Interhemispheric white matter integrity in young people with bipolar disorder and at high genetic risk

G. Roberts[1,2], W.Wen[1,3], A. Frankland[1,2], T. Perich[1,2], E. Holmes-Preston[1,2], F. Levy[1,3], R.K. Lenroot[1,4], D. Hadzi-Pavlovic[1,2], J.I. Nurnberger[5], M. Breakspear[6,7], P.B. Mitchell[1,2,3*]

[1] *School of Psychiatry, University of New South Wales, Sydney, Australia*
[2] *Black Dog Institute, Prince of Wales Hospital, Sydney, Australia*
[3] *Prince of Wales Hospital, Sydney, Australia*
[4] *Neuroscience Research Australia, Sydney, Australia*
[5] *Department of Psychiatry, Indiana University School of Medicine, Indianapolis, Indiana, USA*
[6] *Division of Mental Health Research, Queensland Institute of Medical Research, Brisbane, Australia*
[7] *Royal Brisbane and Women's Hospital, Brisbane, Australia*

[*]Address for correspondence: Scientia Professor Philip Mitchell, School of Psychiatry, University of New South Wales, Prince of Wales Hospital, Randwick, NSW 2031, Australia. (E-mail: phil.mitchell@unsw.edu.au)

Key words: Bipolar disorder; genetic risk; fractional anisotropy; DTI; TBSS.

Word count: abstract 200 words; article body 4,846 words (excluding abstract, acknowledgments, declaration of interest, legends and references).

Manuscript includes 2 tables, 1 figure and 5 supplements

_____

Abstract

Background. White matter (WM) impairments have been reported in patients with bipolar disorder (BD) and those at high familial risk of developing BD. However, the distribution of these impairments has not been well-characterized. Few studies have examined WM integrity in young people early in the course of illness and in individuals at familial risk who have not yet passed the peak age of onset.

Methods. WM integrity was examined in 63 BD subjects, 150 high risk individuals (HR), and 111 participants with no family history of mental illness (CON). All subjects were aged 12 to 30.

Results. This young BD group had significantly lower fractional anisotropy within the genu of the corpus callosum (CC) compared to the CON and HR groups. Moreover, the abnormality in the genu of the CC was also present in HR participants with recurrent major depressive disorder (MDD) (n=16) compared to controls.

Conclusions. Our findings provide important validation of interhemispheric abnormalities in BD patients. The novel finding in HR subjects with recurrent MDD - a group at particular risk of future hypo/manic episodes - suggests that this may potentially represent a trait marker for BD, though this will need to be confirmed in longitudinal follow-up studies.

Introduction

Despite growing evidence for white matter (WM) impairments in bipolar disorder (BD) using diffusion tensor imaging (DTI), findings have been inconsistent (Heng et al., 2010, Nortje et al., 2013, Vederine et al., 2011). Moreover, impaired WM may reflect either genetic liability for the disorder, illness progression or the effects of treatment (Hajek et al., 2005). First-degree relatives of patients with BD are at increased risk of developing this condition, and recent research has demonstrated strong heritability of WM morphology (Chiang et al., 2012). Consequently, studying WM brain abnormalities in individuals at familial risk of developing BD may facilitate understanding of biological risk factors predicting the later development of BD. A number of studies have revealed both widespread and localized reductions in WM integrity in unaffected relatives of patients with BD, but the spatial distribution and extent of these impairments remains uncertain (Versace et al., 2010, Sprooten et al., 2011, Sprooten et al., 2013, Matsuo et al., 2011, Mahon et al., 2013, Linke et al., 2013, Frazier et al., 2007, Emsell et al., 2013, Roybal et al., 2015). For example, whilst lower fractional anisotropy (FA) in the corpus callosum (CC) of high-risk and BD individuals compared to controls has been reported by Sprooten et al. (Sprooten et al., 2013), another study identified this deficit in BD individuals only (Linke et al., 2013). Similarly, reduced WM integrity in the superior longitudinal fasciculus of unaffected relatives has been reported, but whether this effect is intermediate or comparable to the decrease observed in BD patients remains unclear (Frazier et al., 2007, Sprooten et al., 2013), and in the absence of group differences another study reported genetic liability to be associated with FA in this tract (Emsell et al., 2013). Among other regions, pathways involved in emotion regulation involving fronto-temporal lobes such as temporal associative tracts, left orbital frontal tracts, superior longitudinal fasciculus, cingulum (Emsell et al., 2013, Frazier et al., 2007, Mahon et al., 2013, Sprooten et al., 2013, Sprooten et al., 2011, Versace et al., 2010), pathways connecting fibres to frontal regions such as the internal capsule (Sprooten et al., 2011), and intra-hemispheric tracts of the CC responsible for inter-hemispheric communication (Roybal et al., 2015, Sprooten et al., 2013) have been implicated in subjects at increased familial risk of developing BD.

3

DTI allows for sensitive exploration of WM microstructural features. FA is perhaps the most widely accepted DTI measure and reflects directionally-constrained diffusion of water along nerve fibers (Thomason and Thompson, 2011). Tract-based spatial statistics (TBSS) is a popular means of quantifying FA that restricts the evaluation of diffusion parameters to a WM skeleton (Smith et al., 2006) and addresses prior methodological concerns regarding smoothing and normalization (Jones et al., 2002, Chao et al., 2009). Using TBSS, three previous studies have investigated FA in subjects at increased genetic risk of developing BD compared to controls (Sprooten et al., 2011, Versace et al., 2010, Roybal et al., 2015).

A limitation of several studies investigating WM abnormalities in unaffected relatives has been the lack of a comparator BD group. The inclusion of such a group enables clarification of whether FA findings in unaffected relatives are similar to, or distinct from those with the established condition. Hitherto, few studies have investigated (within the same study) FA in patients with established BD, individuals at high genetic risk of developing BD, and controls. To our knowledge, only three previous studies have employed TBSS in all three groups within the same study, each of which had limitations (Linke et al., 2013, Mahon et al., 2013, Sprooten et al., 2013). Two of these three studies had a modest sample size (Linke et al., 2013, Mahon et al., 2013), and the study with the largest sample (n for all groups=170) (Sprooten et al., 2013) recruited unaffected relatives who had already passed the typical age of illness onset (<30 years). Using voxel-based analysis (Chaddock et al., 2009, Frazier et al., 2007), tractography (Emsell et al., 2013), and TBSS (Skudlarski et al., 2013), WM deficits have been detected in both psychotic and non-psychotic BD individuals, as well as their unaffected relatives. In three of these four studies, the mean age of unaffected relatives was over 30 years, again suggesting that many would have already passed the typical onset age of BD illness (Chaddock et al., 2009, Emsell et al., 2013, Skudlarski et al., 2013). Furthermore, DTI findings in studies that included all three groups have been inconsistent in terms of the location and extent of WM abnormalities

Here, using TBSS, we compared FA in a large young sample of: i) 'high risk' (HR) first-degree relatives of patients with BD who have not been diagnosed with BD, ii) patients with BD, and iii) control subjects (CON) without a family history of BD or other mental illness. As those HR subjects with recurrent major depressive disorder (MDD) are at greatest risk for later development of a manic or hypomanic episode

(Hillegers et al., 2005, Duffy, 2010), we also compared this sub-group to both the BD and CON groups. To our knowledge, this DTI study comprises both the largest HR sample, and the largest combined sample across all three groups to date.

Method

Participants

The sample comprised three groups: i) 111 controls (CON) (defined as subjects with no parent or sibling with bipolar I (BD-I) or II (BD-II) disorder, recurrent major depression, schizoaffective disorder, schizophrenia, recurrent substance abuse or any past psychiatric hospitalisation; and no parent with a first degree relative who had a past mood disorder hospitalisation or history of psychosis); ii) 150 subjects genetically 'high risk' for BD who had not yet developed this condition (HR) (defined as children or siblings of a proband with a confirmed DSM-IV diagnosis of BD-I or BD-II); and iii) 63 BD subjects (defined as meeting DSM-IV criteria for either BD-I or BD-II). All participants were aged between 12 and 30 years. Clinical details of the three groups are described elsewhere (Perich et al., 2015). The study was approved by the Human Research Ethics Committee of the University of New South Wales, and complied with the guidelines of the Australian National Health and Medical Research Council.

The lifetime or current presence of psychiatric disorders (including depression and anxiety) besides BD was not an exclusion factor for either HR or CON subjects. This ecological approach has been used by similar studies of individuals at high genetic risk for BD (Nurnberger et al., 2011). All participants are involved in an ongoing longitudinal study of individuals at high risk for BD aged 12-30 years. HR and BD participants were recruited from families who had previously participated in a BD pedigree molecular genetics study or a specialized BD research clinic, or were otherwise recruited from clinicians, mental health consumer organisations and other forms of publicity. Control subjects were recruited via print and electronic media, and noticeboards in universities and local communities.

The subjects aged between 12-21 years are involved in a collaborative high-risk study with a U.S. consortium headed by Dr John Nurnberger which is based at Indiana University, Johns Hopkins University, Washington University in St. Louis, and Michigan University (Nurnberger et al., 2011). As this US-Australian

collaboration involves common clinical assessments for participants aged 12-21years, we report separately on instruments used for the younger (12-21years) and older (22-30 years) age groups in this sample. Both groups shared consensus Best-Estimate DSM-IV current and lifetime diagnoses derived from semi-structured diagnostic interviews. *Brain imaging studies were only undertaken in the Australian sample*. A number of functional imaging studies on this BD HR sample have been previously reported (Breakspear et al., 2015, Roberts et al., 2013).

Proband consensus DSM-IV diagnosis was determined by two independent raters following Best Estimate methodology (Leckman et al., 1982), using information from the Diagnostic Interview for Genetic Studies (DIGS) Version 4, the Family Interview for Genetic Studies (FIGS), and medical records (where available). Confidence rating ranges using the Best Estimate Methodology vary from 1-4, where 1 represents criteria not met for a diagnosis and 4 represents a definite diagnosis. All diagnoses reported have a confidence rating of 3 or 4 (probable or definite). Structured diagnostic interviews were also performed on all HR, CON and BD participants. For those aged between 12 and 21 (CON, n=51; HR group, n=92; BD group = 11), an adapted version of the Schedule for Affective Disorders and Schizophrenia for School-Age Children – Present and Lifetime Version (K-SADS-BP) was developed specifically for use in the US-Australia collaborative study of young people at genetic risk for bipolar disorder (Nurnberger et al., 2011). The K-SADS-BP combines items from the K-SADS Present and Lifetime Version (Kaufman et al., 1997), and uses extended sections on depression, mania and ADHD derived from the Washington University in St Louis K-SADS (WASH-U K-SADS) to elicit detailed information on pre-pubertal mania, rapid-cycling, attentional and sub-threshold bipolar symptoms (Geller et al., 2001). The KSADS-BP was administered to both the child and one parent. For participants aged under 21, clinicians completed the Children's Global Assessment Scale (CGAS).

For participants aged between 22 and 30 (CON, n=60; HR group, n=58; BD group = 52), the DIGS (Version 4) was used to measure the current and lifetime presence of axis I DSM-IV disorders. Consensus DSM-IV diagnoses of the HR, BD, and CON subjects were determined by two independent raters with Best Estimate methodology (Leckman et al., 1982), using the DIGS, the FIGS and medical records (where

available). For participants aged between 22 and 30, clinicians completed the Global Assessment Scale (GAS).

To assess current mood state, for those aged between 12 and 21, the Children's Depression Inventory (CDI) was used, and for participants aged 22-30 years, the Montgomery-Asberg Depression Rating Scale (MADRS) was administered. To assess current manic symptoms the Young Mania Rating Scale (YMRS) (Young et al., 1978) was used for participants aged 22-30 years. Intellectual ability was assessed with the Wechsler Abbreviated Scale of Intelligence (WASI). Participants aged 12-30 completed clinical and neuropsychological assessments on the same day as the scan (McCormack et al., 2016). The majority of parental reports using the K-SADS were also completed on the same day as the scan.

MRI acquisition

Images were acquired with a 3-T Philips Achieva scanner at Neuroscience Research Australia (NeuRA) in Sydney with an 8-channel head coil. Two acquisitions of 32 directional DTI data (b =1000 s/mm^2 with one diffusion weighed b0) were acquired using a single-shot echo planar imaging (EPI) sequence. The imaging parameters were as follows: TR=7638 ms, TE=68 ms, 55 slices, slice thickness=2.5 mm, gap=0 mm, FOV=240×138×240, acquisition matrix size=96×96, flip angle= 90° with an image time of 4 mins 42 secs per acquisition.

Diffusion tensor imaging processing

Details of pre-processing and TBSS processing steps are provided in the supplementary methods. For skeletonisation and ROI extraction from FA images we followed the ENIGMA-TBSS protocols (http://enigma.ini.usc.edu/wp-content/uploads/2014/01/ENIGMA_TBSS_protocol_USC.pdf). We extracted the mean FA across the full skeleton and subsequently calculated average FA values bilaterally for twenty-one intra-hemispheric and four inter-hemispheric tracts, yielding a total of 46 ROIs. The regions analysed are listed in Supplementary Table S1.

Statistical analyses

All statistical analyses were carried out using SPSS v.22 (IBM, Chicago, Illinois). Socio-demographic differences between groups were assessed using univariate analysis of variance (ANOVA) followed by pairwise comparisons for multiple comparisons using the least significant difference (LSD) method. To compare tracts of interest across diagnostic groups, generalised estimating equations (GEE) models were used as they accommodate within-family correlation arising from the inclusion of siblings from within the same family in either the CON, HR or BD groups (families were not split between diagnostic groups). Each base GEE model was run with FA of the region of interest (ROI) as the dependent variable, hemisphere (right and left) as the within-subject factor, diagnostic group as the between-subject factor (CON, HR, and BD), *hemisphere x group* as an interaction, and age and gender as covariates. To address whether the relationship between age and FA of ROI's differed between groups, a second set of GEE models were fitted where the *age x group* interaction was also included as an interaction term in the base model. In order to select the most parsimonious model, for ROIs with no significant *age x group* interaction ($p > 0.05$) the base model was selected.

Correction for multiple testing of the main effects of *group, hemisphere* and the *group x hemisphere* interaction was carried out using false-discovery rate (FDR) q values using the Benjamini & Hochberg (1995) method (Benjamini and Hochberg, 1995). For each dependent variable we used the most parsimonious model and ranked *group*, *hemisphere* and *group x hemisphere* comparisons (*group*, *hemisphere* and *group x hemisphere* comparisons for twenty-one intra-hemispheric tracts, and *group* comparisons for four inter-hemispheric tracts, yielding a total of 67 ranked comparisons). We also compared HR subjects with recurrent MDD (HR-rMDD) (n=16) against the CON and BD groups where we selected the base models and ranked *group*, *hemisphere* and *group x hemisphere* comparisons for twenty-one intra-hemispheric tracts, and *group* comparisons for four inter-hemispheric tracts.

Planned comparisons

Planned post-hoc comparisons were carried out on tracts that revealed either a significant *group* effect or *group x hemisphere* interaction after correction for multiple testing: nine *group x hemisphere* interaction comparisons (three between-group differences within the left and right hemispheres respectively, in addition to three between-hemisphere differences within each study group), and three between-group comparisons. Planned comparisons were carried out using the LSD method, and corrected using (FDR) q values.

Effects of medication and illness severity

Separate analyses were performed to determine whether group differences were related to exposure to lithium, antidepressants, or antipsychotics. To assess associations between FA and measures of illness severity and course, relationships between FA and age of onset of first manic, hypomanic, depressive or any mood episode were explored using Pearson's correlations; and relationships between FA and number of episodes of hypomania, mania or depression relationships were explored using Spearman's correlations (as that data was not normally distributed).

Results

Demographic and clinical data

Demographic and clinical data for each group are provided in Tables 1 and 2. The HR participants were slightly younger than both the CON and BD participants, whilst the BD group was older than the CON group. On the MADRS and YMRS, 22-30 year old BD participants scored significantly higher than both their HR and CON counterparts, with no significant differences observed between the latter two groups. Among the 12-21 year old participants, the BD group had significantly higher scores on the CDI than both the HR and CON participants.

Current and lifetime consensus best-estimate DSM-IV diagnoses in all three groups are also detailed in Table 2. Chi-square comparisons revealed that the lifetime occurrence of at least one major depressive episode (MDE) or any lifetime DSM-IV diagnosis (including recurrent major depressive disorder; rMDD) was significantly higher in the HR ($p < 0.01$) compared to the CON group, consistent with prior reports of high-risk

populations (Nurnberger et al., 2011, Birmaher B et al., 2009). There were also significantly higher rates of anxiety (p<0.05) and behavioural disorders (p<0.01) amongst HR individuals.

Of the 150 HR participants, 81.3% had a parent with bipolar disorder and 18.7% had a sibling with bipolar disorder. Of the 63 BD subjects, 34 had BD-I (54%) and 29 BD-II (46%). The 324 participants belonged to 262 families.

FA results: Group findings

The genu of the corpus callosum was the only tract with significant group differences in FA after correction for multiple comparisons (see Figure 1A). Pairwise comparison between the groups showed that this effect was driven by lower FA in the BD compared to the HR and CON groups (BD<HR p=0.03, BD<CON p=0.003, all FDR corrected). There were no FA differences within the BD group between participants with and without current lithium treatment, between those with and without current antidepressant treatment, and between those with and without current antipsychotic treatment (all p > 0.05 uncorrected). Further, the FA finding remained significant when those with current lithium and/or antipsychotic treatment in the BD group, and those with current MDE and/or current antidepressant treatment current antidepressant use across all three groups were removed (n= 107 CON, 139 HR, 19 BD; Wald χ^2=6.41, df=2, *p*=0.040; BD<CON p=0.025, BD<HR p=0.012; FDR corrected pairwise comparisons).Within the BD group, measures of illness severity and course, and mood state did not correlate with FA (all p > 0.05). Furthermore, there were no differences between BD-I and BD-II patients (all p > 0.05).

To exclude any potential confounding effect of subjects with lifetime major depressive or other DSM-IV disorders in the CON group, we repeated analyses after excluding these subjects. The finding of lower FA in the genu of the corpus callosum remained significant when we compared CON participants with no prior or current depressive episode to our total HR and BD samples (n= 99 CON, 150 HR, 63 BD; Wald χ^2=8.63, df=2, *p*=0.013; BD<CON p=0.003, BD<HR p=0.032), and when we compared CON participants without any lifetime DSM-IV diagnosis to our total HR and CON sample (n= 79 CON, 150 HR, 63 BD; Wald χ^2=8.75, df=2, *p*=0.013; BD<CON p=0.003, BD<HR p=0.035; BD<CON but not BD<HR pairwise comparison survived FDR correction). Further, to exclude any effect of prior depressive episodes in the HR sample, we compared CON

and HR subjects with no prior or current depressive episodes to our BD sample; the differences remained significant (n= 99 CON, 106 HR, 63 BD; Wald χ^2=10.09, df=2, *p*=0.006; BD<CON p=0.002, BD<HR p=0.013). Additionally, there were no significant differences between those with at least one lifetime depressive compared to those without, both within the CON (p=0.798) and HR (p=0.455) groups. Likewise, there were no significant differences between those with or without a lifetime DSM-IV diagnosis within both the CON (p=0.609) and HR (p=0.311) groups.

There were no significant differences in FA in the genu of the corpus callosum between the total HR and CON groups. However, as detailed in the introduction above, when we compared only HR subjects with recurrent MDD (HR-rMDD) (n=16) against the CON and BD groups, we found that the HR-rMDD group had lower FA in the genu of the corpus callosum compared to CON subjects (Wald χ^2=12.53, df=2, *p*=0.002; BD<CON p=0.002, HR-rMDD<CON p=0.012; all FDR corrected) (see Figure 1B). There was no significant difference in FA in this tract between the HR-rMDD and BD groups. This finding remained significant when those with current lithium and/or antipsychotic treatment in the BD group, and those with current MDE and/or current antidepressant use across all three groups were removed (n= 107 CON, 12 HR, 19 BD; Wald χ^2=9.21, df=2, *p*=0.010; BD<CON p=0.012, HR-rMDD<CON p=0.038). Group comparisons of all ROIs are detailed in Supplementary Table S2.

Hemisphere results and group x hemisphere interactions

Hemispheric asymmetries are summarised in Supplementary Table S3 as they are not the focus of this paper. None of the ROI FAs showed a *group* x *hemisphere* interaction at either the corrected or uncorrected level. Supplementary Table S4 summarises all *group x hemisphere* interactions.

Group x age interaction

Only one ROI showed an *age x group* interaction (for sagittal stratum; p=0.046), hence in the GEE models this interaction was only included for the sagittal striatum tract.

Discussion

We found reduced WM integrity in the genu of the corpus callosum (CC) of BD patients compared to both controls and those at high familial risk in a large and relatively young sample, confirming a number of recent reports of reduced WM integrity of the CC in BD (Sprooten et al., 2013, Skudlarski et al., 2013, Linke et al., 2013, Li et al., 2014, Lagopoulos et al., 2013, Emsell et al., 2013, Bauer et al., 2015, Barysheva et al., 2013, Sarrazin et al., 2014, Leow et al., 2013, Sarrazin et al., 2015). Furthermore, and of particular interest, we report the novel finding that whilst the overall HR group showed no difference in this tract compared to controls, there was reduced WM integrity in the CC of those high-risk individuals with recurrent MDD – a finding which remained significant after exclusion of those with a current MDE or antidepressants. This finding is of particular relevance, as recurrent MDD in those at familial high risk for BD is a strong clinical predictor of later development of hypo/manic episodes (Hillegers et al., 2005, Duffy, 2010), suggesting the possibility that lower FA in the genu of the CC may represent a trait marker for BD.

Previous DTI studies of subjects with BD have reported abnormalities in the shape, network organisation, and WM microstructure of the CC (Bellani et al., 2009, Bellani and Brambilla, 2011). A meta-analysis of CC findings in BD revealed a volumetric reduction in the CC of patients, when compared to controls (Arnone et al., 2008). Prior DTI studies have also reported lower FA throughout the CC, whereas others have reported lower FA in various localised sub-regions of the CC, in both youth and adults with BD compared to controls (Barnea-Goraly et al., 2009, Barysheva et al., 2013, Bauer et al., 2015, Chaddock et al., 2009, Emsell et al., 2013, Lagopoulos et al., 2013, Li et al., 2014, Linke et al., 2013, Sarrazin et al., 2015, Sarrazin et al., 2014, Saxena et al., 2012, Skudlarski et al., 2013, Sprooten et al., 2013, Yurgelun-Todd et al., 2007). In accordance with our finding in the genu, four of these studies found significantly reduced FA in the genu of the CC of BD subjects early in the course of illness compared to controls (Lagopoulos et al., 2013, Saxena et al., 2012, Barnea-Goraly et al., 2009, Leow et al., 2013). We found no association in our BD subjects between this FA finding in the genu of the CC and indices of BD illness course and severity, mood state or current psychotropic medications, suggesting that this WM abnormality was not secondary to either illness course or severity, or medication exposure. This is not surprising given that our BD group represented a young population early in the course of their illness, who appeared to be high functioning, as suggested by above-

average IQ scores (Wechsler, 1999) and average GAS scores of over 70 (Shaffer et al., 1983, Endicott et al., 1976).

Relatively few studies have examined WM integrity in the CC of subjects at high familial risk of BD, and results have been less consistent than those in patients with established BD. Using voxel-based morphometry, genetic liability for BD has been found to be significantly associated with WM deficits in, among other regions, the genu of the CC (McDonald et al., 2004). In contrast, in other morphological studies of CC size and shape, abnormalities appeared to be disease-related rather than indicative of genetic vulnerability (Bearden et al., 2011, Walterfang et al., 2009). In TBSS studies which have included the three groups (controls, those at high familial risk of developing BD, and BD subjects), there have been reports of reduced WM integrity in the CC of BD patients but not in unaffected relatives (Linke et al., 2013). However, subtle significant FA reductions in regions such as the splenium and body of the CC have also been observed in unaffected relatives, with more widespread FA reductions evident in BD patients (Sprooten et al., 2013). It is noteworthy that the age ranges in the Linke et al. (Linke et al., 2013) and Sprooten et al. (Sprooten et al., 2013) studies were 18-65 and 18-70 respectively, suggesting that many participants included as being at familial risk of developing BD may have already passed the peak age of onset for BD. Also, in accord with the direction of our FA findings, a prior TBSS study that did not include a BD comparison group reported lower FA in one large cluster that included the CC in young high-risk individuals aged 12-25 compared to controls (Sprooten et al., 2011), whereas in contrast, another TBSS study that also did not include a BD comparison group reported increased FA in the body splenium and genu of the CC of children at high-risk of developing BD compared to controls (Roybal et al., 2015).

As our HR group has not yet passed the peak age of onset for developing BD (< 30 years), this group is more likely than samples of older subjects to include a higher proportion of individuals who will later convert to BD. Although it is not known at the time of this study which of these individuals may later develop BD, those who have already developed recurrent MDD have been found in other studies to be at greater risk of conversion (Hillegers et al., 2005, Duffy, 2010). Our novel finding of reduced WM integrity in the genu of the CC in those who have already developed recurrent MDD in the HR group raises the question of whether this

may represent a biological as well as clinically distinct phenotype that may be at particular high risk of ultimately developing BD.

As the CC is the largest WM fiber tract that connects the brain hemispheres, facilitating communication between left- and right-side brain structures (De Lacoste et al., 1985), our findings support the proposition that abnormal interhemispheric connectivity may represent a trait marker for BD (Bellani et al., 2009). Little is known about the functional role of each sub-region of the CC. However, anterior regions such as the genu connect the frontal cortices (Aboitiz et al., 1992) which play an important role in higher order cognitive and emotional processing (Davidson and Irwin, 1999, Galinowski et al., 2015). Given the frontal abnormalities frequently reported in BD (Phillips, 2014), it is conceivable that microstructural disturbances in the genu of the CC may contribute to the cognitive and emotional deficits characteristic of BD. Bearden et al. (Bearden et al., 2011) has reported a significant association between the genu of the CC and response inhibition, as well as verbal processing speed, in BD probands and their unaffected twins. Recent studies have identified associations between the genu of the CC and multiple neurocognitive domains including working memory, processing speed and psychomotor coordination in BD subjects compared to controls (Ajilore et al., 2015, McKenna et al., 2015, Poletti et al., 2015). Taken together, these findings support the postulation that reduced WM integrity in the genu of CC in our BD and HR-rMDD subjects may be linked to impulsivity and cognitive disturbances.

Limitations

Several methodological issues require consideration. Employment of complementary techniques such as tractography, a technique used to visually represent neural tracts, and whole brain voxel based analyses, would contribute to a fuller understanding of possible WM microstructural abnormalities in this population. Furthermore, while FA is widely accepted to be an index of general WM integrity, specific microstructural defects - such as myelin damage, axonal disorganization, or fiber incoherence - which may underlie changes in this metric, remain to be elucidated (Assaf and Pasternak, 2008). Whilst the current study derived average FA per tract, the majority of the cited TBSS BD high-risk papers have used voxel-based TBSS or whole brain voxel-based analyses, making direct comparison between the present results and previous findings difficult.

Further, as the study groups differed on age, we covaried for the potential effects of age, and *age x group* interactions, in our statistical models. In contrast to a report by Versace et al. (2010), we found no evidence of *age x group* interactions in CC FA. However, as typical WM cortical development is non-linear (Imperati et al., 2011, Lebel et al., 2012) and non-linear effects of age could not be incorporated into our model, we cannot exclude the possibility of non-linear effects of age on our findings.

Our results need to be interpreted with caution given the absence of differences between controls and our total cohort of HR subjects (n=150). Although it is likely that a substantial proportion of our high-risk individuals with recurrent MDD will develop BD, the time course for this is unknown. With all the current participants enrolled in an ongoing longitudinal study, follow-up data will improve our understanding of the underlying mechanisms that contribute to trajectories that result in BD. As our MRI scans are repeated at multiple time points, in future years we will be able to investigate whether abnormalities in the genu of the CC are more frequent in those who subsequently develop BD. Further, in view of recent studies (Bessette et al., 2014, Xu et al., 2013) and a meta-analysis (Liao et al., 2013) which have reported abnormalities in the genu and body of the CC of patients with MDD not at familial risk of BD, it is not possible at present to distinguish between our finding representing a specific risk to future BD or a broader generic risk to mood disorders including MDD.

Conclusions

These results from a large sample provide further evidence that the genu of the CC may play a role in the pathophysiology of BD. Furthermore, in a novel finding, we extend previous CC findings by showing that lower FA in the genu of the CC was also present in those at high familial risk who have recurrent MDD. While it need be acknowledged that the number of those with recurrent MDD was relatively small, and that longitudinal data is needed to definitively determine trajectories that result in BD, our findings nonetheless raise the tantalising possibility that WM integrity of the CC may represent a potential trait marker for BD.

Declaration of interest

J.N. is an investigator for Assurex on a study of the clinical efficacy of pharmacogenetic screening and has consulted with Janssen pharmaceuticals regarding high risk studies in bipolar disorder. All other authors report no biomedical financial interests or other potential conflicts of interest.

References

Aboitiz F, Scheibel AB, Fisher RS & Zaidel E 1992. Fiber composition of the human corpus callosum. *Brain Research,* **598,** 143-153.

Ajilore O, Vizueta N, Walshaw P, Zhan L, Leow A & Altshuler LL 2015. Connectome signatures of neurocognitive abnormalities in euthymic bipolar I disorder. *Journal of Psychiatric Research,* **68,** 37-44.

Arnone D, Mcintosh A, Chandra P & Ebmeier K 2008. Meta-analysis of magnetic resonance imaging studies of the corpus callosum in bipolar disorder. *Acta Psychiatrica Scandinavia,* **118,** 357-362.

Assaf Y & Pasternak O 2008. Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review. *Journal of Molecular Neuroscience,* **34,** 51-61.

Barnea-Goraly N, Chang KD, Karchemskiy A, Howe ME & Reiss AL 2009. Limbic and corpus callosum aberrations in adolescents with bipolar disorder: a tract-based spatial statistics analysis. *Biological Psychiatry,* **66,** 238-244.

Barysheva M, Jahanshad N, Foland-Ross L, Altshuler LL & Thompson PM 2013. White matter microstructural abnormalities in bipolar disorder: A whole brain diffusion tensor imaging study. *Neuroimage Clinical,* **2,** 558-568.

Bauer IE, Ouyang A, Mwangi B, Sanches M, Zunta-Soares GB, Keefe RS, Huang H & Soares JC 2015. Reduced white matter integrity and verbal fluency impairment in young adults with bipolar disorder: a diffusion tensor imaging study. *Journal of Psychiatric Research*, 115-122.

Bearden CE, Van Erp TG, Dutton RA, Boyle C, Madsen S, Luders E, Kieseppa T, Tuulio-Henriksson A, Huttunen M, Partonen T, Kaprio J, Lonnqvist J, Thompson PM & Cannon TD 2011. Mapping corpus callosum morphology in twin pairs discordant for bipolar disorder. *Cerebral Cortex,* **21,** 2415-2424.

Bellani M & Brambilla P 2011. Diffusion imaging studies of white matter integrity in bipolar disorder. *Epidemiology and Psychiatric Sciences,* **20,** 137-140.

Bellani M, Yeh P-H, Tansella M, Balestrieri M, Soares J & Brambilla P 2009. DTI studies of corpus callosum in bipolar disorder. *Biochemical Society Transactions* **37,** 1096.

Benjamini Y & Hochberg Y 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of Royal Statistical Society Series B Statistical Methodology*, 289-300.

Bessette KL, Nave AM, Caprihan A & Stevens MC 2014. White matter abnormalities in adolescents with major depressive disorder. *Brain Imaging and Behavior,* **8,** 531-541.

Birmaher B, Axelson D, Monk K, Kalas C, Goldstein B, Hickey Mb & Et Al 2009. Lifetime psychiatric disorders in school-aged offspring of parents with bipolar disorder: the Pittsburgh Bipolar Offspring study. *Archives of General Psychiatry* **66,** 287-296.

Breakspear M, Roberts G, Green MJ, Nguyen VT, Frankland A, Levy F, Lenroot R & Mitchell PB 2015. Network dysfunction of emotional and cognitive processes in those at genetic risk of bipolar disorder. *Brain,* **138,** 3427-3439.

Chaddock CA, Barker GJ, Marshall N, Schulze K, Hall MH, Fern A, Walshe M, Bramon E, Chitnis XA & Murray R 2009. White matter microstructural impairments and genetic liability to familial bipolar I disorder. *British Journal of Psychiatry,* **194,** 527-534.

Chao T-C, Chou M-C, Yang P, Chung H-W & Wu M-T 2009. Effects of interpolation methods in spatial normalization of diffusion tensor imaging data on group comparison of fractional anisotropy. *Magnetic Resonance Imaging,* **27,** 681-690.

Chiang M-C, Barysheva M, Mcmahon KL, De Zubicaray GI, Johnson K, Montgomery GW, Martin NG, Toga AW, Wright MJ & Shapshak P 2012. Gene network effects on brain microstructure and intellectual performance identified in 472 twins. *Journal of Neuroscience,* **32,** 8732-8745.

Davidson RJ & Irwin W 1999. The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Science,* **3,** 11-21.

De Lacoste MC, Kirkpatrick JB & Ross ED 1985. Topography of the human corpus callosum. *Journal of Neuropathology and Experimental Neurology,* **44,** 578-591.

Duffy A 2010. From predisposition to illness: genetically sensitive intermediate pathways to mood disorders. *British Journal of Psychiatry,* **197,** 341-342.

Emsell L, Chaddock C, Forde N, Van Hecke W, Barker G, Leemans A, Sunaert S, Walshe M, Bramon E & Cannon D 2013. White matter microstructural abnormalities in families multiply affected with bipolar I disorder: a diffusion tensor tractography study. *Psychological Medicine*, 1-12.

Endicott J, Spitzer RL, Fleiss JL & Cohen J 1976. The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry,* **33,** 766-771.

Frazier JA, Breeze JL, Papadimitriou G, Kennedy DN, Hodge SM, Moore CM, Howard JD, Rohan MP, Caviness VS & Makris N 2007. White matter abnormalities in children with and at risk for bipolar disorder. *Bipolar Disorders,* **9,** 799-809.

Galinowski A, Miranda R, Lemaitre H, Paillère Martinot M-L, Artiges E, Vulser H, Goodman R, Penttilä J, Struve M & Barbot A 2015. Resilience and corpus callosum microstructure in adolescence. *Psychological Medicine*, 1-10.

Geller B, Zimerman B, Williams M, Bolhofner K, Craney J, Delbello M & Soutullo C 2001. Reliability of the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. *Journal of the American Academy of Child and Adolescent Psychiatry* **40,** 450-455.

Hajek T, Carrey N & Alda M 2005. Neuroanatomical abnormalities as risk factors for bipolar disorder. *Bipolar Disorders,* **7,** 393-403.

Heng S, Song AW & Sim K 2010. White matter abnormalities in bipolar disorder: insights from diffusion tensor imaging studies. *Journal of Neural Transmission,* **117,** 639-654.

Hillegers MH, Reichart CG, Wals M, Verhulst FC, Ormel J & Nolen WA 2005. Five-year prospective outcome of psychopathology in the adolescent offspring of bipolar parents. *Bipolar Disorders,* **7,** 344-350.

Imperati D, Colcombe S, Kelly C, Di Martino A, Zhou J, Castellanos FX & Milham MP 2011. Differential development of human brain white matter tracts. *PLoS One,* **6,** e23437.

Jones DK, Griffin LD, Alexander DC, Catani M, Horsfield MA, Howard R & Williams SC 2002. Spatial normalization and averaging of diffusion tensor MRI data sets. *Neuroimage,* **17,** 592-617.

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Morec IP & Et Al 1997. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry* **36,** 980-988.

Lagopoulos J, Hermens DF, Hatton SN, Tobias-Webb J, Griffiths K, Naismith SL, Scott EM & Hickie IB 2013. Microstructural white matter changes in the corpus callosum of young people with bipolar disorder: a diffusion tensor imaging study. *PLoS One,* **8,** e59108.

Lebel C, Gee M, Camicioli R, Wieler M, Martin W & Beaulieu C 2012. Diffusion tensor imaging of white matter tract evolution over the lifespan. *Neuroimage,* **60,** 340-352.

Leckman J, Sholomskas D, Thompson W, Belanger A & Weissman M 1982. Best estimate of lifetime psychiatric diagnosis: a methodological study. *Archives of General Psychiatry* **39,** 879-883.

Leow A, Ajilore O, Zhan L, Arienzo D, Gadelkarim J, Zhang A, Moody T, Van Horn J, Feusner J & Kumar A 2013. Impaired inter-hemispheric integration in bipolar disorder revealed with brain network analyses. *Biological Psychiatry,* **73,** 183-193.

Li J, Edmiston EK, Chen K, Tang Y, Ouyang X, Jiang Y, Fan G, Ren L, Liu J & Zhou Y 2014. A comparative diffusion tensor imaging study of corpus callosum subregion integrity in bipolar disorder and schizophrenia. *Psychiatry Research,* **221,** 58-62.

Liao Y, Huang X, Wu Q, Yang C, Kuang W, Du M, Lui S, Yue Q, Chan RC & Kemp GJ 2013. Is depression a disconnection syndrome? Meta-analysis of diffusion tensor imaging studies in patients with MDD. *Journal of Psychiatry and Neuroscience,* **38,** 49-56.

Linke J, King AV, Poupon C, Hennerici MG, Gass A & Wessa M 2013. Impaired anatomical connectivity and related executive functions: differentiating vulnerability and disease marker in bipolar disorder. *Biological Psychiatry,* **74,** 908-916.

Mahon K, Burdick KE, Ikuta T, Braga RJ, Gruner P, Malhotra AK & Szeszko PR 2013. Abnormal temporal lobe white matter as a biomarker for genetic risk of bipolar disorder. *Biological Psychiatry,* **73,** 177-182.

Matsuo K, Kopecek M, Nicoletti M, Hatch J, Watanabe Y, Nery F, Zunta-Soares G & Soares J 2011. New structural brain imaging endophenotype in bipolar disorder. *Molecular Psychiatry,* **17,** 412-420.

Mccormack C, Green MJ, Rowland JE, Roberts G, Frankland A, Hadzi-Pavlovic D, Joslyn C, Lau P, Wright A, Levy F, Lenroot RK & Mitchell PB 2016. Neuropsychological and social cognitive function in young people at genetic risk of bipolar disorder. *Psychological Medicine,* **46,** 745-758.

Mcdonald C, Bullmore ET, Sham PC, Chitnis X, Wickham H, Bramon E & Murray RM 2004. Association of Genetic Risks for Schizophrenia and Bipolar Disorder With Specific and Generic Brain Structural Endophenotypes. *Archives of General Psychiatry* **61,** 974-984.

Mckenna BS, Theilmann RJ, Sutherland AN & Eyler LT 2015. Fusing functional MRI and diffusion tensor imaging measures of brain function and structure to predict working memory and processing speed performance among inter-episode bipolar patients. *Journal of the International Neuropsychological Society* **21,** 330-41.

Nortje G, Stein DJ, Radua J, Mataix-Cols D & Horn N 2013. Systematic review and voxel-based meta-analysis of diffusion tensor imaging studies in bipolar disorder. *Journal of Affective Disorders,* **150,** 192-200.

Nurnberger JI, Mcinnis M, Reich W, Kastelic E, Wilcox HC, Glowinski A, Mitchell P, Fisher C, Erpe M & Gershon ES 2011. A high-risk study of bipolar disorder: Childhood clinical phenotypes as precursors of major mood disorders. *Archives of General Psychiatry,* **68,** 1012-1020.

Perich T, Lau P, Hadzi-Pavlovic D, Roberts G, Frankland A, Wright A, Green M, Breakspear M, Corry J, Radlinska B, Mccormack C, Joslyn C, Levy F, Lenroot R, Nurnberger JIJ & Mitchell PB 2015. What clinical features precede the onset of bipolar disorder? *Journal of Psychiatric Research,* **62,** 71-77.

Phillips MLS, Ha 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.

Poletti S, Bollettini I, Mazza E, Locatelli C, Radaelli D, Vai B, Smeraldi E, Colombo C & Benedetti F 2015. Cognitive performances associate with measures of white matter integrity in bipolar disorder. *Journal of Affective Disorders,* **174,** 342-352.

Roberts G, Green MJ, Breakspear M, Mccormack C, Frankland A, Wright A, Levy F, Lenroot R, Chan HN & Mitchell PB 2013. Reduced inferior frontal gyrus activation during response inhibition to emotional stimuli in youth at high risk of bipolar disorder. *Biological Psychiatry,* **74,** 55-61.

Roybal DJ, Barnea-Goraly N, Kelley R, Bararpour L, Howe ME, Reiss AL & Chang KD 2015. Widespread white matter tract aberrations in youth with familial risk for bipolar disorder. *Psychiatry Research,* **232,** 184-192.

Sarrazin S, D'albis M-A, Mcdonald C, Linke J, Wessa M, Phillips M, Delavest M, Emsell L, Versace A & Almeida J 2015. Corpus callosum area in patients with bipolar disorder with and without psychotic features: an international multicentre study. *Journal of Psychiatry & Neuroscience,* **40,** 352.

Sarrazin S, Poupon C, Linke J, Wessa M, Phillips M, Delavest M, Versace A, Almeida J, Guevara P & Duclap D 2014. A multicenter tractography study of deep white matter tracts in bipolar i disorder: Psychotic features and interhemispheric disconnectivity. *JAMA Psychiatry,* **71,** 388-396.

Saxena K, Tamm L, Walley A, Simmons A, Rollins N, Chia J, Soares JC, Emslie GJ, Fan X & Huang H 2012. A preliminary investigation of corpus callosum and anterior commissure aberrations in aggressive youth with bipolar disorders. *Journal of Child and Adolescent Psychopharmacology,* **22,** 112-119.

Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H & Aluwahlia S 1983. A children's global assessment scale (CGAS). *Archives of General Psychiatry,* **40,** 1228-1231.

Skudlarski P, Schretlen DJ, Thaker GK, Stevens MC, Keshavan MS, Sweeney JA, Tamminga CA, Clementz BA, O'neil K & Pearlson GD 2013. Diffusion tensor imaging white matter endophenotypes in patients with schizophrenia or psychotic bipolar disorder and their relatives. *American Journal of Psychiatry,* **170,** 886-898.

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM & Behrens TE 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage,* **31,** 1487-505.

Sprooten E, Brumbaugh MS, Knowles EE, Mckay DR, Lewis J, Barrett J, Landau S, Cyr L, Kochunov P & Winkler AM 2013. Reduced white matter integrity in sibling pairs discordant for bipolar disorder. *American Journal of Psychiatry,* **170,** 1317-1325.

Sprooten E, Sussmann JE, Clugston A, Peel A, Mckirdy J, Moorhead TWJ, Anderson S, Shand AJ, Giles S & Bastin ME 2011. White matter integrity in individuals at high genetic risk of bipolar disorder. *Biological Psychiatry,* **70,** 350-356.

Thomason ME & Thompson PM 2011. Diffusion imaging, white matter, and psychopathology. *Clinical Psychology,* **7,** 63.

Vederine F-E, Wessa M, Leboyer M & Houenou J 2011. A meta-analysis of whole-brain diffusion tensor imaging studies in bipolar disorder. *Progress in Neuropsychopharmacology and Biological Psychiatry,* **35,** 1820-1826.

Versace A, Ladouceur CD, Romero S, Birmaher B, Axelson DA, Kupfer DJ & Phillips ML 2010. Altered development of white matter in youth at high familial risk for bipolar disorder: a diffusion tensor imaging study. *Journal of the American Academy of Child and Adolescent Psychiatry,* **49,** 1249-1259.

Walterfang M, Wood AG, Barton S, Velakoulis D, Chen J, Reutens DC, Kempton MJ, Haldane M, Pantelis C & Frangou S 2009. Corpus callosum size and shape alterations in individuals with bipolar disorder and their first-degree relatives. *Progress in Neuro-Psychopharmacology and Biological Psychiatry,* **33,** 1050-1057.

Wechsler D 1999. *Wechsler Abbreviated Scale of Intelligence,* New York, The Psychological Corporation: Harcourt Brace & Company.

Xu K, Jiang W, Ren L, Ouyang X, Kong L, Womer F, Liu Z & Blumberg HP 2013. Impaired interhemispheric connectivity in medicationnaive patients with major depressive disorder. *Journal of Psychiatry & Neuroscience,* **38,** 43-48.

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

Yurgelun-Todd DA, Silveri MM, Gruber SA, Rohan ML & Pimentel PJ 2007. White matter abnormalities observed in bipolar disorder: a diffusion tensor imaging study. *Bipolar Disorders,* **9,** 504-512.
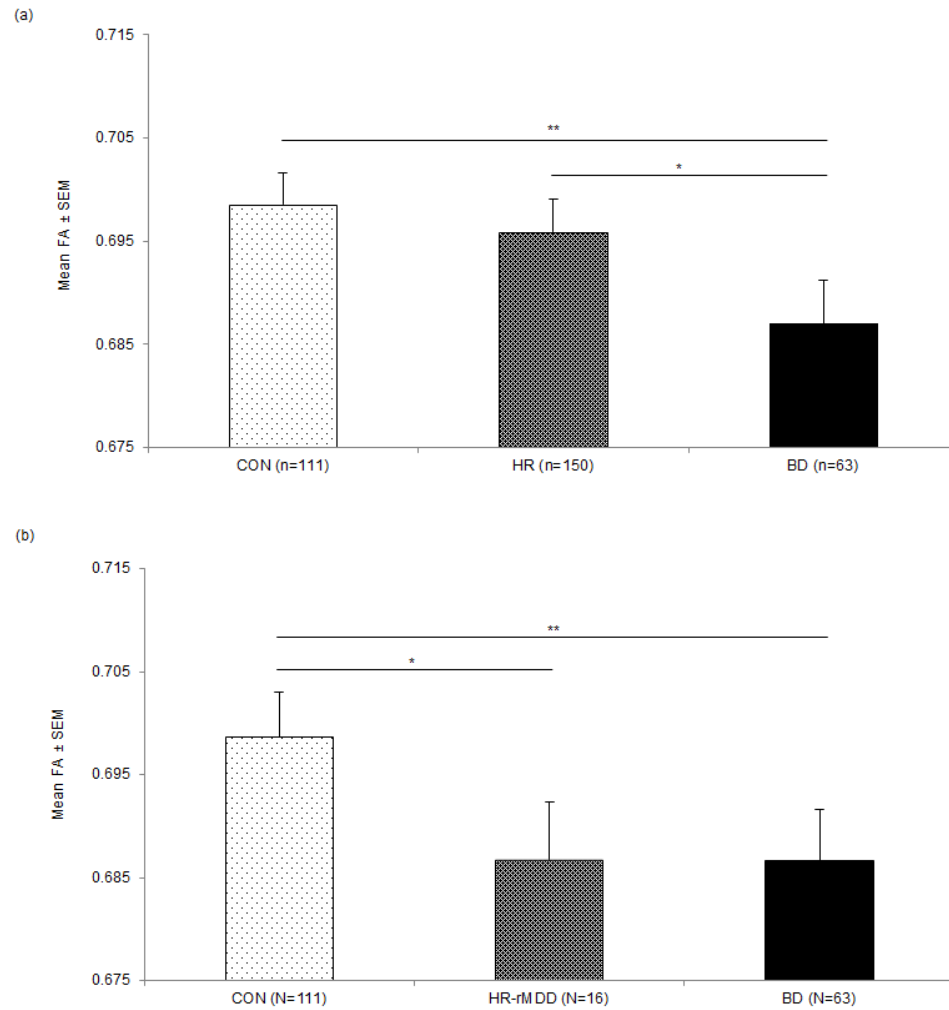
Fig. 1. Group differences in mean FA in the genu of the corpus callosum

(A) The BD group showed lower FA in the genu of the corpus callosum (CC) compared to HR and CON groups. FDR corrected *group* effect and pairwise comparisons. (B) When HR subjects with recurrent MDD (HR-rMDD) (n=16) were selected from our total HR sample (n=150) and compared against the CON and BD groups, the HR-rMDD group had reduced FA in the genu of the CC compared to CON subjects. FDR corrected pairwise comparisons. Abbreviations: CON, controls; HR, high risk; HR-rMDD, high risk subjects with a lifetime diagnosis of recurrent major depressive disorder; BD, bipolar disorder. Estimated means are reported after covarying for age and gender. * p<0.05, **p<0.01.

Table 1. Demographic data for control (CON), high risk (HR), and bipolar disorder (BD) groups

Demographic data	CON (n=111)	HR (n=150)	BD (n=63)	Difference Statistic	p-value	Pairwise Comparisons
Males, *n (%)*	49 (44.1)	65 (43.3)	21 (33.3)	χ^2= 2.25	0.32	
Intelligence Quotient, *mean*	117.26 (1.05)	113.73 (0.96)	116.11 (1.52)	F= 3.11	0.05	CON>HR p<0.05
(SD) Years of Education, *mean*	15.34 (2.77)	12.59 (3.67)	15.48 (2.46)	F=30.95	<0.01	CON>HR p<0.01, BD>HR p<0.01
(SD) Age, *mean (SD)*	21.83 (3.94)	19.73 (5.59)	24.95 (3.83)	F= 27.16	<0.01	CON>HR p<0.01, BD>HR p<0.01, BD>CON p<0.01

Table 2. Clinical and medication data for control (CON), high risk (HR), and bipolar disorder (BD) groups

Lifetime DSM-IV diagnosis	CON (n=111)	HR (n=150)	BD (n=63)	Difference Statistic	p-value	Pairwise Comparisons
Any diagnosis, *n (%)*	32 (28.8)	75 (50.0)	63 (100.0)	$\chi^2 = 82.31$	<0.01	HR>CON p<0.01 BD>CON p<0.01 BD>HR p<0.01
At least one MDE, *n (%)*	12 (10.8)	44 (29.3)	60 (95.2)	$\chi^2 = 131.94$	<0.01	HR>CON p<0.01 BD>CON p<0.01 BD>HR p<0.01
Recurrent MDD, *n (%)*	3 (2.7)	16 (10.7)	-	$\chi^2 = 5.99$	0.01	HR>CON, p<0.05
Any anxiety disorder, *n (%)*	12 (11.7)	32 (22.1)	30 (49.2)	$\chi^2 = 32.63$	<0.01	HR>CON p<0.05 BD>CON p<0.01 BD>HR p<0.01
Any behavioural disorder, *n (%)*	1 (1.0)	16 (11.3)	9 (15.3)	$\chi^2 = 13.33$	<0.01	HR>CON p<0.01 BD>CON p<0.01
Any substance disorder, *n (%)*	7 (6.3)	13 (18.7)	15 (23.8)	$\chi^2 = 14.10$	<0.01	BD>CON p<0.01 BD>HR p<0.01
Current DSM-IV diagnosis						
MDE, *n (%)*	2 (1.8)	3 (2.0)	3 (4.8)	$\chi^2 = 1.61$	0.45	
Hypo/manic episode, *n (%)*	0 (0)	0 (0)	0 (0)	-	-	
Symptom severity scales						
22 to 30 years	n=54	n=49	n=44			
MADRS, *mean (SD)*	2.22 (4.05)	3.00 (4.03)	10.68 (10.65)	F=21.41	<0.01	BD>CON p<0.01 BD>HR p<0.01
YMRS, *mean (SD)*	1.00 (1.71)	0.98 (1.51)	4.61 (4.58)	F=27.27	<0.01	BD>CON p<0.01 BD>HR p<0.01
12 to 21 Years	n=37	n=70	n=9			
CDI, *mean (SD)*	7.13 (4.26)	8.36 (5.98)	19.89 (7.51)	F= 19.36	<0.01	BD>CON p<0.01 BD>HR p<0.01
Clinical Characteristics						
Global Functioning						
GAS/C-GAS, *mean (SD)*	91.01 (5.65)	85.63 (9.88)	78.17 (11.74)	F= 38.86	<0.01	CON>HR p<0.01 CON>BD p<0.01 HR>BD p<0.01

Age at First

MDE, *mean (SD)*	19.18 (2.96)	16.98 (4.77)	15.45 (3.58)	F= 5.95	<0.01	CON>BD p<0.01
Hypomanic episode, *mean (SD)*	-	-	17.96 (4.18)	-	-	-
Manic episode, *mean (SD)*	-	-	19.48 (4.33)	-	-	-
Elevated mood episode, *mean (SD)*	-	-	18.17 (4.21)	-	-	-
Mood episode, *mean (SD)*	19.18 (2.96)	16.98 (4.77)	15.26 (3.99)	F= 7.11	<0.01	CON>BD p<0.01

Number of Episodes

Major depressive episode, *mean (SD)*	1.41 (1.23)	2.44 (2.82)	11.93 (11.68)	F=20.57	<0.01	CON<BD p<0.01
Hypomanic episodes, *mean (SD)*	-	-	10.45 (11.2)	-		HR<BD p<0.01
Manic episode, *mean (SD)*	-	-	3.13 (2.66)	-		-
Any elevated mood episode, *mean (SD)*	-	-	9.95 (11.06)	-		-
Any mood episode, *mean (SD)*	1.41 (1.23)	2.44 (2.82)	21.28 (19.87)	F= 28.04	<0.01	CON<BD p<0.01 HR<BD p<0.01

Current Psychotropic Medication

At least one psychotropic medication, *n (%)*	4 (3.6)	11 (8.0)	46 (76.2)
Anti-depressants, *n (%)*	3 (2.7)	10 (6.7)	36 (57.1)
Mood stabilisers, *n (%)*	-	-	39 (61.9)
Anti-psychotics, *n (%)*	-	-	20 (31.7)
Benzodiazepines, *n (%)*	-	1 (1.1)	2 (1.26)
Stimulants, *n (%)*	1 (1.5)	-	1 (0.63)
Anti-convulsants, *n (%)*	-	-	5 (3.15)

Abbreviations: MADRS, Montgomery-Asberg Depression Rating Scale; YMRS, Youth mania rating Scale; CDI, Children's Depression Inventory: CON, controls; HR, high risk; BD, bipolar disorder; MDE, Major Depressive Episode; MDD, Major Depressive Disorder; GAS, Global Assessment Scale; C-GAS, Children's Global Assessment Scale.   All participants had an IQ above 80. A lifetime major depressive episode is defined as meeting DSM-IV criteria for at least one major depressive episode. Confidence rating ranges using the Best Estimate Methodology vary from 1-4, where 1 represents criteria not met for a diagnosis and 4 represents a definite diagnosis. All diagnoses listed here have a confidence rating of 3 or higher (i.e. probable or definite).

Among the thirteen participants with a behavioural disorder, ten had attention-deficit hyperactivity disorder (five current), three had oppositional defiant disorder (one current) and three had conduct disorder (none current). The average age of onset of recurrent or single episode MDD in our sample was 17.9 (SD 5.0) in the HR group, 19.1 (2.9) in controls and 15.1 (2.8) in the BD group. For anxiety disorders, the figures were 12.7 (6.5), 11.3 (6.7) and 12.5 (6.1) respectively. Of the 39 BD participants taking a mood stabiliser, 15 were taking lithium

Supplementary Methods

Participants

Only one acquisition was collected from ten participants and included in the analysis (five CON, three HR, and two BD). Six additional participants (two CON, three HR, and one BD) were removed because of poor contrast.

Diffusion tensor imaging pre-processing

DTI data were pre-processed offline on a Linux system-based workstation using the FMRIB's Diffusion Toolbox (FDT), which is a part of the FMRIB's Software Library (FSL) (version 5.0)[56]. First, each subject's raw DTI data were visually inspected to exclude any severe head movements and DTI artefacts. Second, both acquisitions were merged and each subject was corrected for eddy current and head motion by registering each participant's 6 diffusion weighted images to their own T2-weighted (b = 0) image using an affine image registration (FLIRT). A binary brain mask was created to remove the non-brain tissue using the Brain Extraction Tool (BET) in FSL. Then, diffusion tensor was reconstructed by fitting a diffusion tensor model to each image voxel of the pre-processed DTI data in ICBM diffusion space, using the DTIfit program included in FSL. Following this, the fractional anisotropy (FA) map was calculated and visually inspected for each participant.

DTI skeletonization and ROI extraction from FA images

In sum, the FA images were non-linearly mapped onto the ENIGMA-DTI FA template to project the skeleton using FSL fnirt (Smith et al., 2007, Smith et al., 2006). Rather than employing the ENIGMA templates provided, registered images from our dataset were combined to create a mask and distance map. Images were visually inspected to ensure FA maps were registered. Images were aligned to the DTI template in ICBM space and mean FA in regions of interest (ROIs) were extracted from the skeletal projections using the TBSS analytic method (Smith et al., 2007, Smith et

al., 2006), where ROIs, were derived from the Johns Hopkins University (JHU) white matter

parcellation atlas (Mori et al., 2008).

The full brain mean FA values were calculated to include all voxels in the ENIGMA-DTI skeleton, not just those voxels in the segmented regions of interest, therefore also including peripheral white matter. This type of analysis has been used previously in both genetics studies and disease related studies (Nir et al., 2013, Kochunov et al., 2011, Jahanshad et al., 2013).

Table S1 Index of regions of interest (ROIs) from the JHU DTI atlas

ROIs	Abbreviation
ACR-L,R	Anterior corona radiata
ALIC-L,R	Anterior limb of internal capsule
CGC-L,R	Cingulum (cingulate gyrus)
CGH-L,R	Cingulum (hippocampus)
CP-L,R	Cerebral peduncle
CST-L,R	Corticospinal tract
EC-L,R	External capsule
FX/ST-L,R	Fornix (cres) / Stria terminalis
GCC	Genu of corpus callosum
ICP-L,R	Inferior cerebellar peduncle
ML-L,R	Medial lemniscus
PCR-L,R	Posterior corona radiata
PLIC-L,R	Posterior limb of internal capsule
PTR-L,R	Posterior thalamic radiation (include optic radiation)
RLIC-L,R	Retrolenticular part of internal capsule
SCP-L,R	Superior cerebellar peduncle
SCR-L,R	Superior corona radiata
SFO-L,R	Superior fronto-occipital fasciculus (could be a part of anterior internal capsule)
SLF-L,R	Superior longitudinal fasciculus
SS-L,R	Sagittal stratum (include inferior longitidinal fasciculus and inferior fronto-occipital fasciculus)
Tap-L,R	Tapetum
UNC-L,R	Uncinate fasciculus
BCC	Body of corpus callosum
FX	Fornix (column and body of fornix)
SCC	Splenium of corpus callosum

Abbreviations: L, left; R, right.

Table S2 Between-Group Effects

	CON	HR	BD		Covariates: Age and Gender	
	Mean (SEM)	Mean (SEM)	Mean (SEM)	p	FDR q	post hoc
ACR	0.465 (0.003)	0.470 (0.003)	0.463 (0.004)	0.113	0.021	na
ALIC	0.567 (0.003)	0.571 (0.003)	0.567 (0.004)	0.423	0.037	na
BCC$^+$	0.648 (0.004)	0.642 (0.004)	0.637 (0.005)	0.081	0.019	na
CGC	0.587 (0.004)	0.581 (0.004)	0.578 (0.006)	0.229	0.029	na
CGH	0.555 (0.006)	0.544 (0.005)	0.554 (0.008)	0.162	0.023	na
CP	0.735 (0.003)	0.735 (0.003)	0.730 (0.003)	0.309	0.033	na
CST	0.595 (0.005)	0.600 (0.005)	0.595 (0.007)	0.864	0.046	na
EC	0.461 (0.003)	0.461 (0.003)	0.460 (0.003)	0.943	0.048	na
FX$^+$	0.531 (0.006)	0.532 (0.007)	0.530 (0.009)	0.980	0.050	na
FX_ST	0.572 (0.004)	0.569 (0.004)	0.565 (0.005)	0.476	0.038	na
GCC$^+$	0.699 (0.003)	0.696 (0.003)	0.687 (0.004)	0.014^{*A}	0.014	BD<HR p=0.03, BD<CON p=0.003^{*A}
ICP	0.593 (0.005)	0.594 (0.004)	0.581 (0.005)	0.030^B	0.015	na
ML	0.666 (0.005)	0.668 (0.005)	0.661 (0.005)	0.559	0.040	na
PCR	0.510 (0.003)	0.510 (0.003)	0.510 (0.004)	0.941	0.048	na
PLIC	0.692 (0.003)	0.695 (0.003)	0.692 (0.003)	0.519	0.040	na
PTR	0.661 (0.004)	0.655 (0.005)	0.651 (0.005)	0.172	0.026	na
RLIC	0.607 (0.003)	0.614 (0.003)	0.613 (0.004)	0.155	0.020	na
SCC$^+$	0.768 (0.003)	0.768 (0.003)	0.762 (0.003)	0.185	0.026	na
SCP	0.650 (0.004)	0.645 (0.004)	0.643 (0.004)	0.385	0.036	na
SCR	0.500 (0.002)	0.4982 (0.002)	0.498 (0.003)	0.634	0.043	na
SFO	0.511 (0.003)	0.512 (0.003)	0.511 (0.004)	0.958	0.049	na
SLF	0.514 (0.002)	0.513 (0.003)	0.510 (0.003)	0.238	0.030	na
SS$^\#$	0.572 (0.003)	0.571 (0.004)	0.565 (0.005)	0.317	0.018	na
TAP	0.577 (0.006)	0.566 (0.006)	0.559 (0.008)	0.041^B	0.015	na
UNC	0.544 (0.005)	0.537 (0.005)	0.538 (0.006)	0.342	0.035	na
Average FA	0.465 (0.002)	0.465 (0.002)	0.461 (0.002)	0.156	n/a	na

$^+$ represents inter-hemispheric regions. ^{B} represents effects that were significant (p<0.05) prior to correcting for multiple tests. *A effects that survive FDR correction at a threshold of p<0.05 (p-values that are smaller than their corresponding q-vakue survived FDR correction). #= only variable where age*group was included as an extra interaction as p<0.05 for this interaction. Abbreviations: CON, controls; HR, high risk; BD, bipolar disorder; na, not-applicable. Estimated means are reported covarying for age and gender.

Table S3 Hemisphere Effects

	Right	Left			
	Mean (SEM)	*Mean (SEM)*	*p*	*FDR q*	*Direction*
ACR	0.470 (0.002)	0.463 (0.002)	<0.001*	0.010	R > L
ALIC	0.569 (0.002)	0.568 (0.002)	0.213	0.028	na
CGC	0.565 (0.004)	0.599 (0.004)	<0.001*	<0.001	R < L
CGH	0.571 (0.005)	0.531 (0.005)	<0.001*	<0.001	R > L
CP	0.745 (0.003)	0.722 (0.002)	<0.001*	0.002	R > L
CST	0.598 (0.005)	0.594 (0.005)	0.076	0.019	na
EC	0.456 (0.002)	0.464 (0.002)	<0.001*	0.008	R < L
FX_ST	0.576 (0.004)	0.562 (0.003)	<0.001*	0.007	R < L
ICP	0.594 (0.004)	0.585 (0.004)	<0.001*	0.007	R > L
ML	0.669 (0.004)	0.661 (0.004)	<0.001*	0.009	R > L
PCR	0.516 (0.003)	0.504 (0.002)	<0.001*	0.003	R > L
PLIC	0.706 (0.002)	0.681 (0.002)	<0.001*	0.004	R > L
PTR	0.662 (0.004)	0.649 (0.004)	<00.01*	0.006	R > L
RLIC	0.611 (0.003)	0.611 (0.003)	0.877	0.050	na
SCP	0.648 (0.005)	0.649 (0.004)	<0.001*	0.012	R > L
SCR	0.497 (0.002)	0.500 (0.002)	0.001*	0.012	R < L

SFO	0.515 (0.003)	0.508 (0.003)	<0.001*	0.011	R > L
SLF	0.513 (0.002)	0.511 (0.002)	0.006*	0.013	ns
SS#	0.578 (0.003)	0.561 (0.003)	<0.001*	0.005	R > L
TAP	0.606 (0.006)	0.528 (0.005)	<0.001*	0.004	R > L
UNC	0.545 (0.004)	0.535 (0.045)	<0.001*	0.010	R > L

*denotes effects that survive FDR correction at a threshold of P<0.05 (p-values that are smaller than their corresponding q-value survived FDR correction). #= only variable where age*group was included as an extra interaction as p<0.05 for this interaction. Abbreviations: ns, non-significant; na, not-applicable. Estimated means are reported covarying for age and gender.

Right hemisphere lateralisation was evident in 13 ROIs and left hemisphere lateralisation was identified in 4 ROIs. Our interhemispheric findings are in accord with a number of previous reports including rightward asymmetry of the anterior and posterior corona radiata, uncinate fasciculus, parts of the posterior thalamic radiation, medial lemniscus, and the posterior limb of the internal capsule (Takao et al., 2011a, Takao et al., 2011b, Takao et al., 2013, Yin et al., 2013), and leftward lateralisation in areas of the external capsule and cingulum (Gong et al., 2005, Takao et al., 2013). As the brain is asymmetric in function (Barrick et al., 2007) and as asymmetry is influenced by a variety of factors such as genes and hormones (Toga and Thompson, 2003), patterns are difficult to interpret and disentangle.

Table S4 Group x Hemisphere Interaction

	CON Left	CON Right	HR Left	HR Right	BD Left	BD Right			Covariates: Age and Gender	
									post hoc	*post hoc*
	Mean (SEM)	*Mean (SEM)*	*Mean (SEM)*	*Mean (SEM)*	*Mean (SEM)*	*Mean (SEM)*	*p*	*FDR q*	*within*	*between*
									hemisphere	*hemisphere*
ACR	0.463 (0.003)	0.468 (0.003)	0.468 (0.003)	0.473 (0.003)	0.458 (0.004)	0.467 (0.004)	0.269	0.032	na	na
ALIC	0.566 (0.003)	0.569 (0.003)	0.571 (0.003)	0.571 (0.003)	0.567 (0.004)	0.567 (0.004)	0.274	0.032	na	na
CGC	0.603 (0.005)	0.571 (0.005)	0.598 (0.005)	0.564 (0.004)	0.595 (0.006)	0.561 (0.006)	0.848	0.045	na	na
CGH	0.536 (0.006)	0.574 (0.006)	0.527 (0.006)	0.562 (0.006)	0.531 (0.008)	0.577 (0.009)	0.239	0.031	na	na
CP	0.724 (0.003)	0.746 (0.003)	0.725 (0.003)	0.746 (0.004)	0.718 (0.004)	0.742 (0.004)	0.328	0.034	na	na
CST	0.591 (0.006)	0.598 (0.006)	0.596 (0.006)	0.600 (0.006)	0.594 (0.007)	0.596 (0.007)	0.611	0.042	na	na
EC	0.463 (0.003)	0.460 (0.003)	0.464 (0.003)	0.459 (0.002)	0.463 (0.003)	0.458 (0.003)	0.562	0.041	na	na
FX_ST	0.567 (0.004)	0.576 (0.004)	0.563 (0.004)	0.576 (0.005)	0.555 (0.005)	0.576 (0.06)	0.092	0.021	na	na
ICP	0.587 (0.004)	0.599 (0.005)	0.589 (0.005)	0.600 (0.004)	0.577 (0.005)	0.584 (0.005)	0.440	0.037	na	na
ML	0.662 (0.005)	0.671 (0.005)	0.663 (0.005)	0.673 (0.005)	0.659 (0.006)	0.663 (0.005)	0.168	0.025	na	na
PCR	0.501 (0.003)	0.517 (0.003)	0.505 (0.003)	0.516 (0.003)	0.504 (0.003)	0.515 (0.004)	0.050	0 .017	na	na
PLIC	0.681 (0.003)	0.704 (0.003)	0.683 (0.003)	0.708 (0.003)	0.679 (0.003)	0.706 (0.003)	0.332	0.035	na	na
rPTR	0.654 (0.005)	0.668 (0.004)	0.646 (0.005)	0.664 (0.005)	0.647 (0.005)	0.655 (0.006)	0.150	0.023	na	rPTR
RLIC	0.608 (0.003)	0.607 (0.004)	0.612 (0.004)	0.616 (0.004)	0.614 (0.004)	0.611 (0.005)	0.163	0.020	na	na
SCP	0.648 (0.004)	0.651 (0.004)	0.645 (0.004)	0.646 (0.004)	0.639 (0.005)	0.646 (0.005)	0.068	0.018	na	na
SCR	0.501 (0.003)	0.499 (0.002)	0.500 (0.002)	0.496 (0.002)	0.499 (0.003)	0.497 (0.003)	0.477	0.039	na	na
SFO	0.507 (0.004)	0.515 (0.003)	0.510 (0.004)	0.513 (0.003)	0.506 (0.005)	0.516 (0.004)	0.196	0.027	na	na
SLF	0.513 (0.002)	0.515 (0.003)	0.512 (0.003)	0.514 (0.003)	0.507 (0.003)	0.510 (0.003)	0.708	0.043	na	na

SS[#]	0.563 (0.004)	0.580 (0.004)	0.563 (0.004)	0.580 (0.004)	0.557 (0.005)	0.573 (0.005)	0.961	0.047	na	na
TAP	0.537 (0.007)	0.617 (0.007)	0.527 (0.006)	0.604 (0.006)	0.520 (0.008)	0.597 (0.009)	0.847	0.044	na	na
UNC	0.540 (0.005)	0.549 (0.005)	0.534 (0.005)	0.540 (0.006)	0.531 (0.007)	0.546 (0.006)	0.228	0.029	na	na

No group*hemisphere were evident at a threshold of $p<0.05$ uncorrected. Abbreviations: CON, controls; HR, high risk; BD, bipolar disorder; na, non-applicable; Estimated means are reported covarying for age and gender.

Additional references

Barrick TR, Lawes IN, Mackay CE & Clark CA 2007. White matter pathway asymmetry underlies functional lateralization. *Cerebral Cortex,* **17,** 591-598.

Gong G, Jiang T, Zhu C, Zang Y, Wang F, Xie S, Xiao J & Guo X 2005. Asymmetry analysis of cingulum based on scale-invariant parameterization by diffusion tensor imaging. *Human Brain Mapping,* **24,** 92-98.

Jahanshad N, Kochunov PV, Sprooten E, Mandl RC, Nichols TE, Almasy L, Blangero J, Brouwer RM, Curran JE & De Zubicaray GI 2013. Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA–DTI working group. *Neuroimage,* **81,** 455-469.

Kochunov P, Glahn DC, Lancaster J, Winkler A, Karlsgodt K, Olvera RL, Curran JE, Carless MA, Dyer TD & Almasy L 2011. Blood pressure and cerebral white matter share common genetic factors in Mexican Americans. *Hypertension,* **57,** 330-335.

Mori S, Oishi K, Jiang H, Jiang L, Li X, Akhter K, Hua K, Faria AV, Mahmood A, Woods R, Toga AW, Pike GB, Neto PR, Evans A, Zhang J, Huang H, Miller MI, Van Zijl P & Mazziotta J 2008. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage,* **40,** 570-82.

Nir TM, Jahanshad N, Villalon-Reina JE, Toga AW, Jack CR, Weiner MW & Thompson PM 2013. Effectiveness of regional DTI measures in distinguishing Alzheimer's disease, MCI, and normal aging. *Neuroimage: Clinical,* **3,** 180-195.

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ & Matthews PM 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage,* **31,** 1487-1505.

Smith SM, Johansen-Berg H, Jenkinson M, Rueckert D, Nichols TE, Miller KL, Robson MD, Jones DK, Klein JC & Bartsch AJ 2007. Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. *Nature Protocols,* **2,** 499-503.

Takao H, Abe O, Yamasue H, Aoki S, Sasaki H, Kasai K, Yoshioka N & Ohtomo K 2011a. Gray and white matter asymmetries in healthy individuals aged 21–29 years: A voxel-based morphometry and diffusion tensor imaging study. *Human Brain Mapping,* **32,** 1762-1773.

Takao H, Hayashi N & Ohtomo K 2011b. White matter asymmetry in healthy individuals: a diffusion tensor imaging study using tract-based spatial statistics. *Neuroscience,* **193,** 291-299.

Takao H, Hayashi N & Ohtomo K 2013. White matter microstructure asymmetry: effects of volume asymmetry on fractional anisotropy asymmetry. *Neuroscience,* **231,** 1-12.

Toga AW & Thompson PM 2003. Mapping brain asymmetry. *Nature Reviews Neuroscience,* **4,** 37-48.

Yin X, Han Y, Ge H, Xu W, Huang R, Zhang D, Xu J, Fan L, Pang Z & Liu S 2013. Inferior frontal white matter asymmetry correlates with executive control of attention. *Human Brain Mapping,* **34,** 796-813.