipital gyrus, middle occipital gyrus, fusiform gyrus and middle temporal gyrus. In sum, genes influencing BD appeared to be primarily associated with regionally thicker cortex whereas unique environmental factors influencing BD were more likely associated with thinner cortices.

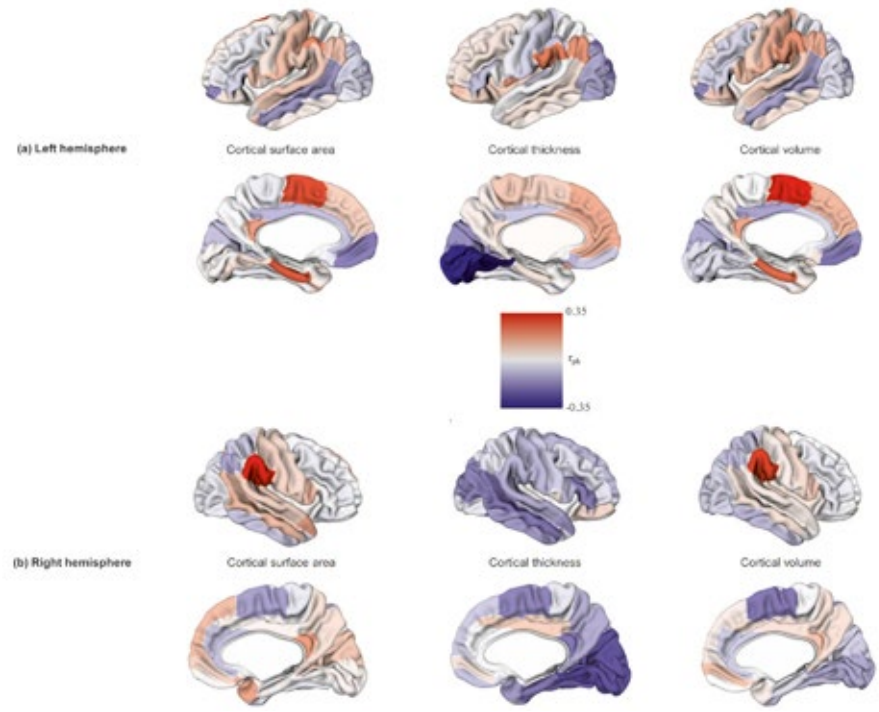
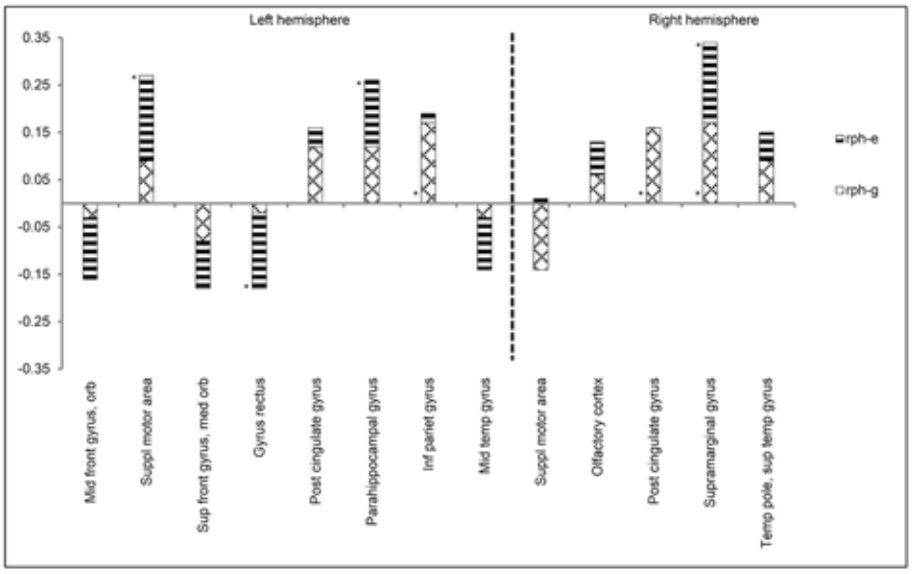


Figure 1. Phenotypic correlations (range $r_{ph:}$ 0.35 to -0.35) between BD liability and cortical surface area/cortical thickness/cortical volume in both hemispheres. Red indicates positive correlation, blue indicates negative correlation and white indicates no correlation.

Association between liability to develop bipolar disorder and regional cortical volume

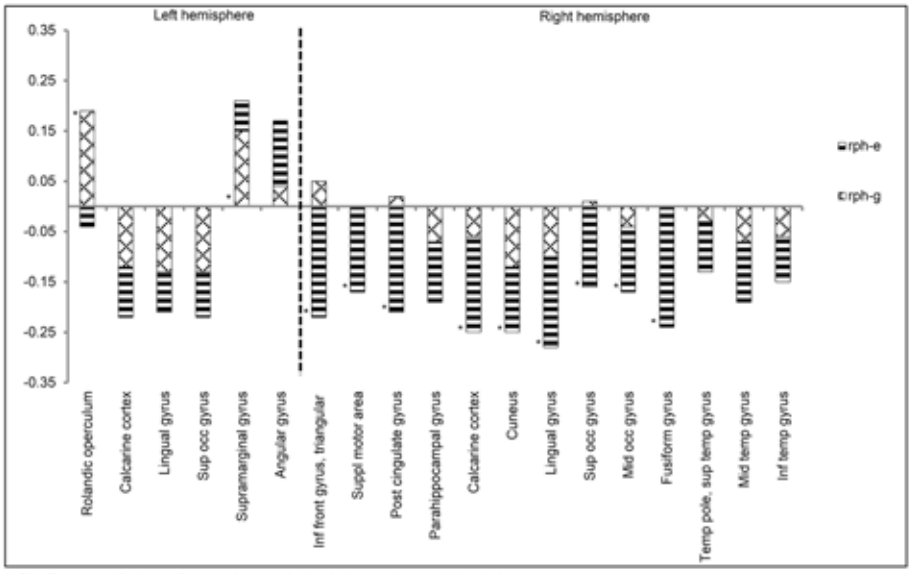
Liability to develop BD was associated with both larger and smaller cortical volumes (see online **Supplementary Table S3** and **Figures 1** and **4**). Unique environmental factors related to BD were significantly associated with larger cortical volumes of the bilateral supramarginal gyri and left supplementary motor area and parahippocampal gyrus, and smaller cortical volume of the left gyrus rectus. There was no significant association between genetic liability to develop BD and regional cortical volume. Therefore, regionally, cortical volume appears chiefly associated with unique environmental factors, probably related to the illness itself, and not with genes influencing BD.

The pattern of phenotypic association between BD liability and cortical volume appeared to a large extent to mimic the pattern of phenotypic association between BD liability and cortical surface area (compare online **Supplementary Tables S1** and **S3**, and see **Figure 1**), which indicates that cortical volume is primarily dependent on cortical surface area, and not cortical thickness.



* Significant contributions of r_{ph-g} and r_{ph-e} to the total phenotypic correlation (r_{ph})

Figure 2. Genetic (r_{ph-g}) and environmental (r_{ph-e}) contributions to the significant phenotypic correlations between liability to develop BD and cortical surface area.



* Significant contributions of r_{ph-g} and r_{ph-e} to the total phenotypic correlation (r_{ph})

Figure 3. Genetic (r_{ph-g}) and environmental (r_{ph-e}) contributions to the significant phenotypic correlations between liability to develop BD and cortical thickness.

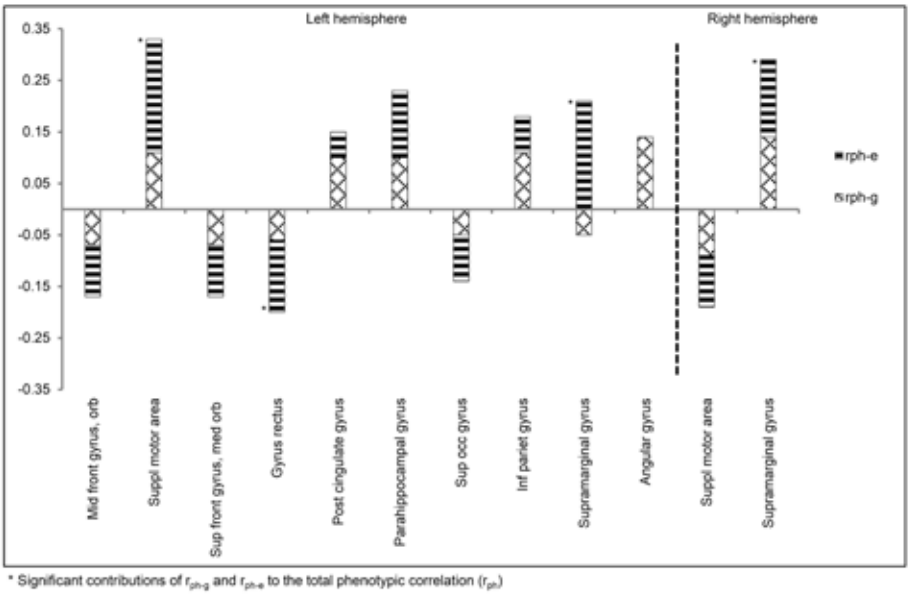


Figure 4. Genetic (r_{ph-g}) and environmental (r_{ph-e}) contributions to the significant phenotypic correlations between liability to develop BD and cortical volume.

In sum, across all measures, the majority of associations between BD liability and the cortex were predominantly influenced by unique environmental factors associated with BD (see **Tables S1, S2, S3** and **S4**).

4.5 DISCUSSION

To the best of our knowledge, this is the first twin study reporting on the influence of genes and environment on the association between BD and cortical surface area. We evaluated to what degree genetic and environmental factors associated with BD were associated with surface area, thickness and volume of the cortex. Furthermore, we studied the extent to which the pattern of association between BD liability and cortical volume would mimic the pattern of association of cortical surface area or cortical thickness with BD liability

Genetic and environmental influences on surface area, thickness and volume of the cortex

Total cortical surface area and cortical volume were highly heritable, and global cortical thickness was moderately heritable, irrespective of disease. In contrast, regional cortical measures were predominantly influenced by unique environmental factors. A similar

reduction in heritability from global to local brain measures has been described earlier and likely reflects an increase in the degree of measurement error (Panizzon et al., 2009; Eyler et al., 2011). In a ROI approach, the magnitude of genetic influence may be underestimated in areas that extend beyond regional boundaries determined by a brain atlas (Eyler et al., 2011).

Cortical surface area in BD

The risk of developing BD was not associated with abnormalities of global cortical surface area. However, larger regional cortical surfaces of bilateral parietal and right limbic regions were associated with the genetic risk of developing BD. Furthermore, unique environmental factors influencing BD were associated with larger surface area in left medial frontal and right parietal regions, and smaller surface area in left orbitofrontal and temporal regions. We also found larger surface area of the left parahippocampal gyrus to be associated with unique environmental factors related to BD. Interestingly, the parahippocampus is part of the limbic system and is involved in emotion processing (Phillips et al., 2003; Emsell and McDonald, 2009; Aldhafeeri et al., 2012; Delvecchio et al., 2013). Deficits in the volume and thickness of this structure have been previously observed in BD (Savitz and Drevets, 2009; van der Schot et al., 2010; Wang et al., 2011; Hulshoff Pol et al., 2012). Abnormal emotion processing appears to be a key feature of BD (Samamé et al.,2012). Further exploration of the link between structural deficits in emotion processing brain regions and the genetic and environmental risk for BD is therefore appropriate.

In summary, genes influencing BD are predominantly associated with regional larger cortical surfaces while unique environmental factors related to BD are associated with both larger and smaller regional cortical surfaces. It has been suggested that surface area size may be dependent on white matter volume, particularly with regard to the degree of myelination of the axons. In the maturing brain, the overlying cortex may stretch as the white matter increases (Seldon, 2005; Hogstrom et al., 2013). Furthermore, cortical surface area appears positively associated with local gyrification, indicating that greater gyrification implies larger cortical surface area (Hogstrom et al., 2013). Also, increased sulcal width and decreased sulcal depth have been suggested to relate to cortical surface shrinkage (Kochunov et al., 2005). Perhaps the larger and smaller surface sizes we found may be relative to regional indices of white matter volume and local gyrification.

Cortical thickness in BD

At the global level, a mean thinner cortex in the right hemisphere was associated with unique environmental factors influencing BD whereas mean cortical thickness of the left hemisphere was not associated to BD liability. Our findings on regional cortical thickness are, not surprisingly, largely in line with our previous vertex-wise analyses (Hulshoff Pol et

al., 2012). Again, we found an association between genetic liability to develop BD and a thicker left supramarginal gyrus, and a phenotypic association between BD liability and a thinner right parahippocampus. A thinner right calcarine cortex was associated with unique environmental factors related to BD. In this study we also found that the genetic liability to develop BD was associated with a thicker left central region.

In this study, using the ROI approach, we now also found that unique environmental factors influencing BD were associated with thinner right frontal, limbic, temporal and occipital cortices. In part, these results are in line with prior findings of a voxel-based morphometry study in this cohort, where we found unique environmental factors influencing BD to be associated with lower grey matter density in right inferior frontal, cingulate, inferior temporal, lingual and limbic regions (van der Schot et al., 2010). Possibly, measures of grey matter density and cortical thickness partly register the same cortical abnormalities, as has been suggested in schizophrenia (Narr et al., 2005).

Cortical volume in BD

We found that smaller total cortical volume of the right hemisphere was associated with unique environmental factors influencing BD, which is in line with the previously found overlap in unique environmental factors influencing both BD and smaller cortical grey matter volume (van der Schot et al., 2009). Regionally, genetic liability to develop BD was not associated with cortical volume. However, unique environmental factors influencing BD were associated with larger cortical volumes of bilateral parietal and left frontal and limbic regions, and smaller left orbitofrontal cortical volume. Only a small number of studies has examined regional brain volumes in subjects with genetic vulnerability for BD, demonstrating associations between genetic risk for BD and a larger volume of left caudate nucleus (Noga et al., 2001), and smaller volumes of the right anterior cingulate gyrus and ventral striatum (McDonald et al., 2004). Remarkably, BD liability was associated with regional cortical volume and cortical surface in many of the same brain regions in this study, in a similar fashion. Moreover, in many brain regions where cortical thickness was associated with unique environmental factors related to BD, cortical volume was not.

Thus our results agree with previous reports that cortical surface area and not cortical thickness primarily promotes cortical volume (Pakkenberg and Gundersen, 1997; Im et al., 2008; Rakic, 2009; Winkler et al., 2010). Importantly, although abnormalities in cortical surface and especially cortical thickness were not always reflected in abnormalities in cortical volume, they were nevertheless present. In the same vein, Fornito et al. (2008) found larger surface and thinner cortex in the anterior cingulate gyrus in schizophrenia patients compared with controls but did not find volume differences in this structure between groups. Consequently, the authors stressed the importance of moving beyond traditional measures of grey matter

volume when assessing cortical abnormalities in clinical populations (Fornito et al., 2008). Critical evaluation of structural neuropathology should therefore include assessment of structural brain elements comprising volume and not be limited to assessment of volume only.

Importantly, the majority of phenotypic associations between liability to BD and cortical measures showed strong unique environmental involvement, indicating overlap of unique environmental factors influencing both BD and the cortex. Possibly, these unique environmental factors are illness-related and dependent on the individual's behavior (i.e. self-generated), for example poor work performance leading to job loss. In bipolar spectrum disorders (not BD type I) and unipolar depression, it has been hypothesized that stress may in part be self-generated which, in turn, may contribute to the maintenance of symptoms (Hammen, 1991; Safford et al., 2007; Bender et al., 2010). However, whether this is the case here and if the cortex is affected as a result require further investigation and is merely speculation at this point.

Limitations

Certain limitations in our study should be considered. First, the issue of statistical power remains challenging to tackle in a single twin MRI study. According to Visscher (2004) and Posthuma and Boomsma (2000), large sample sizes and a larger number of MZ than DZ twin pairs are required to adequately estimate the magnitudes of genetic and environmental sources of variance in both univariate and bivariate designs. This is particularly relevant when heritability is expected to be moderate, as was the case with the regional cortical measures in our study. However, the heritabilities of global cortical measures were much higher and more reliably estimated within our model. Furthermore, in a number of regions the associations between BD and cortical measures were highly significant, in some cases surviving Bonferroni correction for multiple comparisons, indicating that for those regions we had sufficient power to detect an association with BD. In contrast, this may not have been true for other regions, the implication of which is that our estimate of the total number of associations may be conservative. Additionally, regional differences in gyrification or signal strength (causing signal noise variability) may have affected (regional) heritability estimates. Second, although MZ and DZ patient twin pairs were matched to the MZ and DZ control pairs for gender, there were more females than males in both patient and control pairs. Also, DZ patient twin pairs were significantly older than MZ patient and DZ control pairs. Therefore, cortical measures were corrected for gender and age. Third, although we corrected for lithium use in BD patients, correction for antipsychotics, antidepressants or other types of medication was not applied. This was due to unreliable reporting by patients on their use and the irregular frequency of medication use other than lithium. Fourth, the liability threshold

model assumes a dichotomous disease variable, which may not fully capture the complex nature of bipolar phenomenology. Additionally, prevalence and heritability of BD were set *a priori* and not based on this sample. Finally, in our bivariate model assessing genetic and environmental influences on the cortex in BD, possible gene/environment interactions could not be ascertained.

Conclusion

Unique environmental factors, probably illness-related, played a prominent role in driving the majority of phenotypic associations between liability to develop BD and cortical measures. Further evaluation of their influence on the cortical mantle is recommended. Furthermore, phenotypic associations between BD liability and cortical volume followed a distribution similar to that of phenotypic associations between BD liability and cortical surface area, indicating cortical volume to be primarily driven by cortical surface area. However, cortical abnormalities are not restricted to anomalies in surface and volume but also extend to the thickness of the cortex. Assessment of genetic and environmental influences on cortical surface area and cortical thickness could enhance our understanding of the nature of disease-related structural neuropathology and may serve as a basis for judgments on the degree of disease segregation or overlap, particularly in the context of (neuro)development.

SUPPLEMENTARY INFORMATION

Supplementary Method - Acquisition and preprocessing of brain images

Brain imaging

Magnetic resonance images were acquired on a 1.5 Tesla NT scanner (Philips, the Netherlands). T1-weighted 3D fast field echo scans with 160–180 contiguous coronal slices (echo time=4.6 ms, repetition time=30 ms, flip angle=30°, 1x1x1.2 mm^3 voxels) and T2-weighted dual-echo turbo-spin-echo scans with 120 contiguous coronal slices (echo time 1= 14 ms, echo time 2=80 ms, repetition time=6350 ms, flip angle=90°, 1x1x1.6 mm^3 voxels) were obtained (Brans et al., 2008; van der Schot et al., 2009; Brans et al., 2010). In order to eliminate possible between-group biases due to scanner drifts, patient and healthy twin as well as MZ and DZ pairs were included in the study in a random order (i.e. they were randomly assigned to inclusion slots/days). Images were acquired and pre-processed with previously validated protocols (Hulshoff Pol et al., 2004; Brouwer et al., 2010). In brief, the T1-weighted images were put in Talairach orientation and corrected for intensity non-uniformity artefacts (Sled et al., 1998). Separation of brain tissue from cerebrospinal fluid (CSF) and grey from white matter was performed with a partial volume segmentation method that accounts for the non-uniformity of the partial volume distribution that is due to the curvature of the cortex (Brouwer et al., 2010).

Supplementary Table S1. Variance components of cortical surface area and phenotypic, genetic and environmental overlap with liability to develop BD.

	Cortical surface area						
	h^2	e^2	r_{ph}	r_g	r_e	$r_{ph\text{-}g}$	$r_{ph\text{-}e}$
Region	%						
Left hemisphere (ROI)							
Hemisphere, $_{\text{total surface area}}$	**90**	**10**	0.03	0.01	0.14	0.01	0.02
	(82 to 95)	(5 to 18)	(-0.11 to 0.16)	(-0.15 to 0.17)	(-0.44 to 0.69)	(-0.13 to 0.15)	(-0.05 to 0.09)
Middle frontal gyrus, $_{\text{orbital part}}$	11	**89**	**-0.16**	-0.10	-0.37	-0.03	-0.13
	(0 to 33)	(67 to 100)	(-0.28 to -0.04)	(-1 to 1)	(-0.68 to 0)	(-0.17 to 0.11)	(-0.25 to 0)
Supplementary motor area	22	**78**	**0.28***	0.21	**0.53**	0.09	**0.18**
	(0 to 48)	(52 to 100)	(0.15 to 0.39)	(-1 to 1)	(0.16 to 0.81)	(-0.04 to 0.24)	(0.05 to 0.29)
Superior frontal gyrus, $_{\text{medial orbital part}}$	**40**	**60**	**-0.19**	-0.14	-0.35	-0.08	-0.10
	(17 to 59)	(41 to 83)	(-0.31 to -0.06)	(-0.40 to 0.10)	(-0.66 to 0.02)	(-0.22 to 0.06)	(-0.21 to 0.01)
Gyrus rectus	25	**75**	**-0.18**	-0.04	**-0.47**	-0.02	**-0.16**
	(0 to 46)	(54 to 100)	(-0.30 to -0.05)	(-1 to 1)	(-0.77 to -0.10)	(-0.16 to 0.12)	(-0.27 to -0.03)
Posterior cingulate gyrus	**26**	**74**	**0.16**	0.26	0.13	0.12	0.04
	(2 to 48)	(53 to 98)	(0.04 to 0.29)	(-0.04 to 1)	(-0.23 to 0.46)	(-0.02 to 0.26)	(-0.08 to 0.16)
Parahippocampal gyrus	**45**	**55**	**0.26***	0.20	**0.48**	0.12	**0.14**
	(22 to 63)	(37 to 78)	(0.13 to 0.38)	(-0.03 to 0.45)	(0.05 to 0.82)	(-0.02 to 0.27)	(0.02 to 0.24)
Inferior parietal gyrus	**47**	**53**	**0.19**	**0.27**	0.07	**0.17**	0.02
	(23 to 66)	(34 to 77)	(0.06 to 0.32)	(0.03 to 0.54)	(-0.36 to 0.48)	(0.02 to 0.32)	(-0.10 to 0.14)
Middle temporal gyrus	**51**	**49**	**-0.14**	-0.05	**-0.4**	-0.03	-0.11
	(29 to 68)	(32 to 71)	(-0.27 to -0.01)	(-0.27 to 0.18)	(-0.75 to -0.01)	(-0.18 to 0.11)	(-0.21 to 0)
Right hemisphere (ROI)							
Hemisphere, $_{\text{total surface area}}$	**91**	**9**	0.03	0.07	-0.23	0.06	-0.03
	(84 to 96)	(4 to 16)	(-0.1 to 0.17)	(-0.09 to 0.23)	(-0.84 to 0.41)	(-0.08 to 0.20)	(-0.10 to 0.05)
Supplementary motor area	**26**	**74**	**-0.13**	-0.31	0.03	-0.14	0.01
	(2 to 46)	(54 to 98)	(-0.26 to -0.01)	(-1 to 0)	(-0.37 to 0.43)	(-0.28 to 0)	(-0.13 to 0.15)
Olfactory cortex	18	**82**	**0.14**	0.16	0.21	0.06	0.07
	(0 to 42)	(58 to 100)	(0.01 to 0.26)	(-1 to 1)	(-0.23 to 0.61)	(-0.09 to 0.22)	(-0.08 to 0.22)
Posterior cingulate gyrus	**36**	**64**	**0.16**	**0.29**	-0.01	**0.16**	0
	(10 to 57)	(43 to 90)	(0.03 to 0.29)	(0.02 to 0.65)	(-0.42 to 0.37)	(0.01 to 0.31)	(-0.13 to 0.12)
Supramarginal gyrus	**39**	**61**	**0.33***	**0.29**	**0.56**	**0.17**	**0.17**
	(17 to 58)	(42 to 83)	(0.21 to 0.45)	(0.04 to 0.58)	(0.13 to 0.88)	(0.02 to 0.31)	(0.04 to 0.27)
Temporal pole, $_{\text{superior temporal gyrus}}$	**25**	**75**	**0.14**	0.19	0.17	0.09	0.06
	(1 to 46)	(54 to 99)	(0.02 to 0.27)	(-0.14 to 1)	(-0.23 to 0.54)	(-0.06 to 0.23)	(-0.08 to 0.18)

Note. Estimates in bold face are significant at α=0.05.

* Significant effect corrected for multiple comparisons (Bonferroni corrected: α = 0.05/39 regions = 0.0013).

Supplementary Table S2. Variance components of cortical thickness and phenotypic, genetic and environmental overlap with liability to develop BD.

	Cortical thickness						
	h^2	e^2	r_{ph}	r_g	r_e	r_{ph-g}	r_{ph-e}
Region	%						
Left hemisphere (ROI)							
Hemisphere, mean thickness	**51**	**49**	-0.01	0.11	-0.31	0.08	-0.08
	(28 to 69)	(31 to 72)	(-0.14 to 0.12)	(-0.11 to 0.35)	(-0.68 to 0.10)	(-0.07 to 0.22)	(-0.19 to 0.03)
Rolandic operculum	**47**	**53**	**0.15**	**0.30**	-0.14	**0.19**	-0.04
	(24 to 65)	(35 to 76)	(0.02 to 0.27)	(0.07 to 0.57)	(-0.52 to 0.25)	(0.04 to 0.33)	(-0.15 to 0.07)
Calcarine cortex	**36**	**64**	**-0.21**	-0.21	-0.31	-0.12	-0.10
	(10 to 57)	(43 to 90)	(-0.34 to -0.08)	(-0.52 to 0.06)	(-0.68 to 0.11)	(-0.26 to 0.03)	(-0.22 to 0.03)
Lingual gyrus	**53**	**47**	**-0.21**	-0.20	-0.30	-0.13	-0.08
	(32 to 69)	(31 to 68)	(-0.34 to -0.09)	(-0.42 to 0.01)	(-0.66 to 0.10)	(-0.27 to 0.01)	(-0.18 to 0.03)
Superior occipital gyrus	**34**	**66**	**-0.21**	-0.23	-0.28	-0.13	-0.09
	(10 to 54)	(46 to 90)	(-0.33 to -0.09)	(-0.55 to 0.03)	(-0.62 to 0.11)	(-0.27 to 0.02)	(-0.20 to 0.03)
Supramarginal gyrus	**48**	**52**	**0.21**	**0.24**	0.21	**0.15**	0.06
	(23 to 66)	(34 to 77)	(0.08 to 0.33)	(0.01 to 0.51)	(-0.20 to 0.61)	(0.01 to 0.30)	(-0.06 to 0.17)
Angular gyrus	14	**86**	**0.17**	0.12	0.35	0.04	0.13
	(0 to 38)	(62 to 100)	(0.04 to 0.29)	(-1 to 1)	(-0.01 to 0.68)	(-0.10 to 0.18)	(0 to 25)
Right hemisphere (ROI)							
Hemisphere, mean thickness	**48**	**52**	**-0.17***	0.06	**-0.74***	0.04	**-0.21***
	(21 to 67)	(33 to 79)	(-0.29 to -0.04)	(-0.18 to 0.29)	(-1 to -0.3)	(-0.11 to 0.18)	(-0.32 to -0.09)
Inferior frontal gyrus, triangular part	**28**	**72**	**-0.18***	0.10	**-0.68***	0.05	**-0.22***
	(1 to 51)	(49 to 99)	(-0.30 to -0.05)	(-0.19 to 1)	(-0.93 to -0.32)	(-0.09 to 0.18)	(-0.32 to -0.11)
Supplementary motor area	**41**	**59**	**-0.17**	0	**-0.57**	0	**-0.17**
	(14 to 61)	(39 to 86)	(-0.29 to -0.04)	(-0.26 to 0.27)	(-0.93 to -0.13)	(-0.15 to 0.15)	(-0.28 to -0.04)
Posterior cingulate gyrus	**50**	**50**	**-0.19***	0.02	**-0.76***	0.02	**-0.21***
	(26 to 68)	(32 to 74)	(-0.32 to -0.06)	(-0.20 to 0.25)	(-1 to -0.33)	(-0.13 to 0.15)	(-0.33 to -0.09)
Parahippocampal gyrus	13	**87**	**-0.19**	-0.20	-0.33	-0.07	-0.12
	(0 to 39)	(61 to 100)	(-0.31 to -0.06)	(-1 to 1)	(-0.73 to 0.09)	(-0.21 to 0.08)	(-0.26 to 0.03)
Calcarine cortex	**36**	**64**	**-0.25***	-0.11	**-0.62***	-0.06	**-0.19***
	(12 to 56)	(44 to 88)	(-0.37 to -0.13)	(-0.36 to 0.15)	(-0.86 to -0.27)	(-0.19 to 0.07)	(-0.28 to -0.08)
Cuneus	**56**	**44**	**-0.25***	-0.18	**-0.51**	-0.12	**-0.13**
	(34 to 72)	(28 to 66)	(-0.37 to -0.13)	(-0.39 to 0.03)	(-0.84 to -0.10)	(-0.26 to 0.02)	(-0.22 to -0.03)
Lingual gyrus	**29**	**71**	**-0.27***	-0.19	**-0.55**	-0.10	**-0.18**
	(5 to 50)	(50 to 95)	(-0.39 to -0.15)	(-0.52 to 0.09)	(-0.82 to -0.18)	(-0.23 to 0.04)	(-0.28 to -0.06)
Superior occipital gyrus	**34**	**66**	**-0.15**	0.02	**-0.50**	0.01	**-0.16**
	(9 to 54)	(46 to 91)	(-0.27 to -0.03)	(-0.25 to 0.31)	(-0.79 to -0.11)	(-0.13 to 0.14)	(-0.26 to -0.04)
Middle occipital gyrus	**32**	**68**	**-0.17**	-0.07	**-0.42**	-0.04	**-0.13**
	(6 to 54)	(46 to 94)	(-0.29 to -0.04)	(-0.41 to 0.21)	(-0.77 to -0.02)	(-0.18 to 0.10)	(-0.25 to -0.01)
Fusiform gyrus	**29**	**71**	**-0.23***	0.01	**-0.73***	0	**-0.24***
	(3 to 51)	(49 to 97)	(-0.34 to -0.13)	(-0.23 to 0.27)	(-0.97 to -0.37)	(-0.11 to 0.12)	(-0.33 to -0.12)
Temporal pole, superior temporal gyrus	25	**75**	**-0.13**	-0.07	-0.31	-0.03	-0.10
	(0 to 47)	(53 to 100)	(-0.26 to -0.01)	(-1 to 0.29)	(-0.69 to 0.1)	(-0.18 to 0.12)	(-0.23 to 0.03)
Middle temporal gyrus	**48**	**52**	**-0.19**	-0.10	**-0.43**	-0.07	-0.12
	(25 to 66)	(34 to 75)	(-0.31 to -0.06)	(-0.34 to 0.12)	(-0.78 to -0.02)	(-0.21 to 0.08)	(-0.22 to 0)
Inferior temporal gyrus	**51**	**49**	**-0.14**	-0.09	-0.32	-0.06	-0.09
	(27 to 69)	(31 to 73)	(-0.27 to -0.01)	(-0.32 to 0.14)	(-0.71 to 0.12)	(-0.21 to 0.09)	(-0.20 to 0.03)

Note. Estimates in bold face are significant at α=0.05.

* Significant effect corrected for multiple comparisons (Bonferroni corrected: α = 0.05/39 regions = 0.0013).

Supplementary Table S3. Variance components of cortical volume and phenotypic, genetic and environmental overlap with liability to develop BD.

	Cortical volume						
	h^2	e^2	r_{ph}	r_g	r_e	r_{ph-g}	r_{ph-e}
Region	%						
Left hemisphere (ROI) [a]							
Hemisphere, $_{\text{total volume}}$	**88**	**12**	0.05	0.08	-0.10	0.07	-0.01
	(77 to 94)	(6 to 23)	(-0.08 to 0.19)	(-0.09 to 0.24)	(-0.55 to 0.38)	(-0.07 to 0.21)	(-0.08 to 0.05)
Middle frontal gyrus, $_{\text{orbital part}}$	1	**99**	**-0.17**	-1	-0.25	-0.07	-0.10
	(0 to 20)	(78 to 100)	(-0.29 to -0.05)	(-1 to 1)	(-0.57 to 0.10)	(-0.21 to 0.07)	(-0.22 to 0.04)
Supplementary motor area	24	**76**	**0.33**[*]	0.25	**0.65**[*]	0.11	**0.22**[*]
	(0 to 49)	(51 to 100)	(0.21 to 0.45)	(-0.06 to 1)	(0.29 to 0.91)	(-0.02 to 0.25)	(0.10 to 0.32)
Superior frontal gyrus, $_{\text{medial orbital part}}$	**39**	**61**	**-0.17**	-0.12	-0.32	-0.07	-0.10
	(15 to 58)	(42 to 85)	(-0.29 to -0.04)	(-0.38 to 0.13)	(-0.66 to 0.05)	(-0.21 to 0.07)	(-0.20 to 0.02)
Gyrus rectus	**29**	**71**	**-0.20**	-0.11	**-0.43**	-0.06	**-0.14**
	(3 to 51)	(49 to 97)	(-0.32 to -0.07)	(-0.46 to 0.20)	(-0.75 to -0.07)	(-0.20 to 0.08)	(-0.25 to -0.02)
Posterior cingulate gyrus	**32**	**68**	**0.15**	0.20	0.15	0.10	0.05
	(7 to 53)	(47 to 93)	(0.02 to 0.28)	(-0.08 to 0.56)	(-0.23 to 0.50)	(-0.04 to 0.25)	(-0.07 to 0.16)
Parahippocampal gyrus	**34**	**66**	**0.24**[*]	0.19	**0.42**	0.10	0.13
	(10 to 55)	(45 to 90)	(0.11 to 0.36)	(-0.08 to 0.50)	(0.01 to 0.76)	(-0.04 to 0.25)	(0 to 0.24)
Superior occipital gyrus	**40**	**60**	**-0.14**	-0.09	-0.29	-0.05	-0.09
	(16 to 60)	(40 to 84)	(-0.26 to -0.01)	(-0.35 to 0.16)	(-0.63 to 0.10)	(-0.20 to 0.09)	(-0.20 to 0.03)
Inferior parietal gyrus	**46**	**54**	**0.17**	0.17	0.24	0.11	0.07
	(22 to 64)	(36 to 78)	(0.04 to 0.30)	(-0.07 to 0.43)	(-0.18 to 0.62)	(-0.04 to 0.26)	(-0.05 to 0.18)
Supramarginal gyrus	12	**88**	**0.17**	-0.15	**0.59**	-0.05	**0.21**
	(0 to 36)	(64 to 100)	(0.04 to 0.29)	(-1 to 1)	(0.21 to 0.88)	(-0.18 to 0.10)	(0.08 to 0.33)
Angular gyrus	**26**	**74**	**0.13**	0.29	-0.01	0.14	0
	(2 to 47)	(53 to 98)	(0.01 to 0.26)	(-0.01 to 1)	(-0.36 to 0.36)	(-0.01 to 0.27)	(-0.12 to 0.12)
Right hemisphere (ROI) [a]							
Hemisphere, $_{\text{total volume}}$	**86**	**14**	-0.04	0.08	**-0.73**	0.07	**-0.11**
	(76 to 92)	(8 to 24)	(-0.17 to 0.10)	(-0.08 to 0.24)	(-0.96 to -0.30)	(-0.07 to 0.21)	(-0.16 to -0.04)
Supplementary motor area	**35**	**65**	**-0.19**	-0.17	-0.32	-0.09	-0.10
	(12 to 54)	(46 to 88)	(-0.31 to -0.07)	(-0.47 to 0.09)	(-0.69 to 0.10)	(-0.23 to 0.05)	(-0.22 to 0.03)
Supramarginal gyrus	**33**	**67**	**0.30**[*]	0.27	**0.48**	0.14	**0.15**
	(9 to 53)	(47 to 91)	(0.17 to 0.41)	(-0.01 to 0.66)	(0.05 to 0.84)	(0 to 0.29)	(0.02 to 0.27)

Note. Estimates in bold face are significant at α=0.05.

[*] Significant effect corrected for multiple comparisons (Bonferroni corrected: α = 0.05/39 regions = 0.0013).

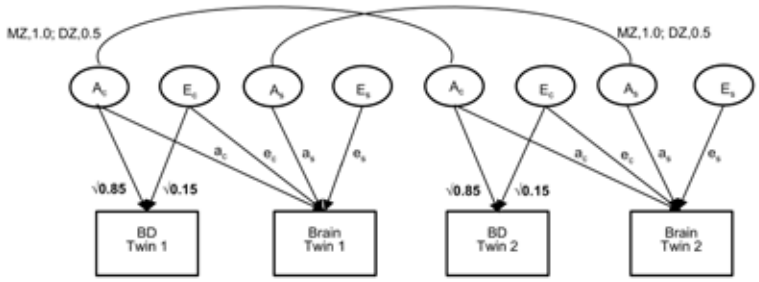
Supplementary Table S4. Cross-twin/within-trait and cross-twin/cross-trait correlations in MZ and DZ twin pairs.

	Cross-twin/within-trait[a]		Cross-twin/cross-trait[b]	
Measure/Region	**MZ**	**DZ**	**MZ**	**DZ**
Cortical surface				
Left hemisphere (ROI)				
Middle frontal gyrus, $_{\text{orbital part}}$	0.09	0.09	-0.04	-0.02
Supplementary motor area	**0.36**	-0.19	0.04	0.12
Superior frontal gyrus, $_{\text{medial orbital part}}$	**0.34**	**0.46**	-0.10	-0.01
Gyrus rectus	0.23	**0.60**	-0.03	0.19
Posterior cingulate gyrus	0.26	0.14	0.14	-0.02
Parahippocampal gyrus	**0.46**	0.18	0.02	0.18
Inferior parietal gyrus	**0.42**	0.3	0.13	0.10
Middle temporal gyrus	**0.53**	0.23	-0.07	-0.01
Right hemisphere (ROI)				
Supplementary motor area	0.25	0.28	**-0.19**	0.06
Olfactory cortex	0.17	0.20	0.04	0.05
Posterior cingulate gyrus	**0.40**	0.12	0.18	-0.03
Supramarginal gyrus	**0.38**	0.26	0.17	**0.21**
Temporal pole, $_{\text{superior temporal gyrus}}$	0.26	0.28	0.1	-0.04
Cortical thickness				
Left hemisphere (ROI)				
Rolandic operculum	**0.49**	0.26	0.17	0.1
Calcarine cortex	**0.37**	0.18	-0.06	-0.11
Lingual gyrus	**0.53**	**0.39**	-0.18	0.01
Superior occipital gyrus	**0.36**	0.21	-0.06	**-0.19**
Supramarginal gyrus	**0.50**	0.15	**0.21**	0.12
Angular gyrus	0.24	-0.13	0.12	0
Right hemisphere (ROI)				
Inferior frontal gyrus, $_{\text{triangular part}}$	**0.32**	-0.06	0	0.11
Supplementary motor area	**0.35**	0.23	0.04	-0.09
Posterior cingulate gyrus	**0.53**	**0.11**	0.04	-0.12
Parahippocampal gyrus	0.26	-0.11	0.03	**-0.19**
Calcarine cortex	0.24	**0.43**	-0.07	0.04
Cuneus	**0.57**	0.27	-0.14	-0.08
Lingual gyrus	0.24	**0.41**	-0.05	-0.02
Superior occipital gyrus	**0.37**	0.14	0.03	-0.08
Middle occipital gyrus	**0.42**	0.15	-0.03	-0.14
Fusiform gyrus	**0.42**	0.06	-0.03	0.12
Temporal pole, $_{\text{superior temporal gyrus}}$	**0.32**	0.12	0	-0.14
Middle temporal gyrus	**0.48**	0.28	-0.05	-0.15
Inferior temporal gyrus	**0.57**	0.19	-0.06	-0.09
Cortical volume				
Left hemisphere (ROI)				
Middle frontal gyrus, $_{\text{orbital part}}$	-0.05	-0.13	-0.07	-0.04
Supplementary motor area	**0.35**	-0.26	0.08	0.10
Superior frontal gyrus, $_{\text{medial orbital part}}$	**0.34**	**0.39**	-0.08	-0.03
Gyrus rectus	0.26	**0.42**	-0.06	0.09
Posterior cingulate gyrus	**0.34**	0.10	0.14	-0.07
Parahippocampal gyrus	**0.32**	0.08	-0.01	**0.22**
Superior occipital gyrus	**0.42**	0.21	-0.02	-0.08
Inferior parietal gyrus	**0.39**	**0.35**	0.07	0.05
Supramarginal gyrus	0.15	0.09	-0.05	0.01
Angular gyrus	0.24	0.18	0.04	**0.29**
Right hemisphere (ROI)				
Supplementary motor area	**0.37**	0.30	-0.14	0.10
Supramarginal gyrus	**0.36**	0.10	0.13	0.17

Note. Estimates in bold face are significant at α=0.05.

[a]Depicts the cross-twin correlation of cortical measure (trait 2 of both twins)

[b]Depicts the correlation between trait 1 (liability to BD) of twin 1 with trait 2 (brain measure) of twin 2.



Supplementary Figure S1. Example of a bivariate Cholesky decomposition for liability to BD and cortical phenotype. Latent additive genetic (A) and unique environmental (E) factors influence disease and brain, as indicated by arrows. The additive genetic factors (A) of monozygotic (MZ) twins are perfectly correlated (1.0), whereas those of dizygotic (DZ) twins are correlated at 0.5; unique environmental influences (E) are always uncorrelated between twins. Path coefficients (a_c and a_s) quantify the effects of genetic influences on the brain, where a_c represents genetic influences that also influence BD and a_s represents genetic influences that are unique for brain phenotypes. Similarly, path coefficients e_c and e_s quantify the effect of unique environmental (E) influences on brain phenotypes. Genetic variance for BD is fixed to 85%. Unique environmental factors account for 15% of the variance in BD.

Chapter 5

A study of genetic and environmental contributions to structural brain changes over time in twins concordant and discordant for bipolar disorder

Florian Bootsman, Rachel M. Brouwer, Hugo G. Schnack, Sanne M. Kemner, Manon H.J. Hillegers, Gayane Sarkisyan, Astrid C. van der Schot, Ronald Vonk, Hilleke E. Hulshoff Pol, Willem A. Nolen, René S. Kahn, Neeltje E.M. van Haren

In press

5.1 ABSTRACT

This is the first longitudinal twin study examining genetic and environmental contributions to the association between liability to bipolar disorder (BD) and changes over time in global brain volumes, and global and regional measures of cortical surface area, cortical thickness and cortical volume.

A total of 50 twins from pairs discordant or concordant for BD (monozygotic: 8 discordant and 3 concordant pairs, and 1 patient and 3 co-twins from incomplete pairs; dizygotic: 6 discordant and 2 concordant pairs, and 1 patient and 7 co-twins from incomplete pairs) underwent magnetic resonance imaging twice. In addition, 57 twins from healthy twin pairs (15 monozygotic and 10 dizygotic pairs, and 4 monozygotic and 3 dizygotic subjects from incomplete pairs) were also scanned twice. Mean follow-up duration for all twins was 7.5 years (standard deviation: 1.5 years). Data were analyzed using structural equation modeling software OpenMx.

The liability to BD was not associated with global or regional structural brain changes over time. Although we observed a subtle increase in cerebral white matter in BD patients, this effect disappeared after correction for multiple comparisons. Heritability of brain changes was generally low to moderate.

Structural brain changes appear to follow similar trajectories in BD patients and healthy controls. Existing brain abnormalities in BD do not appear to progressively change over time, but this requires additional confirmation. Further study with large cohorts is recommended to assess genetic and environmental influences on structural brain abnormalities in BD, while taking into account the influence of lithium on the brain.

Keywords: Bipolar disorder, longitudinal twin study, magnetic resonance imaging, brain volume, cortex, heritability

5.2 INTRODUCTION

In bipolar disorder (BD), structural brain abnormalities such as smaller volumes of prefrontal and temporal grey matter, corpus callosum and cerebral white matter, and enlarged volumes of the ventricles and striatum (Emsell and McDonald, 2009; Selvaraj et al., 2012) have been found. In addition, a few longitudinal case-control studies have reported prefrontal, anterior cingulate and subgenual grey matter volume loss, and less consistent findings in volume change in temporal and subcortical regions (Lim et al., 2013). Importantly, abnormal brain volume may be explained by deficits in cortical surface area and/or cortical thickness, structural brain elements comprising volume (Panizzon et al., 2009). Indeed, a number of case-control studies have demonstrated thinner prefrontal, temporal and parietal cortices in BD patients compared with healthy controls (Rimol et al., 2010; Foland-Ross et al., 2011). Cortical surface abnormalities in BD had not yet been shown (Fornito et al., 2009; Rimol et al., 2012) until we recently demonstrated subtle cortical surface abnormalities in our BD twin study (Bootsman et al., 2015).

In BD, the genetic contribution (heritability) to the disease is high and estimated to be around 85% (McGuffin et al., 2003). Nonetheless, the concordance rate of BD in monozygotic twins is only 40-70% (Craddock and Jones, 1999), indicating that environmental factors are likely involved as well. A number of neuroimaging twin studies have addressed the degree to which genetic and environmental factors contributing to BD are associated with deficits in brain volume, cortical surface area and cortical thickness. On a global level, the genetic risk for developing BD is associated with smaller white matter volume (Kieseppä et al., 2003; van der Schot et al., 2009) whereas we previously found smaller cerebral grey matter volume to be associated with unique environmental factors related to BD (van der Schot et al., 2009). Furthermore, we demonstrated that genes contributing to BD were associated with larger cortical surfaces in limbic and parietal regions, and thicker cortex in central and parietal regions. Moreover, unique environmental factors influencing BD were strongly associated with larger frontal, parietal and limbic, and smaller orbitofrontal surface area, as well as thinner frontal, limbic and occipital cortex (Bootsman et al., 2015).

Up to now, in BD, the genetic and environmental contributions to changes over time in global brain volume, and cortical surface area and cortical thickness remain to be determined. Therefore, in this longitudinal twin study, we assessed whether BD is associated with progressive changes in global brain volume, cortical surface area and cortical thickness and, if so, to what extent genetic and environmental factors associated with BD contribute to these brain changes.

5.3 MATERIALS AND METHODS

5.3.1 Subjects

A total of 50 twins from pairs discordant or concordant for BD (monozygotic (MZ): 8 discordant and 3 concordant pairs, and 1 patient and 3 co-twins from incomplete pairs; dizygotic (DZ): 6 discordant and 2 concordant pairs, and 1 patient and 7 co-twins from incomplete pairs) underwent magnetic resonance imaging (MRI) twice with a mean interval of 7.7 (standard deviation, SD: 1.5) years. In addition, 57 twins from healthy twin pairs (15 MZ and 10 DZ pairs, and 4 MZ and 3 DZ subjects from incomplete pairs) were scanned twice with a mean interval of 7.4 (SD: 1.6) years. The control twins were group-wise matched to the bipolar twins for zygosity, gender, age and parental education. At baseline, subjects were between 18 and 60 years of age (see **Table 1** for demographic information of twins that were included in all analyses, i.e. those where separation of grey and white matter was possible, see section on extraction of MRI phenotypes). Baseline MRI measures of our BD twin sample have been published previously (van der Schot et al., 2009; Bootsman et al., 2015).

Clinical assessment was carried out as previously described (van der Schot et al., 2009; Bootsman et al., 2015). Confirmation of clinical diagnosis of Axis I psychiatric disorders and Axis II personality disorders was achieved with the Structured Clinical Interview for DSM-IV (SCID) (First et al., 1997) and the Structured Interviews for DSM-IV Personality (SIDP and SCID-II) (First et al., 1997; Pfohl et al., 1997) respectively, and through consulting available medical records. At baseline, none of the subjects had drug or alcohol dependency in the last six months prior to inclusion while at follow-up 5 BD patients and 2 co-twins met diagnostic criteria for substance abuse or dependency in the past six months. With a medical history inventory it was verified that none had severe medical illness. Current mood state of patients was assessed at both measurements using the Young Mania Rating Scale (YMRS) (Young et al., 1978), the Inventory for Depressive Symptomatology (IDS) (Beck et al., 1961) and the Hamilton Depression Rating Scale (HDRS, second measurement only) (Hamilton, 1960). At baseline, most patients were euthymic with a YMRS score of 4 or less and an IDS score of 12 or less, except for 3 BD patients and 1 co-twin of a patient who were mildly to severely depressed (IDS scores of 17, 20, 38 and 14, respectively). At follow-up, 2 patients were mildly hypomanic (both had YMRS scores of 14) and 2 patients were mildly depressed (HDRS scores of 12 and 18 respectively). All other twins were euthymic. Of the total group, 12 patients had used lithium at both measurements, 5 patients had never used lithium, 5 patients stopped using lithium during the scan interval, and 4 patients started using lithium during the scan interval.

At baseline, none of the control twins had a history of Axis I psychiatric disorder or Axis II personality disorder according to DSM-IV criteria (confirmed with the SCID and SIDP, respectively), nor did they have a history of severe medical illness. At follow-up, 3 control

twins were diagnosed with specific phobia, 2 control twins were diagnosed with adjustment disorder, 2 control twins had experienced a major depressive episode and 3 control twins met diagnostic criteria for alcohol abuse. Healthy control twins had no first-degree relatives with a history of a major Axis I psychiatric disorder (DSM-IV) such as BD, schizophrenia, psychotic disorder, major mood disorder, anxiety disorder or substance-related disorder. Family histories of twin pairs were collected with the Family Interview Genetic Studies (FIGS) (Nurnberger jr. et al., 1994), performed with both twins. Determination of zygosity was achieved in the laboratory of the Division Biomedical Genetics, University Medical Center Utrecht (UMCU) with DNA fingerprinting using high polymorphic microsatellite markers 9-11. The study was approved by the medical ethics review board of the UMCU. Written informed consent was obtained from all participants after full explanation of the study aims and procedures (van der Schot et al., 2009).

5.3.2 Image acquisition

Magnetic resonance images were acquired on Philips 1.5 Tesla scanners (at baseline Intera, at follow-up Achieva, Philips, the Netherlands). Previously, both scanners were used in a large multicenter study, attesting to their cross-scanner reliability (Hibar et al., 2015). Furthermore, we calculated intraclass correlations (ICC) between measures of the intracranium, total brain, and grey and white matter in a test set of 5 subjects that were scanned on both scanners on the same day. These ICC's were found to be very high (> 0.99), but the small size of the test sample should be noted. We ensured that imaging parameters were identical across scanners and measurements: T1-weighted 3D fast field echo scans with 160–180 contiguous coronal slices (echo time=4.6 ms, repetition time=30 ms, flip angle=30°, 1x1x1.2 mm^3 voxels) (van der Schot et al., 2009). Within measurements, we tested whether a scanner drift was present by regressing change rates on scanning date, including age and sex and patient status as covariates. There was no significant effect of scanner date on the global brain measures. Locally, there were 26 regions (out of 468 tested: 3 types of measures for 39 local regions per hemisphere, across two time points [3x39x2x2]) that showed a significant effect of scanner, which is around the number that is expected to occur by chance alone. These regions did not overlap with those listed as having a significant genetic or environmental correlation with BD (see sections 'Extraction of MRI phenotypes' and 'Results'). At both assessments, MZ and DZ patient twins and healthy twins were randomly assigned to MRI slots, eliminating between-group biases due to scanner drifts.

5.3.3 Extraction of MRI phenotypes

To estimate global brain volumes, images were processed with previously described and validated protocols (Brouwer et al., 2010; Hulshoff Pol et al., 2004, see **Supplementary**

Table 1. Demographic and clinical characteristics of the bipolar twin pairs and matched healthy control twin pairs for which analysis of all brain measures was possible.

	Bipolar patients and their co-twins (n=46)		**Control twins** (n=55)							
	MZ[a] (n=24)	**DZ** (n=22)	**MZ**[b] (n=34)	**DZ** (n=21)	χ^2	**Z**	**F**	**df**	**p**	$\Delta_{post\text{-}hoc}$
$Age_{baseline}$, y[c]	36.9 (12.4)	41.6 (8.1)	40.1 (8.9)	39.0 (6.7)	-	-	1.1	3, *97*	0.36	
Follow-up duration[d]	7.4 (1.4)	8.0 (1.5)	7.3 (1.5)	7.6 (1.7)	-	-	1	3, *97*	0.41	
Parental educ., (highest) y[e]	12.2 (3)	12.1 (3.5)	10.5 (3.5)	12.1 (3.9)	-	-	1.8	3, *97*	0.16	
Education,y[f]	12.1 (2.3)	13.9 (1.9)	13.5 (2.6)	13.6 (1.7)	-	-	3.1	3, *97*	**0.03**	$MZ_{bd} < DZ_{bd}$, MZ/DZ_{ctrl}
Gender, m/f[g]	6/18	11/11	10/24	9/12	4.2	-	-	3	0.24	
Handedness[h] (right/left/both)	16/5/3	21/0/1	30/3/1	20/1/0	11.8	-	-	6	0.07	

	Patient (n=13)	**Co** (n=11)	**Patient** (n=11)	**Co** (n=11)		χ^2	**Z**	**F**	**df**	**p**	$\Delta_{post\text{-}hoc}$
Birth order, $_{1st/2nd}$[i]	5/8	8/3	4/7	7/4		-	-0.1	-	-	0.92	
Onset, age	26 (10.2)	-	31 (9.5)	-		-	-	1.5	1, *22*	0.23	
IDS_{base}[j]	4.4 (3.1)	3.2 (4.8)	8.8 (11.9)	2.6 (4)		-	-	1.8	3, *40*	0.16	
$YMRS_{base}$[k]	0.6 (0.9)	0.3 (0.6)	0.9 (1.4)	0.3 (0.9)		-	-	1	3, *39*	0.41	
$HDRS_{follow\text{-}up}$[l]	3.1 (3.5)	2.1 (2.7)	4.4 (5.3)	1.1 (1.9)		-	-	1.6	3, *36*	0.22	
$YMRS_{follow\text{-}up}$[m]	1.5 (1.6)	-	5.8 (5)	-		-	-	7.5	1, *20*	**0.01**	$MZ_{bd} < DZ_{bd}$
$GAF_{follow\text{-}up}$[n]	69.4. (15.1)	79.6 (8.2)	67 (14.2)	84.6 (16.3)		-	-	4	3, *41*	**0.01**	$DZ_{bd} < MZ_{co}$, DZ_{co} $MZ_{bd} < DZ_{co}$
$Hospitalizations_{follow\text{-}up}$[o]	0.3 (0.7)	-	1.2 (1.8)	-		-	-	2.4	1, *21*	0.14	
Psychotic sympt.[p]	6	1	7	2		0.7	-	-	1	0.39	
Lithium[q]						2.7	-	-	3	0.44	
-both time points	8	-	4	-							
-none at either time point	1	-	3	-							
-baseline only	3	-	2	-							
-follow-up only	1	-	2	-							
Substance abuse[r]	3	2	1	-	1						

Table 1. Legend
Abbreviations: MZ, Monozygotic; DZ, Dizygotic
n= number of individuals
Univariate analysis of variance (ANOVA) was performed for all demographic variables, except where noted otherwise.
[a] 7 MZ discordant pairs and 2 MZ concordant pairs (plus 2 MZ patients and 4 MZ co-twins from incomplete pairs); 6 DZ discordant pairs and 2 DZ concordant pairs (plus 1 DZ patient and 5 DZ co-twins from incomplete pairs)
[b] 15 MZ and 9 DZ healthy control pairs (plus 4 MZ and 3 DZ healthy controls from incomplete pairs)
[c] Baseline age did not differ significantly between MZ/DZ patient and control groups
[d] Follow-up duration did not differ significantly between MZ/DZ patient and control groups
[e] Parental education did not differ significantly between MZ/DZ patient and control groups
[f] MZ patients had significantly fewer years of education than the DZ patient group and MZ/DZ control groups
[g] Gender was not significantly different among groups (chi^2 test)
[h] Handedness was not significantly different among groups (chi^2 test)
[i] Birth order was compared between MZ and DZ patients only (Mann-Whitney U test)
[j] Baseline IDS score was not determined for 2 MZ co-twins
[k] Baseline YMRS score was not determined for 1 DZ patient and 2 MZ co-twins
[l] Follow-up HDRS score was not determined for 3 MZ patients, 1 DZ patient, 1 MZ co-twin and 1 DZ co-twin
[m] Follow-up YMRS score was not determined for 2 MZ patients or for any of the co-twins; MZ patients had significantly lower YMRS scores at follow-up than DZ patients
[n] GAF score was not determined for 1 MZ patient; DZ patients had lower GAF scores than MZ and DZ co-twins, and MZ patients had lower GAF scores than DZ co-twins
[o] Mean number of hospitalizations between measurements was not determined for 1 MZ patient
[p] Number of patients who had ever experienced psychotic symptoms was not significantly different among MZ and DZ patients
[q] Number of patients who had used lithium at a particular measurement was not significantly different among MZ and DZ patients
[r] 6 months prior to follow-up measurement, in the group for which all brain measures were available, 3 patients met diagnostic criteria for alcohol abuse and/or dependency (and one of them also for abuse of cannabis, sedatives and morphine), 1 patient had a cocaine dependency, 2 co-twins of patients met criteria for alcohol abuse and 1 control twin met diagnostic criteria for alcohol abuse.

Method). Here, we obtained volumes of the total brain, lateral and third ventricles, cerebrum and cerebral grey and white matter.

As applied previously (Bootsman et al., 2015), estimates of cortical surface area, cortical thickness and cortical volume were obtained with the CLASP (Constrained Laplacian Anatomic Segmentation Using Proximity) algorithm in a custom implementation of CIVET, developed at the McConnell Brain Imaging Centre of the Montreal Neurological Institute (MacDonald et al., 2000; Kabani et al., 2001; Kim et al., 2005). We used the grey/white/CSF segments from our own segmentation algorithm as inputs for the original CLASP algorithm. For each hemisphere, a 3-dimensional surface consisting of the 81,920 polygons and 40,962 vertices was fitted to the grey/white matter intersection, thereby creating the inner cortical surface. Subsequently, the outer cortical surface was then constructed by expanding the inner surface to fit the grey matter/CSF intersection (MacDonald et al., 2000; Kim et al., 2005). The surfaces are modeled as nets of polygons (triangles) with vertices being the points where the polygons meet (i.e. the angular points of the nets). For regional comparison of cortical thickness, the surfaces of the individual subjects were registered to an average surface

(Lyttelton et al., 2007). All polygon vertices on the outer surface had their counterparts on the inner surface; cortical thickness was defined as the distance between these inner and outer surface polygons. Using automated anatomical labeling (AAL), parcellations of 39 bilateral cortical regions that encompass these vertices were rendered, leaving out 6 bilateral subcortical and 13 bilateral cerebellar regions (Tzourio-Mazoyer et al., 2002). Cortical surface area was estimated from the mid surface, positioned half-way the outer and inner surfaces. Via multiplication of cortical surface area with cortical thickness, the cortical volume of each region was calculated. In addition, cortical surface area, mean cortical thickness and cortical volume of the entire cortex were estimated for baseline and follow-up. Please refer to **Supplementary Table S1** for specific cortical regions of which the surface area, thickness and volumes were extracted and assessed longitudinally.

For all brain measures, annual brain change was calculated by subtracting the value at baseline from the value at follow-up and dividing the difference by the number of years between measurements (for each subject individually). We used only these annual brain change measures for our analyses.

Of the initial group that underwent MR imaging twice, 6 subjects were excluded from analysis of grey and white matter volume, cortical surface, cortical thickness and cortical volume because grey matter could not be separated from white matter reliably due to poor image quality. However, these subjects were included for volume analysis of the total brain, cerebrum and ventricles.

5.3.4 Lithium use

Although the influence of lithium on brain volume has been noted (Hafeman et al., 2012), the small sample sizes of the respective lithium using groups ('lithium use at both measurements', 'started using lithium during interval', 'quit using lithium during interval' and 'no lithium use at either measurement') hindered reliable assessment of the influence of lithium on brain measures (**Table 1**). Therefore, we conducted our twin analyses without correcting for lithium use. Nonetheless, we reran our analyses including only those individuals who were either on or off lithium at both measurements to assess whether the nature of results changed.

5.3.5 Statistical analyses

Univariate analysis of (co)variance and correlations of brain changes with clinical measures

Within-group ('BD patients', 'co-twins of patients' and 'healthy controls') annual change in global brain measures was tested using a linear regression model. Here, we assessed whether annual global brain change in each measure as dependent variable was significantly different from zero, correcting for age at baseline, gender and handedness. For comparison between groups, univariate analysis of variance was performed with group ('BD patients'

versus 'healthy controls' and 'co-twins of patients' versus 'healthy controls') as between-subject variable and annual change in each brain measure as the dependent variable, after the effects of age, gender and handedness had been regressed out. These analyses were repeated with intracranial volume (ICV) as added covariate, for the global brain measures. Analyses of local annual cortical change were repeated with the appropriate global measures included as a covariate (i.e., local surface change with total surface, local thickness change with average thickness and local volumetric change with ICV) to correct for a possible global effect.

In addition, we assessed whether longitudinal annual brain changes in BD patients correlated with number of hospitalizations during scan interval, lifetime psychotic experiences or global functioning (GAF, global assessment of functioning) in those regions that showed a significant phenotypic correlation with BD.

Model fitting

Genetic model fitting procedures were similar to those previously described (Bootsman et al., 2015). A bivariate liability threshold model was chosen to estimate genetic and environmental contributions to the association between annual brain changes and liability to BD. We implemented this model in structural equation modeling software OpenMx (Kenny et al., 2009), which runs integrally under the statistical programming environment R (R development Core Team, 2008). Prior to model fitting, the variance in global and regional annual brain change measures attributable to age, gender and handedness was linearly regressed out in the whole group. Subsequently, a bivariate Cholesky decomposition was fitted to the obtained standardized residuals to estimate additive genetic (A), and unique environmental (E) variance components of annual brain changes and phenotypic, genetic and environmental overlap between BD and the annual brain change measures (**Supplementary Figure S1**). There was no evidence of shared environmental factors (C) influencing selected traits and these were therefore dropped from the model. In the bivariate model, disease status was dichotomous. It is assumed to represent an underlying continuous liability with a mean (SD) of 0 (1). BD patients will have high values on the liability scale and thereby cross a certain threshold (patient status=1). Discordant co-twins of patients and control twins will not cross the critical threshold (patient status = 0). Because we selected subjects based on disease status, estimates of the critical threshold and heritability (the relative contribution of genetic variance to total variance) for the underlying liability to BD would have been affected and not reflected population values. Therefore, prevalence and heritability of BD were not based on this sample but were fixed to population values; prevalence was set to 1% (Regeer et al., 2004) and heritability was set to 85% (McGuffin et al., 2003). To apply the threshold model to the brain measures, the standardized residuals of annual brain change

measures were used to construct a five-category ordinal scale identical for all subjects and put in the model. Thresholding was based on normality plots. Here, the boundaries of the 'outer' two categories were set at -1.5 SD and 1.5 SD respectively, whereas the middle three categories were 1 SD wide. By creating an ordinal scale of change we lose some information, but also noise that can be expected to be present in the annual change rates.

The phenotypic correlation (r_{ph}) indicates the magnitude of the association between phenotypes (e.g. liability to BD and annual brain change) and was based on calculations of within-twin/cross-trait correlations. Heritability (h^2) and influence of unique environment (e^2) was based on polychoric cross-twin/within-trait correlations whereas decomposing the observed correlation between liability to BD and brain change measures into genetic and environmental components was based on cross-twin/cross-trait correlations within MZ and DZ groups (Neale and Miller, 1997). The heritability of annual brain changes was determined within the bivariate model. Here, a larger correlation between traits in MZ twins than in DZ twins is indicative of genetic involvement in the trait, since MZ twins are genetically identical whereas DZ twins only share on average 50% of their segregating genes. However, a larger influence of shared environmental factors is more likely when MZ and DZ correlations are similar (Boomsma et al., 2002). The genetic (r_g) and unique environmental (r_e) correlations respectively indicate the degree of overlap in either genes or unique environment influencing both phenotypes. The phenotypic correlation is comprised of the genetic correlation weighted by the square root of the heritabilities of the two traits ($r_g * h_{BD} * h_{brain}$) and the environmental correlation weighted by the square root of environmental variance associated with the two traits ($r_e * e_{BD} * e_{brain}$). These components can be written as $r_{ph\text{-}g}$ and $r_{ph\text{-}e}$ (Toulopoulou et al., 2007). To ascertain the significance of variance components, different nested models were fitted to the data. Their goodness of fit was determined using Akaike's Information Criterion (AIC). A saturated model estimating the means, variances and correlations served as the baseline model to which more restrictive models were compared. As the AE-model consistently had the best fit, it was applied in all ROIs. 95% confidence intervals (CI) were obtained to determine the significance of parameter estimates and correlations (Neale and Miller, 1997).

Correction for multiple testing

Results were corrected for multiple testing by dividing the alpha of 0.05 by the estimated number of independent tests performed (Nyholt et al., 2004; Li and Ji, 2005). This number was computed using matrix decomposition of the correlation matrix of the relevant variables: for the 12 global measures [total brain, lateral and third ventricles, cerebrum, cerebral grey/white matter, left/right total cortical surface area, left/right cortical thickness and left/right cortical volumes], this yielded a significance threshold of $\alpha = 0.05/7.2 = 0.007$.

For the regional cortical areas we corrected for the total number of measures, for each type of measure (i.e. surface, thickness or volume) separately. This resulted in a significance threshold of α=0.05/66.7=0.0008 for surface area annual change rate. The threshold for annual thickness change and regional volume annual change rate were 0.05/48.8=0.001 and 0.05/66.5 = 0.0008 respectively. For the clinical analyses, a Bonferroni threshold was set at α=0.05/3=0.017 (3 clinical measures). In all tables and figures significance at α=0.05 and after correction for multiple testing is indicated.

5.4 RESULTS

5.4.1 Demographic and clinical characteristics

Demographic and clinical characteristics of subjects are presented in **Table 1** (for twins that were included in all analyses)**.** There were no significant differences between patient and control twins in age at baseline, follow-up duration or parental education. Univariate analysis of variance revealed group differences in years of education [$F_{3, 97}$=3.1, p <0.05] where post-hoc analysis showed that MZ BD twins had significant fewer years of education than DZ BD twins and control twins. Furthermore, MZ BD twins had lower YMRS scores at follow-up than DZ BD twins [$F_{1, 20}$=7.5, p<0.05].

5.4.2 Global annual brain changes

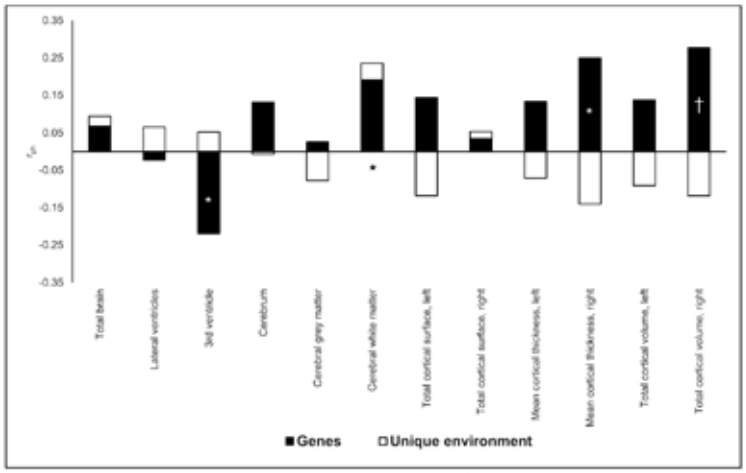
Univariate analysis of variance

Supplementary Table 2 shows the uncorrected annual global brain changes of BD patients, co-twins of patients and healthy controls. There were significant within-group annual brain changes in all subject-groups in the majority of global brain measures, including loss of cerebral grey matter volume and increase of cerebral white matter volume over time. No significant between-group differences were found for global brain measures, except for a more pronounced increase in cerebral white matter volume in patients as compared to healthy controls [$F_{1,77}$=5.61, p=0.02] (**Supplementary Figure S2**). However, this effect was no longer significant after correction for multiple testing, as the threshold for significance based on the Nyholt correction was set to α=0.007. Furthermore, we found an excessive increase in third ventricular volume over time in healthy controls as compared to co-twins [$F_{1,79}$=16.27, p<0.007]. Correction for ICV did not change the nature of results and was therefore not applied.

Genetic and environmental contributions to global annual brain changes

Table 2 shows the genetic (h^2) and environmental (e^2) contributions to annual brain change in global volumes and global cortical measures. The heritability of annual volume change

of the lateral ventricles was high (h^2: 71%) while heritability of annual volume changes in total brain, cerebrum, cerebral white matter, and in total cortical surface area, mean cortical thickness and total cortical volume was generally low to moderate (range h^2: 9% [right cortical volume] to 36% [right cortical thickness]). Genes and unique environment approximately equally influenced annual volume changes in the third ventricle and cerebral grey matter.



Note. Please refer to table 2 for the 95% confidence intervals of the total phenotypic correlation (r_{ph}). * Phenotypic, genetic or environmental correlation between BD and global brain change significant at α =0.05; †Phenotypic, genetic or environmental correlation between BD and global brain change significant after correction for multiple testing (Hyholt).

Figure 1. Genetic ($r_{ph\text{-}g}$) and environmental ($r_{ph\text{-}e}$) contributions to the significant phenotypic correlations between liability to BD and global brain change measures.

Liability to bipolar disorder and changes in global brain measures

BD was significantly associated with cerebral white matter volume increase (r_{ph}: 0.24) at α=0.05 but not at α=0.007. Although genes contributing to BD overlapped with those influencing cerebral white matter volume increase ($r_{ph\text{-}g}$:0.19) and explained most of the phenotypic correlation between BD liability and cerebral white matter increase, this genetic association was not significant (p = 0.07).

Liability to BD was neither significantly associated with annual changes in other global volume measures nor with annual changes in global cortical measures of surface area, thickness or volume (**Table 2 and Figure 1**).

5.4.3 Regional annual brain changes

The liability to BD was not significantly associated with regional annual changes in surface area, thickness and volume of the cortex after correction for multiple comparisons. Please refer to **Supplementary Table S3** and **Supplementary Figures S3 and S4** for findings that

Table 2. Genetic/environmental influences (with 95% confidence intervals) on global brain change, and phenotypic, genetic and environmental correlations with bipolar disorder

Measure (change) [a]	h^2	e^2	r_{ph}	r_g	r_e	$r_{ph\text{-}g}$	$r_{ph\text{-}e}$
	%						
Total brain	31	**69**†	0.09	0.13	0.08	0.07	0.03
	(0 to 70)	(30 to 100)	(-0.09 to 0.27)	(-1 to 1)	(-0.40 to 0.51)	(-0.13 to 0.26)	(-0.12 to 0.19)
Lateral ventricles	**71**†	**29**†	0.04	-0.03	0.31	-0.02	0.07
	(42 to 87)	(13 to 58)	(-0.15 to 0.23)	(-0.29 to 0.23)	(-0.20 to 0.76)	(-0.22 to 0.18)	(-0.04 to 0.17)
3rd ventricle	**50**†	**50**†	-0.17	**-0.34**	0.19	**-0.22**	0.05
	(17 to 73)	(27 to 83)	(-0.34 to 0.02)	(-0.76 to -0.03)	(-0.32 to 0.61)	(-0.40 to -0.02)	(-0.09 to 0.18)
Cerebrum	15	**85**†	0.12	0.37	-0.02	0.13	-0.01
	(0 to 55)	(45 to 100)	(-0.06 to 0.30)	(-1 to 1)	(-0.49 to 0.43)	(-0.07 to 0.32)	(-0.17 to 0.16)
Cerebral grey matter	**45**†	**55**†	-0.05	0.04	-0.27	0.03	-0.08
	(10 to 70)	(30 to 90)	(-0.24 to 0.14)	(-0.31 to 0.37)	(-0.75 to 0.25)	(-0.17 to 0.22)	(-0.22 to 0.08)
Cerebral white matter	16	**84**†	**0.24**	0.52	0.13	0.19	0.04
	(0 to 52)	(48 to 100)	(0.05 to 0.41)	(-0.03 to 1)	(-0.34 to 0.57)	(-0.01 to 0.38)	(-0.12 to 0.21)
Total cortical surface area, *left*	30	**70**†	0.03	0.28	-0.36	0.14	-0.11
	(0 to 67)	(33 to 100)	(-0.16 to 0.21)	(-1 to 1)	(-0.91 to 0.17)	(-0.06 to 0.34)	(-0.27 to 0.06)
Total cortical surface area, *right*	23	**77**†	0.05	0.08	0.05	0.03	0.02
	(0 to 61)	(39 to 100)	(-0.13 to 0.24)	(-1 to 1)	(-0.47 to 0.51)	(-0.16 to 0.24)	(-0.15 to 0.18)
Mean cortical thickness, *left*	18	**82**†	0.06	0.34	-0.20	0.13	-0.07
	(0 to 53)	(47 to 100)	(-0.13 to 0.25)	(-1 to 1)	(-0.66 to 0.30)	(-0.06 to 0.32)	(-0.23 to 0.11)
Mean cortical thickness, *right*	**36**†	**64**†	0.11	**0.45**	-0.45	**0.25**	-0.14
	(4 to 62)	(38 to 96)	(-0.08 to 0.29)	(0.09 to 1)	(-0.78 to 0.01)	(0.05 to 0.43)	(-0.26 to 0)
Total cortical volume, *left*	21	**79**†	0.05	0.33	-0.26	0.14	-0.09
	(0 to 58)	(42 to 100)	(-0.14 to 0.23)	(-1 to 1)	(-0.73 to 0.23)	(-0.06 to 0.33)	(-0.25 to 0.08)
Total cortical volume, *right*	**9**	**91**†	0.16	**1**†	-0.32	**0.28**†	-0.12
	(1 to 40)	(60 to 99)	(-0.03 to 0.34)	(0.28 to 1)	(-0.74 to 0.16)	(0.08 to 0.45)	(-0.26 to 0.06)

Note. Columns show variance components of genetic and unique environmental influences on brain measures (h^2, e^2) and the phenotypic (r_{ph}), genetic (r_g) and unique environmental (r_e) correlations between BD and brain change measure, as well as the genetic ($r_{ph\text{-}g}$) and unique environmental ($r_{ph\text{-}e}$) contributions to the total phenotypic correlation.

† significant after correction for multiple testing (Nyholt threshold: α=0.007), estimates in bold face are significant at α=0.05.

[a] The majority of global brain measures were analyzed in 26 patients, 24 co-twins of patients and 57 healthy controls. However, in 6 subjects (2 in each group) separation of grey and white matter was not possible due to bad image quality and were therefore excluded from analysis of cerebral grey and white matter, cortical surface, cortical thickness and cortical volume.

are uncorrected for multiple testing. Adding global measures as covariate did not change the results.

5.4.4 Association between brain changes and clinical measures

In BD patients, annual brain changes in white matter volume, and regional cortical surface area (left middle occipital gyrus, and right middle orbitofrontal gyrus), cortical thickness (left superior occipital gyrus, right precentral gyrus and right calcarine cortex) and cortical volume (left inferior frontal opercular gyrus, left lingual gyrus, right middle orbitofrontal gyrus, right inferior orbitofrontal gyrus and right calcarine gyrus) were not significantly associated with number of hospitalizations, lifetime experience of psychotic symptoms or GAF (at p=0.05). However, patients with lifetime experience of psychotic symptoms showed an increase in surface area of the left postcentral gyrus compared to nonpsychotic patients (p=0.03) but this difference disappeared after correction for multiple testing.

5.4.5 Lithium use

When the analyses were performed in only those patients who were either on or off lithium at both measurements, results did not change.

5.5 DISCUSSION

In this longitudinal twin study, we focused on the genetic and environmental contributions to global and regional annual structural brain changes in twins concordant and discordant for BD.

In this study, BD was not associated with annual changes in global brain volume nor with annual changes in global measures of cortical surface, cortical thickness and cortical volume over time, after correction for multiple comparisons. Here, we did observe a subtle cerebral white matter increase in BD patients at α=0.05 (supplementary figure S2). Cerebral white matter increase was negatively significantly associated with an increase in lateral ventricular volume at the individual level in the whole group ($r = -0.46$, $p < 0.01$), as well as separately in co-twins ($r= -0.74$, $p<0.01$), controls ($r= -0.46$, $p<0.01$), and, at trend level in patients ($r= -0.39$, $p=0.057$) (data available upon request). In our previous cross-sectional study, the genetic risk for BD was associated with smaller white matter volume (van der Schot et al., 2009). In previous longitudinal studies, white matter volume increase in BD patients has also been reported (Farrow et al., 2005), but so have decreases or no changes in white matter (Dickey et al., 2004; Moore et al., 2009). Interestingly, a review of 56 longitudinal structural MRI studies showed that in healthy subjects white matter volume increases until 45-50 years

of age (Westlye et al., 2010; Hedman et al., 2012). It appears white matter volume change in adulthood follows an inverted U-shaped pattern, with protracted increase until middle adulthood and a decrease after that, as opposed to grey matter change which appears to decrease linearly with age (Bartzokis et al., 2001; Taki et al., 2009; Westlye et al., 2010; Hedman et al., 2012). We speculate that the presently observed subtle increase in cerebral white matter volume in BD patients relative to healthy controls might indicate accelerated aging. However, in healthy individuals the peak white matter volume is reached around 50 years of age (Westlye e.a., 2010), and patients and control subjects included in the current study were on average younger than that. This issue requires confirmation in future longitudinal studies. As there may be a genetic link between BD and white matter volume (Kieseppä et al., 2003; van der Schot et al., 2009), identification of genes conferring risk for white matter pathology may aid in targeting genes associated with BD.

In a previous cross-sectional study we found that unique environmental factors associated with BD influenced smaller cerebral grey matter volume (van der Schot et al., 2009). In this study, we did not find an association between BD liability and structural changes over time in cerebral grey matter nor in total brain, cerebrum, lateral and third ventricles, total cortical surface, mean cortical thickness or total cortical volume. Thus it appears that structural changes over time in these brain measures are not different between BD patients and healthy controls. However, these findings should be interpreted with caution given their limited generalizability due to the small sample size of this study. Moreover, previously, both increases and decreases of grey matter volume in the prefrontal cortex in BD patients compared to healthy controls have been demonstrated (Lim et al., 2013), so findings so far remain inconclusive. Importantly, lithium use in patients influences brain volume (Hafeman et al., 2012) but is or is not accounted for in the different studies, which may contribute to the inconsistent findings concerning cerebral grey matter change over time in BD patients. Preservation of total brain volume in BD is supported by earlier findings as is stability of ventricular volume over time, despite ventricular enlargement at baseline (Lim et al., 2013).

Regionally, the liability to BD was not associated with changes over time in cortical surface area, cortical thickness or cortical volume, after correction for multiple testing. Previously demonstrated increased left and decreased right orbitofrontal volume, and grey matter loss in the bilateral anterior cingulate cortex (Lim et al., 2013), areas involved in emotion regulatory processing (Phillips et al., 2003), were not replicated in our study. However, those studies included pediatric and adolescent subjects whereas we studied adult subjects, which might account for the differences in findings. In addition, BD was not associated with annual change in the limbic and orbitofrontal regions that showed structural alterations in BD at baseline (Bootsman et al., 2015). So regional annual change may be similar among BD patients, co-twins of patients and healthy controls.

Finally, our findings in BD contrast with those reported in schizophrenia, where lateral ventricular enlargement, volumetric decreases in total brain, grey and white matter (Olabi et al., 2011), widespread excessive cortical thinning (van Haren et al., 2011) and regional volumetric decreases (Vita et al., 2012) have been consistently observed.

A number of limitations of this study should be considered. First, two different MRI scanners were used at different time points in this study, therefore the presence of scanner effects on our change measures cannot be completely ruled out. However, we ensured that scanner field strength (1.5 Tesla), imaging parameters and (pre)processing algorithms were equal for all subjects in both measurements. Moreover, all subjects were scanned at baseline with the Intera scanner and at follow-up with the Achieva scanner, so group differences and differences between zygosity groups could not be caused by scanner effects. Furthermore, the genetic analyses depend on relative differences of brain changes between members of twin pairs, rather than absolute changes. Therefore, a possible shift in the data due to the use of different scanners will not influence conclusions. Second, at baseline, subjects were between 18 and 60 years of age. Structural annual brain changes may be different for younger and older age groups. Third, although MZ and DZ BD twin pairs were matched to the MZ and DZ control pairs for gender, females were overrepresented. Therefore, brain measures were corrected for gender. Fourth, we did not correct for the influence of lithium use on brain change measures because the respective sample sizes of groups of patients who had used lithium at both measurements, those who had never used lithium or those who switched on or off lithium in between measurements were very small. This reduced statistical power to reliably assess the influence of lithium on brain measures. Although we chose not to correct for lithium use, we cannot rule out that lithium may have had an effect on brain measures in patients, particularly on measures of annual grey matter volume change. Previous studies have indicated increased grey matter in lithium treated patients (Hafeman et al., 2012). However, the influence of pharmacotherapy on grey matter volume remains a topic of debate as no influence of lithium or other psychotropic medication on grey matter tissue changes has also been reported (Moorhead et al., 2007; Kalmar et al., 2009). Furthermore, we did not assess the influence of other medication on brain measures (e.g. antipsychotics and antidepressants) due to the unreliable reporting by patients on their use and, importantly, the fact that their influence on the brain in BD appears to be limited (Hafeman et al., 2012). Fifth, although our BD twin cohort is the largest twin cohort that underwent longitudinal structural MRI described in the literature to date, the sample size still is relatively small. As a result, statistical power might have been suboptimal, increasing the risk for type I and type II errors. Therefore, conclusions drawn from the data should be regarded as provisional. In addition, some parameter estimates showed large confidence intervals, which may in part be due to noise in the data. Sixth, at follow-up, some of the

control subjects had met diagnostic criteria for DSM-IV psychopathology, which is to be expected based on population studies showing that a percentage of people will develop psychopathology in the general non-clinical population (e.g. Steel et al., 2014). It is hard to statistically remove the possible effects of including these subjects because of their small number and eliminating these subjects from analysis would have seriously affected statistical power. Therefore correction for psychopathology was not carried out.

In summary, no global or regional structural brain changes over time were found in BD after correction for multiple testing, suggesting that previously reported smaller grey and white matter volume may remain stable in adulthood. However, with larger cohorts, white matter volume increase may be detected in BD, possibly reflecting accelerated aging. Furthermore, the influence of lithium use on brain measures should be accounted for in future studies. Our findings contrast with the progressive widespread global and regional structural brain changes observed in schizophrenia in other studies. Given the limited number of longitudinal (twin) neuroimaging studies in BD, further assessment of the genetic and environmental contributions to brain changes over time in BD is warranted.

SUPPLEMENTARY INFORMATION

Supplementary Method

In brief, the T1-weighted images were automatically put in Talairach orientation and corrected for intensity non-uniformity artefacts (Sled et al., 1998). Separation of brain tissue from cerebrospinal fluid (CSF) and grey from white matter was performed with a partial volume segmentation method that accounts for the non-uniformity of the partial volume distribution that is due to the curvature of the cortex (Brouwer et al., 2010). Ventricular segmentation was performed as previously described (Schnack et al., 2001).

Supplementary Table S1. Regional **c**ortical brain regions that obtained with the CLASP algorithm for each hemisphere.

Region
1. Precentral gyrus
2. Superior frontal gyrus
3. Superior frontal gyrus, orbital part
4. Middle frontal gyrus
5. Middle frontal gyrus, orbital part
6. Inferior frontal gyrus, opercular part
7. Inferior frontal gyrus, triangular part
8. Inferior frontal gyrus, orbital part
9. Rolandic operculum
10. Supplementary motor area
11. Olfactory cortex
12. Superior frontal gyrus, medial
13. Superior frontal gyrus, medial orbital
14. Gyrus rectus
15. Insula
16. Anterior cingulate gyrus
17. Median cingulate gyrus
18. Posterior cingulate gyrus
19. Parahippocampal gyrus
20. Calcarine cortex
21. Cuneus
22. Lingual gyrus
23. Superior occipital gyrus
24. Middle occipital gyrus
25. Inferior occipital gyrus
26. Fusiform gyrus
27. Postcentral gyrus
28. Superior parietal gyrus
29. Inferior parietal gyrus
30. Supramarginal gyrus
31. Angular gyrus
32. Precuneus
33. Paracentral lobule
34. Heschl gyrus
35. Superior temporal gyrus
36. Temporal pole: superior temporal gyrus
37. Middle temporal gyrus
38. Temporal pole: middle temporal gyrus
39. Inferior temporal gyrus

Supplementary Table S2. Mean annual change in global brain volume, cortical surface area, cortical thickness and cortical volume for BD patients, co-twins of patients and healthy controls.

	Mean (SD) within-group annual change			Patients versus controls				Co-twins versus controls			
	Patients[+]	Co-twins	Healthy controls[++]	F	df	p	Δ	F	df	p	Δ
Measure	(n=26)	(n=24)	(n=57)								
Total brain (volume, in ml/year)	**1.75 (3.78)**	0.33 (3.48)	0.28 (2.32)	3.59	1,81	0.06		0	1,79	0.98	
Lateral ventricles (volume, in ml/year)	**0.33 (0.57)**[†]	**0.35 (0.42)**[†]	**0.39 (0.42)**[†]	0.03	1,81	0.88		0.78	1,79	0.38	
3[rd] ventricle (volume, in ml/year)	**0.04 (0.04)**[†]	**0.04 (0.03)**[†]	**0.06 (0.03)**[†]	2.28	1,81	0.14		**16.27**	**1,79**	**p<0.007**[†]	**ctrl>co**
Cerebrum (volume, in ml/year)	1.29 (3.35)	0.30 (3.20)	-0.05 (2.13)	3.62	1,81	0.06		0.28	1,79	0.6	
Cerebral grey matter (volume, in ml/year)[a]	**-2.97 (3.17)**[†]	**-2.67 (1.68)**[†]	**-2.54 (1.90)**[†]	0.27	1,77	0.61		0.13	1,75	0.72	
Cerebral white matter (volume, in ml/year)[a]	**4.38 (3.41)**[†]	**3.36 (2.85)**[†]	**2.52 (2.73)**[†]	**5.61**	**1,77**	**0.02**	**pt>ctrl**	1.6	1,75	0.21	
Cortical surface area (mm^2/year)[a]											
- left	**253.66 (212.78)**[†]	**272.38 (191.51)**[†]	**239.33 (155.56)**[†]	0.19	1,77	0.67		0.58	1,75	0.45	
- right	**344.59 (246.62)**[†]	**301.25(172.41)**[†]	**309.67 (171.83)**[†]	0.72	1,77	0.40		0.08	1,75	0.78	
Cortical thickness (mm/year)[a]											
- left	**-0.020 (0.016)**[†]	**-0.020 (0.012)**[†]	**-0.022 (0.011)**[†]	0.38	1,77	0.54		0.67	1,75	0.42	
- right	**-0.017 (0.019)**[†]	**-0.016 (0.012)**[†]	**-0.022 (0.011)**[†]	1.75	1,77	0.19		3.65	1,75	0.06	
Cortical volume (ml/year)[a]											
- left	**-0.96 (1.47)**[†]	**-0.82 (0.82)**[†]	**-1.1 (0.84)**[†]	0.5	1,77	0.48		1.56	1,75	0.22	
- right	-0.44 (1.44)	**-0.46 (0.81)**	**-0.89 (0.80)**[†]	3.04	1,77	0.09		3.56	1,75	0.06	

Note. Table depicts uncorrected annual global volume, cortical surface area, cortical thickness and cortical volume changes. Volumes of the total brain, ventricles, cerebrum (which includes subcortical volume) and cerebral grey and white matter were obtained with our own pipeline whereas measures of total cortical surface area, mean cortical thickness and total cortical volume were obtained with the CIVET pipeline. Univariate analysis of variance was performed with group (patients versus healthy controls and co-twins versus healthy controls) as between-group variable and annual brain change for each region as the dependent variable, after the effects of age, gender and handedness had been regressed out.

[+] Including concordant pairs, [++] including both twins from complete pairs.

[†] significant after correction for multiple testing (Nyholt correction, threshold: α =0.007), all bold face estimates are significant at α=0.05.

[a] Cerebral grey/white matter and cortical surface area, cortical thickness and cortical volume were analyzed in 24 patients, 22 co-twins of patients and 55 healthy controls.

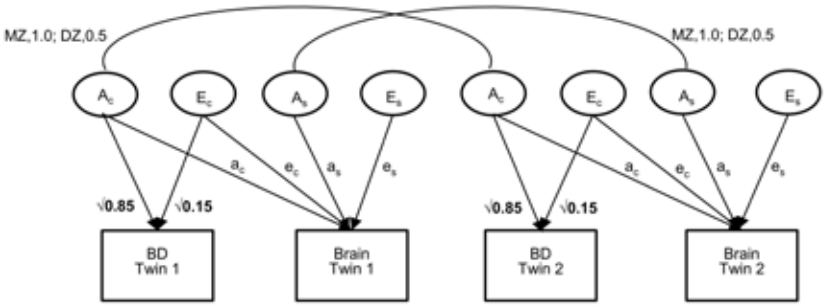
Supplementary Table S3. Genetic/environmental influences (with 95% confidence intervals) on changes in regional cortical surface area, cortical thickness and cortical volume, and phenotypic, genetic and environmental correlations with bipolar disorder

	h^2	e^2	r_{ph}	r_g	r_e	r_{ph-g}	r_{ph-e}
Measure/region (change)	%						
Cortical surface area							
Left middle occipital gyrus	**27**	**73**†	**0.21**	**0.60**	-0.23	**0.29**	-0.08
	(2 to 63)	(37 to 98)	(0.02 to 0.39)	(0.14 to 1)	(-0.68 to 0.27)	(0.08 to 0.48)	(-0.23 to 0.09)
Left postcentral gyrus ↓	0	**100**†	**-0.19**	1	**-0.61**	0.04	**-0.24**
	(0 to 24)	(76 to 100)	(-0.37 to -0.01)	(-1 to 1)	(-0.95 to -0.19)	(-0.14 to 0.22)	(-0.37 to -0.07)
Right middle frontal gyrus, *orbital part*	3	**97**†	**0.21**	1	0.14	0.15	0.05
	(0 to 23)	(77 to 100)	(0.02 to 0.39)	(-1 to 1)	(-0.32 to 0.55)	(-0.03 to 0.34)	(-0.12 to 0.21)
Cortical thickness							
Left superior occipital gyrus ↓	3	**97**†	**-0.23**	-0.48	-0.40	-0.07	-0.15
	(0 to 41)	(59 to 100)	(-0.41 to -0.03)	(-1 to 1)	(-0.78 to 0.09)	(-0.27 to 0.12)	(-0.30 to 0.03)
Right precentral gyrus ↓	**55**†	**45**†	**0.21**	**0.38**	-0.20	**0.26**	-0.05
	(21 to 78)	(22 to 79)	(0.01 to 0.39)	(0.08 to 0.76)	(-0.65 to 0.33)	(0.06 to 0.45)	(-0.18 to 0.09)
Right calcarine cortex	33	**67**†	**0.25**	0.32	0.25	0.17	0.08
	(0 to 62)	(38 to 100)	(0.05 to 0.43)	(-0.07 to 1)	(-0.29 to 0.74)	(-0.03 to 0.36)	(-0.09 to 0.24)
Cortical volume							
Left inferior frontal gyrus, *opercular part*	11	**89**†	**0.21**	0.46	0.19	0.14	0.07
	(0 to 48)	(52 to 100)	(0.01 to 0.39)	(-1 to 1)	(-0.33 to 0.65)	(-0.07 to 0.34)	(-0.12 to 0.24)
Left lingual gyrus	6	**94**†	**0.23**	**1**	0.01	**0.22**	0.01
	(0 to 22)	(78 to 100)	(0.05 to 0.39)	(0.31 to 1)	(-0.42 to 0.42)	(0.05 to 0.39)	(-0.16 to 0.16)
Right middle frontal gyrus, *orbital part*	7	**93**†	**0.23**	**1**	-0.02	**0.24**	-0.01
	(0 to 31)	(69 to 100)	(0.04 to 0.41)	(0.24 to 1)	(-0.49 to 0.44)	(0.04 to 0.42)	(-0.18 to 0.17)
Right inferior frontal gyrus, *orbital part* ↓	**8**	**92**†	**0.24**	**1**	-0.08	**0.27**	-0.03
	(1 to 37)	(63 to 99)	(0.05 to 0.41)	(0.28 to 1)	(-0.48 to 0.35)	(0.08 to 0.44)	(-0.17 to 0.13)
Right calcarine cortex	11	**89**†	**0.21**	**0.72**	-0.02	**0.22**	-0.01
	(0 to 56)	(44 to 100)	(0.03 to 0.39)	(0.06 to 1)	(-0.46 to 0.42)	(0.02 to 0.40)	(-0.17 to 0.16)

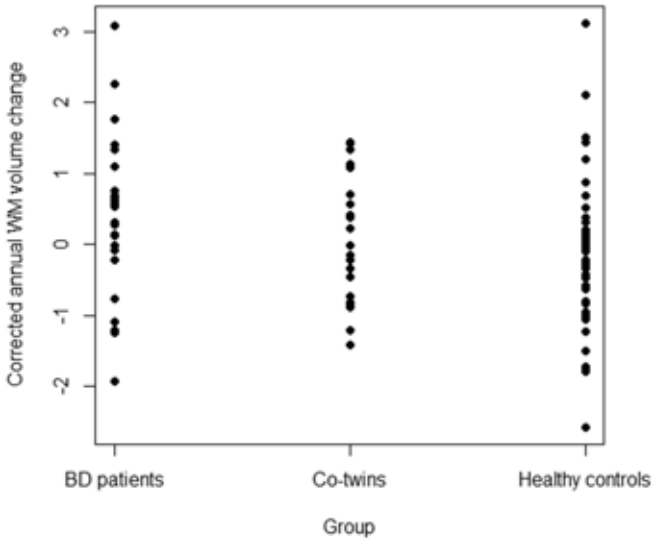
Note. Columns show variance components of genetic and unique environmental influences on brain measures (h^2, e^2) and the phenotypic (r_{ph}), genetic (r_g) and unique environmental (r_e) correlations between BD and brain change measure, as well as the genetic (r_{ph-g}) and unique environmental (r_{ph-e}) contributions to the total phenotypic correlation.

† significant after correction for multiple testing (Nyholt threshold for regional cortical measures, cortical surface area:α=0.0008; cortical thickness: α=0.001; cortical volume: α=0.0008), estimates of correlations in bold face are significant at α=0.05.

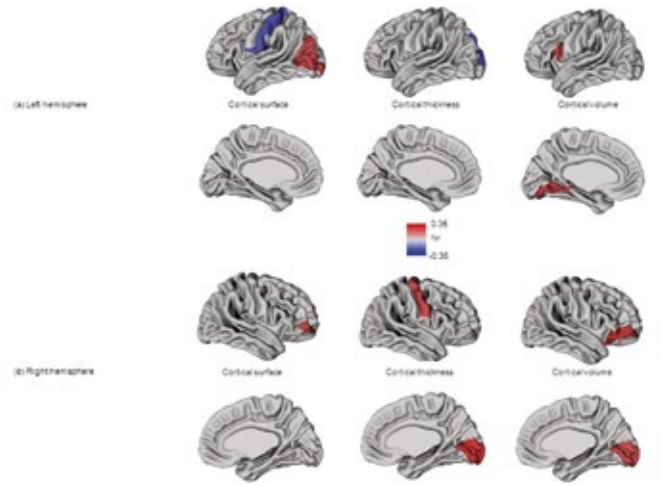
↓ Structure shows raw annual decrease in patients. A positive correlation may indicate increase or a smaller decrease than in controls or co-twins, which is why arrow was added.



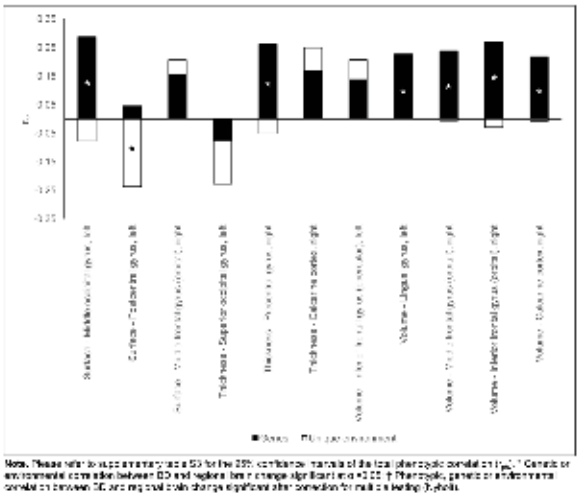
Supplementary Figure S1. Example of a bivariate Cholesky decomposition for liability to BD and brain change phenotype. Latent additive genetic (A) and unique environmental (E) factors influence disease and brain change, as indicated by arrows. The additive genetic factors (A) of monozygotic (MZ) twins are perfectly correlated (1.0), whereas those of dizygotic (DZ) twins are correlated at 0.5; unique environmental influences (E) are always uncorrelated between twins. Path coefficients (a_c and a_s) quantify the effects of genetic influences on the brain change measures, where a_c represents genetic influences that also influence BD and a_s represents genetic influences that are unique for brain change phenotypes. Similarly, path coefficients e_c and e_s quantify the effect of unique environmental (E) influences on brain change phenotypes. Genetic variance for BD is fixed to 85%. Unique environmental factors account for 15% of the variance in BD.



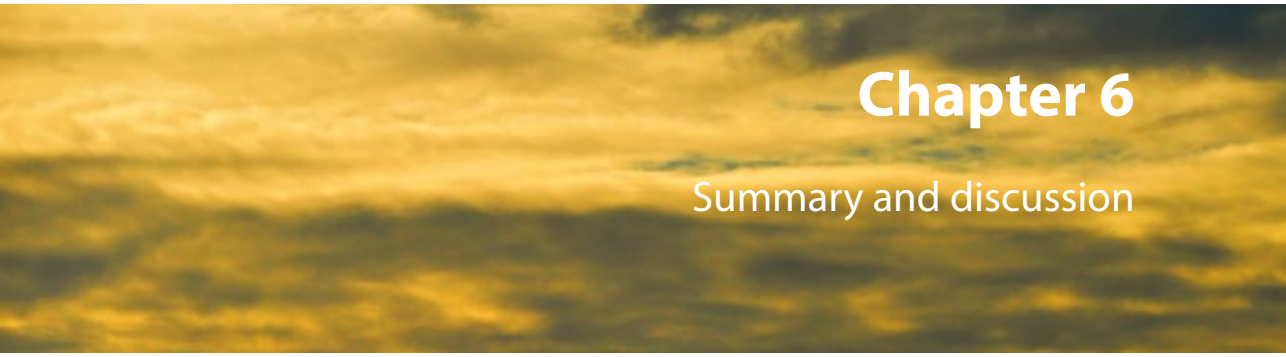
Supplementary Figure S2. Scatterplot of annual white matter volume change, corrected for age at baseline, gender and handedness, for each group.



Supplementary Figure S3. Phenotypic correlations (range $r_{ph:}$ 0.35 to -0.35) between BD liability and changes in cortical surface area/cortical thickness/cortical volume in both hemispheres. Red indicates positive correlation, blue indicates negative correlation and white indicates no correlation.



Supplementary Figure S4. Genetic (r_{ph-g}) and environmental (r_{ph-e}) contributions to the significant phenotypic correlations between liability to BD and regional brain change measures.



Chapter 6

Summary and discussion

"There are no facts, only interpretations"

Friedrich Nietzsche (notebooks, 1886-1887)

6.1 THE STUDY

From the outset, the goal of this longitudinal twin study was to determine whether bipolar disorder (BD), a severe and disabling mental illness associated with extreme episodic elevations in mood ((hypo)mania and depression), is associated with structural brain abnormalities at baseline and over time in adults. Second, using the classical twin design, its purpose was to assess the extent to which genetic and environmental factors contribute to brain structure and to the association between BD and brain structure, both cross-sectionally and longitudinally.

An important question was if BD is a condition that shows progressive brain structure loss over time. This question is particularly relevant in light of the Kraepelinian dichotomy that distinguishes between dementia praecox - now reformulated as schizophrenia and which, according to Kraepelin at the time, may have a progressive trajectory, fulminating in a final state of mental deterioration - and manic-depressive psychosis (presently defined as bipolar disorder), which has an episodic nature and may be less severe (Kraepelin, 1899; American Psychiatric Association, 1980, 1994, 2000). Although the neuroanatomy of schizophrenia and BD was not directly compared in this thesis and the focus was solely on determining structural brain deficits in the latter, the findings may be of help in fueling the discussion concerned with the degree to which the illnesses can be distinguished. Schizophrenia and BD are diagnostically distinct but do share some clinical characteristics (American Psychiatric Association, 1994, 2000; Lin and Mitchell, 2008). Moreover, studies where the genetic bases of both illnesses were assessed indicate a certain degree of genetic overlap between the two (Craddock and Owen, 2010), but also acknowledge small differences in possible genetic mechanisms underlying each disease. Here, schizophrenia appears to be associated with more copy number variants (CNV's) than BD, which may have implications for differences in neurodevelopmental trajectories between the disorders (Lee et al., 2012; Cardno and Owen, 2014). Regarding, neuroanatomy, it has been well documented that schizophrenia patients show widespread structural abnormalities, particularly in fronto-temporal brain regions, both at baseline and over time (van Haren et al., 2011; Shepherd et al., 2012; Vita et al., 2012;). In contrast, findings in BD are much less conclusive and

inconsistent across studies (Emsell and McDonald, 2009; Savitz and Drevets, 2009; Lim et al., 2013).
The studies discussed in this thesis were carried out with BD patients, their co-twins and healthy control twins. Subcortical and cortical brain volumes, as well as cortical surface area and cortical thickness, were assessed at baseline and over time in all subjects, using MRI. Data were analyzed with structural equation modeling software OpenMx (Kenny et al., 2009). Findings are discussed in light of some methodological limitations and implications.

6.2 THE ASSOCIATION BETWEEN HIPPOCAMPAL VOLUME AND LIFE EVENTS IN HEALTHY TWINS

In **chapter 2**, a study assessing the association between hippocampal volume and life events is described. The hippocampus is involved in negative feedback signaling to the hypothalamus during the hypothalamic-pituitary-adrenal (HPA) activated response to environmental stress, a process which results in secretion of glucocorticoids that are vital to short-term survival (Jankord and Herman, 2008). However, chronic stress may cause disruption in feedback signaling and consequent glucocorticoid toxicity in the hippocampus itself, which, in turn, may show atrophy and diminished neurogenesis (Sapolsky et al., 1990; Brown et al., 2004; Conrad, 2006; Mirescu and Gould, 2006; Gianaros et al., 2007; McEwen, 2007). Therefore, in this study, the association between hippocampal volume and the life events with different degrees of severity was ascertained. Moreover, the extent to which genes and environment influenced the association was determined.
Here, smaller hippocampal volume was related to higher severe life event load. These findings are in line with previous studies indicating an association between hippocampal volume loss and higher levels of stress (Gianaros et al., 2007; Papagni et al., 2011). Moreover, these findings were extended in the present study as we showed that particularly severe life event load was associated with smaller hippocampal volume and not total or mild life event loads. Furthermore, environmental factors that were shared between twins fully explained the association between smaller hippocampal volume and higher severe life event load. This suggests that severe life events that are experienced by both twins, for example losing one or both parents, may have the largest impact on the hippocampus. In addition, hippocampal volume was primarily influenced by genes whereas life event measures were predominantly influenced by shared and unique environmental factors. Based on these results, it is recommended to assess the influence of life events that affect twins similarly and to distinguish between severe and mild life events when assessing the relation between hippocampal volume and life stress.

Main findings chapter 2
Smaller hippocampal volume was strongly associated with higher load of severe life events that were shared between twins. The heritability of hippocampal volume was high, whereas life event loads were predominantly influenced by shared and unique environmental factors.

6.3 CONTRIBUTION OF GENES AND UNIQUE ENVIRONMENT TO CROSS-SECTIONAL AND LONGITUDINAL MEASURES OF SUBCORTICAL VOLUMES IN BIPOLAR DISORDER

In **chapter 3**, results of a longitudinal assessment of subcortical volumes in BD are discussed. As abnormal processing of emotion is considered a key feature of BD (Goodwin et al., 2007), many studies assessing brain abnormalities in BD have devoted their attention to a network of subcortical brain regions presumed to be involved in emotion processing, including the amygdala, hippocampus, thalamus and (ventral) striatum (Phillips et al., 2003, 2008). Indeed, studies have indicated abnormalities in some or most of these regions, cross-sectionally and longitudinally, but disagree on the extent and direction of the abnormalities, which is likely due to the influence of lithium use, age, familial load, mood status and variability in imaging methodology on brain measures (Emsell and McDonald, 2009; Savitz and Drevets, 2009; Rimol et al., 2010; Hallahan et al., 2011; Hajek et al., 2012; Lim et al., 2013; Phillips and Swartz, 2014). Nonetheless, subcortical volumetric abnormalities appear to be present to at least some degree in BD. However, the extent to which genes and environment contribute to the associations between BD and cross-sectional and longitudinal measures of subcortical volume remains largely unknown. Therefore, in this study, subcortical volumes were assessed at baseline and over time in twins concordant and discordant for BD.

At baseline, BD was associated with smaller volumes of the thalamus, putamen and nucleus accumbens, but only when correction for lithium use in BD patients was applied. Moreover, genes influencing both BD and these volumes showed substantial overlap. This indicates that volumetric abnormalities in subcortical brain regions may at least in part be influenced by genes that are involved in BD. This result is in line with a few other studies that have noted a shared genetic variance between BD and subcortical brain abnormalities in the ventral striatum, anterior putamen and thalamus (McDonald et al., 2004; McIntosh et al., 2004). Cross-sectional case-control studies have shown inconsistent findings regarding abnormalities in the thalamus, putamen and nucleus accumbens, reporting smaller, larger and preserved volumes (Almeida et al., 2009; Emsell and McDonald, 2009; Rimol et al., 2010; Hallahan et al., 2011; Haller et al., 2011; Hibar et al., 2013; Womer et al., 2014). Importantly, the associations between smaller volumes of the thalamus, putamen and nucleus accumbens and BD were only found after correction

for lithium use in BD patients, indicating that lithium has neuroprotective properties. It has been suggested previously that lithium use may mask true structural abnormalities in mood regulation brain regions, particularly in the hippocampus and amygdala (Hallahan et al., 2011, Hafeman et al., 2012; Hajek et al., 2012). In addition, heritability of subcortical brain volumes was high, which has been found previously (den Braber et al., 2013; Bohlken et al., 2014).
In contrast to the reported subcortical volume abnormalities in BD patients at baseline, change in subcortical volumes over time was not significantly different between BD patients, their co-twins and healthy controls. Irrespective of disease, heritability of subcortical volume change was low, indicating larger influence of unique environmental factors. Based on this study, BD does not appear to be associated with progressive subcortical volume loss but these results should be regarded as provisional as the sample size for the longitudinal assessment was relatively small, limiting statistical power to detect effects. Although there have not been any other longitudinal twin studies addressing subcortical volume change in BD, a few longitudinal case-control studies have been carried out. These show mixed results, noting both volumetric increase, decrease and stability in the hippocampus (Lim et al., 2013) but volume increases in the thalamus and caudate nucleus (Lisy et al., 2011). Further study, with larger cohorts, is needed to better characterize structural subcortical brain changes in BD. Furthermore, as lithium may influence subcortical brain volume, particularly in the hippocampus and amygdala, future studies could benefit from correcting for its use (Hallahan et al., 2011; Hajek et al., 2012). A failure to account for the influence of lithium on the brain may have resulted in underestimation of the nature of subcortical volume changes in this study and several others. Further study of the brain network involved in emotion processing and the genetic and environmental contributions to abnormalities in this network is recommended, in order to determine what mechanisms potentially underlie BD and how these may arise.

Main findings chapter 3
At baseline, BD patients showed significantly smaller volumes of the thalamus, putamen and nucleus accumbens, subcortical volumes involved in emotion processing. These smaller volumes were strongly influenced by genes that also contribute to the disease itself. BD was not associated with subcortical volume change in any of the examined regions. The heritability of subcortical brain regions at baseline was high, whereas change in subcortical regions showed low heritability.

6.4 GENETIC AND ENVIRONMENTAL INFLUENCES ON CORTICAL SURFACE AREA AND CORTICAL THICKNESS IN BIPOLAR DISORDER

In **chapter 4,** the association between BD and cortical surface area, cortical thickness and cortical volume was assessed. Traditionally, structural imaging studies devoted to brain

structure in BD have focused on volume or density as singular indices of brain anatomy. However, as surface area and thickness are structural brain elements that may either singly or both contribute to volume (Rakic, 1988, 1995; Pakkenberg and Gundersen, 1997; Jansen and Andermann, 2005; Fornito et al., 2008; Im et al., 2008; Pontious et al., 2008; Panizzon et al., 2009; Rakic et al., 2009; Winkler et al., 2010; Eyler et al., 2011), their assessment may aid in understanding subtle volumetric abnormalities that are sometimes encountered in BD. Therefore, in this study, the relation between BD and surface area, thickness and volume of the cortex was investigated. Here, not only the degree to which genes and environment overlapped between BD and these cortical brain measures was determined, but it was also ascertained whether the pattern of association between BD and cortical volume mimicked that of the association between BD and surface and/or thickness.

At the global level, unique environmental factors influencing BD were associated with mean thinner and smaller cortical volume of the right hemisphere. Regionally, genes contributing to BD overlapped with those influencing larger cortical surfaces in parietal and limbic regions, and thicker cortex in central and parietal regions. Unique environmental factors associated with BD overlapped with those influencing smaller frontal and temporal cortical surfaces, and thinner cortices across frontal, temporal and occipital regions, as well as larger and smaller frontal cortical volumes and larger parietal cortical volumes. It appears that particularly unique environmental factors drive the associations between BD and cortical surface, cortical thickness and cortical volume. Perhaps these influences concern illness-related factors contributing to both BD and cortical measures. Thinner prefrontal and temporal cortices have been found in BD in other studies (Rimol et al., 2010; Foland-Ross et al., 2011). Furthermore, in this study, the pattern of phenotypic association between BD and cortical volume mimicked that of the association between BD and cortical surface, indicating that cortical volume may be particularly dependent on cortical surface and not cortical thickness, which has been suggested previously (Pakkenberg and Gundersen, 1997; Im et al., 2008; Rakic, 2009; Winkler et al., 2010). In future studies, it is recommended to not only assess cortical volume but also investigate the surface and thickness of the cortex, as subtle deficits in these structural brain elements may be overlooked when assessing volume only. This approach has also been suggested by others (Fornito et al., 2008). Moreover, further assessment of the genetic and environmental contributions to cortical surface area and cortical thickness is recommended in order to determine the nature of disease-related structural neuropathology which, in turn, could serve to assist in ascertaining the degree of disease segregation or overlap.

Main findings chapter 4
At baseline, BD was associated with smaller and larger regional cortical surfaces and cortical volumes. BD was also associated with subtle deficits in cortical thickness. Unique environmental factors predominantly influenced cortical measures as well as the association between BD and cortical deficits. The pattern of association between BD and cortical volume mimicked that of the associations between BD and cortical surface, indicating that cortical volume is particularly dependent on cortical surface area and less on cortical thickness.

6.5 A STUDY OF GENETIC AND ENVIRONMENTAL CONTRIBUTIONS TO STRUCTURAL BRAIN CHANGES OVER TIME IN TWINS CONCORDANT AND DISCORDANT FOR BIPOLAR DISORDER

In **chapter 5**, changes over time in global brain volumes and in global and regional measures of cortical surface area, cortical thickness and cortical volume in twins concordant and discordant for BD and healthy control twins were investigated. Specifically, the purpose was to assess whether subtle abnormalities that were present at baseline (van der Schot et al., 2009; Bootsman et al., 2015) would show progressive changes over time in BD twins.
In the study described in this chapter, there were no significant changes over time in either global brain volumes or global and regional measures of cortical surface area, cortical thickness or cortical volume in BD patients compared to their co-twins or healthy control twins after correction for multiple testing. Applying no correction for multiple testing yielded a subtle significant increase in cerebral white matter which showed trend level genetic overlap with BD and might reflect accelerated aging. Furthermore, all three groups (BD patients, their co-twins and healthy control twins) showed significant within-group change over time in the majority of global brain volumes. In addition, heritability of structural brain changes was generally low to moderate. An apparent absence of global volume change has been reported before in longitudinal case-control studies (Lim et al., 2013) but longitudinal studies of brain change in BD are rare and those that are available show conflicting results, particularly concerning frontal and temporal, and white matter brain changes in BD, where stability of volumes and increases and decreases have all been reported (Lim et al., 2013). Our findings contrast with the progressive widespread global and regional structural brain changes observed in schizophrenia in other studies. To address the limited understanding we have of structural brain changes in BD and the inconsistencies in findings, larger studies that carefully account for the influence of medication use (particularly lithium use) on the brain should be carried out. Such studies may confirm preliminary findings reported in this study and others, allowing for better understanding of the trajectory of structural brain changes in BD.

Main findings chapter 5
Over time, BD did not show progressive change in global brain volumes or global and regional cortical brain structures. It appears that structural abnormalities that were found at baseline do not progressively worsen over time. Furthermore, change in global and regional cortical measures showed generally low heritability.

6.6 DISTINCTION BETWEEN SCHIZOPHRENIA AND BIPOLAR DISORDER

In this study, no structural brain changes over time were found in BD. The absence of structural brain changes in BD in this study contrasts with studies in schizophrenia, where cortical structural deficits over time have been described extensively and more consistently than in BD (van Haren et al., 2011; Olabi et al., 2011; Vita et al., 2012). Moreover, the genetic liability to schizophrenia was found to be associated with progressive volume loss in the total brain, and frontal and temporal regions in a sample of approximately the same size as the one used in this thesis (Brans et al., 2008). Although the diseases might show partly overlapping cortical and subcortical abnormalities, the neurodevelopmental pattern of structural brain abnormalities may differ between the two diseases, with schizophrenia patients generally showing greater deficits than BD patients, both at disease onset and later in adulthood (Demjaha et al., 2012; Hulshoff Pol et al., 2012, Shepherd et al., 2012; Hibar et al., 2013; Arango et al., 2014; Thompson et al., 2014; van Erp et al., 2015), suggesting the two diseases are indeed at least partly distinguishable based on neuroanatomical markers. Interestingly, although schizophrenia and BD show genetic and clinical overlap to some degree, significant differences in underlying genetic architecture, clinical manifestation and cognitive functioning have certainly been noted as well, with schizophrenia patients, again, showing more impairment than BD patients (Demjaha et al., 2012; Lee et al., 2012; Arango et al., 2014; Cardno and Owen, 2014), which also might suggest the two diseases to be at least partly separate entities. However, more cross-sectional and longitudinal studies directly comparing schizophrenia and BD are required to determine the degree of disease overlap or segregation based on brain morphology, as there is little consensus about the extent of the overlap between the two diseases, a discussion that has been taking place for over a century.

6.7 METHODOLOGICAL CONSIDERATIONS

The findings of the studies described in this thesis should be viewed in light of a few limitations.

6.7.1 Sample size

Although the BD twin cohort described in this thesis constitutes the largest intensively studied (particularly with respect to structural neuroimaging) BD twin sample followed over time in the literature to date, the size of the sample is still relatively small. Therefore, findings should be interpreted with caution, particularly those concerning the absence of changes in brain structure over time, as there is a possibility of a false-negative finding due to the small size of the sample. However, some of the structural brain abnormalities that were reported in this thesis proved to be quite robust. Particularly at baseline, where heritabilities for the various brain measures were expected to be high, statistical power might have been less of an issue than at follow-up, where heritabilities generally were low and difficult to reliably estimate.

6.7.2 Imaging methodology

In this study, two different MRI scanners (both Philips scanners operating at 1.5 Tesla) were used for the two measurements. Therefore the presence of scanner effects on the brain measures cannot be completely ruled out. However, all baseline scans were obtained on one scanner (Philips Intera) while all follow-up scans were obtained on the other scanner (Philips Achieva), thereby controlling for within-twin effects. In addition, it was ensured that scanner field strength (1.5 Tesla), imaging parameters and (pre)processing algorithms were equal for all subjects across all measurements.

6.7.3 Gender

The majority of subjects were of the female gender. As the differences in brain volumes between males and females have been documented extensively, with males having larger brains than females (Ruigrok et al., 2014), correction for gender was necessary to limit its influence as a confounder.

6.7.4 Age

As the included subjects showed quite a wide age range, statistically controlling for age was appropriate. However, particularly in the change measures, age might have played a role and influenced brain change differently in different age groups. To adjust for this effect the best way possible with a sample of this size, all change measures were divided by the number of years in each subject and corrected for age at baseline (and gender, see above).

6.7.5 Medication use

The influence of medication, especially lithium, on the morphology of the brain has been reported repeatedly. According to a recent review of studies assessing the impact of lithium,

antipsychotics and antidepressants on the brain reported lithium had the most clear effect on brain volume (Hafeman et al., 2012). Here, the effect of lithium appears quite robust in the hippocampus and amygdala, with BD patients using lithium showing larger volumes than BD patients not using lithium (Hallahan et al., 2011; Hafeman et al., 2012).
In the studies described in this thesis, correction for lithium was applied where possible by adding the difference in means between non-using patients and using patients to the individual brain measures of the latter group, thereby normalizing those measures to the level of that of the non-using group of patients. This was only possible in the baseline measures as the sample sizes of the respective lithium using groups in the longitudinal cohort ('no lithium at either measurement', 'used at both time points', 'started in the interval', 'stopped during the interval') were too small to reliably correct for lithium use. Therefore, an effect of lithium on the change measures cannot be ruled out completely.

6.7.6 Prevalence and heritability

In the twin models that were used, prevalence and heritability of BD were set a priori and not based on this sample, as doing so would not have resulted in representative population values. However, we tested whether varying the heritability and prevalence in our models would affect results. It did not.

6.8 FUTURE DIRECTIONS

Although the present study has limited direct implications for clinical practice, it could contribute to the strengthening of our knowledge about the neural mechanisms associated with BD. However, there is an additional, substantial need for better characterization of genetic and environmental factors that influence the structural brain in BD, especially over time. Therefore, in future studies, it might be worthwhile to further assess the influence of genes and (particularly) unique environmental factors (e.g. severe stressful life events such as losing a parent or loved one or recurrent hospitalization) on circumscribed subcortical and cortical areas, especially those involved in emotion processing. Here, assessment of both structural and functional connectivity within and between subcortical and cortical networks in genetically informative cohorts would be a step forward in determining the genetic and environmental contributions to brain mechanisms underlying affective dysregulation.
It would also be recommended to set up studies with larger cohorts and ascertain with higher statistical power the degree of genetic and environmental influences on the brain in BD, in both cross-sectional and longitudinal designs. Fortunately, large international collaborations are already being formed, allowing for assessment of brain structure in BD

in data sets with thousands of subjects. The Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) consortium is an example of such a large multicenter collaboration investigating brain structure in healthy subjects and psychiatric patients (Thompson et al., 2014). In a large sample, they showed, for example, reductions in thalamic, hippocampal and amygdala in BD patients (Hibar et al., 2013) but also investigated the contributions of specific gene variants to subcortical volume in thousands of healthy subjects (Hibar et al., 2015). In addition, if it is true that structural brain abnormalities arise earlier in life and remain stable during adulthood, studies involving younger subjects than those used in this study may be particularly beneficial to enhancing our understanding of the development of the disease itself as well as its associated brain deficits. One way would be to set up prospective imaging studies that follow large cohorts of subjects from childhood into adulthood, to obtain a better understanding of the (neuro)development of BD.

If future studies succeed in determining the genetic and environmental contributions to brain structure in BD, intervention strategies aimed at preventing or delaying the onset of the disease or reducing the burden in those who do progress to clinical manifestation of symptoms may be developed. For example, addressing coping styles in psychosocial interventions in those individuals that have experienced severe life events could prove beneficial both to their mental health as well as to their brain, the hippocampus in particular.

6.9 SUMMARY AND GENERAL CONCLUSION

In this longitudinal twin study, the goal was to ascertain the extent to which BD shows abnormalities in subcortical and cortical brain regions at baseline and over time. Baseline assessments in this study revealed that BD was associated with smaller volumes of the thalamus, putamen and nucleus accumbens. These associations were strongly influenced by genes influencing both BD and smaller volumes of these regions, indicating that BD may share some genetic background with structural abnormalities in brain regions involved in emotion processing. However, BD showed no significant relation with structural abnormalities in the hippocampus, a subcortical brain structure that is often implicated in the disease and the atrophy of which has been linked to life stress in healthy individuals. Furthermore, at baseline, BD was associated with subtle deficits in cortical surface area, cortical thickness and cortical volume. Here, unique environmental factors primarily drove the associations between BD and cortical measures, particularly those involving cortical thinning. In contrast, although brain structure deficits may arise during development, these appear to remain relatively stable over time in adulthood, as structural brain changes in cortical and subcortical brain regions were not different among BD patients, co-twins and

healthy controls. This suggests that BD is not a disease that shows progressive structural brain atrophy and may be specifically associated with subtle brain abnormalities that arise earlier in life, perhaps even during neurodevelopment. This finding contrasts with results obtained in schizophrenia, where significant progressive brain atrophy during adulthood, particularly in fronto-temporal regions, has been observed and was found to be linked to the genetic liability for the disease. Therefore, schizophrenia and BD may show different neuroanatomical trajectories, with the former showing more severe abnormalities than the latter. However, more longitudinal studies with larger cohorts than the one described here are necessary to confirm the absence of structural brain change in BD, especially since findings of structural brain changes in prefrontal cortical and subcortical structures conflict considerably between studies. In addition, many factors may affect brain measures and could contribute to inconsistencies across studies, such as lithium use, age, familial load, mood status, number of episodes, number of hospitalizations and variability in imaging methodology.

Finally, to account more comprehensively for the genetic and environmental contributions to structural brain abnormalities in BD, cross-sectionally and longitudinally, (prospective) longitudinal studies with larger cohorts (such as those included by the ENIGMA consortium) are required. More insight into which specific genetic and/or environmental influences contribute to the development of BD and associated brain anomalies, particularly in brain networks subserving emotion processing, could serve as a guide in the developing intervention strategies directed at reducing disease burden or even preventing disease onset.

References

Adler, C.M., Delbello, M.P., Jarvis, K., Levine, A., Adams, J., Strakowski, S.M. (2007).Voxel-based study of structural changes in first-episode patients with bipolar disorder. *Biological Psychiatry* 61, 776-781.

Adriano, F., Caltagirone, C., Spalletta, G. (2012). Hippocampal volume reduction in first-episode and chronic schizophrenia: a review and meta-analysis. *Neuroscientist* 18, 180-200.

Akiskal, H.S. (2006). Special issue on circular insanity and beyond: historic contributions of French psychiatry to contemporary concepts and research on bipolar disorder. *Journal of Affective Disorders* 96, 141-3.

Aldhafeeri, F.M., Mackenzie, I., Kay, T., Alghamdi, J., Sluming, V. (2012). Regional brain responses to pleasant and unpleasant IAPS pictures: different networks. *Neuroscience Letters* 512, 94-98.

Almeida, J.R., Akkal, D., Hassel, S., Travis, M.J., Banihashemi, L., Kerr, N., Kupfer, D.J., Phillips, M.L., (2009). Reduced gray matter volume in ventral prefrontal cortex but not amygdala in bipolar disorder: significant effects of gender and trait anxiety. *Psychiatry Research* 171, 54-68.

American Psychiatric Association, (1980). Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition. Washington, DC: Author.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders, 4th ed. Washington, DC: Author.

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders, 4th ed., text rev. Washington, DC: Author.

American Psychiatry Association (2013). Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Arlington: American Psychiatric Publishing.

Arango, C., Fraguas, D., Parellada, M. (2014). Differential neurodevelopmental trajectories in patients with early-onset bipolar and schizophrenia disorders. *Schizophrenia Bulletin* 40 (Suppl 2):S138–S146.

Baaré, W.F., Hulshoff Pol, H.E., Boomsma, D.I., Posthuma, D., de Geus, E.J., Schnack, H.G., van Haren, N.E., van Oel, C.J., Kahn, R.S. (2001). Quantitative genetic modeling of variation in human brain morphology. *Cerebral Cortex* 11, 816-824.

Bartzokis, G., Beckson, M., Lu, P.H., Nuechterlein, K.H., Edwards, N., Mintz, J. (2001). Age-related changes in frontal and temporal lobe volumes in men: a magnetic resonance imaging study. *Archives of General Psychiatry* 58,461–5.

Bearden, C.E., Thompson, P.M., Dalwani, M., Hayashi, K.M., Lee, A.D., Nicoletti, M., Trakhtenbroit, M., Glahn, D.C., Brambilla, P., Sassi, R.B., Mallinger, A.G., Frank, E., Kupfer, D.J., Soares, J.C. (2007).Greater cortical gray matter density in lithium-treated patients with bipolar disorder. *Biological Psychiatry* 62, 7-16.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry* 4, 561-571.

Bender, R.E., Alloy, L.B., Sylvia, L.G., Urosevic, S., Abramson, L.Y. (2010). Generation of life events in bipolar spectrum disorders: a re-examination and extension of the stress generation theory. *Journal of Clinical Psychology* 66, 907-926.

Bird, C.M. and Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nature Reviews Neuroscience* 9, 182-194.

Bleuler, E. (1908). "Die Prognose der Dementia Praecox -- Schizophreniegruppe". Allgemeine Zeitschrift fur Psychiatrie 65, 436–434.

Blokland, G.A., de Zubicaray, G.I., McMahon, K.L., Wright, M.J. (2012). Genetic and environmental influences on neuroimaging phenotypes: a meta-analytical perspective on twin imaging studies. *Twin Research Human Genetics* 15, 351-371.

Blond, B.N., Fredericks, C.A., Blumberg, H.P. (2012). Functional neuroanatomy of bipolar disorder: structure, function, and connectivity in an amygdala-anterior paralimbic neural system. *Bipolar Disorders* 14, 340-355.

Bohlken, M.M., Brouwer, R.M., Mandl, R.C.W., van Haren, N.E.M., Brans, R.G.H., van Baal, G.C.M., de Geus, E.J.C., Boomsma, D.I., Kahn, R.S., Hulshoff Pol, H.E. (2014). Genes contributing to subcortical volumes and intellectual ability implicate the thalamus. *Human Brain Mapping* 35, 2632-2642.

Boomsma, D.I. (1998). Twin registers in Europe: an overview. *Twin Research* 1,34-51.

Boomsma, D., Busjahn, A., Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics* 3, 872-882.

Bora, E., Fornito, A., Yücel, M., Pantelis, C. (2010). Voxelwise meta-analysis of gray matter abnormalities in bipolar disorder. *Biological Psychiatry* 67, 1097-1105.

Bootsman, F., Brouwer, R.M., Schnack, H.G., van Baal, G.C.M., van der Schot, A.C., Vonk, R., Hulshoff Pol, H.E., Nolen, W.A., Kahn, R.S., van Haren, N.E.M. (2015). Genetic and environmental influences on cortical surface area and cortical thickness in bipolar disorder. *Psychological Medicine* 45, 193-204.

Brambilla, P., Bellani, M., Yeh, P-H., Soares, J.C. and Tansella M. (2009) White matter connectivity in bipolar disorder. *International Review of Psychiatry* 21, 380-386.

Brans, G.H., van Haren, N.E.M., van Baal, G.C.M., Schnack, H.G., Kahn, R.S., Hulshoff Pol, H.E. (2008). Heritability of changes in brain volume over time in twin pairs discordant for schizophrenia. *Archives of General Psychiatry* 65, 1259-1268.

Brans, R.G., Kahn, R.S., Schnack, H.G., van Baal, G.C., Posthuma, D., van Haren, N.E., Lepage, C., Lerch, J.P., Collins, D.L., Evans, A.C., Boomsma, D.I., Hulshoff Pol, H.E. (2010). Brain plasticity and intellectual ability are influenced by shared genes. *Journal of Neuroscience* 30, 5519-5524.

Brouwer, R.M., Hedman, A.M., van Haren, N.E.M., Schnack, H.G., Brans, R.G., Smit, D.J., Kahn, R.S., Boomsma, D.I., Hulshoff Pol, H.E. (2014). Heritability of brain volume change and its relation to intelligence. *Neuroimage* 100, 676-683.

Brouwer, R.M., Hulshoff Pol, H.E., Schnack, H.G. (2010). Segmentation of MRI brain scans using nonuniform partial volume densities. *Neuroimage* 49, 467-477.

Brown, G.H., Harris, T.O. (1978). Social Origins of Depression: A study of Psychiatric Disorder in Woman. London: Tavistock Publications.

Brown, G.H., Harris, T.O. (1982). Fall-off in the reporting of life events. *Social Psychiatry* 17,23-28.

Brown, G.H., Harris, T.O. (1989). Life Events and Illness. New York: Guilford Press.

Brown, E.S., Woolston, J., Frol, A., Bobadilla, L., Khan, D.A., Hanczyc, M., Rush AJ, Fleckenstein, J., Babcock, E., Cullum, C.M. (2004). Hippocampal volume, spectroscopy, cognition, and mood in patients receiving corticosteroid therapy. *Biological Psychiatry* 55, 538-545.

Cardno, A.G. and Owen, M.J. (2014). Genetic relationships between schizophrenia, bipolar disorder and schizoaffective disorder. *Schizophrenia Bulletin* 40, 504-515.

Chen, C-H., Suckling, J., Lennox, B.R., Ooi, C., Bullmore, E.T. (2011). A quantitative meta-analysis of fMRI studies in bipolar disorder. Bipolar Disorders 13, 1–15.

Conrad, C.D. (2006). What is the functional significance of chronic stress-induced CA3 dendritic retraction within the hippocampus? *Behavioral Cognitive Neuroscience Reviews* 5, 41-60.

Craddock, N., Jones, I. (1999). Genetics of bipolar disorder. *Journal of Medical Genetics* 36, 585-594.

Craddock, N., Owen, M.J. (2010). The Kraepelinian dichotomy – going, going... but still not gone. *British Journal of Psychiatry* 196, 92–95.

Craddock, N., Sklar, P. (2013). Genetics of bipolar disorder. *The Lancet* 381, 1654-62.

den Braber, A., Bohlken, M.M., Brouwer, R.M., van 't Ent, D., Kanai, R., Kahn, R.S., de Geus, E.J.C., Hulshoff Pol, H.E., Boomsma, D.I. (2013). Heritability of subcortical brain measures: A perspective for future genome-wide association studies. *Neuroimage* 83, 98-102.

DelBello, M.P., Zimmerman, M.E., Mills, N.P., Getz, G.E., Strakowski, S.M. (2004). Magnetic resonance imaging analysis of amygdala and other subcortical brain regions in adolescents with bipolar disorder. *Bipolar Disorders* 6, 43-52.

Delvecchio, G., Sugranyes, G., Frangou, S. (2013). Evidence of diagnostic specificity in the neural correlates of facial affect processing in bipolar disorder and schizophrenia: a meta-analysis of functional imaging studies. *Psychological Medicine* 43, 553-569.

Demjaha, A., MacCabe, J.H., Murray, R.M. (2012). How Genes and Environmental Factors Determine the Different Neurodevelopmental Trajectories of Schizophrenia and Bipolar Disorder. *Schizophrenia Bulletin* 38, 209–214.

Dickey C.C., Salisbury, D.F., Nagy, A.I., Hirayasu, Y., Lee, C.U., McCarley, R.W., Shenton, M.E. (2004). Follow-up MRI study of prefrontal volumes in first-episode psychotic patients. *Schizophrenia Research* 71, 349-351.

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

Emsell, L., McDonald, C. (2009). The structural neuroimaging of bipolar disorder. *International Review of Psychiatry* 21, 297-313.

Eyler, L.T., Prom-Wormley, E., Panizzon, M.S., Kaup, A.R., Fennema-Notestine, C., Neale, M.C., Jernigan, T.L., Fischl, B., Franz, C.E., Lyons, M.J., Grant, M., Stevens, A., Pacheco, J., Perry, M.E., Schmitt, J.E., Seidman, L.J., Thermenos, H.W., Tsuang, M.T., Chen, C.H., Thompson, W.K., Jak, A., Dale, A.M.,Kremen, W.S. (2011).Genetic and environmental contributions to regional cortical surface area in humans: a magnetic resonance imaging twin study. *Cerebral Cortex* 21, 2313-2321.

Falret, J.P. (1851). Marche de la folie. Gazette Hopitaux. 24, 18–19.

Farrow, T.F., Whitford, T.J., Williams, L.M., Gomes, L., Harris, A.W. (2005). Diagnosis-related regional gray matter loss over two years in first episode schizophrenia and bipolar disorder. *Biological Psychiatry* 58, 713-723.

First, M.B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., Benjamin, L.S. (1997). Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II). American Psychiatric Press, Washington, DC.

First, M., Spitzer, R., Gibbon, M., Williams, J. (1997). User's guide for the Structured Clinical Interview for DSM-IV Axis I Disorders - Clinical Version (SCID-CV). American Psychiatric Press , Washington, DC.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale AM. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex* 14, 11–22.

Fjell, A.M., Westlye, L.T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Salat, D.H., Greve, D.N., Fischl, B., Dale, A.M., Walhovd, K.B. (2009). Minute effects of sex on the aging brain: a multisample magnetic resonance imaging study of healthy aging and Alzheimer's disease. *Journal of Neuroscience* 29, 8774-8783.

Foland-Ross, L.C., Thompson, P.M., Sugar, C.A., Madsen, S.K., Shen, J.K., Penfold, C., Ahlf, K., Rasser, P.E., Fischer, J., Yang, Y., Townsend, J., Bookheimer, S.Y., Altshuler, L.L. (2011). Investigation of cortical thickness abnormalities in lithium-free adults with bipolar I disorder using cortical pattern matching. *American Journal of Psychiatry* 168, 530-539.

Fornito, A., Yucel, M., Wood, S.J., Adamson, C., Velakoulis, D., Saling, M.M., McGorry, P.D., Pantelis, C. (2008). Surface-based morphometry of the anterior cingulate cortex in first episode schizophrenia. *Human Brain Mapping* 29, 478-489.

Fornito, A., Yucel, M., Wood, S.J., Bechdolf, A., Carter, S., Adamson, C., Velakoulis, D., Saling, M.M., McGorry, P.D., Pantelis, C. (2009). Anterior cingulate cortex abnormalities associated with a first psychotic episode in bipolar disorder. *British Journal of Psychiatry* 194, 426-433.

Fusar-Poli, P., Howes, O., Bechdolf, A., Borgwardt, S. (2012). Mapping vulnerability to bipolar disorder: a systematic review and meta-analysis of neuroimaging studies. *Journal of Psychiatry and Neuroscience* 37, 170–184.

Gianaros, P.J., Jennings, J.R., Sheu, L.K., Greer, P.J., Kuller, L.H., Matthews, K.A. (2007). Prospective reports of chronic life stress predict decreased grey matter volume in the hippocampus. *Neuroimage* 35, 795-803.

Gilbertson, M.W., Shenton, M.E., Ciszewski, A., Kasai, K., Lasko, N.B., Orr, S.P., Pitman, R.K. (2002). Smaller hippocampal volume predicts pathologic vulnerability to psychological trauma. *Nature Neuroscience* 5, 1242-1247.

Goodwin, F.K., Jamison, K.R., Ghaemi, S.N. (2007). Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression. 2nd ed. New York. Oxford University Press, New York

Gu, Y., Janoschka, S., Ge, S. (2013). Neurogenesis and hippocampal plasticity in adult brain. *Current Topics in Behavioral Neuroscience* 15, 31-48.

Hafeman, D.M., Chang, K.D., Garrett, A.S., Sanders, E.M., Phillips, M.L. (2012). Effects of medication on neuroimaging findings in bipolar disorder: an updated review. *Bipolar Disorders* 14, 375-410.

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

Hallahan, B., Newell, J., Soares, J.C., Brambilla, P., Strakowski, S.M., Fleck, D.E., Kieseppä, T., Altshuler, L.L., Fornito, A., Malhi, G.S., McIntosh, A.M., Yurgelun-Todd, D.A., Labar, K.S., Sharma, V., MacQueen, G.M., Murray, R.M., McDonald, C. (2011). Structural magnetic resonance imaging in bipolar disorder: An international collaborative mega-analysis of individual adult patient data. *Biological Psychiatry* 69, 326-335.

Haller, S., Xekardaki, A., Delaloye, C., Canuto, A., Lövblad, K.O., Gold, G., Giannakopoulos, P. (2011). Combined analysis of grey matter voxel-based morphometry and white matter tract-based spatial statistics in late-life bipolar disorder. *Journal of Psychiatry and Neuroscience* 36, 391-401.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry* 23, 56-62.

Hammen, C. (1991). Generation of stress in the course of unipolar depression. *Journal of Abnormal Psychology* 100, 555-561

Harris, T. (2001). Recent developments in understanding the psychosocial aspects of depression. *British Medical Bulletin* 57, 17-32.

Hedman, A.M., van Haren, N.E.M., Schnack, H.G., Kahn, R.S., Hulshoff Pol, H.E. (2012). Human brain changes across the life span: a review of 56 longitudinal magnetic resonance imaging studies. *Human Brain Mapping* 33, 1987-2002.

Hibar, D.P., van Erp, T.G.M., Turner, J.A., Haukvik, U.K., Agartz, I., et al. (2013). Meta-analysis of structural brain differences in bipolar disorder: the ENIGMA-Bipolar Disorder Project. Seattle, WA: OHBM.

Hibar, D.P., Stein, J.L., Renteria, M.E., Arias-Vasquez, A., Desrivières, S., Jahanshad, N., Toro, R., Wittfeld, K. et al. (2015). Common genetic variants influence human subcortical brain structures. *Nature* 520, 224-229.

Hillegers, M.H., Burger, H., Wals, M., Reichart, C.G., Verhulst, F.C., Nolen, W.A., Ormel, J. (2004). Impact of stressful life events, familial loading and their interaction on the onset of mood disorders: study in a high-risk cohort of adolescent offspring of parents with bipolar disorder. *British Journal of Psychiatry* 185, 97-101.

Hogstrom, L.J., Westlye, L.T., Walhovd, K.B., Fjell, A.M. (2013). The structure of the cerebral cortex across adult life: age-related patterns of surface area, thickness and gyrification. *Cerebral Cortex* **23**, 2521-2530.

Hulshoff Pol, H.E., Brans, R.G., van Haren, N.E.M., Schnack, H.G., Langen, M., Baaré, W.F., van Oel, C.J., Kahn, R.S. (2004). Gray and white matter volume abnormalities in monozygotic and same-gender dizygotic twins discordant for schizophrenia. *Biological Psychiatry* 55, 126-130.

Hulshoff Pol, H.E., van Baal, G.C., Schnack, H.G., Brans, R.G., van der Schot, A.C., Brouwer, R.M., van Haren, N.E., Lepage, C., Collins, D.L., Evans, A.C., Boomsma, D.I., Nolen, W., Kahn, R.S. (2012). Overlapping and segregating structural brain abnormalities in twins with schizophrenia or bipolar disorder. *Archives of General Psychiatry* 69, 349-359.

Im, K., Lee, J.M., Lyttelton, O., Kim, S.H., Evans, A.C., Kim, S.I. (2008). Brain size and cortical structure in the adult human brain. *Cerebral Cortex* 18, 2181-2191.

Inano, S., Takao, H., Hayashi, N., Yoshioka, N., Mori, H., Kunimatsu, A., Abe, O., Ohtomo, K. (2013). Effects of age and gender on neuroanatomical volumes. *Journal of Magnetic Resonance Imaging* 37, 1072-1076.

Jamison, K.R. (2000). Suicide and bipolar disorder. *Journal of Clinical Psychiatry* 61, 47-51.

Jancke, L., Merillat, S., Liem, F., Hanggi, J. (2015). Brain size, sex, and the aging brain. *Human Brain Mapping* 36, 150-169.

Jankord, R., Herman, J.P. (2008). Limbic regulation of hypothalamo-pituitary-adrenocortical function during acute and chronic stress. *Annals of the New York Academy of Sciences* 1148, 64-73.

Jansen, A., Andermann, E. (2005). Genetics of the polymicrogyria syndromes. *Journal of Medical Genetics* 42, 369-378.

Johnson, S.L. (2005). Life events in bipolar disorder: towards more specific models. *Clinical Psychology Review* 25, 1008-1027.

Jung, C.G. (1903). On Manic mood disorder. In: Psychiatric Studies Vol-1, Collected works, second edition. Routledge and Kegan Paul, pp.109-111.

Kabani, N., Le, G.G., MacDonald, D., Evans, A.C. (2001). Measurement of cortical thickness using an automated 3-D algorithm: a validation study. *Neuroimage* 13, 375-380.

Kalmar, J.H., Wang, F., Spencer, L., Edmiston, E., Lacadie, C.M., Martin, A., Constable, R.T., Duncan, J.S., Staib, L.H., Papademetris, X., Blumberg, H.P. (2009). Preliminary evidence for progressive prefrontal abnormalities in adolescents and young adults with bipolar disorder. *Journal of the International Neuropsychological Society* 15, 476-481.

Karl, A., Schaefer, M., Malta, L.S., Dorfel, D., Rohleder, N., Werner, A. (2006). A meta-analysis of structural brain abnormalities in PTSD. *Neuroscience and Biobehavioral Reviews* 30, 1004-1031.

Kempton, M.J., Geddes, J.R., Ettinger, U., Williams, S.C.R., Grasby, P.M. (2008). Meta-analysis, Database, and Meta-regression of 98 Structural Imaging Studies in Bipolar Disorder. *Archives of General Psychiatry* 65, 1017-1032.

Kempton, M.J., Salvador, Z., Munafo, M.R., Geddes, J.R., Simmons, A., Frangou, S., Williams, S.C. (2011). Structural neuroimaging studies in major depressive disorder. Meta-analysis and comparison with bipolar disorder. *Archives of General Psychiatry* 68, 675-690.

Kendler, K.S., Baker J.H. (2007). Genetic influences on measures of the environment: a systematic review. Psychological Medicine 37, 615-626.

Kenny, S., Andric, M., Boker, S.M., Neale, M.C., Wilde, M., Small, S.L. (2009). Parallel workflows for data-driven structural equation modeling in functional neuroimaging. *Frontiers in Neuroinformatics* 3,34.

Kessler, R.C., Berglund, P., Demler, O., Jin, R., Merikangas, K.R., Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* 62,593–602.

Kieseppä, T., van Erp, T.G., Haukka, J., Partonen, T., Cannon, T.D., Poutanen, V.P., Kaprio, J., Lönnqvist, J. (2003). Reduced left hemispheric white matter volume in twins with bipolar I disorder. *Biological Psychiatry* 54, 896-905.

Kieseppa T, Partonen T, Haukka J, Kaprio, J., Lönnqvist, J. (2004). High concordance of bipolar I disorder in a nationwide sample of twins. *American Journal of Psychiatry* 161, 1814–21.

Kim, J.S., Singh, V., Lee, J.K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J.M., Kim, S.I., Evans, A.C. (2005). Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *Neuroimage* 27, 210-221.

Kochunov, P., Mangin, J.F., Coyle, T., Lancaster, J., Thompson, P., Rivière, D., Cointepas, Y., Regis, J., Schlosser, A., Royall, D.R., Zilles, K., Mazziotta, J., Toga, A., Fox, P.T. (2005). Age-related morphology trends of cortical sulci. *Human Brain Mapping* 26, 210-220.

Kraepelin, E. (1899). Manic depressive insanity and paranoia. E.S. Livingston, Edinburgh.

Kuhn, S., Gallinat, J. (2013). Gray matter correlates of posttraumatic stress disorder: a quantitative meta-analysis. *Biological Psychiatry* 73, 70-74.

Kupka, R., Goossens, P., van Bendegem, M., Daemen, P., Daggenvoorde, T., Daniels, M., Dols, A., Hillegers, M., Hoogelander, A., ter Kulve, E., Peetoom, T., Schulte, R., Stevens, A. (2015). Multidisciplinary guideline bipolar disorders (Multidisciplinaire richlijn bipolaire stoornissen). 3rd revised ed. De Tijdstroom uitgeverij BV, Utrecht, Nederland.

Lee, K.W., Woon, P.S., Teo, Y.Y., Sim, K. (2012). Genome wide association studies (GWAS) and copy number variation (CNV) studies of the major psychoses: what have we learnt? *Neuroscience and Biobehavioral Reviews* 36, 556-571.

Li, J., Ji, L. (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 95, 221-227

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 and Biobehavioral Reviews* 37, 418-435.

Lin, P., Mitchell, B.D. (2008). Approaches for unraveling the joint genetic determinants of schizophrenia and bipolar disorder. *Schizophrenia Bulletin* 34, 791-797.

Lisy, M.E., Jarvis, K.B., Delbello, M.P., Mills, N.P., Weber, W.A., Fleck, D., Strakowski, S.M., Adler, C.M. (2011). Progressive neurostructural changes in adolescent and adult patients with bipolar disorder. *Bipolar Disorders* 13, 396-405.

Lyoo, I.K., Sung, Y.H., Dager, S.R., Friedman, S.D., Lee, J.Y., Kim, S.J., Kim, N., Dunner, D.L., Renshaw, P.F. (2006). Regional cerebral cortical thinning in bipolar disorder. *Bipolar Disorders* 8, 65-74.

Lyttelton, O., Boucher, M., Robbins, S., Evans, A. (2007). An unbiased iterative group registration template for cortical surface analysis. *Neuroimage* 34, 1535-1544.

MacDonald, D., Kabani, N., Avis, D., Evans, A.C. (2000). Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage* 12, 340-356.

Marchand, W.R., Yurgelun-Todd, D. (2010). Striatal structure and function in mood disorders: a comprehensive review. *Bipolar Disorders* 12, 764-785.

McDonald, C., Bullmore, E.T., Sham, P.C., Chitnis, X., Wickham, H., Bramon, E., Murray, R.M. (2004). Association of genetic risks for schizophrenia and bipolar disorder with specific and generic brain structural endophenotypes. *Archives of General Psychiatry* 61, 974-984.

McDonald, C., Marshall, N., Sham, P.C., Bullmore, E.T., Schulze, K., Chapple, B., Bramon, E., Filbey, F., Quraishi, S., Walshe, M., Murray, R.M. (2006). Regional brain morphometry in patients with schizophrenia or bipolar disorder and their unaffected relatives. *American Journal of Psychiatry* 163, 478-87.

McEwen, B.S. (2007). Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiological Reviews* 87, 873-904.

McGuffin, P., Rijsdijk, F., Andrew, M., Sham, P., Katz, R., Cardno, A. (2003). The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Archives of General Psychiatry* 60, 497-502.

McIntosh, A.M., Job, D.E., Moorhead, W.J., Harrison, L.K., Forrester, K., Lawrie, S.M., Johnstone, E.C. (2004). Voxel-based morphometry of patients with schizophrenia or bipolar disorder and their unaffected relatives. *Biological Psychiatry* 56, 544-552.

McIntosh, A.M., Job, D.E., Moorhead, W.J., Harrison, L.K., Whalley, H.C., Johnstone, E.C., Lawrie, S.M. (2006). Genetic liability to schizophrenia or bipolar disorder and its relationship to brain structure. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetic.)* 141B, 76-83.

Meda, S.A., Gill, A., Stevens, M.C., Lorenzoni, R.P., Glahn, D.C., Calhoun, V.D., Sweeney, J.A., Tamminga, C.A., Keshavan, M.S., Thaker, G., Pearlson, G.D. (2012). Differences in resting-state functional magnetic resonance imaging functional network connectivity between schizophrenia and psychotic bipolar probands and their unaffected first-degree relatives. *Biological Psychiatry* 71, 881–889.

Merikangas, K.R., Jin, R., He, J-P., Kessler, R.C., Lee, S., Sampson, N.A., Barkov, Z. (2011). Prevalence and correlates of bipolar spectrum disorder in the World Mental Health Survey Initiative. *Archives of General Psychiatry* 68, 241–251.

Milner, B., Squire, L.R., Kandel, E.R. (1998). Cognitive neuroscience and the study of memory. *Neuron* 20, 445-468.

Mirescu, C., Gould, E. (2006). Stress and adult neurogenesis. *Hippocampus* 16, 233-238.

Moore, G.J., Cortese, B.M., Glitz, D.A., Zajac-Benitez, C., Quiroz, J.A., Uhde, T.W., Drevets, W.C., Manji, H.K. (2009). A longitudinal study of the effects of lithium treatment on prefrontal and subgenual prefrontal gray matter volume in treatment-responsive bipolar disorder patients. *Journal of Clinical Psychiatry* 70, 699-705.

Moorhead, T.W., McKirdy, J., Sussmann, J.E., Hall, J., Lawrie, S.M., Johnstone, E.C., McIntosh, A.M., (2007). Progressive gray matter loss in patients with bipolar disorder. *Biological Psychiatry* 62, 894-900.

Mountcastle, V.B. (1997).The columnar organization of the neocortex. *Brain* 120 (Pt 4), 701-722.

Mühleisen, T.W., Leber, M., Schulze, T.G., Strohmaier, J., Degenhardt, F., Treutlein, J., et al. (2014). Genome-wide association study reveals two new risk loci for bipolar disorder. *Nature Communications* 5:3339. doi: 10.1038/ncomms4339

Narr, K.L., Bilder, R.M., Toga, A.W., Woods, R.P., Rex, D.E., Szeszko, P.R., Robinson, D., Sevy, S., Gunduz-Bruce, H., Wang, Y.P., DeLuca, H., Thompson, P.M. (2005). Mapping cortical thickness and gray matter Concentration in first episode schizophrenia. *Cerebral Cortex* 15, 708-719.

Neale, M.C., Cardon, L.R. (1992). Methodology for Genetic Studies of Twins and Families. Kluwer Academic Publishers, Dordrecht.

Neale, M.C., Miller, M.B. (1997). The use of likelihood-based confidence intervals in genetic models. *Behavioral Genetics* 27, 113-120.

Noga, J.T., Vladar, K., Torrey, E.F. (2001). A volumetric magnetic resonance imaging study of monozygotic twins discordant for bipolar disorder. *Psychiatry Research* 106, 25-34.

Nurnberger jr., J.I., Blehar, M.C., Kaufmann, C.A., York-Cooler, C., Simpson, S.G., Harkavy-Friedman, J., Severe, J.B., Malaspina, D., Reich, T. (1994). Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. *Archives of General Psychiatry* 51, 849-859.

Nyholt, D.R. (2004). A Simple Correction for Multiple Testing for Single-Nucleotide Polymorphisms in Linkage Disequilibrium with Each Other. *American Journal of Human Genetics* 74(4), 765–769.

Olabi, B., Ellison-Wright, I., McIntosh, A.M., Wood, S.J., Bullmore, E., Lawrie, S.M. (2011). Are there progressive brain changes in schizophrenia? A meta-analysis of structural magnetic resonance imaging studies. *Biological* Psychiatry 70, 88-96.

Ormel, J., Oldehinkel, A.J., Brilman, E.I. (2001). The interplay and etiological continuity of neuroticism, difficulties, and life events in the etiology of major and subsyndromal, first and recurrent depressive episodes in later life. *American Journal of Psychiatry* 158, 885-891.

Pakkenberg, B., Gundersen, H.J. (1997).Neocortical neuron number in humans: effect of sex and age. *Journal of Comparative Neurology* 384, 312-320.

Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., Jernigan, T.L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M.J., Grant, M.D., Franz, C.E., Xian, H., Tsuang, M., Fischl, B., Seidman, L., Dale, A., Kremen, W.S. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cerebral Cortex* 19, 2728-2735.

Papagni, S.A., Benetti, S., Arulanantham, S., McCrory, E., McGuire, P., Mechelli, A. (2011). Effects of stressful life events on human brain structure: a longitudinal voxel-based morphometry study. *Stress* 14, 227-232.

Paykel, E.S. (1997). The Interview for Recent Life Events. *Psychological Medicine* 27, 301-310.

Peper, J.S., Brouwer, R.M., Boomsma, D.I., Kahn, R.S., Hulshoff Pol, H.E. (2007). Genetic influences on human brain structure: a review of brain imaging studies in twins. *Human Brain Mapping* 28, 464-473.

Pfohl, B., Blum, N., Zimmerman, M. (1997). Structured Interview for DSM-IV Personality SIDP-IV. American Psychiatric Press, Washington DC.

Phillips, M.L., Drevets, W.C., Rauch, S.L., Lane, R. (2003). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biological Psychiatry* 54, 504-514.

Phillips, M.L., Ladouceur, C.D., Drevets, W.C. (2008). A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Molecular Psychiatry* 13, 833- 857.

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.

Pichot, P. (2004). "150e anniversaire de la Folie Circulaire" [Circular insanity, 150 years on]. Bulletin De l'Académie Nationale De Médecine (in French) 188, 275–284.

Pontious, A., Kowalczyk, T., Englund, C., Hevner, R.F. (2008).Role of intermediate progenitor cells in cerebral cortex development. *Developmental Neuroscience* 30, 24-32.

Posthuma, D., Boomsma, D.I. (2000). A note on the statistical power in extended twin designs. *Behavioral Genetics* 30, 147-158.

R Development Core Team (2008). R: A language and environment for statistical computing. Available from http://www.r-project.org/.

Rabl, U., Meyer, B.M., Diers, K., Bartova, L., Berger, A., Mandorfer, D., Popovic, A., Scharinger, C., Huemer, J., Kalcher, K., Pail, G., Haslacher, H., Perkmann, T., Windischberger, C., Brocke, B., Sitte, H.H., Pollak, D.D., Dreher, J.C., Kasper, S., Praschak-Rieder, N., Moser, E., Esterbauer, H., Pezawas, L. (2014). Additive gene-environment effects on hippocampal structure in healthy humans. *Journal of Neuroscience* 34, 9917-9926.

Rakic, P. (1988). Specification of cerebral cortical areas. *Science* 241, 170-176.

Rakic, P. (1995). A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends in Neurosciences* 18, 383-388.

Rakic, P., Ayoub, A.E., Breunig, J.J., Dominguez, M.H. (2009).Decision by division: making cortical maps. *Trends in Neurosciences* 32, 291-301.

Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. *Nature Reviews Neuroscience* 10, 724-735.

Regeer, E.J., ten Have, M., Rosso, M.L., Hakkaart-van Roijen, L., Vollebergh, W., Nolen, W.A. (2004). Prevalence of bipolar disorder in the general population: a Reappraisal Study of the Netherlands Mental Health Survey and Incidence Study. *Acta Psychiatrica Scandinavica* 110, 374-382.

Rimol, L.M., Hartberg, C.B., Nesvag, R., Fennema-Notestine, C., Hagler jr., D.J., Pung, C.J., Jennings, R.G., Haukvik, U.K., Lange, E., Nakstad, P.H., Melle, I., Andreassen, O.A., Dale, A.M., Agartz, I. (2010). Cortical thickness and subcortical volumes in schizophrenia and bipolar disorder. *Biological Psychiatry* 68, 41-50.

Rimol, L.M., Nesvag, R., Hagler jr., D.J., Bergmann, O., Fennema-Notestine, C., Hartberg, C.B., Haukvik, U.K., Lange, E., Pung, C.J., Server, A., Melle, I., Andreassen, O.A., Agartz, I., Dale, A.M. (2012). Cortical volume, surface area, and thickness in schizophrenia and bipolar disorder. *Biological Psychiatry* 71, 552-560.

Rijsdijk, F.V., Sham, P.C. (2002). Analytic approaches to twin data using structural equation models. *Brief Bioinformatics* 3, 119-33.

Roberts, M., Hanaway, J., Morest, D.K. (1987). Atlas of the Human Brain in Section*., 2nd ed.* Philadelphia: Lea, Febiger.

Ruigrok, A.N.V., Salimi-Khorshidi, G., Lai, M.C., Baron-Cohen, S., Lombardo, M.V., Tait, R.J., Suckling, J. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience and Biobehavioral Reviews* 39, 34–50

Safford, S.M., Alloy, L.B., Abramson, L.Y., Crossfield, A.G. (2007). Negative cognitive style as a Predictor of negative life events in depression-prone individuals: a test of the stress generation hypothesis. *Journal of affective disorders* 99, 147-154.

Samamé, C., Martino, D.J., Strejilevich, S.A. (2012). Social cognition in euthymic bipolar disorder: systematic review and meta-analytic approach. *Acta Psychiatrica Scandinavica* 125, 266-280.

Sapolsky, R.M., Uno, H., Rebert, C.S., Finch, C.E. (1990). Hippocampal damage associated with prolonged glucocorticoid exposure in primates. *Journal of Neuroscience* 10, 2897-2902.

Savitz, J., Drevets, W.C. (2009). Bipolar and major depressive disorder: neuroimaging the developmental-degenerative divide. *Neuroscience and Biobehavioral Reviews* 33, 699-771.

Schmitt, J.E., Eyler, L.T., Giedd, J.N., Kremen, W.S., Kendler, K.S., Neale, M.C. (2007). Review of twin and family studies on neuroanatomic phenotypes and typical neurodevelopment. *Twin Research and Human Genetics* 10, 683-694.

Schnack, H.G., Hulshoff Pol, H.E., Baaré, W.F., Viergever, M.A., Kahn, R.S. (2001). Automatic segmentation of the ventricular system from MR images of the human brain. *Neuroimage* 14, 95-104.

Seldon, H.L. (2005). Does brain white matter growth expand the cortex like a balloon? Hypothesis and consequences. *Laterality* 10, 81-95.

Selvaraj, S., Arnone, D., Job, D., Stanfield, A., Farrow, T.F., Nugent, A.C., Scherk, H., Gruber, O., Chen, X., Sachdev, P.S., Dickstein, D.P., Malhi, G.S., Ha, T.H., Ha, K., Phillips, M.L., McIntosh, A.M. (2012). Grey matter differences in bipolar disorder: a meta-analysis of voxel-based morphometry studies. *Bipolar Disorders* 14, 135-145.

Sham, P.C. (1998). Statistics in Human Genetics. Oxford University Press, New York.

Shepherd, A.M., Laurens, K.R., Matheson, S.L., Carr, V.J., Green, M.J. (2012). Systematic meta-review and quality assessment of the structural brain alterations in schizophrenia. *Neuroscience and Biobehavioral Reviews* 36, 1342-1356.

Sled, J.G., Zijdenbos, A.P., Evans, A.C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging* 17, 87-97.

Spalding, K.L., Bergmann, O., Alkass, K., Bernard, S., Salehpour, M., Huttner, H.B., Boström, E., Westerlund, I., Vial, C., Buchholz, B.A., Possnert, G., Mash, D.C., Druid, H., Frisén, J. (2013). Dynamics of hippocampal neurogenesis in adult humans. *Cell* 153, 1219-1227.

Steel, Z., Marnane, C., Iranpour, C., Chey, T., Jackson, J.W., Patel, V., Silove, D. (2014). The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. I*nternational Journal of Epidemiology* 43(2),476-93.

Strakowski, S.M., DelBello, M.P., Adler, C.M. (2005). The functional neuroanatomy of bipolar disorder: a review of neuroimaging findings. *Molecular Psychiatry* 10, 105-116.

Strakowski, S.M., Adler, C.M., Almeida, J., Altshuler, L.L., Blumberg, H.P., Chang, K.D., DelBello MP, Frangou, S., McIntosh, A., Phillips, M.L., Sussman, J.E., Townsend, J.D. (2012). The functional neuroanatomy of bipolar disorder: a consensus model. *Bipolar Disorders* 14, 313-325.

Sullivan, G.M., Ogden, R.T., Oquendo, M.A., Kumar, J.S., Simpson, N., Huang, Y.Y., Mann, J.J., Parsey R.V. (2009). Positron emission tomography quantification of serotonin-1A receptor binding in medication-free bipolar depression. *Biological Psychiatry* 66, 223-230.

Taki, Y., Kinomura, S., Sato, K., Goto, R., Kawashima, R., Fukuda, H., (2011). A longitudinal study of gray matter volume decline with age and modifying factors. *Neurobiology of Aging* 32, 907-915.

ten Have, M., Vollebergh, W., Bijl, R., Nolen, W.A. (2002). Bipolar disorder in the general population in The Netherlands (prevalence, consequences and care utilisation): results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Journal of Affective Disorders* 68, 203-213.

Thompson, P.M., Cannon, T.D., Narr, K.L., van Erp T., Poutanen, V.P., Huttunen, M., Lönnqvist, J., Standertskjold-Nordenstam, C.G., Kaprio, J., Khaledy, M., Dail, R., Zoumalan, C.I., Toga, A.W. (2001). Genetic influences on brain structure. *Nature Neuroscience* 4, 1253-1258.

Thompson, P.M., Stein, J.L., Medland, S.E., Hibar, D.P., Vasquez, A.A., Renteria, M.E., Toro, R., Jahanshad, N., Schumann, G., Franke, B., Wright, M.J., Martin, N.G., Agartz, I., Alda, M., Alhusaini, S., Almasy, L., Almeida, J., Alpert, K., Andreasen, N.C., Andreassen, O.A., Apostolova, L.G., Appel, K., Armstrong, N.J., Aribisala, B., Bastin, M.E., Bauer, M., Bearden, C.E., Bergmann, O., Binder, E.B., Blangero, J., et al. (2014). The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging and Behavior* 8, 153–182.

Toulopoulou, T., Picchioni, M., Rijsdijk, F., Hua-Hall, M., Ettinger, U., Sham, P., Murray, R. (2007). Substantial genetic overlap between neurocognition and schizophrenia: genetic modeling in twin samples. *Archives of General Psychiatry* 64, 1348-1355.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer B., Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273-289.

van der Schot, A.C., Vonk, R., Brans, R.G., van Haren, N.E., Koolschijn, P.C., Nuboer, V., Schnack, H.G., van Baal, G.C., Boomsma, D.I., Nolen, W.A., Hulshoff Pol, H.E., Kahn, R.S. (2009). Influence of genes and environment on brain volumes in twin pairs concordant and discordant for bipolar disorder. *Archives of General Psychiatry* 66, 142-151.

van der Schot, A.C., Vonk, R., Brouwer, R.M., van Baal, G.C., Brans, R.G., van Haren, N.E., Schnack, H.G., Boomsma, D.I., Nolen, W.A., Hulshoff Pol, H.E., Kahn, R.S. (2010). Genetic and environmental influences on focal brain density in bipolar disorder. *Brain* 133, 3080-3092.

van Erp T.G.M., Hibar D.P., Rasmussen J.M., Glahn D.C., Pearlson G.D., Andreassen O.A., Agartz I., Westlye L.T., Haukvik U.K., Dale A.M., Melle I., Hartberg C.B., Gruber O., Kraemer B., Zilles D., Donohoe G., Kelly S., McDonald C., Morris D.W., Cannon D.M., Corvin A., Machielsen M.W.J., Koenders L., de Haan L., Veltman D.J., Satterthwaite T.D., Wolf D.H., Gur R.C., Gur R.E., Potkin S.G., Mathalon D.H., Mueller B.A. (2015). Subcortical brain volume abnormalities in 2,028 individuals with schizophrenia and 2,540 healthy controls via the ENIGMA Consortium. Molecular Psychiatry.

van Erp, T.G., Saleh, P.A., Huttunen, M., Lönnqvist, J., Kaprio, J., Salonen, O., Valanne, L., Poutanen, V.P., Standertskjöld-Nordenstam, C.G., Cannon, T.D. (2004). Hippocampal volumes in schizophrenic twins. *Archives of General Psychiatry* 61, 346-353.

van Haren, N.E.M., Schnack, H.G., Cahn, W., van den Heuvel, M.P., Lepage, C., Collins, L., Evans, A.C., Hulshoff Pol, H.E., Kahn, R.S. (2011). Changes in cortical thickness during the course of illness in schizophrenia. *Archives of General Psychiatry* 68, 871-880.

Vinkhuyzen, A.A., van der Sluis, S., de Geus, E.J., Boomsma, D.I., Posthuma, D. (2010). Genetic influences on 'environmental' factors. *Genes, Brain and Behavior* 9, 276-287.

Visscher, P.M. (2004). Power of the classical twin design revisited. *Twin Research* 7, 505-512.

Vita, A., De Peri, P.L., Deste, G., Sacchetti, E. (2012). Progressive loss of cortical gray matter in schizophrenia: a meta-analysis and meta-regression of longitudinal MRI studies. *Translational Psychiatry* 2, e190.

Wang, A.Y., Lohmann, K.M., Yang, C.K., Zimmerman, E.I., Pantazopoulos, H., Herring, N., Berretta, S., Heckers, S., Konradi, C. (2011). Bipolar disorder type 1 and schizophrenia are accompanied by decreased density of parvalbumin- and somatostatin-positive interneurons in the parahippocampal region. *Acta Neuropathologica* 122, 615-626.

Westlye, L.T., Walhovd, K.B., Dale, A.M., Bjørnerud, A., Due-Tønnessen, P., Engvig, A., Grydeland, H., Tamnes, C.K., Østby Y., Fjell, A.M. (2010). Life-span changes of the human brain white matter: Diffusion tensor imaging (DTI) and volumetry. *Cerebral Cortex* 20, 2055-2068

Winkler, A.M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P.T., Duggirala, R., Glahn, D.C. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage* 53, 1135-1146.

Womer, F.Y., Wang, L., Alpert, K.I., Smith, M.J., Csernansky, J.G., Barch, D.M., Mamah, D. (2014). Basal ganglia and thalamic morphology in schizophrenia and bipolar disorder. *Psychiatry Research: Neuroimaging* 223, 75-83.

Wray, N., Visscher, P. (2008). Estimating trait heritability. *Nature Education* **1,** 29.

Young, R.C., Biggs, J.T., Ziegler, V.E., Meyer, D.A. (1978). A rating scale for mania: reliability, validity and sensitivity. *British Journal of Psychiatry* 133, 429-435.

List of publications

Bootsman F, Brouwer RM, Kemner S, Schnack HG, van der Schot AC, Vonk R, Hillegers MHJ, Boomsma DI, Hulshoff Pol HE, Nolen WA, Kahn RS, van Haren NEM (2015). Contribution of genes and unique environment to cross-sectional and longitudinal measures of subcortical volumes in bipolar disorder. *European Neuropsychopharmacology* 25, 2197-2209.

Bootsman F, Brouwer RM, Schnack HG, van Baal GCM, van der Schot AC, Vonk R, Hulshoff Pol HE, Nolen WA, Kahn RS, van Haren NEM (2015). Genetic and environmental influences on cortical surface area and cortical thickness in bipolar disorder. *Psychological Medicine* 45, 193-204.

Bootsman F, Brouwer RM, Schnack HG, Kemner SM, Hillegers MHJ , Sarkisyan G, van der Schot AC, Vonk R, Hulshoff Pol HE, Nolen WA, Kahn RS, van Haren NEM. A study of genetic and environmental contributions to structural brain changes over time in twins concordant and discordant for bipolar disorder. *In press.*

Bootsman F, Kemner SM, Hillegers MHJ, Brouwer RM, Vonk R, van der Schot AC, Hulshoff Pol HE, Nolen WA, Kahn RS, van Haren NEM. The association between hippocampal volume and life events in healthy twins. *In press.*

Kemner SM, van Haren NEM, Bootsman F, Eijkemans MJC, Vonk R, van der Schot AC, Nolen WA, Hillegers MHJ (2015). The influence of life events on first and recurrent admissions in bipolar disorder. *International Journal of Bipolar Disorders*, doi: 10.1186/s40345-015-0022-4.

Haarman BCM, Riemersma-van der Lek RF, Burger H, Netkova M, Drexhage RC, Bootsman F, Mesman E, Hillegers MHJ, Spijker AT, Hoencamp E, Drexhage HA, Nolen WA (2014). Relationship between clinical features and inflammation related monocyte gene expression in bipolar disorder – towards a better understanding of psychoimmunological interactions. *Bipolar Disorders* 2, 137-150.

Daalman K, van Zandvoort M, Bootsman F, Boks M, Kahn RS, Sommer I (2011). Auditory verbal hallucinations and cognitive functioning in healthy individuals. *Schizophrenia Research* 132, 203-207.

Abramovic L, Bootsman F, van Haren NEM (2015). Hoofdstuk 17: Bipolaire stoornis. In: Aleman A & Hulshoff Pol HE. *Handboek Beeldvorming van het Brein – Imaging voor Psychiaters en Psychologen*. Uitgeverij De Tijdstroom.

Summary in dutch

Nederlandse samenvatting

ACHTERGROND EN GLOBALE OPZET STUDIE

In dit proefschrift worden studies beschreven die zijn uitgevoerd met tweelingen. Dit waren zowel eeneiige als twee-eiige tweelingparen waarbij ofwel een van de twee een bipolaire stoornis (voorheen manisch-depressieve stoornis) had, of beiden, of geen van beiden. Het grootste gedeelte van de tweelingparen heeft twee keer meegedaan aan de studie, die als doel had om veranderingen in de structuur van de hersenen – denk daarbij aan het volume, de oppervlakte of dikte van de totale hersenen als ook specifieke hersengebieden – te bestuderen in patiënten met een bipolaire stoornis en die eventuele veranderingen te vergelijken met mensen die geen bipolaire stoornis hadden.

DE BIPOLAIRE STOORNIS

De bipolaire stoornis is een ernstige psychiatrische stoornis waarbij de persoon in kwestie afgebakende episodes van verhoogde (manische) dan wel verlaagde (depressieve) stemming ervaart - soms tegelijkertijd (gemengde episode) - die al dan niet gepaard kunnen gaan met zogenaamde psychotische verschijnselen, zoals bijvoorbeeld het horen van onverklaarbare stemmen of het hebben van waanideeën. Wanneer men spreekt van verhoogde stemming in deze context dan heeft dat betrekking op *manie*: een periode (minimaal 1 week) van eufore, uitgelaten stemming waarbij de persoon een opgeblazen zelfvertrouwen heeft, barst van de energie, weinig slaap nodig heeft, veel en snel spreekt, een stroom van snel afwisselende gedachten ervaart en zich dikwijls in activiteiten stort die potentieel gevaarlijk zijn of negatieve consequenties kunnen hebben (e.g. hard rijden met de auto, spontaan al het spaargeld uitgeven). Manische episodes worden door de patiënt op het moment zelf vaak niet als vervelend of nadelig ervaren maar zijn wel zorgelijk omdat men het contact met de realiteit dreigt kwijt te raken. Bij de bipolaire stoornis wordt onderscheid gemaakt tussen manie en hypomanie. Een *hypomanie* is een mildere vorm van manie waarbij nog steeds sprake is van een periode (minimaal 4 dagen) van verhoogde stemming met bijbehorend druk gedrag maar waarbij het functioneren minder lijkt te zijn aangetast dan bij een manie. De *depressieve* episodes laten zich kenmerken door een periode (minimaal 2 weken) van verlaagde stemming waarbij de persoon somber is, geen plezier meer kan beleven aan activiteiten die voorheen als plezierig werden ervaren, lusteloosheid en energieverlies, veranderingen in eetlust, slaapproblemen, cognitieve problemen (bijvoorbeeld problemen met het vasthouden van de aandacht), gevoelens van waardeloosheid en/of schuld en dikwijls gedachten aan de dood. *Psychotische* episodes kunnen voorkomen tijdens een manie of depressie (niet bij hypomanie) en laten zich vaak kenmerken door de aanwezigheid

van hallucinaties – meestal auditief, e.g. het horen van onverklaarbare stemmen – en wanen, ideeën of overtuigingen die niet gedeeld worden door de meeste mensen en die vaak een bizar of zeer ongewoon karakter hebben. Een voorbeeld van een waan is dat iemand het idee heeft achtervolgd te worden door mensen op straat zonder dat daar een directe aanleiding voor is. Die achtervolgingswaan kan zich als een olievlek uitbreiden in dat de persoon na verloop van tijd denkt dat iedereen het op hem/haar gemunt heeft, en dat bijvoorbeeld zelfs de regering of FBI daar achter zit.

Er zijn verschillende typen bipolaire stoornis te onderscheiden. **Bipolaire stoornis type I** wordt gekenmerkt door manische episoden en dikwijls ook depressieve episoden al zijn die laatste niet strikt noodzakelijk voor de diagnose. **Bipolaire stoornis type II** wordt gekenmerkt door hypomane en depressieve episoden. De depressie staat hier vaak voorop. **Bipolaire stoornis niet anderszins omschreven** is een restcategorie waarin mensen vallen die wel duidelijke episodische stemmingsschommelingen hebben maar die niet voldoen aan de criteria voor bipolaire stoornis type I of II. Een diagnose van **cyclothymie** wordt gesteld wanneer de persoon in kwestie hypomane episodes ervaart en depressieve episodes die qua ernst niet voldoen aan de criteria voor een ernstige depressieve episode. Patiënten met een bipolaire stoornis kunnen *rapid cycling* ervaren: het hebben van vier of meer (hypo) manische, depressieve en/of gemengde episoden binnen 1 jaar.

De bipolaire stoornis manifesteert zich over het algemeen voor het eerst in de late adolescentie tot het 25e levensjaar. Ongeveer 1-3% van de bevolking lijdt aan de ziekte die door de Wereld Gezondheidsorganisatie (WHO) wordt beschouwd als de zesde meest invaliderende ziekte die er is. Er is geschat dat 25%-50% van de patiënten minstens eenmaal een suïcidepoging doet in hun leven.

Hoewel de oorzaak van de bipolaire stoornis (nog) niet volledig begrepen wordt zijn er sterke aanwijzingen dat genen hierin een rol spelen. De behandeling van de stoornis beslaat vaak psychotherapeutische interventie al dan niet in combinatie met medicatiegebruik, met als primair doel het terugbrengen van de frequentie en ernst van de stemmingsepisodes en het stimuleren van optimaal herstel tussen de episodes.

BEELDVORMING VAN DE HERSENEN

Alle proefpersonen die hebben deelgenomen aan deze studie zijn in een MRI-scanner geweest. MRI (magnetische resonantie imaging) maakt gebruik van de magnetische eigenschappen van water in het lichaam om contrasten tussen weefsels te kunnen maken. Op die manier kunnen verschillende weefsels in beeld gebracht worden. In deze studie zijn uitsluitend MRI-scans gemaakt van de hersenen. Met behulp van computersoftware is het

mogelijk om bijvoorbeeld het volume, de oppervlakte en de dikte maar ook de functie van de hersenen te bepalen (i.e. wat de hersenen *doen*). Deze studie heeft zich echter uitsluitend gericht op de structuur van de hersenen en niet de functie.
Een groot gedeelte van de proefpersonen is op twee momenten in de scanner geweest, met een tussenliggende periode van iets meer dan 7 jaar. Dat maakt deze studie een longitudinale studie en heeft de onderzoekers in staat gesteld om niet alleen naar de hersenen op een bepaald moment te kijken maar ook te kunnen bestuderen of de hersenen veranderen over de tijd.

TWEELINGEN

De reden dat er voor deze studie uitsluitend tweelingen zijn geïncludeerd is omdat door te kijken naar de (genetische) verschillen tussen eeneiige en twee-eiige tweelingen, de relatieve bijdragen van genen enerzijds en omgevingsfactoren anderzijds aan een bepaalde trek (of *fenotype,* in dit geval de hersenen) goed geschat konden worden. Omdat eeneiige tweelingen 100% van hun genen delen en twee-eiige tweelingen ongeveer 50%, zou een grotere overeenkomst in bijvoorbeeld hersenvolume bij die eerste groep in vergelijking met de tweede groep kunnen betekenen dat vooral genen van grote invloed zijn op hersenvolume. Daarnaast kon er ook berekend worden in hoeverre genen of omgevingsfactoren die samenhangen met de bipolaire stoornis tevens de hersenen beïnvloedden.
Die berekeningen werden gedaan met behulp van speciale statistische modellen. Het klassieke tweelingmodel veronderstelt dat een bepaald kenmerk (bijvoorbeeld hersenvolume) onderhevig kan zijn aan genetische en omgevingsfactoren. Een manier om de relatieve bijdrage van die factoren te schatten zonder daarbij specifieke genen of omgevingsfactoren te benoemen is door te kijken naar in hoeverre eeneiige tweelingen verschillen van twee-eiige tweelingen. Dit gebeurt door naar de correlaties in bijvoorbeeld hersenvolume binnen tweelingparen te kijken en die correlaties te vergelijken tussen eeneiige en twee-eiige tweelingen. Immers, als het vooral omgevingsfactoren zouden zijn die hersenvolume beïnvloeden dan zou er geen verschil tussen eeneiige en twee-eiige tweelingen moeten zijn in de grootte van de hersenen. Eventuele verschillen die er zijn moeten dan dus wel met genen samenhangen want dat is het enige wat de twee typen tweelingen onderscheidt. Door van een grote groep eeneiige en twee-eiige tweelingen de gemeten hersenmaten te vergelijken kan er zo berekend worden hoeveel procent van een fenotype ('hersenvolume' bijvoorbeeld) beïnvloed wordt door genen (de zogenaamde heritabiliteit, symbool: h^2) en hoeveel procent beïnvloed wordt door gedeelde (symbool:

c^2) en unieke (symbool: e^2) omgevingsfactoren. Ook kan bepaald worden in hoeverre genen of omgevingsfactoren die met het ene fenotype samenhangen (bijvoorbeeld de diagnose 'bipolaire stoornis') ook van invloed zijn op een ander fenotype (bijvoorbeeld hersenvolume). Dit laatste wordt bepaald aan de hand van de grootte van de zogenaamde genetische (symbool: r_g) en omgevingscorrelaties (symbolen: r_c, r_e), die een waarde tussen -1 en 1 kunnen aannemen. De fenotypische correlatie (symbool: r_{ph}) zegt iets over de totale samenhang tussen fenotypen, los van genetische of omgevingsfactoren. Een positieve correlatie die een waarde heeft tussen de 0 en 1 laat bijvoorbeeld zien dat de bipolaire stoornis samenhangt met een groter hersenvolume. Als de correlatie echter negatief is en een waarde tussen 0 en -1 heeft dan hangt een bipolaire stoornis eerder samen met een kleiner hersenvolume. Voor de genetische en omgevingscorrelaties gelden specifiek dat bij een positieve waarde genen of omgevingsfactoren die met de stoornis samenhangen ook samenhangen met een groter hersenvolume. Een negatieve waarde betekent juist dat genen of omgevingsfactoren die met de stoornis samenhangen tevens samenhangen met een kleiner hersenvolume.

CENTRALE VRAGEN VAN DEZE STUDIE

Een van de belangrijke vragen die ten grondslag lag aan het opzetten van deze studie was of de bipolaire stoornis een ziekte is die gepaard gaat met progressieve hersenveranderingen over de tijd. Met andere woorden, laten patiënten een grotere afname in bijvoorbeeld hersenvolume over de tijd zien dan gezonde controleproefpersonen? En zijn die hersenveranderingen vooral in samenhang met genetische of omgevingsfactoren? En zoja, in hoeverre zijn dat dezelfde genetische of omgevingsfactoren die ook samenhangen met de ziekte zelf? En indien het omgevingsfactoren zijn die vooral een rol spelen, zijn dat dan omgevingsfactoren die uniek zijn voor het individu (bijvoorbeeld trauma, schulden, ziekenhuisopnames) of omgevingsfactoren die door beide leden van een tweelingpaar gedeeld worden (bijvoorbeeld overlijden of scheiding van de ouders)? Deze vragen lagen ten grondslag aan de in dit proefschrift beschreven studies. Tevens kunnen de uit dit onderzoek verkregen resultaten en bevindingen de discussie voeden die reeds lange tijd gevoerd wordt over in hoeverre de bipolaire stoornis overlap vertoont met schizofrenie. De twee ziektes lijken enige genetische en symptomatologische overlap te hebben. In termen van neuroanatomie is bij schizofrenie redelijk consistent aangetoond dat de ziekte gepaard gaat met progressieve reductie in het volume van frontale en temporale hersengebieden maar de vraag is of er bij de bipolaire stoornis een zelfde patroon te ontdekken is.

RESULTATEN

In dit proefschrift zijn vier studies opgenomen die ieder een eigen vraag of set vragen trachtten te beantwoorden.

In **hoofdstuk 2** is een studie beschreven die is uitgevoerd in enkel gezonde tweelingen. Het doel van deze studie was om te analyseren in hoeverre het volume van de hippocampus – een zeer belangrijk en prominent hersengebied dat betrokken is bij stressregulatie en geheugen, gelegen diep in de hersenen – samenhing met stressvolle levensgebeurtenissen. Een van de onderliggende hypothesen was dat stressvolle levensgebeurtenissen de hippocampus negatief beïnvloeden waardoor mogelijk zijn vermogen tot het reguleren van stress beperkt wordt. Bij deze studie is echter vooral bestudeerd hoe stressvolle levensgebeurtenissen samenhingen met het volume van de hippocampus en niet zozeer met zijn functie. Uit de resultaten bleek dat met name zeer ingrijpende levensgebeurtenissen negatief correleerden met hippocampaal volume. Dat wil zeggen dat de groep mensen in deze studie die relatief meer ingrijpende levensgebeurtenissen hadden meegemaakt een kleiner volume van de hippocampus lieten zien. Verder bleek dat het vooral ingrijpende levensgebeurtenissen die gedeeld werden door tweelingen binnen een paar samenhingen met een kleiner hippocampaal volume en niet zozeer gebeurtenissen die uniek waren voor het individu. Wel is er nog meer studie nodig om dit te bevestigen, zeker aangezien de groep mensen die getest is relatief klein was.

In **hoofdstuk 3** wordt een studie beschreven met gezonde tweelingparen en tweelingparen met een bipolaire stoornis. De studie richtte zich op het volume van hersengebieden die betrokken zijn bij het reguleren en verwerken van emoties. De gedachte is dat de symptomen van de bipolaire stoornis – de stemmingsschommelingen – samen zouden kunnen hangen met verstoring in dit emotie-verwerkende netwerk van hersengebieden. De gebieden die in deze studie zijn onderzocht betroffen de thalamus, caudate nucleus, putamen, globus pallidus, hippocampus, amygdala en nucleus accumbens. Deze gebieden bevinden zich diep in de hersenen en worden ook wel subcorticale hersengebieden genoemd. Het bijzondere aan deze studie is dat de proefpersonen op twee momenten gescand zijn. Daarom was het niet alleen mogelijk om het volume van genoemde hersengebieden te beoordelen op één moment maar ook om naar de verandering in volume over de tijd te kijken. De resultaten wezen uit dat de bipolaire stoornis geassocieerd was met een kleiner volume van de thalamus, putamen en nucleus accumbens maar niet met verandering in volume in een van deze of de andere hersengebieden. Dat suggereert dat hoewel bipolaire patiënten mogelijk kleinere volumes hebben van hersengebieden die betrokken zijn bij het verwerken van emoties, deze hersengebieden niet progressief verslechteren in termen van volumeverlies of –toename over de tijd in de volwassenheid. Wat interessant is, is dat de kleinere volumes

op baseline een sterke genetische associatie met de bipolaire stoornis hadden. Dat suggereert dat genen die betrokken zijn bij de bipolaire stoornis tevens een sterke invloed hebben op de betrokken hersengebieden, meer dan omgevingsfactoren. Mogelijk dat in de vroege ontwikkeling de volumes van patienten achter blijven in vergelijking met gezonde controles en dat dit vooral onder genetische invloed staat maar dat in de volwassenheid de volumeveranderingen relatief stabiel blijven en niet verder (progressief) achter uit gaan, vergelijkbaar met de volumeveranderingen in gezonde volwassenen.

Hoofdstuk 4 handelt over de relatie tussen de bipolaire stoornis en de oppervlakte, dikte en het volume van corticale hersengebieden. Corticale hersengebieden zijn gebieden die zich aan de buitenkant van de hersenen bevinden. Van zowel de totale cortex zijn de oppervlakte, dikte en volume bepaald als ook die van afzonderlijke regionale corticale hersengebieden. Mogelijk dat bij mensen met een bipolaire stoornis subtiele afwijkingen in de cortex zichtbaar zijn in vergelijking met gezonde personen. Daarnaast is er ook gekeken naar in hoeverre corticaal volume afhankelijk was van corticale oppervlakte en dikte: wordt volume vooral bepaald door het een of door de ander? Dit laatste kan vooral implicaties hebben voor het type afwijkingen dat men wel gevonden heeft in het volume van de cortex bij bipolaire patiënten. De vraag is of die afwijkingen dan een reflectie zijn van een onderliggende afwijking in de oppervlakte of juist de dikte, of beiden. Uit de resultaten bleek dat mensen met een bipolaire stoornis subtiele verschillen in de oppervlakte, dikte en volume lieten zien ten opzichte van broers/zussen en gezonde controles, in hersengebieden die verspreid zijn over de hele cortex. Zowel grotere als kleinere en dikkere als dunnere gebieden werden geobserveerd in de patiëntengroep. Het bleek dat vooral unieke omgevingsfactoren een grote rol speelden in het beïnvloeden van de oppervlakte, dikte en volume, bij alle proefpersonen, niet alleen bij patiënten. Maar de verschillen tussen patiënten en mensen zonder bipolaire stoornis werden sterk beïnvloed door unieke omgevingsfactoren die samenhingen met de ziekte, meer dan door genen. Een tweede bevinding was dat corticaal volume meer afhankelijk was van corticale oppervlakte dan dikte. Daarom is het juist belangrijk om ook naar de dikte te kijken in patiënten want als men enkel op het volume af gaat dan kunnen afwijkingen in de dikte over het hoofd gezien worden.

In **hoofdstuk 5** wordt een longitudinale studie beschreven waarbij expliciet gekeken is naar de verandering over de tijd in de hersenen. Hier zijn zowel totale hersenvolumes als ook de totale en regionale oppervlaktes, diktes en volumes bestudeerd over de tijd in patiëntenkoppels en controlekoppels. Ook hier was de centrale vraag of de bipolaire stoornis geassocieerd is met progressieve hersenveranderingen, in vergelijking met mensen zonder een bipolaire stoornis. Dat leek niet het geval. Er was een klein verschil in toename van de witte stof bij patiënten in vergelijking met gezonde controles maar na een noodzakelijke statistische correctie viel dit verschil weg. In geen van de andere onderzochte

hersengebieden waren er noemenswaardige verschillen tussen patiënten, broers/zussen van patiënten en gezonde controles. Ook hier was de conclusie dat de bipolaire stoornis niet geassocieerd was met progressieve veranderingen in de hersenen. Verschillen in de grijze en de witte stof die eerder op baseline werden gerapporteerd tussen patiënten en mensen zonder bipolaire stoornis leken stabiel te blijven over de tijd.

DE VERGELIJKING TUSSEN DE BIPOLAIRE STOORNIS EN SCHIZOFRENIE

Uit de resultaten van dit onderzoek komt het beeld naar voren dat de bipolaire stoornis niet geassocieerd lijkt met grote veranderingen in de hersenen over de tijd. Hoewel deze resultaten met enige voorzichtigheid geïnterpreteerd dienen te worden, onder andere vanwege de relatief kleine groep die bestudeerd is en het feit dat bijvoorbeeld de invloed van medicatie niet goed te bepalen was, lijkt het er wel op dat de bipolaire stoornis minder sterke atrofie in de hersenen laat zien dan bij patiënten met schizofrenie. Dit zou suggereren dat de twee ziektes in termen van neuroanatomie beperkte overlap tonen, als die er al is.

IMPLICATIES VAN BEVINDINGEN EN VOORUITBLIK

De bevindingen van deze studie hebben geen directe implicaties voor de klinische praktijk. Het doel van het onderzoek was vooral om meer inzicht te vergaren in de relatie tussen de bipolaire stoornis en hersenstructuur. Daarbij was de centrale vraag of de bipolaire stoornis geassocieerd was met pathologie in het hersenweefsel, op baseline maar ook over de tijd. Een tweede vraag was wat de relatieve bijdragen van genen en omgeving waren aan eventuele hersenafwijkingen. Wanneer de wetenschap er in slaagt om de rol van genen en omgeving bij hersengebieden die betrokken zijn bij affectieve regulatie beter te duiden, kan dit leiden tot meer inzicht in het ontstaan van de symptomen die horen bij de bipolaire stoornis. In dezelfde lijn zou men zich in de toekomst (nog) meer kunnen toeleggen op het bestuderen van de functionele en structurele verbindingen tussen hersengebieden, in plaats van alleen te kijken naar de gebieden zelf. Hersengebieden werken immers in netwerken met elkaar samen en het zou kunnen dat een deel van de symptomatologie van de bipolaire stoornis toe te schrijven is aan inefficiënte of slechte communicatie tussen gebieden. Het zou verder de moeite waard zijn om dit in jonge groepen patiënten of kinderen van patiënten te doen, zodat men de vroege ontwikkeling van met de bipolaire stoornis geassocieerde hersenafwijkingen – indien aanwezig – kan bestuderen, hetgeen

implicaties kan hebben voor vroege detectie en interventie bij de bipolaire stoornis. Meer inzicht in de rol van genen en omgeving daarbij zou kunnen leiden tot de ontwikkeling van nieuwe en betere interventiestrategieën. Wanneer onderzoekers een duidelijk verband kunnen aantonen tussen bijvoorbeeld specifieke levensgebeurtenissen en de hersenen en symptomen van patiënten met een bipolaire stoornis dan wordt het wellicht mogelijk om gedragstherapeutische interventies beter af te stemmen op de individuele klachten en situatie, bijvoorbeeld door middel van het aanleren van effectieve copingstijlen in de omgang met traumatische gebeurtenissen.

CONCLUSIE

De resultaten van dit onderzoek wezen naar subtiele verschillen tussen patiënten met een bipolaire stoornis en mensen zonder een bipolaire stoornis in zowel corticale als subcorticale hersengebieden. Unieke omgevingsfactoren speelden in deze studies een voorname rol in de beïnvloeding van de corticale gebieden, genen speelden een sterkere rol in de beïnvloeding bij de subcorticale gebieden, op baseline. De subtiele verschillen zijn vermoedelijk ontstaan gedurende de ontwikkeling en lijken relatief stabiel in de volwassenheid gehandhaafd te worden. Dat wil zeggen dat er vooralsnog geen duidelijke aanwijzingen zijn voor progressieve veranderingen over de tijd in de structuur van de hersenen van patiënten met een bipolaire stoornis. Dit staat in contrast met schizofrenie waar progressieve afname in hersenvolume, voornamelijk frontaal en temporaal, is waargenomen over de tijd.

Acknowledgements

Dankwoord

Het is gedaan. De tweede meting van deze longitudinale tweelingstudie is afgerond. Het is een langdurig traject geweest voor alle betrokkenen. Hoewel ik veel plezier heb beleefd aan het uitvoeren van deze studie en dit promotietraject, geldt voor mijzelf dat ik tijdens het proces – vooral op het eind – het ook zwaar heb gevonden en een groot gedeelte van het stemmingsspectrum zelf heb ervaren, hetzij in mindere mate dan veel van de patiënten die ik gesproken heb. Ik heb dan ook versteld gestaan van de doorzettingskracht en het optimisme van de deelnemers aan deze studie, mensen die soms vreselijke dingen hebben meegemaakt of hebben gevoeld, die toch weer de moed vonden om te proberen hun leven op de rails te krijgen. Dat gold voor zowel de patiënten, hun broers of zussen als de gezonde controlepersonen.

Het zijn die mensen die ik dan ook in de eerste plaats wil bedanken voor hun deelname aan deze studie, die er zonder hen niet was geweest. Dank jullie wel voor jullie openhartigheid en betrokkenheid bij het onderzoek!

Dan zijn er nog een hele hoop mensen die de afgelopen jaren expliciet dan wel impliciet hebben bijgedragen aan de totstandkoming van dit proefschrift die ik hier zou willen bedanken.

Mijn co-promotor, **dr. Neeltje van Haren**. Beste Neeltje, ik heb het getroffen met jou als co-promotor. Ik heb jouw begeleiding als een steun in de rug ervaren, waarbij je op een gebalanceerde, zuivere, heldere maar ook betrokken manier richting hebt gegeven aan mijn proces. Ook als ik het moeilijk had of vast zat was je er om wat meer lucht te geven aan mijn soms wat ernstige inborst. Het maakte niet uit wanneer, je had altijd tijd en ruimte voor me. Dat heb ik erg gewaardeerd, dank je wel!

Mijn andere co-promotor, **dr. Rachel Brouwer**. Beste Rachel, hoewel jij pas last-minute formeel bent toegevoegd aan het begeleidingsteam was jij natuurlijk tijdens het hele traject al betrokken bij deze studie. Ik heb veel aan je gehad. Jouw creativiteit in het vinden van oplossingen en can-do attitude hebben deze studie en mij veel gebracht. Dank je wel!

Mijn eerste promotor, **prof. dr. René Kahn**. Beste René, bedankt dat ik de mogelijkheid heb gekregen deze mooie studie te mogen doen! Het is een leerzaam proces geweest en ik denk dat ik nu betere artikelen schrijf dan aan het begin van dit traject.

Mijn tweede promotor, **prof. dr. Willem Nolen**. Beste Willem, bedankt voor je vertrouwen in mij. Ik heb de samenwerking als erg prettig ervaren! Ik kon altijd rekenen op je snelle en nuttige feedback op de artikelen.

De leescommissie: **prof. dr. Geert-Jan Biessels**, **prof. dr. Iris Sommer**, **prof. dr. Brenda Penninx**, **prof. dr. Ralph Kupka** en **prof. dr. Jeroen Hendrikse**. Hartelijk dank voor het kritisch lezen van het manuscript en het zitting nemen in de corona!

Dit proefschrift had er ook niet gelegen zonder de voor mij voornaamste twee bipo-dames: **Esther Mesman** en **Sanne Kemner**. Samen hebben we de klus mooi geklaard! Het was een lange weg die wij met zijn drieën hebben afgelegd maar ik ben er trots op dat we alle drie de eindstreep gehaald hebben! Dank jullie wel voor de gezelligheid en de updates over Sylvie en Rafael!

De bipo-onderzoekers van het eerste uur: **Ronald Vonk** en **Astrid van der Schot**. Zonder jullie werk was het niet mogelijk geweest een tweede meting te doen! Bedankt!

Bipo-onderzoeker **Manon Hillegers**. Manon, ik heb veel geleerd van onze klinische overleggen en daar wil ik je hierbij voor bedanken. Ook voor de gezelligheid tijdens etentjes, congressen en borrels.

Ook wil ik **Hugo Schnack** bedanken voor de ondersteuning tijdens het hele project. Hugo, jouw kritische en precieze blik heeft menig manuscript en analyse beter gemaakt, bedankt!

Hilleke Hulshoff Pol wil ik bedanken voor ons gesprek voorafgaand aan mijn OIO-periode, dat er onder meer toe heeft geleid dat ik op dit project kon beginnen. Ook bedankt voor de snelle en nuttige feedback op de artikelen!

Mijn paranimfen, **Daan** en **Josje**. Met zijn vieren, met **Willemijn** erbij, hebben we het vaak gehad over onze respectievelijke promotieperikelen, vooral de keerzijdes er van. Het was voor mij altijd heel fijn om jullie in de buurt te hebben om bij te kunnen spuien of juist successen mee te kunnen delen. Ik geniet van onze avonden met kaas, wijn en goede gesprekken (ook met **Mikaatje** erbij), die gelukkig niet alleen over promoveren gaan maar ook over de niet-triviale trivialiteiten van het leven. Het is voor mij niet meer dan logisch dat jullie mijn paranimfen zijn maar ik vind het bijzonder dat jullie dat ook wilden. Dank jullie wel!

De stagiair(e)s, die allen keihard gewerkt hebben en met mooie scripties de deur uit zijn gegaan: **Erik**, **Hans**, **Afanaisa**, **Nanja**, **Misha** en **Raissa**, bedankt voor jullie harde werk!

Dan zijn er natuurlijk nog de collega's en oud-collega's met wie ik in wandelgangen, scannerruimtes, op kantoortuinen, congressen, retraites, summerschools en (promotie)

feestjes het nodige te stellen heb gehad in mijn tijd bij het UMC: **Kelly** (wanneer kom je weer eens in NL ronddiederen?), **Kirstin**, **Marc B.**, **Bram**, **Mariët**, **Cé**, **Matthijs**, **Remko**, **Bas**, **Janna Marie**, **Thalia**, **Sven**, **Tjerk**, **Mirjam B.** (die zoektocht naar de sleutels bij de scannerruimte tot in de onchristelijke uurtjes…), **Marloes**, **Ruby**, **Antoin**, **Sanne van R.** (ik kijk met plezier terug op de Apeldoorn-retraites!), **Mitzy**, **Dienke**, **Annabel**, **Juliette**, **Sarai**, **Janna**, **Lucija** (het lollige tripje naar de Chinese muur staat me nog helder voor de geest!), **Sanne V.**, **Anna K.**, **Anna H.**, **Anouk M.**, **Mireille**, **Thomas S.** (met je dieren-dobbel, of wat was het ook al weer? Biggen?), **Maria**, **Tamar**, **Patrick**, **Jessica**, **Gayane**, **Anouk v/d W.**, **Kim**, **Monica**, **Inge**, **Jiska**, **Martijn K.**, **Vincent**, **Branko**, **Remmelt**, **Max**, **Arija**, **Lara**, **Iris**, **Annet**, **Elemi**, **Wanda**, **Marinka**, **de IOP opblaaspelikaan** (weet iemand waar die gebleven is, trouwens? **Marc**?), **Marieke B.**, **Lotte**, **Sanne K.**, **Martijn v/d H.**, **de Mandl, Rachel (Brans), Caroline** (bedankt voor je hulp bij het eerste paper en je immer kalme predispositie)**, Fiona, Heleen, Wiepke, Sarah, Marieke L., Sander, Yumas, het immuno-lab in Rotterdam: Hemmo, Harm, Angelique, Annemarie en collega's**, **Marco, Steven, de dames van het secretariaat: Carla, Annick en Yvonne** (bedankt voor alle goede zorgen rondom het aanmelden en opvangen van de deelnemers!), **Tjen** (altijd bereid om te helpen met logistieke zaken, dank je wel!), **Emmy**, **Elly**, **Mirjam van Male** (bedankt voor je hulp bij de laatste fase!), **Paula** (bedankt voor je betrokkenheid bij het project en je bereidheid te helpen als er weer eens iets bij Radiologie geregeld moest worden!) en **de laboranten van Radiologie: Guus, Erik, Esther, Els, Niels, Gerrit, Evita, DoDo, Saskia, Marieke** en **anderen**, dank voor jullie hulp!

De collega's van Rivierduinen/EDIT te Leiden: **Rianne**, **Esther**, **Lieke**, **Greet**, **Sarah**, **Marga**, **Hans**, **Gerdie**, **Lisa**, **Kristel, Astrid** en **Tinie**. Na vertrek bij het UMC kwam ik in Leiden in een warm bad terecht. Bedankt voor de leuke tijd!

De collega's van het WODC: **Liza**, **Katy**, **Daphne A.**, **Rik**, **Nikolaj**, **Suzan**, **Rianne, Daphne B.**, **Jonas**, **Iris**, **Lise**, **Charlotte**, **Maite**, **Ronald**, **Bart**, **Jan-Jaap**, **Teun**, **Chrissie**, **Ap, Frans** en **alle anderen.** Als neuropsycholoog in jullie midden voelde ik me aanvankelijk een vreemde eend in de bijt maar dat veranderde al snel. Bedankt voor de leuke en gezellige tijd die ik tot dusver heb gehad in Den Haag!

Mijn dierbare vrienden van het Jordan: **Nieke**, **Ren**, **Jasper**, **Tos**, **Mar**, **Daan**, **Mathijs** en **Lot.** Ruim twintig jaar geleden leerden we elkaar kennen en hebben sinds die tijd heel erg veel meegemaakt samen. Ieder neemt zijn eigen route maar de bodem die wij hebben gelegd in Zeist maakt dat het iedere keer weer voelt als thuis komen als ik jullie zie. Op naar de volgende twintig! Bedankt dat jullie er zijn!

Jeroen en **Anton**, mafkezen van het eerste uur. Wij zijn nu zo'n dertig jaar onderweg met zijn drieën. We doen verschillende dingen, soms in verschillende landen, maar verliezen elkaar niet uit het oog. Ik ben bang dat ik toch niet meer van jullie af kom, dus proost, op naar de volgende dertig!

Dan een hele hoop mensen die het leven gezelliger maken, in de kroeg, op festivals, op (kinder)feestjes, op huwelijken, op vakanties: **Marc(us), Netty**, **Joeke, Gitte, Josje, Mika, H-G, Bas, Meg, Saskia, Coen, Dianne, Sanne Bo, Wessel, Sanne Bij, Michael vd A., Merkske, Olga, Abel, Jette, Jordi, Ben, Madelijne, Jornt, Bert, Rita, Marieke, Lise, Tom, José, Dagmar, Sander, Naomi, Jos, Arjan, Suus, Jannes, Marlies, Hilde, Ruud, Hanneke, Maartje, Jacqueline, Joost, Peter, Liset, Femke, Sjaak, Thijs, Elin, Maaike, Dave, Elleke, Lennert, Rina, Michiel, Stuti, Roccio, Roos, Oran, Michael D., Jan-Kees** en **nog een boel anderen.**

Mijn tweede familie: **Kees**, **Margreet**, **Bernard** en **Floor**. Ik heb het getroffen met zulke schoontjes! Jullie zijn ook erg betrokken geweest bij mijn hele traject en hebben mij daarbij heel goed gesteund. Dank jullie wel! Ook mensen van de bredere schoonfamilie wil ik bedanken voor de belangstelling en gezelligheid tijdens familiefeestjes, etentjes, huwelijken of verhuizingen, in het bijzonder: **Harry**, **Mieke**, **Janneke**, **Esther**, **Pieter**, **Oma**, **Joost, Maarten, Emiel, Elisabeth, Eloise, Siebe, Heleen, Berend, Annet, Nina, Vera, Leendert**, **Anneke**, **Myrthe**, **Isa**, **Jolijn**, **Pali**, **Beatrice**, **Luigi, Mario** en ook de **rest van de familie**.

Mijn **moeder**. Lieve mam, jij bent er altijd voor mij en houdt goed mijn emotionele behoeften in de gaten. Je hebt me heel erg gesteund in periodes dat ik het zwaar had. Ik ben blij dat je nu zo op je plek bent in je nieuwe huis. En weer lekker dichtbij!

Mijn **vader** en **Pien**. Lieve Pap, jouw niet-aflatende interesse in de wetenschap is vaak de benzine voor onze gesprekken, ook al komen die regelmatig op een totaal ander onderwerp uit. Ik geniet van onze uitstapjes de laatste jaren, laten we dat blijven doen! Pien, bedankt voor je betrokkenheid en interesse!

Dan de siblings. **Rem**, ik vind het knap hoe jij je leven vorm geeft en je eigen pad kiest. Voor mij ben je soms nog het meisje met het roze brilletje en het fonteintje op je hoofd – dat is een beeld wat een grote broer maar moeilijk uit zijn hoofd kan zetten – maar daarnaast ben je een volwassen, zelfstandige jonge vrouw die weet wat ze wil! En ook nog eens een leuk en lief mens! **Barbara** en **Mick**, ik geniet er iedere keer van als we allemaal samen in een ruimte zijn en de boel op stelten zetten. Het is fijn dat er mensen zijn die net zulke cynische

en zwarte humor hebben als ik, dat voelt bevrijdend. Daarnaast zijn jullie lieve mensen bij wie ik mij altijd thuis voel! Bedankt voor jullie betrokkenheid! Dat was ook altijd zo met en bij **Frodo**, die wij allemaal heel erg missen.

En dan de rest van de familie: **Oma, Robert (Jan)**, **Jantine**, **Hielke**, **Anne**, **Danique**, **Jolynn**, **John**, **Justin**, **Nick** en **Niek**. Bedankt voor de gezelligheid die jullie brengen. Ik heb een leuke familie!

En dan mijn lieve meissie **Willemijn**, die dit promotietraject net zo intensief beleefd heeft als ik. Jouw kalmte, relativering en optimisme hebben mij ontzettend geholpen in de laatste fase, dank je wel daarvoor! Ik ben zo blij dat jij samen met mij het leven door wil; je maakt het leuker en interessanter, en de kleuren feller. Een bijzonder mens ben je, iets wat ik niet helemaal in woorden kan of wil vangen. Ik hou van je!

Curriculum vitae

Florian Bootsman was born in Zeist, on July 1st 1982. In 2002 he enrolled in the bachelor's program Psychology, majoring in Neuropsychology, obtaining the degree in 2007. Subsequently, he enrolled in the master's program Neuropsychology during which he combined a clinical and research internship at the University Medical Center Utrecht, department of Psychiatry, under supervision of prof. dr. Iris Sommer, dr. Ron Hijman and dr. Martine van Zandvoort. Here he was responsible for the neuropsychological assessment of psychiatric patients and reporting on the results, for the benefit of their treatment. In addition, he did his thesis research on the neuropsychological underpinnings of auditory verbal hallucinations. In 2008, he also started working as a research assistant and MRI flowmanager, on the one hand responsible for the inclusion of patients with schizophrenia and their family members in the GROUP-study and on the other hand responsible for carrying out MRI scans in patients who were participating in different studies at the department of Psychiatry. He obtained his master's degree in 2008 and in 2009 he enrolled in the PhD program Clinical and Experimental Neuroscience at the Brain Center Rudolf Magnus, where he studied longitudinal brain changes in twins concordant and discordant for bipolar disorder, under supervision of prof. dr. René Kahn, prof. dr. Willem Nolen, dr. Neeltje van Haren and dr. Rachel Brouwer. Between 2014 and 2015 he worked as a diagnostician at GGZ Rivierduinen in Leiden, in the Early Detection and Intervention Team (EDIT), responsible for guiding adolescent and young adults to preventative treatment in case of risk for psychosis. Since May of 2015 he works for the Research and Documentation Center (WODC) at the Ministry of Security and Justice in The Hague, researching treatment options for juvenile offenders.