Manuscript Details

Manuscript number PSY_2016_933

Title The association of sleep and physical activity with integrity of white matter microstructure in bipolar disorder patients and healthy controls

Abstract

We investigate how the sleep disruptions and irregular physical activity levels that are prominent features of bipolar disorder (BD) relate to white matter microstructure in patients and controls. Diffusion tension imaging (DTI) and 14-day actigraphy recordings were obtained in 51 BD I patients and 55 age-and-gender-matched healthy controls. Tract-based spatial statistics (TBSS) was used for voxelwise analysis of the association between fractional anisotropy (FA) and sleep and activity characteristics in the overall sample. Next, we investigated whether the relation between sleep and activity and DTI measures differed for patients and controls. Physical activity was related to increased integrity of white matter microstructure regardless of bipolar diagnosis. The relationship between sleep and white matter microstructure was more equivocal; we found an expected association between higher FA and effective sleep in controls but opposite patterns in bipolar patients. Confounding factors such as antipsychotic medication use are a likely explanation for these contrasting findings and highlight the need for further study of medication-related effects on white matter integrity.

Submission Files Included in this PDF

File Name [File Type]

Correction Letter_Psychiatry_Research_mb.docx [Response to reviewers]

DTI_Sleep_and_Activity_Manuscript_Version12_clean.docx [Manuscript]

Figure 1_v2_cropped.pdf [Figure]

Figure 2_v2_cropped.pdf [Figure]

Figure 3_v2_cropped.pdf [Figure]

Figure 4_v2_cropped.pdf [Figure]

Tables.docx [Table]

Abstract.docx [Abstract]

Highlights.docx [Highlights]

Supplementary Figure 1_v2_cropped.pdf [Supporting File]

Supplementary Figure 2_v2_cropped.pdf [Supporting File]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Dear prof. Buchsbaum,

Thank you for reviewing our paper *The association of sleep and physical activity with integrity of white matter microstructure in bipolar disorder patients and healthy controls* and the opportunity to revise. We would like to re-submit this paper. Please find below our address of all points raised by the reviewers. We have indicated the changes in the response letter in italic and in the manuscript in yellow. We feel the changes constitute important improvements and we look forward to your further assessment of suitability for publication in your journal.

Reviewer 1:

General: the manuscript is fairly well-written and the topic is of relevance for the field of brain stimulation and quality of life in bipolar disorder. The strength of the findings is, however, questionable. To start with the DTI component of this study is lacking. Using FA without mentioning RD, AD (and to a certain extent MD) is restrictive as these parameters may provide information on myelination and other white matter alterations. The authors do not mention ROIs or white matter tracks where differences in FA between BD and HC were observed. No FA skeleton map is provided and they use an average FA value of the brain without explaining the rationale for this approach. Little is said about previous work and data correlating sleep/physical activity and DTI. Further, the authors should have addressed the potential biological correlates between the proposed variables. The term “physical activity” appears in the title and in the manuscript but is not well defined. The authors obviously do not mean “physical exercise”, are they talking about movements while individuals sleep? Although promising the current findings are too general, hard to interpret and do not advance knowledge in this field.

Response: We thank the reviewer for his/her comments and the positive assessment of the writing.

Major comments:

Introduction

1. Provide additional information on biological mechanisms that may link DTI changes and sleep or physical activity (oxygen-related, cardiovascular health etc.).

Response: We agree with the reviewer that our proposed link between DTI and sleep/physical activity can be further elucidated by the following additions to the introduction:

Manuscript: Animal studies showed that oligodendrocyte proliferation and myelin-related gene expression strongly increase during sleep and reduce during wake time and experimental sleep deprivation (Bellesi et al., 2013). Conversely, chronic sleep loss has been associated with down-regulation of the expression of myelin-related genes coding for plasmolipin and a membrane protein expressed in the mature myelin sheath (Cirelli et al., 2006).

Manuscript: Possible underlying mechanisms involve the increased expression and release of brain-derived neurotropic factor and oxygen supply as a result of increased physical activity (Burzynska et al., 2014).

2. Provide additional information on previous studies on DTI and sleep/physical activity and explain why DTI is an ideal technique to study associations between these variables.

Response: We thank the reviewer for this suggestion. As of yet, few studies have addressed the link between DTI measures and sleep/physical activity. We have extended our introduction with recent articles on the topic of DTI and sleep/activity:

Manuscript: Human studies using DTI measures found that variability in sleep duration during week-days was associated with reduced FA in the parietocortical, frontocortical and frontostriatal tracts (Telzer et al., 2015). Furthermore, a sleep deprivation study in adult males found widespread changes in DTI measures after just one day of sleep deprivation (Elvsåshagen et al., 2015). These findings suggest that effective sleep relates to white matter microstructure and that sleep disruptions could lead to changes in oligodendrocytes, myelination, and axonal integrity and organisation.

3. Please explain, even if briefly, what FA changes mean in biological and functional connectivity terms.

Response: Yes, we have added the following statement to our introduction.

Manuscript: More organized and myelinated axon will cause water molecules to diffuse in a particular direction rather than randomly in all directions. Therefore, FA is thought to reflect integrity of white matter with lower FA values pointing to sparse, poorly myelinated, or divergent fibers (Beaulieu, 2002).

Methods

4. When reading the methods it was unclear to me if this was a cross-sectional or longitudinal study, and whether BD or siblings and BD were included. Please make it clear (for instance by adding "we used a subsample of participants recruited as part of the DBC study..", "Participants included HC and BDs..".

Response: This is a good point raised by the reviewer. We have edited this paragraph so that it is now clear we used a cross-sectional design and recruited a subsample of bipolar patients and healthy controls from the DBC study.

Manuscript: A total of 106 participants (51 bipolar patients, 55 controls) were recruited from the Dutch Bipolar Cohort (DBC) study, which is a collaboration between the University Medical Center Utrecht (UMCU), various health care institutes in the Netherlands and the University of California Los Angeles (UCLA). The objective of the DBC study is to collect a deep phenotype characterization of bipolar I patients, their first-degree relatives and controls. After completion of the DBC protocol, a subgroup of patients and controls were invited by telephone and mail to participate in the actigraphy and MRI measurements.The medical ethical committee of the UMCU approved the DBC study and the actigraphy

and MRI protocols. Written informed consent was obtained from all participants prior to participation. The majority of participants underwent MRI scanning prior to the two-week actigraphy protocol, with the exception of 6 participants who started with actigraphy measurements. The mean time difference between MRI and actigraphy measurements was 1.4 years.

5. Explain why a.only BD I, b.why 3 Dutch grandparents are needed, c.what are your inclusion/exclusion criteria for neurological, cardiovascular, substance use disorders

Response: We see how the 3 Dutch grandparents raises confusion, as this is an inclusion criterion specifically for the DBC study and not for the actigraphy and MRI protocols. We therefore decided to delete it from the method section. We did extend the paragraph with details regarding the BD I diagnosis and the exclusion criteria. We did not exclude participants with cardiovascular or substance use disorder, which we have now added to our limitation section.

Manuscript, Methods: Inclusion criterion for patients was a diagnosis of bipolar disorder type I, in order to create a clinically homogeneous sample. Controls with a bipolar or psychotic diagnosis were excluded, as were control subjects who had a first or second degree relative with a bipolar or psychotic diagnosis. Any other type of psychiatric disorder was not an exclusion criterion in order to avoid the recruitment of a 'super-normal control sample', that is not representative of the general population. Exclusion criteria for all participants were sleep apnoea, pregnancy, head trauma and neurological disorders.

Manuscript, Discussion: Also, we need to point out that we did not exclude participants diagnosed with conditions other than sleep apnoea, head trauma or neurological disorders.

6. Was the data used in this manuscript collected in this manuscript data from the subgroup tested after 1.4 years?

Response: The time difference of 1.4 years is between the actigraphy and MRI measurements. We have edited the method section to make this more clear:

Manuscript: The majority of participants underwent MRI scanning prior to the two-week actigraphy protocol, with the exception of 6 participants who started with actigraphy measurements. The mean time difference between MRI and actigraphy measurements was 1.4 years.

7. physical activity: please define what you measured exactly. Did you collect information on lifestyle and dietary habits too?

Response: We thank the reviewer for the opportunity to clarify this. Physical activity is measured with an accelerometer (the Actiwatch-2) and scored as the average wrist movements per minute. We have added this to our introduction and further clarified our methods. We did not analyze lifestyle and dietary habits.

Manuscript, Introduction: The current study investigates the link between white matter microstructure and objectively measured physical activity and sleep patterns of bipolar patients and controls. Both physical activity and sleep are measured using actigraphy.

Manuscript, Methods: The raw mean of activity was calculated over 24-hr period, followed by the mean of activity in four 6-hr periods corresponding to morning, afternoon, evening, night time: midnight to 6 AM (24-6 hr), 6 AM to noon (6 – 12 hr), noon to 6 PM (12 – 18 hr), and 6 PM to midnight (18 – 24 hr).

8. how did the authors correct for a potential "Hawthorne bias"? did they record information on usual sleep activity etc.?

Response: No objective information regarding usual sleep and activity patterns were available. We did collect self-report measures of habitual sleep patterns, which were only moderately correlated with the objective actigraphy measures ($r=0.39$, $p<0.01$ for work days and $r=0.22$, $p<0.05$ for free days). These moderate correlations have previously been reported and are thought to reflect a systemic bias in self reports of sleep, rather than a Hawthorn effect (Lauderdale et al., 2008). However, we don't expect that this bias resulting from a possible Hawthorne was systemic and has affected our results.

Statistical analyses

9. was the TBSS pipeline performed on entire FA dataset or did the authors perform analyses in specific ROIs?

Response: The TBSS analyses were performed on the entire FA dataset, according to the standard TBSS analysis pipeline as specified in references 38 and 39 (Smith et al., 2002; 2006). We have adapted the following sentence in our methods to further clarify this:

Manuscript: Whole-brain voxelwise analyses were carried out using the 'randomise' procedure, a tool for permutation-based non-parametric testing using the general linear model (Nichols and Holmes, 2002).

10.Please provide maps showing FA differences between BD and HC or at the very least mention which tracts could be of interest. 10b. have the authors considered entering the avarge FAs of significant tracts when conducting their regression analyses?

Response: This is an excellent suggestion and we have now added maps showing differences between patients and controls on FA to our results section (figure 1 and table 2). Given that these difference between patients and controls were found in numerous tracts, we decided not to use the average FA of the significant voxels in the subsequent regression analysis because this would limit the interpretability of our findings. Instead, we choose for the data-driven and hypothesis-free approach of a voxelwise analysis.

11. did the authors analyse RD, AD and MD too? if not please discuss or mention in the manuscript why you did not.

Response: We thank the reviewer for the opportunity to clarify this. We choose to focus on FA because this measures is the most studied diffusion parameter compared to AD, RD and MD. Moreover, FA correlates well with microscopic changes in white matter, whereas the biological underpinnings of AD, RD and MD are less well understood. We have added the following to our method section:

Manuscript: Compared to the other DTI metric (i.e. axial diffusivity, radial diffusivity and mean diffusivity), FA is the most well-studied diffusion parameter and correlates well with microscopic changes in white matter such as myelination, axonal diameter, density and organization (Beaulieu, 2009).

12. as part of your regression analyses please define method of entry of relevant variables e.g. stepwise, enter? And did you enter variables in blocks? If so please explain which variables were included in each block. Could you please clarify if you conducted an ordinal regression including a dummy variable for BD/HC?

Response: All linear regression analyses were performed using the Enter method in one block. We have added this to our method section. Mean FA was never categorized, so no ordinal regression analyses were performed. The dummy variable was used in the overall group analyses and the interaction analyses.

Manuscript: In case these variables showed significant associations with both FA and actigraphy measures, they were considered as potential confounders and separately entered as covariates in the stratified patient regression analyses, using the Enter method, with all variables entered in one block.

13. what was your statistical p-threshold?

Response: We have further clarified our p-value in our method section.

Manuscript: Our statistical threshold was a threshold-free cluster enhancement (TFCE) and family-wise error corrected P-value of 0.05.

14. the number of confounding variables is elevated. Did you enter confounding variables all at the same time? Did you consider the negative impact of confounding variables on the statistical power of this study? please discuss in the conclusions e.g. limitations.

Response: We agree with the reviewer that entering all confounding variables at once would negatively impact the power of our analyses. Instead, we studied the effect per confounding variable in separate models. We have adjusted our method section to clarify this issue:

Manuscript: In case these variables showed significant associations with both FA and actigraphy measures, they were considered as potential confounders and separately entered as covariates in the stratified patient regression analyses, using the Enter method, with all variables entered in one block.

15. which software did you use for statistical analyses?

Response: We thank the reviewer for pointing out this omission in our method section. We added the following:

Manuscript: In case of significant voxels in the overall sample, we created scatterplots to visualize the association between average FA of the significant clusters and actigraphy measures in IBM SPSS Statistics 21.0.

16. Table 1. Please provide measures of IQ, education, and current mood (HAMD, YMRS etc.) and please specific if BDs were euthymic or remitted etc. Please provide F and p-values too.

Response: These are excellent additions to our sample description and have been added to our results section and table 1.

Manuscript: Although none of the patients reported being in a mood episode, patients reported significantly more current depressive symptoms.

	Patients	**Controls**	F / χ^2	p-value
Age M (sd)	49.5 (11.4)	45.5 (15.8)	2.18	0.14
Gender M/F (% male)	28/23 (54.9%)	25/30 (45.5%)	0.44	0.22
Handedness R/L/B (% right)	42/6/3 (82.4%)	44/11/0 (80.0%)	4.37	0.11
IQ M (sd)	99.6 (15.7)	105.7 (14.4)	4.18	0.04*
Level of education n (%)			6.32	0.28
1. Low education	5 (9.3%)	4 (7.8%)		
2. Intermediate secondary education	4 (7.4%)	9 (17.6%)		
3. Intermediate professional education	7 (13.0%)	10 (19.6%)		
4. High prepatory vocational / pre-university	13 (24.1%)	6 (11.8%)		
5. Bachelor degree	11 (20.4%)	13 (25.5%)		
6. Master or PhD degree	14 (25.9%)	9 (17.6%)		
Inventory of Depressive Symptoms M (sd)	14.3 (10.1)	5.2 (3.9)	35.60	<0.001*
Altman Self-Rating Mania Scale M (sd)	1.9 (1.8)	1.7 (2.3)	0.41	0.53
Body Mass Index M (sd)	27.4 (4.3)	25.2 (3.3)	8.45	0.01*
Nr of workdays M (sd)	4.2 (3.5)	5.7 (3.7)	4.11	0.05*
Illness duration/ years M (sd)	16.5 (14.2)	-		
Number of episodes M (sd)	10.0 (10.8)	-		
Age at onset M (sd)	31.7 (12.2)	-		
History of psychotic symptoms N (%)	35 (70.0%)	-		

* Significant group difference at p<0.05

17. Table 2. What does “mean activity 0 to 24h” refer to?

Response: Table 2 describes the four 6-hour time windows in which 24 hour activity levels are subdivided. In the method section we described these windows in AM/PM times, so we understand why this does not directly translate to the terminology in table 2. We added this to our method section:

Manuscript: The raw mean of activity was calculated over 24-hr period, followed by the mean of activity in four 6-hr periods corresponding to morning, afternoon, evening, night time: midnight to 6 AM (24-6 hr), 6 AM to noon (6 – 12 hr), noon to 6 PM (12 – 18 hr), and 6 PM to midnight (18 – 24 hr).

18. please provide a table with FA statistical maps showing differences between BD and HC. this table should include size of ROI, name of tract (cite name of atlas too), MNI coordinates of ROI, F and p values

Response: The table showing differences between patients and controls on FA has been added to the manuscript, along with the maps.

Manuscript, Table 2:

Analysis	Cluster size (number of voxels)	Corrected p-value	T-value	MNI coordinates of center of gravity (x; y; z)	Anatomical region
Significant group difference in FA between patients and controls	44232	<0.001	3.08	2.37; -21.4; 16.8	Genu of corpus callosum Body of corpus callosum Splenium of corpus callosum R cerebral peduncle L cerebral peduncle R posterior limb of internal capsule L posterior limb of internal capsule L retrolenticular part of internal capsule R anterior corona radiata L anterior corona radiata R superior corona radiata Lsuperior corona radiata Right posterior corona radiata L posterior corona radiata R posterior thalamic radiation (include optic radiation) L posterior thalamic radiation (include optic radiation) R external capsule L external capsule L cingulum (cingulate gyrus) R superior longitudinal fasciculus L superior longitudinal fasciculus

Manuscript, methods: The MNI coordinates of the center of gravity of significant clusters was calculated and the anatomical regions of these clusters were identified using the 'JHU ICBM-DTI-81 White-Matter Labels' atlas, similar to Liu et al., 2016.

19. please provide a summary table of the results of your regressions including B, b value, p, R2 and also N after including confounding variables. I was a bit confused by the N cited on page 10.

Response: The results from the regression analyses have been summarized in table 5. After receiving the comment from one of the other reviewers that reporting on these regression analyses may be considered as double dipping (http://www.nature.com/neuro/journal/v12/n5/abs/nn.2303.html), we now only report the results of the post-hoc stratified patient analyses in order to only use this data as a means of sensitivity analysis. The N cited on page 10 solely included patients who were not using antipsychotic medication, which we have further clarified:

Manuscript: When the TBSS analysis for sleep duration was restricted to patients not using antipsychotic medication (N=25 patients and 55 controls), the interaction effect between sleep duration and group was no longer significant providing further support for a confounding effect of mediation.

	β	**R^2**	**p-value**	**N**
Sleep duration	-0.39	0.32	0.004	51
Sleep onset	0.49	0.37	<0.001	51
Sleep inertia	0.34	0.33	0.008	51

20. Discussion: the authors do not mention any region in particular where differences were found between BD and HC and it's therefore hard to understand what the current findings mean. Also it's hard to know why the authors used a whole-brain FA value? This is way too general and doesn't inform the reader on potential mechanistic or connectivity alterations in the brain. Please discuss or mention this in your introduction.

Response: The reviewer has a point that regarding the localization of the results improvements are needed. We have now added that the differences between patients and controls reflect several major white matter tracts. We agree that a measure of whole-brain FA would not be informative as this encompasses a very general measure of white matter integrity. We therefore report results from our voxel-wise analyses, which are data-driven and suit our hypothesis-free approach. We do include average FA in our regression analyses, but these measures concern the average FA of the significant

voxels from the corresponding TBSS analyses and not an whole-brain average. We have specified this in our method section.

Manuscript, Discussion: Bipolar patients were characterized by widespread reductions in fractional anisotropy (FA) in several major white matter tracts, including the corpus collosum, cerebral peduncle, corona radiata, internal and external capsule, thalamic radiation and superior longitudinal fasciculus.

Manuscript, Methods: We analysed the confounding effect of illness-specific variables (number of episodes, history of psychotic symptoms, illness duration and use of psychotropic medication) by testing the correlation with average FA of the significant cluster and the actigraphy measures.

21.Include a paragraph discussing biological changes (e.g.VO2 max and FA, blood flow increase-activity and FA, BDNF, reduced volume loss

Response: we agree that our discussion can be further improved by adding biological mechanisms that could relate to our findings:

Manuscript: As a clinical study previously demonstrated, sleep positively affects both the number and proliferation rate of olygodendrocytes, while spontaneous and forced wake states reduced oligodendrocyte proliferation by half (Bellesi et al., 2013). Longer sleep duration, earlier sleep onset and shorter sleep inertia could therefore increase olygodendrocyte proliferation and myelination, ultimately leading to increased integrity of white matter.

Manuscript: This may be partly explained by an up-regulation of the expression of brain-derived neurotropic factor (BDNF) by physical activity. This neurotrophin has consistently been related to increases in gray matter volume and evidence also points to a relation with myelinogenesis and white matter microstructure (Du et al., 2003; Tost et al., 2013). Alternatively, higher levels of cardiorespiratory fitness lead to improvements in the cerebral blood flow, resulting in increased brain volume and white matter integrity (Chen et al., 2013; Zhu et al., 2015).

22.Figures: please improve your captions to include reference to L-R, coordinates, name of tracts highlighted in the picture

Response: The names of the highlighted tracts are presented in table 2. The following has been added to the captions of our figures:

Manuscript figure 1, 2 and 3: Significant correlations are shown in yellow to red, depending on the significance level. The white matter skeleton is shown in green and overlaid on the MNI152 template brain. Images are shown in radiological convention (the left is displayed on the right) and corresponding Montreal Neurological Institute (MNI) coordinates are displayed below.

Minor comments

23. Title: consider removing or rewording physical activity

Response: To aid the interpretation of the term physical activity, as used in the title and throughout the manuscript, we have added a widely used definition to our introduction.

Manuscript: Physical activity, defined as bodily movements produced by skeletal muscles (Caspersen et al., 1985), is thought to positively affect brain characteristics

24. Abstract: please provide essential demographic information such as age, gender and N. Please provide basic information on statistical analyses. Names of regions or tracts where significant FA differences were found should be mentioned.

Response: The guidelines of Psychiatry Research allow an abstract of no more than 200 words with the request not to report statistics or p-values. We agree that the abstract lacked in detail so we now report our N and the fact that the groups were matched on age and gender. As the difference in FA between patients and controls is not the major focus of our paper, we feel that this is beyond the scope of the abstract.

Manuscript: Diffusion tension imaging (DTI) and 14-day actigraphy recordings were obtained in 51 BD I patients and 55 age-and-gender-matched healthy controls.

Reviewer 2

The authors investigate the association between white matter (DTI) and sleep and activity measures (14 days of actigraphy) in 51 BD I patients and 55 healthy controls. Physical activity was associated to increased integrity of white matter microstructure in both patients and controls. Increased integrity of white matter was also associated to mean physical activity levels and effective sleep in controls, but inversely related in BD I patients. This latter association could be explained by antipsychotic medication.

1. On page 2, at the end of the first paragraph, how is the term "sleep quality" defined in the sentence: "These findings suggest that sleep quality relates to white matter microstructure and that sleep disruptions could lead to reduced integrity of white matter"? Sleep quality or duration? In the abstract and highlights, the authors use the term "effective sleep". Consider deciding on one term and using it consistently, as sleep quality, sleep duration and sleep efficiency have different meanings.

Response: we thank the reviewer for pointing out this inconsistency. We decided to only use the term 'effective sleep' as a description of undisturbed sleep patterns instead of using 'sleep quality'. As sleep duration and sleep efficiency are both measurable sleep variables, we kept these terms in the manuscript.

Manuscript: *These findings suggest that effective sleep relates to white matter microstructure and that sleep disruptions could lead to reduced integrity of white matter.*

2. Also on page 2, when referring to the meta-analysis by Sexton et al, in whom were these findings of physical exercise leading to improvements in white matter microstructure made? Consider specifying this.

Response: This meta-analysis focused on the aging brain of healthy adults. We added this to our introduction.

Manuscript: Indeed, a meta-analysis of 29 studies on the association between white matter and physical activity in healthy older adults tentatively concluded that physical exercise leads to both global and local improvements in white matter (micro)structure (Sexton et al., 2015).

3. On page 3, paragraph "2.1 Sample": Could the authors describe the participants (both patients and controls) in greater detail? How were they recruited in the DBC study? What separated patients from controls?

Response: We have added a more detailed sample description to our manuscript. Table 2 now includes a description of the current mood symptomatology, cognitive functioning, level of education and BMI. Moreover, we added the following to our method section:

Manuscript: After completion of the DBC protocol, a subgroup of patients and controls were invited by telephone and mail to participate in the actigraphy and MRI measurements. (…)Inclusion criterion for patients was a diagnosis of bipolar disorder type I, in order to create a clinically homogeneous sample. Controls with a bipolar or psychotic diagnosis were excluded, as were control subjects who had a first or

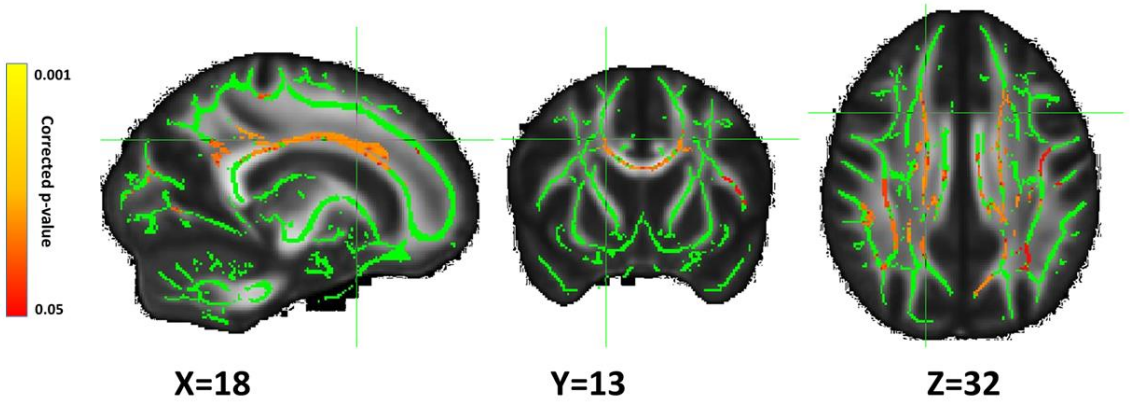
second degree relative with a bipolar or psychotic diagnosis. Any other type of psychiatric disorder was not an exclusion criterion in order to avoid the recruitment of a 'super-normal control sample', that is not representative of the general population.

4. On page 4, end of paragraph 2.2: day-to-day variations in sleep and activity have also been found to be a prominent feature of bipolar subjects in several actigraphy studies, as in the referenced Pagani study, where reduced interdaily stability (methods described by van Someren) was reported. Would it be possible to include one or more measures of variability to add interest to the association between white matter structure and actigraphy findings?

Response: We agree with the reviewer that interdaily stability (IS) is an interesting phenotype reflecting variability in the rest-activity rhythm and is worth studying in bipolar disorder (although the paper by Pagani and colleagues concludes that it is unrelated to the disorder). We now include IS in our method section and analyses.

Manuscript, methods: To measure the degree of day-to-day stability of 24-h activity rhythm, we calculated interdaily stability (IS), which is the ratio between the variance of the average 24-h pattern around the mean and the overall variance (Van Someren et al., 1996). IS ranges from 0-1, with higher values indicating more stable activity rhythms.

Manuscript, results: Also, a more stable 24-h activity rhythm was associated with increased FA.



Supplementary figure 1

5. On page 7, "3.1 Participants": What about BMI? Minor somatic disorders that did not qualify as an exclusion criterion? Comorbidity? Any information about such variables? In what state of bipolar disorder were patients upon actigraphy monitoring vs. MRI? Any mood rating scales available? The difference in time between actigraphy and DTI could preferably be presented more clearly earlier in the manuscript. And the results about medication could be easier to understand if presented in a table.

Response: we thank the reviewer for these suggestions. BMI has now been added to our table of sample characteristics as well as mood rating during the actigraphy measurements. Also, we described the time

difference between actigraphy and MRI acquisition more clearly.We did not rate mood during MRI acquisition or minor somatic disorders. We furthermore created a table for our medication results.

Manuscript, Method: The majority of participants underwent MRI scanning prior to the two-week actigraphy protocol, with the exception of 6 participants who started with actigraphy measurements. The mean time difference between MRI and actigraphy measurements was 1.4 years.

Table 1 – Sample characteristics

	Patients	**Controls**	F / χ^2	p-value
Age M (sd)	49.5 (11.4)	45.5 (15.8)	2.18	0.14
Gender M/F (% male)	28/23 (54.9%)	25/30 (45.5%)	0.44	0.22
Handedness R/L/B (% right)	42/6/3 (82.4%)	44/11/0 (80.0%)	4.37	0.11
IQ M (sd)	99.6 (15.7)	105.7 (14.4)	4.18	0.04*
Level of education n (%)			6.32	0.28
1. Low education	5 (9.3%)	4 (7.8%)		
2. Intermediate secondary education	4 (7.4%)	9 (17.6%)		
3. Intermediate professional education	7 (13.0%)	10 (19.6%)		
4. High prepatory vocational / pre-university	13 (24.1%)	6 (11.8%)		
5. Bachelor degree	11 (20.4%)	13 (25.5%)		
6. Master or PhD degree	14 (25.9%)	9 (17.6%)		
Inventory of Depressive Symptoms M (sd)	14.3 (10.1)	5.2 (3.9)	35.60	<0.001*
Altman Self-Rating Mania Scale M (sd)	1.9 (1.8)	1.7 (2.3)	0.41	0.53
Body Mass Index M (sd)	27.4 (4.3)	25.2 (3.3)	8.45	0.01*
Nr of workdays M (sd)	4.2 (3.5)	5.7 (3.7)	4.11	0.05*
Illness duration/ years M (sd)	16.5 (14.2)	-		
Number of episodes M (sd)	10.0 (10.8)	-		
Age at onset M (sd)	31.7 (12.2)	-		
History of psychotic symptoms N (%)	35 (70.0%)	-		

* Significant group difference at p<0.05

Table 4 – Correlation (*r*) between sleep parameters and psychotropic medication, and average FA of the significant clusters and psychotropic medication. P < 0.05 indicated with one asterisk, p<0.001 indicated with dubble asterisk.

	Lithium	*Other mood stabilizers*	*Benzodiazepines*	*Antipsychotics*
Sleep duration	0.10	0.12	0.07	0.47*
FA sleep duration	0.34 *	-0.33 *	-0.03	-0.31*
Sleep onset	-0.10	-0.04	0.13	-0.40**
FA sleep onset	0.28*	-0.15	0.20	-0.26
Sleep inertia	0.17	-0.14	0.01	-0.23
FA sleep inertia	0.38**	-0.34*	0.05	-0.17

Minor errors:

Response: All suggestions below have been included in the manuscript. We thank the reviewer for pointing out these typos.

6. Last paragraph, page 1: "…disturbances in sleep patterns are considered AS? core symptoms of bipolar mood episodes."

7. Page 2: "MechanismS of the link…"

8. Page 2: If oligodendrocyte proliferation and myelin-related gene expression are two separate processes, the verbs should be changed to "increase" and "reduce".

9. On page 8, Bonferroni is misspelled.

10. In the Figure 1 legend, "significantly correlations" should be changed to "significant correlations."

11. On page 15, 2nd paragraph of the discussion, "we extent" should be changed to "we extend."

Reviewer 3

The authors combine actigraphy and DTI to examine the relationships between sleep, activity and white matter measures in a study of 51 bipolar I patients and 55 healthy controls. They report that physical activity was related to white matter measures across both groups, that higher FA was associated with sleep in the expected direction in controls, but associated with sleep in the opposite direction in bipolar patients. I enjoyed reading the paper, and the findings will be of interest to the field. However, I have a number of major concerns:

Major

1. The sentence "The normalized standard deviation of the diffusivities is referred to as fractional anisotropy (FA) and ranges from 0 to 1, in which *higher values indicate faster diffusivity* (Assaf and Pasternak, 2008)" is misleading and needs to be reworded. Higher values do not necessarily indicate faster diffusivity. Rather, higher values indicate anisotropic diffusion.

Response: We thank the reviewer for pointing out this error in our manuscript. We have reworded the sentence in our manuscript to:

Manuscript: Water diffuses more easily along white matter tracts than perpendicular to them. This directional dependence is usually quantified with fractional anisotropy (FA), a scalar variable which ranges from 0 to 1, in which higher values indicate a higher directional dependence of diffusion.

2. Could the authors please add their main hypotheses to the final paragraph of the introduction.

Response: Yes, we extended the introduction with our hypotheses, although the difference in relation between patients and controls regarding the FA – Sleep/Activity associations was an explorative analysis.

Manuscript: We hypothesize that both sleep and physical activity are positively related to increased integrity of white matter, and we further explore whether these associations are similar for patients and controls.

3. Was there any cognitive characterization of participants, and can this be reported and considered? E.g. MMSE?

Response: We included a global measure of intelligence (IQ) to our table of sample characteristics (table 1)

4. There is currently no information regarding current symptoms, or their severity. As the authors state in their introduction – depressive and manic episodes typically differ in terms of physical activity, and

this variation could have a big impact on results. Were any measures taken, and can these be reported? If not, this should be added as a limitation of the study.

Response: Yes, we have also included details regarding current symptom level in our table of sample characteristics (table 1).

5. The control group included 1 participant currently receiving an antidepressant, and 1 participant receiving a benzodiazepine. In case-control studies, any psychiatric illness is often an exclusion criteria for control participants. Can analyses be repeated excluding these 2 participants to examine if they are clouding any differences between the groups?

Response: we did not exclude control subject with any kind of psychiatric diagnosis, but only excluded controls with a bipolar or psychotic disorders. We feel that excluding any kind of disorder leads to a unrepresentative sample of controls. We now state this more clearly in our method section:

Manuscript: Control subjects with a bipolar or psychotic diagnosis were excluded, as were control subjects who had a first or second degree relative with a bipolar or psychotic diagnosis. Any other type of psychiatric disorder was not an exclusion criteria in order to avoid the recruitment of a 'super-normal control sample', that is not representative of the general population.

We do agree with the reviewer that it is worth exploring whether our results are (mainly) driven by these two controls. We repeated our case-control comparison and found that our results were very similar (see table below). We therefore decided to report the original sample in our manuscript.

Analysis	**Cluster size (number of voxels)**	**Corrected p-value**	**MNI coordinates of center of gravity (x; y; z)**	**White matter tracts**
Significant group difference in FA after exclusion two control subjects who used benzodiazepines or antidepressants	46156	<0.001	-0.65; -21.2; 18.4	Genu of corpus callosum Body of corpus callosum Splenium of corpus callosum Fornix (column and body of fornix) Left Corticospinal tract Right Corticospinal tract Right Superior cerebellar peduncle Left Superior cerebellar peduncle Right Cerebral

				peduncle Left Cerebral peduncle Right Anterior limb of internal capsule Left Anterior limb of internal capsule Right Posterior limb of internal capsule Left Posterior limb of internal capsule Right Retrolenticular part of internal capsule Left Retrolenticular part of internal capsule Right Anterior corona radiata Left Anterior corona radiata Right Superior corona radiata Left Superior corona radiata Right Posterior corona radiata Left Posterior corona radiata L Right Posterior thalamic radiation (include optic radiation) Left Posterior thalamic radiation (include optic radiation) Right Sagittal stratum (include inferior longitidinal fasciculus and inferior fronto-occipital fasciculus) Left Sagittal stratum (include inferior longitidinal fasciculus and inferior fronto-occipital fasciculus) Right External capsule Left External capsule Right Cingulum

				(cingulate gyrus) Left Cingulum (cingulate gyrus) Right Cingulum (hippocampus) Right Fornix (cres) / Stria terminalis Left Fornix (cres) / Stria terminalis Right Superior longitudinal fasciculus Left Superior longitudinal fasciculus Right Superior fronto-occipital fasciculus Left Uncinate fasciculus Right Uncinate fasciculus Left Tapetum Right Tapetum

6. Group differences between bipolar and control participants in DTI metrics is currently not reported. Such differences (or lack of) would aid in the interpretation of the findings. Could these be reported, or if published elsewhere, referred to and discussed?

Response: We agree that a description of patient-control differences in FA would be of added value to our manuscript. We have included these in figure 1 and table 2 and our discussion:

Manuscript, Discussion: Bipolar patients were characterized by widespread reductions in fractional anisotropy (FA) in several major white matter tracts, including the corpus collosum, cerebral peduncle, corona radiata, internal and external capsule, thalamic radiation and superior longitudinal fasciculus.

7. I recommend that Section 3.2 is re-ordered so that it follows the same pattern as described in the methods – i.e. voxelwise results reported first.

Response: Yes, we agree that we should start section 3.2 with our voxelwise analyses:

*Manuscript: Voxelwise correlations in the overall sample (corrected for age and gender, and patient/control group) revealed positive associations between mean activity throughout the day, activity between noon and 6 PM, and between 6 PM and midnight (**Figure 1**).*

8. I am concerned about 'double-dipping' with the results (http://www.nature.com/neuro/journal/v12/n5/abs/nn.2303.html) - in that statistical analysis is currently run on regions significant in voxelwise analysis. This approach leads to inflated R and p values. The authors could instead use anatomically defined ROI (not just limited to significant regions).

Response: This is a valid point raised by the reviewer and we agree that the p-value and R2 from the regression analyses may not correspond with the previous TBSS results. We therefore decided to show the scatterplot for visualization purposes only and delete the test statistics. For our confounder analyses we did use the mean FA value, as this is a post-hoc analysis performed in the patient sample only. We did not have specific hypotheses regarding regions of interest so we preferred this method over ROI analyses.

9. Long sleep duration is not necessarily indicative of a more healthy sleep pattern. Long sleep has been associated with a wide range of poor health outcomes, including with increased WMH (Ramos et al 2014 - Sleep duration is associated with white matter hyperintensity volume in older adults: the Northern Manhattan Study). From the scatterplots, it appears that several participants in the bipolar I group have sleep duration of 9 or more hours, but not in the control group. This difference in distribution of sleep durations could underlie differences in relationships between groups. I would recommend the authors revise their analyses on sleep duration. For example, to be sensitive to inverse-U shaped relationships with sleep duration, where either long or short sleep is considered detrimental. Or, alternatively, classifying participants as short, normal or long sleepers.

Response: The reviewer suggests a very interesting perspective on our relationship between sleep duration and FA. We analyzed this possible U-shaped relation and found that squaring, or centering plus squaring the sleep duration variable yielded similar results:

Untransformed sleep duration: β=-0.36, p=0.005, R^2 =0.32
Squared sleep duration: β=-0.38, p=0.004, R^2 =0.32
Centered and squared sleep duration: β=-0.39, p=0.004, R^2 =0.32

There seems to be no U-shaped relation between sleep and FA, but this is possibly due to the fact that only 4 patients slept less than 6 hours. Next, we tested whether there is a difference in association between the 22 patients that could be classified as long sleepers (> 8 hours) vs 25 normal sleepers (6-8 hours). We found that the association between sleep duration and FA in 'normal sleepers' (β=-0.27, p=0.10, R2=0.53) was lower compared to long sleepers (β=-0.43, p=0.06, R^2 =0.21), but still in the same direction, indicating that the relation between sleep duration and FA is linear.

Minor

10. Could "somatic illness" be further defined please.

Response: Yes, the following has been added to our method section:

Manuscript: Exclusion criteria for all participants were sleep apnoea, pregnancy, head trauma and neurological disorders.

11. There is a typo in section 2.5 – "actighraphy"

Response: Thank you for pointing out this error.

12. Can the authors please include a justification for dividing physical activity into 6 h time periods?

Response: The overall mean activity level gives a very general indication of activity levels but includes the complete 24-h cycle. To give a more detailed picture regarding activity per part of day (i.e. morning, afternoon, evening, night), we divided the activity measurements in 6-h time windows.

Manuscript: The raw mean of activity was calculated over 24-hr period, followed by the mean of activity in four 6-hr periods corresponding to morning, afternoon, evening, night time: midnight to 6 AM (24-6 hr), 6 AM to noon (6 – 12 hr), noon to 6 PM (12 – 18 hr), and 6 PM to midnight (18 – 24 hr).

13. Can the section on exclusion criteria be re-organised slightly. It starts with exclusion criteria for all participants (fine), then additional criteria for control subjects (fine). It's not clear whether subsequent criteria (e.g. head trauma, neurological illness) apply to just control participants or both groups though.

Response: We agree that this section needs some modification. We now describe criteria for patients first, followed by criteria for controls and end with general criteria.

Manuscript: Inclusion criterion for patients was a diagnosis of bipolar disorder type I, in order to create a clinically homogeneous sample. Controls with a bipolar or psychotic diagnosis were excluded, as were control subjects who had a first or second degree relative with a bipolar or psychotic diagnosis. Any other type of psychiatric disorder was not an exclusion criterion in order to avoid the recruitment of a 'super-normal control sample', that is not representative of the general population. Exclusion criteria for all participants were sleep apnoea, pregnancy, head trauma and neurological disorders.

14. In table 2 and in figures, can units be given please.

Response: Yes, we have now improved our captions with the following statements:

Manuscript, Table 3: Sleep duration, sleep onset latency, sleep inertia and WASO are measured in minutes. Sleep onset and sleep offset describe minutes prior or past midnight. Sleep efficiency is a percentage ranging from 1-100. All activity variables are activity counts averaged per minute.

Manuscript, Figure 1: TBSS was used to assess the association of activity counts (per minute) with integrity of white matter microstructure.

*Manuscript, Figure 3: Sleep duration (measured in minutes) is negatively associated with FA in patients and positively associated with FA in controls (**A**). Sleep onset (minutes prior or post midnight) is positively associated with FA in patients and negatively associated with FA in controls (**B**). Similarly, sleep inertia (measured in minutes) is positively associated with FA in patients and negatively associated with FA in controls (**C**).*

15. Can a reference be given to support the claim that earlier onset of sleep is an indicator of a more healthy sleep pattern.

Response: We have added a reference that shows the association between a late chronotype ('evening type') and sleep complains, such as insufficient sleep, insomnia symptoms and nightmares.

Manuscript: A longer sleep duration, earlier onset of sleep and shorter sleep inertia are regarded as indicators of a more healthy sleep pattern (Merikanto et al., 2012).

16. With regard to "a longer sleep duration, earlier sleep onset and shorter sleep inertia were related to higher FA, which is consistent with previous reports (Bellesi et al., 2013; Kopp et al., 2006; Kreutzmann et al., 2015; Mirescu et al., 2006)", could the authors add in a clarification of how results are consistent, i.e. in supporting a role between sleep and brain measures, not links with FA directly

Response: Yes, we see how this statement could be improved. We now describe the previous studies that underscore our results in more detail.

Manuscript: In controls, a longer sleep duration, earlier sleep onset and shorter sleep inertia were related to higher FA, which is in line with previous studies reporting on the effects of sleep on cell proliferation, synaptic plasticity and neurogenesis (Bellesi et al., 2013; Kopp et al., 2006; Kreutzmann et al., 2015; Mirescu et al., 2006).

17. It may be useful to include a discussion of Yaffe et al 2016 (Sleep Duration and White Matter Quality in Middle-Aged Adults) – this paper may not have been published at the time of submission.

Response: This article was not yet available at the time of writing, but we agree that this is an excellent addition to our discussion.

Manuscript: *A recent study by Yaffe et al. (2016) found that short sleep duration (i.e. ≤ 6h) correlated with higher mean diffusivity and lower fractional anisotropy although the latter was no longer significant after adjustments for confounding variables.*

18. Typo "Here, we extent these findings"

Response: Thank you for pointing out this error, we have corrected the typo.

Manuscript: Here, we extend these findings, and show that physical activity and white matter integrity are similarly related in bipolar disorder patients.

19. Could a brief discussion of how medications used in the bipolar I group typically affect sleep patterns be added to the discussion.

Response: That would indeed be valuable information and has now been included in our discussion:

Manuscript: Several widely prescribed types of medication affect sleep-wake patterns; lithium normalizes the 24-hour circadian cycle, whereas benzodiazepines and antipsychotic medication can promote sleep onset and/or maintenance (Klemfuss, 1992; Monti and Monti, 2004).

1. Background

Bipolar disorder is a severe and recurrent psychiatric disorder, characterized by alternating depressive and manic mood episodes (Belmaker, 2004). Although the aetiology of bipolar disorder remains poorly understood, there is an extensive body of evidence suggesting that brain abnormalities are a core feature of the disorder. Previously, structural imaging studies found marked anatomical differences in grey matter volume between patients and healthy controls (Hallahan et al., 2011; Hibar et al., 2016). Moreover, volumes of several areas (i.e. amygdala, globus pallidus) were associated with polygenic risk scores for bipolar disorder and are affected in first-degree relatives (Caseras et al., 2015; Sandoval et al., 2014). In addition to volumetric measures, diffusion tensor imaging (DTI) studies have shown that bipolar disorder is also characterized by changes in white matter microstructure (Mahon et al., 2010). DTI quantifies the integrity and coherence of white matter tracts by measuring the diffusion of water parallel and perpendicular along white matter fibre tracts. Water diffuses more easily along white matter tracts than perpendicular to them. This directional dependence is usually quantified with fractional anisotropy (FA), a scalar variable which ranges from 0 to 1, in which higher values indicate a higher directional dependence of diffusion. More organized and myelinated axon will cause water molecules to diffuse in a particular direction rather than randomly in all directions. Therefore, FA is thought to reflect integrity of white matter with lower FA values pointing to sparse, poorly myelinated, or divergent fibers (Beaulieu, 2002). A voxel-based meta-analysis of 18 DTI studies found that bipolar patients have lower FA in the genu of the corpus callosum and left cingulum white matter, suggesting poor integrity of white matter (Wise et al., 2016). This finding has also been established in antipsychotic and mood-stabilizer naïve patients (Yip et al., 2013).

In addition to these brain changes, disturbances in sleep patterns are considered as core symptoms of bipolar mood episodes (American Psychiatric Association, 2013). The majority of patients experience a reduced need for sleep during manic phases of the disorder, and hypersomnia or insomnia during

depressive episodes (Harvey, 2008). Sleep has several properties which are crucial for normal brain functioning and integrity (Diekelmann and Born, 2010; Kang et al., 2009; Ooms et al., 2014; Tononi and Cirelli, 2006; Wang et al., 2011; Xie et al., 2013). In clinical studies of sleep disorders several neurophysiological alterations have been observed. For example, insomnia has been linked to reductions in orbitofrontal and parietal grey matter and larger rostral anterior cingulate cortex volumes (Altena et al., 2010; Winkelman et al., 2013), and narcolepsy and obstructive sleep apnoea are associated with widespread reductions in grey matter volume and thickness (Joo et al., 2011, 2010, 2009). Furthermore, animal studies have demonstrated that sleep deprivation interferes with long term synaptic potentiation and neurogenesis (Kopp et al., 2006; Kreutzmann et al., 2015; Mirescu et al., 2006). Mechanisms of the link between sleep and integrity of white matter remain largely unknown. Animal studies showed that oligodendrocyte proliferation and myelin-related gene expression strongly increase during sleep and reduce during wake time and experimental sleep deprivation (Bellesi et al., 2013). Conversely, chronic sleep loss has been associated with down-regulation of the expression of two myelin-related genes coding for plasmolipin and a membrane protein expressed in the mature myelin sheath (Cirelli et al., 2006). Human studies using DTI measures found that variability in sleep duration during week-days was associated with reduced FA in the parietocortical, frontocortical and frontostriatal tracts (Telzer et al., 2015). Furthermore, a sleep deprivation study in adult males found widespread changes in DTI measures after just one day of sleep deprivation (Elvsåshagen et al., 2015). These findings suggest that effective sleep relates to white matter microstructure and that sleep disruptions could lead to changes in oligodendrocytes, myelination, and axonal integrity and organisation.

Not only sleep, but also altered levels of physical activity are a hallmark of bipolar mood episodes: while depression is characterized by a marked reduction in overall physical activity, manic episodes stand out for their increased goal-directed behaviour and psychomotor agitation (American Psychiatric Association, 2013). Physical activity, defined as bodily movements produced by skeletal muscles

(Caspersen et al., 1985), is thought to positively affect brain characteristics; higher levels of physical activity and physical fitness have been associated with increased grey matter volumes, both cross-sectionnally, longitudinally, and after a physical exercise intervention (Bherer et al., 2013; Erickson et al., 2014). Indeed, a meta-analysis of 29 studies on the association between white matter and physical activity in healthy older adults tentatively concluded that physical exercise leads to both global and local improvements in white matter (micro)structure (Sexton et al., 2015). Possible underlying mechanisms involve the increased expression and release of brain-derived neurotropic factor and oxygen supply as a result of increased physical activity (Burzynska et al., 2014). Results from an exercise intervention study in schizophrenia patients indicated that after six months of training, the connectivity in white matter fibre tracts associated with motor functioning improved equally in the schizophrenia patients and a control group (Svatkova et al., 2015). Until now, similar studies in bipolar patients are absent.

The current study investigates the link between white matter microstructure and objectively measured physical activity and sleep patterns of bipolar patients and controls. Both physical activity and sleep are measured using actigraphy. We hypothesize that both sleep and physical activity are positively related to increased integrity of white matter, and we further explore whether these associations are similar for patients and controls.

2. Methods

2.1 Sample

A total of 106 participants (51 bipolar patients, 55 controls) were recruited from the Dutch Bipolar Cohort (DBC) study, which is a collaboration between the University Medical Center Utrecht (UMCU), various health care institutes in the Netherlands and the University of California Los Angeles (UCLA). The objective of the DBC study is to collect a deep phenotype characterization of bipolar I patients, their first-degree relatives and controls. After completion of the DBC protocol, a subgroup of patients and controls were invited by telephone and mail to participate in the actigraphy and MRI measurements.The medical ethical committee of the UMCU approved the DBC study and the actigraphy and MRI protocols. Written informed consent was obtained from all participants prior to participation. The majority of participants underwent MRI scanning prior to the two-week actigraphy protocol, with the exception of 6 participants who started with actigraphy measurements. The mean time difference between MRI and actigraphy measurements was 1.4 years.

All participants had a minimum age of 18 years old. Inclusion criterion for patients was a diagnosis of bipolar disorder type I, in order to create a clinically homogeneous sample. Controls with a bipolar or psychotic diagnosis were excluded, as were control subjects who had a first or second degree relative with a bipolar or psychotic diagnosis. Any other type of psychiatric disorder was not an exclusion criterion in order to avoid the recruitment of a 'super-normal control sample', that is not representative of the general population. Exclusion criteria for all participants were sleep apnoea, pregnancy, head trauma and neurological disorders. An independent radiologist evaluated all MRI scans. Participants with any clinical findings were excluded from further analyses. Patient diagnoses were verified using the Structured Clinical Interview for DSM-IV (SCID) and controls subjects were assessed using the Mini-International Neuropsychiatric Interview (M.I.N.I) (First et al., 1997; Sheehan et al., 1998).

2.2 Actigraphy recordings and analyses

Circadian rhythmicity and sleep-wake measurements were recorded with an Actiwatch (the Actiwatch 2; Philips Respironics Inc, Murrysville, PA, USA). The Actiwatch has a solid-state piezo-electric accelerometer and a lithium rechargeable battery. It records wrist movements and the sum of wrist movements is scored in epochs of 1 minute. All participants were instructed to wear the Actiwatch for a period of 14 consecutive days on their non-dominant wrist and only to remove it when exposed to water for long periods of time (e.g. swimming). During the 14-day recording period, participants kept a sleep diary in which bed times, nap times and off-wrist periods were noted.

To calculate the sleep and activity patterns, a series of algorithms in R statistical package (R Development Core Team, 2014) were used according to the procedure developed by Pagani et al. (2015). These algorithms resulted in 7 sleep measures: sleep duration, timing of sleep onset (i.e. minutes prior or after midnight), timing of sleep offset (i.e. minutes after midnight), sleep onset latency, sleep efficiency (i.e. minutes asleep divided by minutes in bed), wake after sleep onset (WASO) and sleep inertia (i.e. wake interval between sleep offset and time out of bed). The raw mean of activity was calculated over 24-hr period, followed by the mean of activity in four 6-hr periods corresponding to morning, afternoon, evening, night time: midnight to 6 AM (24-6 hr), 6 AM to noon (6 – 12 hr), noon to 6 PM (12 – 18 hr), and 6 PM to midnight (18 – 24 hr). To measure the degree of day-to-day stability of 24-h activity rhythm, we calculated interdaily stability (IS), which is the ratio between the variance of the average 24-h pattern around the mean and the overall variance (Van Someren et al., 1996). IS ranges from 0-1, with higher values indicating more stable activity rhythms.

2.3 MRI acquisition

Structural Magnetic Resonance Images were acquired on a 3 Tesla Philips Achieva scanner (Philips

Healthcare, Best, the Netherlands), equipped with an 8-channel SENSE-headcoil. Fast field echo scans with 200 contiguous sagittal slices (TE=4.6 ms, TR=10 ms, flip angle=8˚, FOV=240 mm, 0.75 x 0.75 x 0.80 mm^3 voxels) were obtained. Processing was done on the neuroimaging computer network of the University Medical Center Utrecht - Brain Center Rudolf Magnus, Utrecht, the Netherlands. All MRI analyses were conducted using the FMRIB Software Library (FSL v5.0) (Jenkinson et al., 2012). Diffusion weighted images were pre-processed using FMRIB's Diffusion Toolbox (FDT). First, *topup* was used to correct for susceptibility induced distortions (distortions caused by the magnetic susceptibility inhomogeneities in the subject's head), using two non-diffusion weighted (b-value=0) images with opposite phase-encoding directions (anterior to posterior and opposite), which thus have distortions going in opposite directions. The susceptibility-induced off-resonance field (distortions caused by inhomogeneities to magnetic susceptibility of the subject's head) was estimated using a method similar to that described in Andersson et al. (Andersson et al., 2003) and the two images were combined into a single corrected one. Next, *eddy* was used to correct for eddy-current distortion and head movements, and *bet* (brain-extraction tool) was used to exclude any non-brain tissue (Smith, 2002). Finally, *dtifit* was used to fit a diffusion tensor model at each voxel.

2.4 TBSS analysis

Tract-based spatial statistics (TBSS) was used for voxelwise comparison of fractional anisotropy (FA), between subjects (Smith et al., 2007, 2006). Compared to the other DTI metric (i.e. axial diffusivity, radial diffusivity and mean diffusivity), FA is the most well-studied diffusion parameter and correlates well with microscopic changes in white matter such as myelination, axonal diameter, density and organization (Beaulieu, 2009). First, the FA maps of all subjects were non-linearly aligned to a standard FA template (FMRIB55_FA; the average of 55 good quality healthy subjects) after which the average was computed. The resulting average FA map was thinned to create a white matter 'skeleton', representing

the centre of white matter tracts common to all subjects. This white matter skeleton was thresholded at FA>0.2 to exclude non-white matter and voxels in extremities where there is too much cross-subject variability in alignment, resulting in a white matter skeleton of 141,230 voxels. Next, each subject's FA data was projected onto the white matter skeleton; for each subject the highest FA value perpendicular to each voxel of the skeleton (i.e. the individual's local white matter tract centre) was projected onto the white matter skeleton.

2.5 Statistical analyses

Whole-brain voxelwise analyses were carried out using the 'randomise' procedure, a tool for permutation-based non-parametric testing using the general linear model (Nichols and Holmes, 2002). All statistical tests were performed with 5000 permutations. Our statistical threshold was a threshold-free cluster enhancement (TFCE) and family-wise error corrected P-value of 0.05. (Smith and Nichols, 2009). First, we investigated the difference in FA between patients and controls for each voxel. The MNI coordinates of the centre of gravity of significant clusters was calculated and the anatomical regions of these clusters were identified using the 'JHU ICBM-DTI-81 White-Matter Labels' atlas, similar to Liu et al. (2016). Next, the we analysed the correlation between actigraphy measures and FA for each voxel, in the overall sample of patients and healthy controls. Furthermore, we investigated whether actigraphy measures were differently associated with DTI measures in patients and controls, using an interaction term for Group*Sleep/Activity variable. In case of significant voxels in the overall sample, we created scatterplots to visualize the association between average FA of the significant clusters and actigraphy measures in IBM SPSS Statistics 21.0. For variables with significant interaction effects, we created stratified scatterplots to determine the direction of association between actigraphy measures and the average FA of significant clusters. We analysed the confounding effect of illness-specific variables (number of episodes, history of psychotic symptoms, illness duration and use of psychotropic

medication) by testing the correlation with average FA of the significant cluster and the actigraphy measures. In case these variables showed significant associations with both FA and actigraphy measures, they were considered as potential confounders and separately entered as covariates in the stratified patient regression analyses, using the Enter method, with all variables entered in one block.

3. Results

3.1 Participants

A total of 106 participants were included (51 patients, 55 controls). See table 1 for sample characteristics. No differences between patients and controls were found for age, gender and handedness. Patients worked on average fewer days during the actigraphy measurement period and had higher BMI levels as compared with controls. Although none of the patients reported being in a mood episode, patients reported significantly more current depressive symptoms.

Table 1 – Sample characteristics

	Patients	**Controls**	F / χ^2	p-value
Age M (sd)	49.5 (11.4)	45.5 (15.8)	2.18	0.14
Gender M/F (% male)	28/23 (54.9%)	25/30 (45.5%)	0.44	0.22
Handedness R/L/B (% right)	42/6/3 (82.4%)	44/11/0 (80.0%)	4.37	0.11
IQ M (sd)	99.6 (15.7)	105.7 (14.4)	4.18	0.04*
Level of education n (%)			6.32	0.28
1. Low education	5 (9.3%)	4 (7.8%)		
2. Intermediate secondary education	4 (7.4%)	9 (17.6%)		
3. Intermediate professional education	7 (13.0%)	10 (19.6%)		
4. High prepatory vocational / pre-university	13 (24.1%)	6 (11.8%)		
5. Bachelor degree	11 (20.4%)	13 (25.5%)		
6. Master or PhD degree	14 (25.9%)	9 (17.6%)		
Inventory of Depressive Symptoms M (sd)	14.3 (10.1)	5.2 (3.9)	35.60	<0.001*
Altman Self-Rating Mania Scale M (sd)	1.9 (1.8)	1.7 (2.3)	0.41	0.53
Body Mass Index M (sd)	27.4 (4.3)	25.2 (3.3)	8.45	0.01*
Nr of workdays M (sd)	4.2 (3.5)	5.7 (3.7)	4.11	0.05*
Illness duration/ years M (sd)	16.5 (14.2)	-		
Number of episodes M (sd)	10.0 (10.8)	-		
Age at onset M (sd)	31.7 (12.2)	-		
History of psychotic symptoms N (%)	35 (70.0%)	-		

* Significant group difference at p<0.05

With the exception of 1, all patients used between 1-4 types of psychotropic medication during the actigraphy measurements; 24 patients used lithium and 17 patients used other mood stabilizing agents. Furthermore, 15 patients received antipsychotic medication, 9 patients used antidepressants and 19 patients used benzodiazepines. Of the control subjects, 1 person received an antidepressant and 1 a benzodiazepine.

Patients showed widespread reductions in FA compared to controls (**Figure 1** and **Table 2**). These differences were located in the corpus collosum, cerebral peduncle, internal capsule, corona radiate, thalamic radiation, external capsule and the posterior limb of internal capsule.

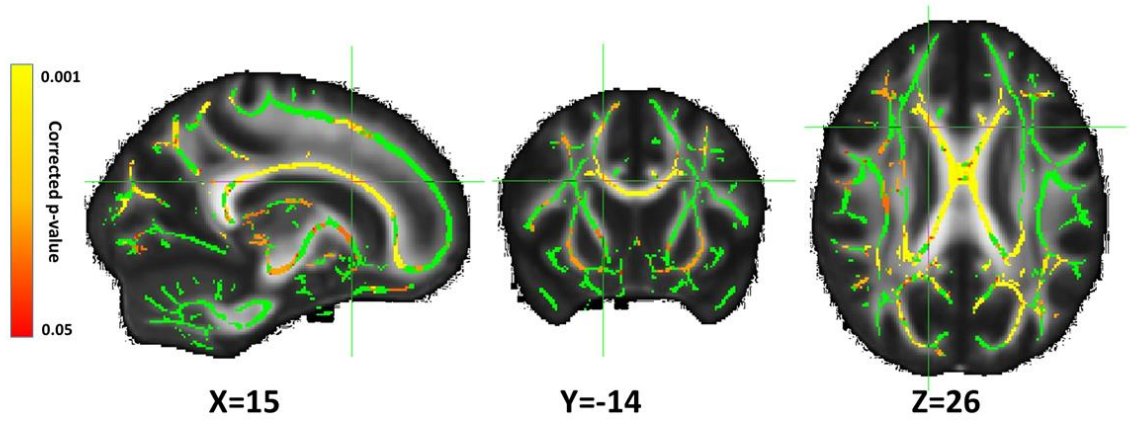


Figure 1 – TBSS was used to assess the difference in FA between patients and controls. Significant correlations are shown in yellow to red, depending on the significance level. The white matter skeleton is shown in green and overlaid on the MNI152 template brain. Images are shown in radiological convention (the left is displayed on the right) and corresponding Montreal Neurological Institute (MNI) coordinates are displayed below. P-values are TFCE and FWE corrected for multiple comparisons across space, and corrected for age and gender.

Table 2 - Cluster size (number of voxels), p-value, t-value, MNI coordinates of center of gravity and anatomical location of significant differences in FA between patients and controls

Analysis	**Cluster size**	**Corrected p-value**	**T-value**	**MNI coordinates (x; y; z)**	**Anatomical region**
Significant group difference in FA between patients and controls	44232	<0.001	3.08	2.37; -21.4; 16.8	Genu of corpus callosum Body of corpus callosum Splenium of corpus callosum R cerebral peduncle L cerebral peduncle R posterior limb of internal capsule L posterior limb of internal capsule L retrolenticular part of internal capsule R anterior corona radiata L anterior corona radiata R superior corona radiata Lsuperior corona radiata Right posterior corona radiata L posterior corona radiata R posterior thalamic radiation (include optic radiation) L posterior thalamic radiation (include optic radiation) R external capsule L external capsule L cingulum (cingulate gyrus) R superior longitudinal fasciculus L superior longitudinal fasciculus

Table 3 – Difference between sleep and activity measures for patients and controls

	Patients	**Controls**	p-value
Sleep duration M (sd)	475.2 (86.3)	447.2 (40.4)	0.01
Sleep onset M (sd)	-0.50 (75.3)	17.6 (73.0)	0.22
Sleep onset latency M (sd)	5.0 (5.7)	4.9 (5.2)	0.89
Sleep offset M (sd)	474.7 (88.5)	464.8 (74.8)	0.33
Sleep intertia M (sd)	6.8 (7.2)	5.4 (6.3)	0.37
Wake after sleep onset M (sd)	59.6 (27.0)	48.5 (19.5)	0.03
Sleep efficiency M (sd)	84.8 (8.1)	87.3 (4.9)	0.12
Mean Activity M (sd)	204.7 (89.3)	234.4 (54.5)	0.08
Activity 0 to 6 h M (sd)	42.1 (36.6)	45.0 (41.5)	0.82
Activity 6 to 12 h M (sd)	227.3 (127.8)	260.8 (91.6)	0.12
Activity 12 to 18 h M (sd)	333.4 (143.1)	361.8 (93.9)	0.34
Activity 18 to 24 h M (sd)	215.6 (109.6)	270.4 (84.9)	0.02

Means (M) and standard deviations (sd) on sleep and activity measures. Models are corrected for age and gender. Sleep duration, sleep onset latency, sleep inertia and WASO are measured in minutes. Sleep onset and sleep offset describe minutes prior or past midnight. Sleep efficiency is a percentage ranging from 1-100. All activity variables are activity counts averaged per minute.

Mean scores and standard deviations on the 7 sleep measures and 5 activity measures are shown in table 3. Patients had on average a longer sleep duration ($F[1,102]=6.6$, p=0.01, a longer wake time after sleep onset ($F[1,102]=5.1$, p=0.03) and lower activity counts between 6 PM and midnight ($F[1,102]=6.1$, p=0.02). However, after corrections for multiple testing, none of the group difference remained significant at a bonferroni-corrected significance level of 0.004.

3.2 Activity measures

Voxelwise correlations in the overall sample (corrected for age and gender, and patient/control group) revealed positive associations between mean activity throughout the day, activity between noon and 6 PM, and between 6 PM and midnight (**Figure 1**). Also, a more stable 24-h activity rhythm was associated with increased FA (**Supplementary figure 1)**. These correlations were found in largely overlapping regions of the brain, most of all in the genu and body of the corpus callosum and the right anterior corona radiate (**Table 2**). There were no significant interactions between group and activity measures with any of the DTI measures, suggesting that the association between activity measures and DTI measures was not significantly different for patients and controls.

3.3. Sleep measures

Voxelwise correlations in the overall sample (corrected for age, gender, and patient/control group) revealed no significant correlations between sleep measures and FA. Analyses of the interaction between group and sleep on FA revealed that the association of three sleep measures (i.e. sleep duration, sleep onset, and sleep inertia) with FA differed between patients and healthy controls in several regions of the brain (**Figure 2**). These interaction effects of sleep duration and sleep inertia with FA were primarily localized in the genu, body and splenium of the corpus collosum and the bilateral corona radiata (**Table 2**). The interaction effect of sleep onset with group on FA was localized in a small cluster in the left superior longitudinal fasciculus. After stratifying for patients and controls, we found that mean FA of the voxels with a significant interaction effect was positively correlated with sleep duration in controls, while mean FA was negatively associated with sleep duration in patients (**Figure 3A**). Similar contrasting relations were found for sleep onset and sleep inertia, with positive associations in patients and negative associations in controls (**Figure 3B and 3C**). A longer sleep duration, earlier

onset of sleep and shorter sleep inertia are regarded as indicators of a more healthy sleep pattern (Merikanto et al., 2012).

3.4 Confounder analyses

Stratified analyses of the sleep parameters in patients only allowed us to examine the possible confounding effects of medication and illness characteristics by analyzing the association with both sleep parameter and white matter integrity outcomes. First, we analyzed the association between the sleep parameters with average FA of the significant clusters resulting from the TBSS analysis (**Table 4**). Sleep onset and average FA were significantly associated with the number of episodes (**Table 5**). History of psychotic symptoms was also significantly associated with sleep onset and average FA (**Table 5**). However, adding number of episodes and history of psychotic symptoms as covariate to the association between sleep onset and average FA did not alter the results (**Supplementary table 1**). Antipsychotic drug use was associated with sleep duration and average FA (**Table 5**). Subsequent addition of antipsychotic drug use as covariate to the model rendered the previous association between sleep duration and average FA in patients non-significant (β = -0.26, t=-1.64, p=0.11). When the TBSS analysis for sleep duration was restricted to patients not using antipsychotic medication (N=25 patients and 55 controls), the interaction effect between sleep duration and group was no longer significant providing further support for a confounding effect of medication. The interaction effect between sleep duration and group remained significant when the TBSS analysis was performed in patients using antipsychotic medication, albeit in a smaller area (**Supplementary figure 2**).

Table 4 - Cluster size (number of voxels), p-value, t-value, MNI coordinates of center of gravity and anatomical location of significant associations between sleep and activity parameters with FA

Analysis	Cluster size	Corrected p-value	T-value	MNI coordinates (x; y; z)	Anatomical region
*Significant Group*Sleep duration interaction with FA*	651	0.04	3.74	10.2; 20.5; 19.5	Genu of corpus callosum
	16	0.05	2.34	-6.25; 12.7; 22	Body of corpus callosum
*Significant Group*Sleep onset interaction with FA*	292	0.039	4.36	37.6; -11.9; 29.4	R superior longitudinal fasciculus
	266	0.038	4.85	-37; -15.6; 27.2	L superior longitudinal fasciculus
	134	0.047	3.39	2.29; 24.9; 13.2	Genu of corpus callosum
	75	0.046	3.78	18.1; -51.6; 30.4	Splenium of corpus callosum R posterior corona radiata
	45	0.047	3.63	12.2; 37.9; 21	Splenium of corpus callosum
	33	0.047	4.96	38; -40.5; 24.8	R superior longitudinal fasciculus
*Significant Group*Sleep inertia interaction with FA*	21826	0.006	3.46	12.2; -13.3; 17.2	Genu of corpus callosum Body of corpus callosum Splenium of corpus callosum
	227	0.047	3.10	-11.6; 8.42; -7.77	Cerebral peduncle Posterior limb of left internal capsule
	202	0.049	2.19	8.01; -26; -28.8	R corticospinal tract Pontine crossing tract R cerebral peduncle
	162	0.046	3.96	-7.18; -28.3; -30.7	L corticospinal tract Pontine crossing tract L cerebral peduncle
	104	0.049	2.63	-29.4; -23.4; -1.7	Retrolenticular part of left internal capsule Sagittal stratum Fornix (cres) / Stria terminalis
	78	0.049	2.58	-37.9; -47.6; 1.23	Posterior thalamic radiation Sagittal stratum
	38	0.05	3.57	-35.1; -58.4; -5.16	Posterior thalamic radiation Sagittal stratum
	15	0.05	1.93	-40.6; -25.7; -4.6	Sagittal stratum
	13	0.05	2.39	-32; -203; 0.769	L external capsule
Significant correlation of mean activity with	3153	0.019	2.84	9.01; 19.2; 18.7	Genu of corpus callosum Body of corpus callosum R anterior corona radiata

FA					
Significant correlation of activity from 12h to 18h with FA	2466	0.028	3.24	7.37; 17; 19.7	Genu of corpus callosum Body of corpus callosum R anterior corona radiata
Significant correlation of activity from 18h to 24h with FA	333	0.044	3.46	3.64; 25.7; 10.9	Genu of corpus callosum Body of corpus callosum

Table 5 – Association between average FA and sleep parameters, controlled for age and gender

	β	**R^2**	**p-value**	**N**
Sleep duration	-0.39	0.32	0.004	51
Sleep onset	0.49	0.37	<0.001	51
Sleep inertia	0.34	0.33	0.008	51

Table 6 – Correlation (*r*) between sleep parameters and psychotropic medication, and average FA of the significant clusters and psychotropic medication. P < 0.05 indicated with one asterisk, p<0.001 indicated with double asterisk.

	Lithium	*Other mood stabilizers*	*Benzodiazepines*	*Antipsychotics*
Sleep duration	0.10	0.12	0.07	0.47*
FA sleep duration	0.34 *	-0.33 *	-0.03	-0.31*
Sleep onset	-0.10	-0.04	0.13	-0.40**
FA sleep onset	0.28*	-0.15	0.20	-0.26
Sleep inertia	0.17	-0.14	0.01	-0.23
FA sleep inertia	0.38**	-0.34*	0.05	-0.17

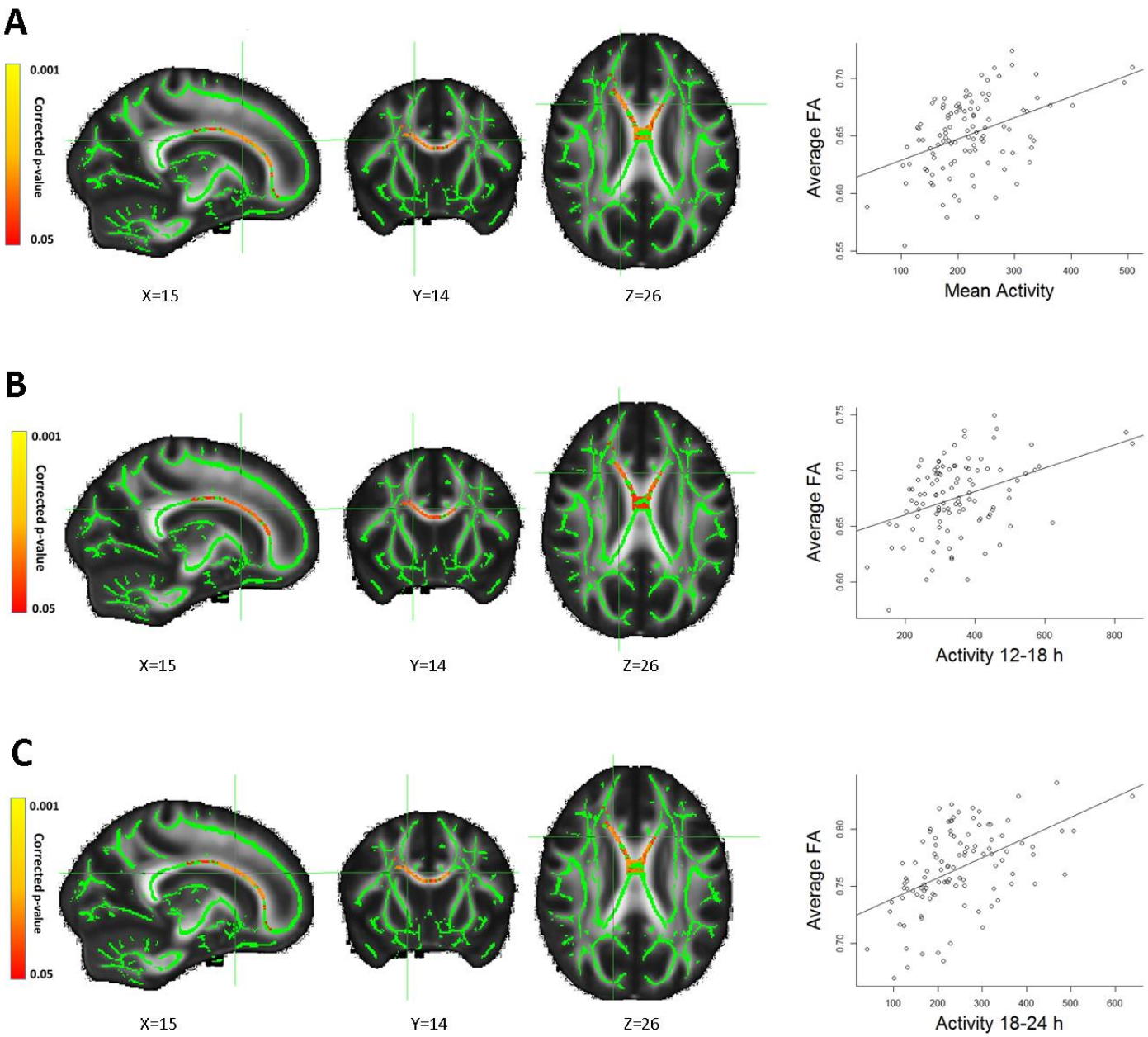


Figure 2: Daytime activity is positively associated with integrity of white matter microstructure. TBSS was used to assess the association of activity counts (per minute) with integrity of white matter microstructure. In the combined sample (patients and healthy controls) we found significant positive correlations of mean activity of the whole day (**A**), activity between noon and 6PM (**B**) and activity between 6PM and midnight (**C**). Significant correlations are shown in yellow to red, depending on the significance level. The white matter skeleton is shown in green and overlaid on the MNI152 template brain. Images are shown in radiological convention (the left is displayed on the right) and corresponding Montreal Neurological Institute (MNI) coordinates are displayed below. P-values are TFCE and FWE corrected for multiple comparisons across space and corrected for age, gender and (patient/control) group.

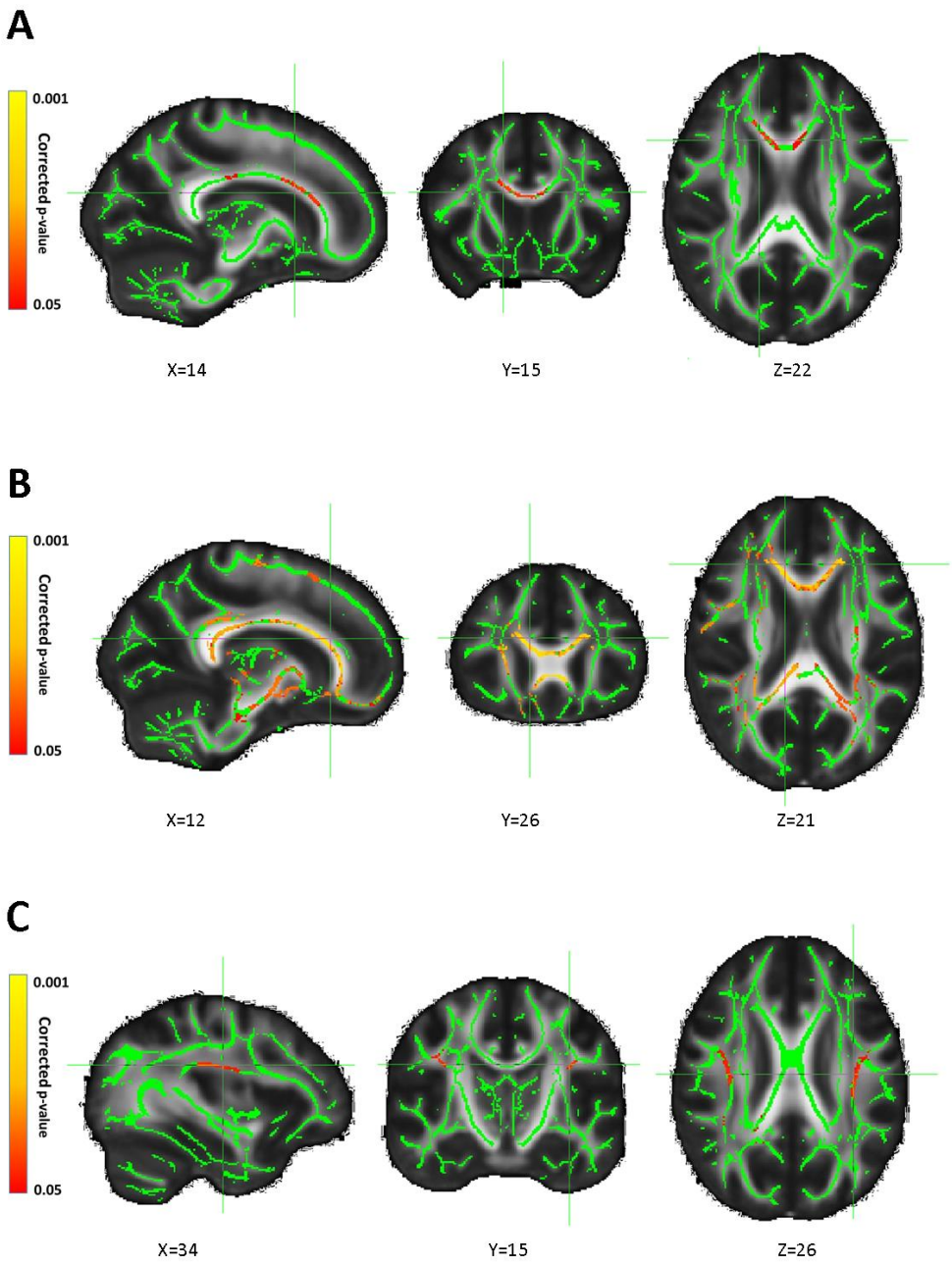


Figure 3: The association of sleep measures with integrity of white matter microstructure differs between patients and controls. TBSS was used to assess the interaction effects of sleep variables and group (Group * Sleep) with integrity of white matter microstructure. There is a significant negative interaction between sleep duration*group and FA (**A**), a positive interaction between sleeponset*group and FA (**B**), a positive interaction between sleep inertia and FA (**C**). Positive interaction effects are defined as interactions where the slope between sleep variables and DTI measures is higher for patients than for controls. Significant correlations are shown in yellow to red, depending on the significance level. The white matter skeleton is shown in green and overlaid on the MNI152 template brain. Images are shown in radiological convention (the left is displayed on the right) and corresponding Montreal Neurological Institute (MNI) coordinates are displayed below. P-values are TFCE and FWE corrected for multiple comparisons across space, and corrected for age and gender.

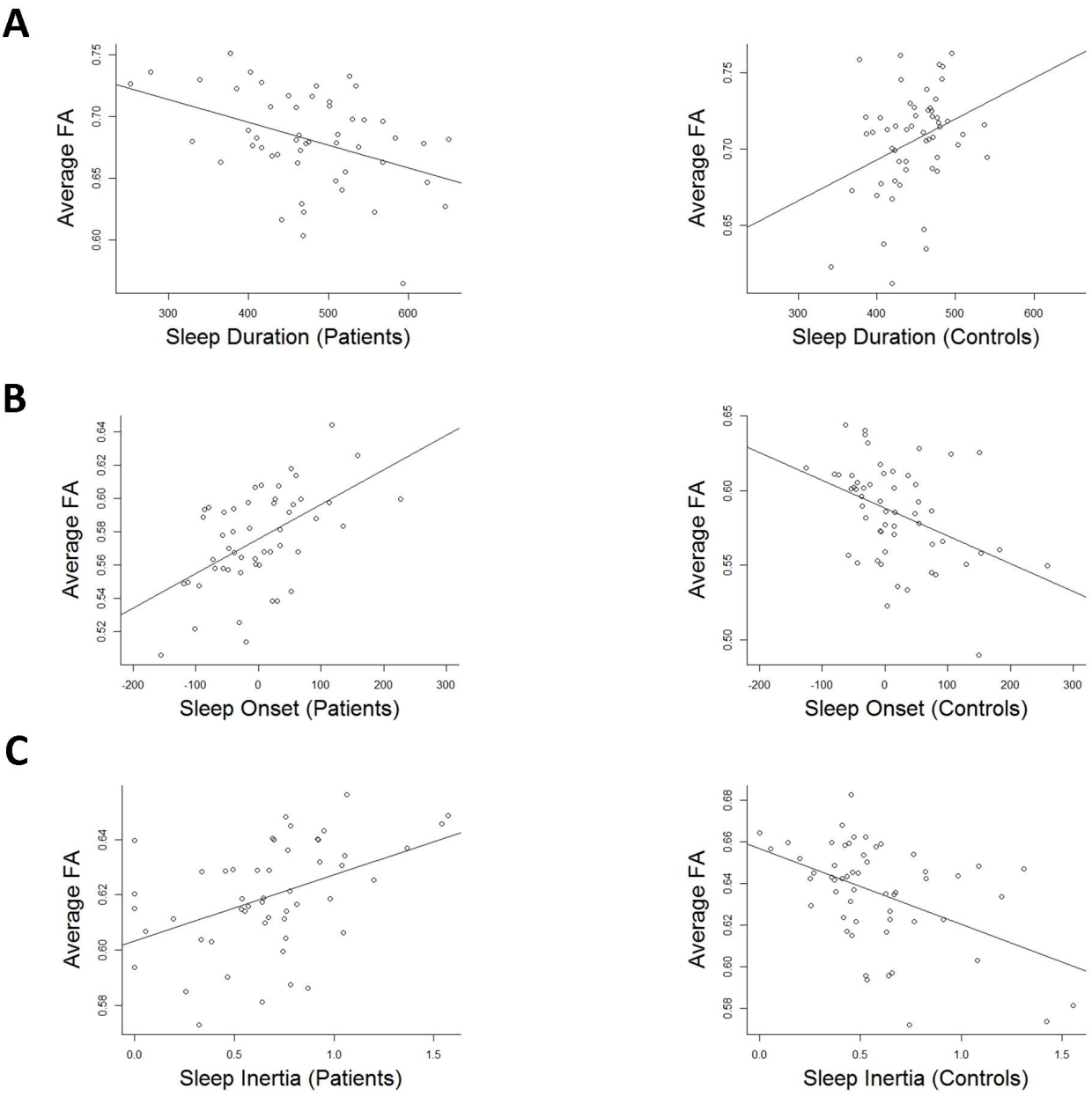


Figure 4: Sleep measures are associated with integrity of white matter microstructure within the patient group as well as within the control group. Scatterplots showing the association of sleep measures with average FA of the voxels with a significant interaction effect. Sleep duration (measured in minutes) is negatively associated with FA in patients and positively associated with FA in controls (**A**). Sleep onset (minutes prior or post midnight) is positively associated with FA in patients and negatively associated with FA in controls (**B**). Similarly, sleep inertia (measured in minutes) is positively associated with FA in patients and negatively associated with FA in controls (**C**).

4. Discussion

The current study investigated the association between sleep and physical activity patterns and integrity of white matter microstructure in bipolar disorder. Bipolar patients were characterized by widespread reductions in fractional anisotropy (FA) in several major white matter tracts, including the corpus collosum, cerebral peduncle, corona radiata, internal and external capsule, thalamic radiation and superior longitudinal fasciculus. Furthermore, we found that daytime and evening physical activity correlated with higher FA in bipolar patients and controls. In controls, a longer sleep duration, earlier sleep onset and shorter sleep inertia were related to higher FA, which is in line with previous studies reporting on the effects of sleep on cell proliferation, synaptic plasticity and neurogenesis (Bellesi et al., 2013; Kopp et al., 2006; Kreutzmann et al., 2015; Mirescu et al., 2006). However, several measures of sleep showed opposing associations with integrity of white matter in patients. These differences between patients and controls are most likely the result of confounding factors, such as use of antipsychotic drugs in the patient group.

Our results are in keeping with evidence suggesting that physical activity is positively related to measures of white matter integrity in healthy subjects (Sexton et al., 2015; Voelcker-Rehage and Niemann, 2013). This may be partly explained by an up-regulation of the expression of brain-derived neurotropic factor (BDNF) by physical activity. This neurotrophin has consistently been related to increases in gray matter volume and evidence also points to a relation with myelinogenesis and white matter microstructure (Du et al., 2003; Tost et al., 2013). Alternatively, higher levels of cardiorespiratory fitness lead to improvements in the cerebral blood flow, resulting in increased brain volume and white matter integrity (Chen et al., 2013; Zhu et al., 2015). Here, we extend these findings, and show that physical activity and white matter integrity are similarly related in bipolar disorder patients. This is relevant as integrity of white matter is hypothesized to be a factor in the etiology of bipolar disorder (Heng et al., 2010; Nortje et al., 2013). It raises the question whether symptoms and disease progression

of bipolar disorder can be ameliorated by improving integrity of white matter through physical fitness therapy or behavioral activation. Although the effect of such therapies has not yet been studied in a bipolar sample, a first study in schizophrenia patients found that an exercise intervention increased the level of white matter integrity (Svatkova et al., 2015).

A longer sleep duration, earlier onset of sleep and shorter sleep inertia are all indicative of a more healthy sleep pattern, so the finding that this relates to better integrity of white matter in controls confirms the hypothesis that sleep and white matter integrity are positively associated. Previous studies that measured the link between sleep and human brain function concluded that sleep is associated with synaptic homeostasis, clearance of amyloid beta, gray matter volume and cortical thickness (Altena et al., 2010; Diekelmann and Born, 2010; Joo et al., 2011, 2010, 2009; Kang et al., 2009; Ooms et al., 2014; Tononi and Cirelli, 2006; Wang et al., 2011; Winkelman et al., 2013; Xie et al., 2013). Our study is one of the first to expand these findings to better integrity of white matter microstructure. A recent study by Yaffe et al. (2016) found that short sleep duration (i.e. ≤ 6h) correlated with higher mean diffusivity and lower fractional anisotropy although the latter was no longer significant after adjustments for confounding variables. As a clinical study previously demonstrated, sleep positively affects both the number and proliferation rate of olygodendrocytes, while spontaneous and forced wake states reduced oligodendrocyte proliferation by half (Bellesi et al., 2013). Longer sleep duration, earlier sleep onset and shorter sleep inertia could therefore increase oligodendrocyte proliferation and myelination, ultimately leading to increased integrity of white matter. In our data bipolar patients show a reversed association between sleep and integrity of white matter. We show that a likely explanation for this counterintuitive association is the use of psychotropic medication. Several widely prescribed types of medication affect sleep-wake patterns; lithium normalizes the 24-hour circadian cycle, whereas benzodiazepines and antipsychotic medication can promote sleep onset and/or maintenance (Klemfuss, 1992; Monti and Monti, 2004). We found that in bipolar patients who received antipsychotic medication longer sleep

duration was associated with lower FA. A previous study already reported that chronic use of antipsychotic medication without mood stabilizing effect decreases myelin/oligodendrocyte related gene expression in white matter (Narayan et al., 2007). Moreover, a recent study suggests that sleep may be influenced by antipsychotics via mTORC1- proteins synthesis that in turn is involved in neuronal function, but further studies are required (Bowling et al., 2014). Our findings do warrant further study of the effect per medication type on sleep and white matter microstructure. Randomized clinical trials studying the effect of medication on white matter microstructure could address this by incorporating sleep measures. We found no relationship with other types of medication, but given that the majority of our bipolar sample used more than one type of medication, we cannot reliably distinguish between the effects of medication types (Brambilla et al., 2009). Also, confounding by indication may be at play here, whereby more severely affected patients may receive particular types of medication.

A further limitation of the current study is the average time difference between the MRI acquisition and actigraphy measurement of 1.4 years. Changes in exercise behavior and sleep patterns, might have resulted in variation in white matter microstructure that were now left undetected. However, the time between MRI and actigraphy measurements was similar for patients and controls, and there is no reason to assume that such variation was systematic and has biased the results. Also, we need to point out that we did not exclude participants diagnosed with conditions other than sleep apnoea, head trauma or neurological disorders.

In conclusion, the current study found that higher levels of physical activity were related to better integrity of white matter, regardless of bipolar diagnosis. In controls, less disturbed sleep was associated with better integrity of white matter, but opposing associations were found in bipolar patients, most likely due to extraneous variables such as use of antipsychotic medication.

Conflict of interest

Dr. van Haren reports personal fees for educational activities from Eli Lilly and Janssen. All other authors declare that they have no competing interests.

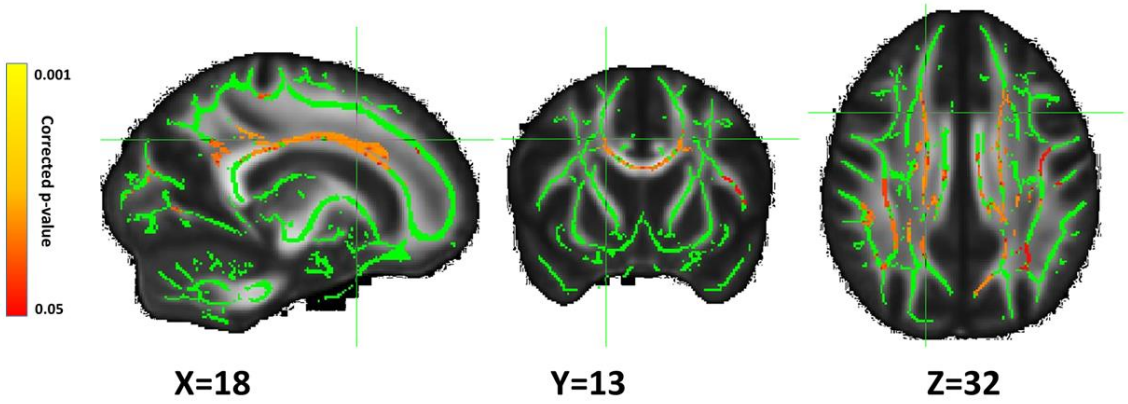
Acknowledgements

The authors would like to thank dr. René Mandl for his input during the process of data analysis. This study was funded by the National Institute of Mental Health [Grant number: R01 MH090553 to RAO]. The National Institute of Mental Health had no role in the design and conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

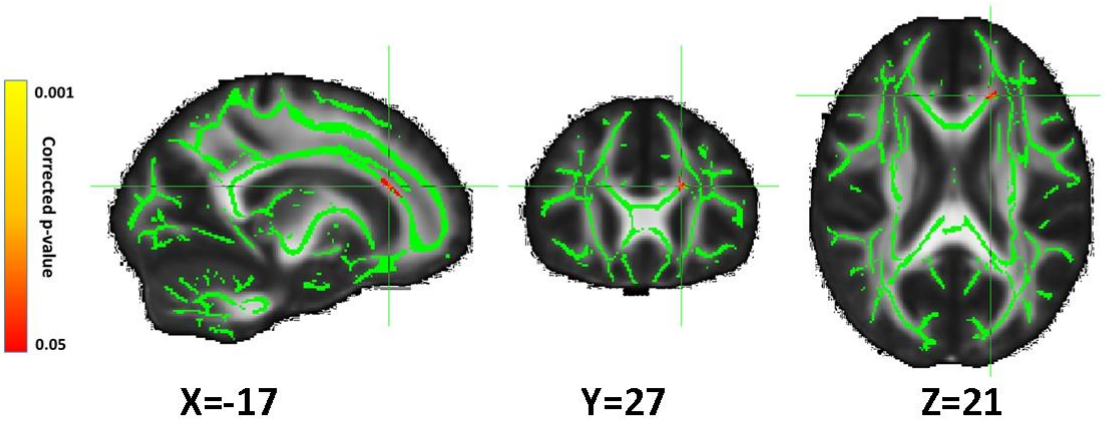
Supplementary material

Suppl. Table 1- Association between sleep onset and average FA while controlling for history of psychotic symptoms or number of episodes

	Beta	**P-value**	**R^2**	N
Model + psychotic symptoms	0.41	0.002	0.35	50
Model + number of episodes	0.35	0.02	0.40	41



Supplementary figure 1 - TBSS was used to assess the association of interdaily stability with integrity of white matter microstructure. We found positive significant correlations in the combined sample (patients and healthy controls) which are shown in yellow to red, depending on the significance level. The white matter skeleton is shown in green and overlaid on the MNI152 template brain. Images are shown in radiological convention (the left is displayed on the right) and corresponding Montreal Neurological Institute (MNI) coordinates are displayed below. P-values are TFCE and FWE corrected for multiple comparisons across space and corrected for age, gender and (patient/control) group.



Suppl. Figure 2: The association of sleep duration with integrity of white matter microstructure differs between patients using antipsychotics and controls. TBSS was used to assess the interaction effects of sleep duration and group (Group * Sleep) with integrity of white matter microstructure. There is a significant negative interaction between sleep duration*group and FA. A negative interaction is defined as interactions where the slope between sleep variables and DTI measures is lower for patients than for controls. . Significant interactions are shown in yellow to red, depending on the significance level. The white matter skeleton is shown in green and overlaid on the MNI152 template brain. Images are shown in radiological convention (the left is displayed on the right) and corresponding Montreal Neurological Institute (MNI) coordinates are displayed below. P-values are TFCE and FWE corrected for multiple comparisons across space, and corrected for age and gender.

Suppl. Table 2 - Cluster size (number of voxels), p-value, t-value, MNI coordinates of center of gravity and anatomical location of significant associations between interdaily stability and sleepduration (latter analysis restricted to patients using antipsychotics) with FA

Analysis	Cluster size	Corrected p-value	T-value	MNI coordinates (x; y; z)	White matter tracts
Significant correlation of interdaily stability with FA	9935	0.018	3.12	6.22; -25; 28.6	Genu of corpus callosum Body of corpus callosum Splenium of corpus callosum Right anterior corona radiata Left anterior corona radiata Right superior corona radiata Left superior corona radiata Right posterior corona radiata Left posterior corona radiata Right posterior thalamic radiation (include optic radiation) Left cingulum (cingulate gyrus) Right superior longitudinal fasciculus
	989	0.043	3.07	-40.2; -5.46; 27.3	Left superior longitudinal fasciculus
	480	0.044	3.83	38.1; -22.4; -4.51	Retrolenticular part of right internal capsule Sagittal stratum Right external capsule Fornix (cres) / Stria terminalis
	311	0.043	3.81	48.4; -32.8; -10.6	Sagittal stratum
*Significant Group*Sleep duration interaction with FA in patient using antipsychotics and controls*	53	0.048	3.95	-16.9; 27.2; 21.7	Genu of corpus callosum Left anterior corona radiata

References

Altena, E., Vrenken, H., Van Der Werf, Y.D., van den Heuvel, O. a., Van Someren, E.J.W., 2010. Reduced Orbitofrontal and Parietal Gray Matter in Chronic Insomnia: A Voxel-Based Morphometric Study. Biol. Psychiatry 67, 182–185. doi:10.1016/j.biopsych.2009.08.003

American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders: DSM-5, Washington DC: American Psychiatric Association. doi:10.1176/appi.books.9780890425596.744053

Andersson, J.L.R., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. Neuroimage 20, 870–888. doi:10.1016/S1053-8119(03)00336-7

Beaulieu, C., 2009. The biological basis of diffusion anisotropy, in: Diffusion MRI. pp. 105–126. doi:10.1016/B978-0-12-374709-9.00006-7

Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system - a technical review 15, 435–455. doi:10.1002/nbm.782

Bellesi, M., Pfister-Genskow, M., Maret, S., Keles, S., Tononi, G., Cirelli, C., 2013. Effects of sleep and wake on oligodendrocytes and their precursors. J. Neurosci. 33, 14288–300. doi:10.1523/JNEUROSCI.5102-12.2013

Belmaker, R.H., 2004. Bipolar disorder. N. Engl. J. Med. 351, 476 – 486. doi:10.1056/NEJMra035354

Bherer, L., Erickson, K.I., Liu-Ambrose, T., 2013. A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. J Aging Res 2013, 657508. doi:10.1155/2013/657508

Bowling, H., Zhang, G., Bhattacharya, A., Pérez-Cuesta, L.M., Deinhardt, K., Hoeffer, C. a, Neubert, T. a, Gan, W., Klann, E., Chao, M. V, 2014. Antipsychotics activate mTORC1-dependent translation to enhance neuronal morphological complexity. Sci. Signal. 7, ra4. doi:10.1126/scisignal.2004331

Brambilla, P., Bellani, M., Yeh, P.H., Soares, J.C., 2009. Myelination in bipolar patients and the effects of mood stabilizers on brain anatomy. Curr.Pharm.Des 15, 2632–2636.

Burzynska, A.Z., Chaddock-Heyman, L., Voss, M.W., Wong, C.N., Gothe, N.P., Olson, E.A., Knecht, A., Lewis, A., Monti, J.M., Cooke, G.E., Wojcicki, T.R., Fanning, J., Chung, H.D., Awick, E., McAuley, E., Kramer, A.F., 2014. Physical activity and cardiorespiratory fitness are beneficial for white matter in low-fit older adults. PLoS One 9. doi:10.1371/journal.pone.0107413

Caseras, X., Tansey, K.E., Foley, S., Linden, D., 2015. Association between genetic risk scoring for schizophrenia and bipolar disorder with regional subcortical volumes. Transl. Psychiatry 5, e692. doi:10.1038/tp.2015.195

Caspersen, C.J., Powell, K.E., Christenson, G.M., 1985. Physical Activity , Exercise , and Physical Fitness : Definitions and Distinctions for Health-Related Research. Public Health Rep. 100, 126–131.

Chen, J.J., Rosas, H.D., Salat, D.H., 2013. The Relationship between Cortical Blood Flow and Sub-Cortical White-Matter Health across the Adult Age Span. PLoS One 8. doi:10.1371/journal.pone.0056733

Cirelli, C., Faraguna, U., Tononi, G., 2006. Changes in brain gene expression after long-term sleep deprivation. J. Neurochem. 98, 1632–1645. doi:10.1111/j.1471-4159.2006.04058.x

Diekelmann, S., Born, J., 2010. The memory function of sleep. Nat.Rev.Neurosci. 11, 114–126. doi:10.1038/nrn2762

Du, Y., Fischer, T.Z., Lee, L.N., Lercher, L.D., Dreyfus, C.F., 2003. Regionally specific effects of BDNF on oligodendrocytes. Dev. Neurosci. 25, 116–126. doi:10.1159/000072261

Elvsåshagen, T., Norbom, L.B., Pedersen, P.Ø., Quraishi, S.H., Bjørnerud, A., Malt, U.F., Groote, I.R., Westlye, L.T., 2015. Widespread Changes in White Matter Microstructure after a Day of Waking and Sleep Deprivation 1–15. doi:10.1371/journal.pone.0127351

Erickson, K.I., Leckie, R.L., Weinstein, A.M., 2014. Physical activity, fitness, and gray matter volume. Neurobiol. Aging 35, S20–S28. doi:10.1016/j.neurobiolaging.2014.03.034

First, M.B. et, Spitzer, R.L., Gibbon, M., Williams, J.B.W., 1997. Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I). Biometric Research Department, New York.

Hallahan, B., Newell, J., Soares, J.C., Brambilla, P., Strakowski, S.M., Fleck, D.E., Kiesepp, 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. Biol. Psychiatry 69, 326–335. doi:10.1016/j.biopsych.2010.08.029

Harvey, A.G., 2008. Sleep and circadian rhythms in bipolar disorder: Seeking synchrony, harmony, and regulation. Am. J. Psychiatry 165, 820–829. doi:10.1176/appi.ajp.2008.08010098

Heng, S., Song, A.W., Sim, K., 2010. White matter abnormalities in bipolar disorder: insights from diffusion tensor imaging studies. J Neural Transm 117, 639–654. doi:10.1007/s00702-010-0368-9

Hibar, D.P., Westlye, L.T., van Erp, T.G.M., Rasmussen, J., Leonardo, C.D., Faskowitz, J., Haukvik, U.K., Hartberg, C.B., Doan, N.T., Agartz, I., Dale, A.M., Gruber, O., Krämer, B., Trost, S., Liberg, B., Abé, C., Ekman, C.J., Ingvar, M., Landén, M., Fears, S.C., Freimer, N.B., Bearden, C.E., Sprooten, E., Glahn, D.C., Pearlson, G.D., Emsell, L., Kenney, J., Scanlon, C., McDonald, C., Cannon, D.M., Almeida, J., Versace, A., Caseras, X., Lawrence, N.S., Phillips, M.L., Dima, D., Delvecchio, G., Frangou, S., Satterthwaite, T.D., Wolf, D., Houenou, J., Henry, C., Malt, U.F., Bøen, E., Elvsåshagen, T., Young, A.H., Lloyd, A.J., Goodwin, G.M., Mackay, C.E., Bourne, C., Bilderbeck, A., Abramovic, L., Boks, M.P., van Haren, N.E.M., Ophoff, R.A., Kahn, R.S., Bauer, M., Pfennig, A., Alda, M., Hajek, T., Mwangi, B., Soares, J.C., Nickson, T., Dimitrova, R., Sussmann, J.E., Hagenaars, S., Whalley, H.C., McIntosh, A.M., Thompson, P.M., Andreassen, O.A., 2016. Subcortical volumetric abnormalities in bipolar disorder. Mol. Psychiatry 1–7. doi:10.1038/mp.2015.227

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage. doi:10.1016/j.neuroimage.2011.09.015

Joo, E.Y., Jeon, S., Lee, M., Kim, S.T., Yoon, U., Koo, D.L., Lee, J.-M., Hong, S.B., 2011. Analysis of Cortical Thickness in Narcolepsy Patients with Cataplexy. Sleep 34, 1357–64. doi:10.5665/SLEEP.1278

Joo, E.Y., Tae, W.S., Kim, S.T., Hong, S.B., 2009. Gray matter concentration abnormality in brains of narcolepsy patients. Korean J. Radiol. 10, 552–558. doi:10.3348/kjr.2009.10.6.552

Joo, E.Y., Tae, W.S., Lee, M.J., Kang, J.W., Park, H.S., Lee, J.Y., Suh, M., Hong, S.B., 2010. Reduced brain gray matter concentration in patients with obstructive sleep apnea syndrome. Sleep 33, 235–241.

Kang, J.-E., Lim, M.M., Bateman, R.J., Lee, J.J., Smyth, L.P., Cirrito, J.R., Fujiki, N., Nishino, S., Holtzman, D.M., 2009. Amyloid-b Dynamics Are Regulated by Orexin and the Sleep-Wake Cycle 326, 1005–1007. doi:10.1126/science.1180962

Klemfuss, H., 1992. Rhythms and the pharmacology of lithium. Pharmacol. Ther. 56, 53–78.

Kopp, C., Longordo, F., Nicholson, J.R., Lüthi, A., 2006. Insufficient sleep reversibly alters bidirectional synaptic plasticity and NMDA receptor function. J. Neurosci. 26, 12456–65.

doi:10.1523/JNEUROSCI.2702-06.2006

Kreutzmann, J.C., Havekes, R., Abel, T., Meerlo, P., 2015. Sleep deprivation and hippocampal vulnerability: Changes in neuronal plasticity, neurogenesis and cognitive function. Neuroscience 309, 173–190. doi:10.1016/j.neuroscience.2015.04.053

Liu, D., Wang, Z., Shu, H., Zhang, Z., 2016. Disrupted white matter integrity is associated with cognitive deficits in patients with amnestic mild cognitive impairment: An atlas-based study. SAGE open Med. 4, 1–8. doi:10.1177/2050312116648812

Mahon, K., Burdick, K.E., Szeszko, P.R., 2010. A role for white matter abnormalities in the pathophysiology of bipolar disorder. Neurosci. Biobehav. Rev. 34, 533–554. doi:10.1016/j.neubiorev.2009.10.012

Merikanto, I., Kronholm, E., Peltonen, M., Laatikainen, T., Lahti, T., Partonen, T., 2012. Relation of chronotype to sleep complaints in the general Finnish population. Chronobiol. Int. 29, 311–317. doi:10.3109/07420528.2012.655870

Mirescu, C., Peters, J.D., Noiman, L., Gould, E., 2006. Sleep deprivation inhibits adult neurogenesis in the hippocampus by elevating glucocorticoids. Proc. Natl. Acad. Sci. U. S. A. 103, 19170–19175. doi:10.1073/pnas.0608644103

Monti, J.M., Monti, D., 2004. Sleep in schizophrenia patients and the effects of antipsychotic drugs. Sleep Med. Rev. 8, 133–48. doi:10.1016/S1087-0792(02)00158-2

Narayan, S., Kass, K.E., Thomas, E.A., 2007. Chronic haloperidol treatment results in a decrease in the expression of myelin/oligodendrocyte-related genes in the mouse brain. J. Neurosci. Res. 85, 757–765. doi:10.1002/jnr.21161

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Hum. Brain Mapp. 15, 1–25. doi:10.1002/hbm.1058

Nortje, G., Stein, D.J., Radua, J., Mataix-Cols, D., Horn, N., 2013. Systematic review and voxel-based meta-analysis of diffusion tensor imaging studies in bipolar disorder. J. Affect. Disord. 150, 192–200. doi:10.1016/j.jad.2013.05.034

Ooms, S., Overeem, S., Besse, K., Rikkert, M.O., Verbeek, M., Claassen, J. a H.R., 2014. Effect of 1 Night of Total Sleep Deprivation on Cerebrospinal Fluid β-Amyloid 42 in Healthy Middle-Aged Men: A Randomized Clinical Trial. JAMA Neurol. 71, 971–977. doi:10.1001/jamaneurol.2014.1173

Pagani, L., St. Clair, P. a., Teshiba, T.M., Service, S.K., Fears, S.C., Araya, C., Araya, X., Bejarano, J., Ramirez, M., Castrillón, G., Gomez-Makhinson, J., Lopez, M.C., Montoya, G., Montoya, C.P., Aldana, I., Navarro, L., Freimer, D.G., Safaie, B., Keung, L.-W., Greenspan, K., Chou, K., Escobar, J.I., Ospina-Duque, J., Kremeyer, B., Ruiz-Linares, A., Cantor, R.M., Lopez-Jaramillo, C., Macaya, G., Molina, J., Reus, V.I., Sabatti, C., Bearden, C.E., Takahashi, J.S., Freimer, N.B., 2016. Genetic contributions to circadian activity rhythm and sleep pattern phenotypes in pedigrees segregating for severe bipolar disorder. Proc. Natl. Acad. Sci. 113, E754–E761. doi:10.1073/pnas.1513525113

R Development Core Team, R., 2014. R: A Language and Environment for Statistical Computing. doi:10.1007/978-3-540-74686-7

Sandoval, H., Soares, J.C., Mwangi, B., Asonye, S., Alvarado, L.A., Zavala, J., Ramirez, M.E., Sanches, M., Enge, L.R., Escamilla, M.A., 2014. Confirmation of MRI anatomical measurements as endophenotypic markers for bipolar disorder in a new sample from the NIMH Genetics of Bipolar disorder in Latino Populations study. Psychiatry Res. - Neuroimaging 247, 34–41. doi:10.1016/j.pscychresns.2015.11.004

Sexton, C.E., Betts, J.F., Demnitz, N., Dawes, H., Ebmeier, K.P., Johansen-Berg, H., 2015. A systematic review of MRI studies examining the relationship between physical fitness and activity and the white matter of the ageing brain. Neuroimage. doi:10.1016/j.neuroimage.2015.09.071

Sheehan, D. V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59, 22–33. doi:10.1016/S0924-9338(99)80239-9

Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155. doi:10.1002/hbm.10062

Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.J., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487–505. doi:10.1016/j.neuroimage.2006.02.024

Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44, 83–98. doi:10.1016/j.neuroimage.2008.03.061

Smith, S.M., Smith, S.M., Johansen-berg, H., Johansen-berg, H., Jenkinson, M., Jenkinson, M., Rueckert, D., Rueckert, D., Nichols, T.E., Nichols, T.E., Miller, K.L., Miller, K.L., Robson, M.D., Robson, M.D., Jones, D.K., Jones, D.K., Klein, J.C., Klein, J.C., Bartsch, A.J., Bartsch, A.J., Behrens, T.E.J., Behrens, T.E.J., 2007. Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. Nat. Protoc. 2, 499–503. doi:10.1038/nprot.2007.45

Svatkova, A., Mandl, R.C.W., Scheewe, T.W., Cahn, W., Kahn, R.S., Hulshoff Pol, H.E., 2015. Physical Exercise Keeps the Brain Connected: Biking Increases White Matter Integrity in Patients with Schizophrenia and Healthy Controls. Schizophr. Bull. 41, 869–878. doi:10.1093/schbul/sbv033

Telzer, E.H., Goldenberg, D., Fuligni, A.J., Lieberman, M.D., Gálvan, A., 2015. Sleep variability in adolescence is associated with altered brain development. Dev. Cogn. Neurosci. 14, 16–22. doi:10.1016/j.dcn.2015.05.007

Tononi, G., Cirelli, C., 2006. Sleep function and synaptic homeostasis. Sleep Med. Rev. 10, 49–62. doi:10.1016/j.smrv.2005.05.002

Tost, H., Alam, T., Geramita, M., Rebsch, C., Kolachana, B., Dickinson, D., Verchinski, B. a, Lemaitre, H., Barnett, A.S., Trampush, J.W., Weinberger, D.R., Marenco, S., 2013. Effects of the BDNF Val66Met polymorphism on white matter microstructure in healthy adults. Neuropsychopharmacology 38, 525–32. doi:10.1038/npp.2012.214

Van Someren, E.J.W., Hagebeuk, E.E.O., Lijzenga, C., Scheltens, P., de Rooij, S.E.J.A., Jonker, C., Pot, A., Mirmiran, M., Swaab, D.F., 1996. Circadian rest-activity rhythm disturbances in Alzheimer ' s Disease. Biol. Psychiatry 40, 259–270. doi:10.1016/0006-3223(95)00370-3

Voelcker-Rehage, C., Niemann, C., 2013. Structural and functional brain changes related to different types of physical activity across the life span. Neurosci. Biobehav. Rev. 37, 2268–2295. doi:10.1016/j.neubiorev.2013.01.028

Wang, G., Grone, B., Colas, D., Appelbaum, L., Mourrain, P., 2011. Synaptic plasticity in sleep: Learning, homeostasis and disease. Trends Neurosci. 34, 452–463. doi:10.1016/j.tins.2011.07.005

Winkelman, J.W., Plante, D.T., Schoerning, L., Benson, K., Buxton, O.M., O'Connor, S.P., Jensen, J.E., Renshaw, P.F., Gonenc, A., 2013. Increased Rostral Anterior Cingulate Cortex Volume in Chronic

Primary Insomnia. Sleep 36, 991–998. doi:10.5665/sleep.2794

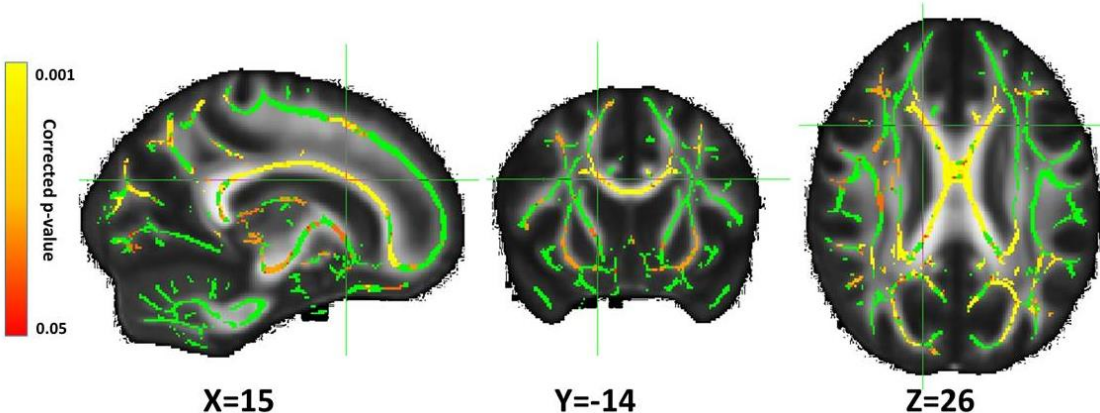
Wise, T., Radua, J., Nortje, G., Cleare, A.J., Young, A.H., Arnone, D., 2016. Voxel-based meta-Analytical evidence of structural disconnectivity in major depression and bipolar disorder. Biol. Psychiatry 79, 293–302. doi:10.1016/j.biopsych.2015.03.004

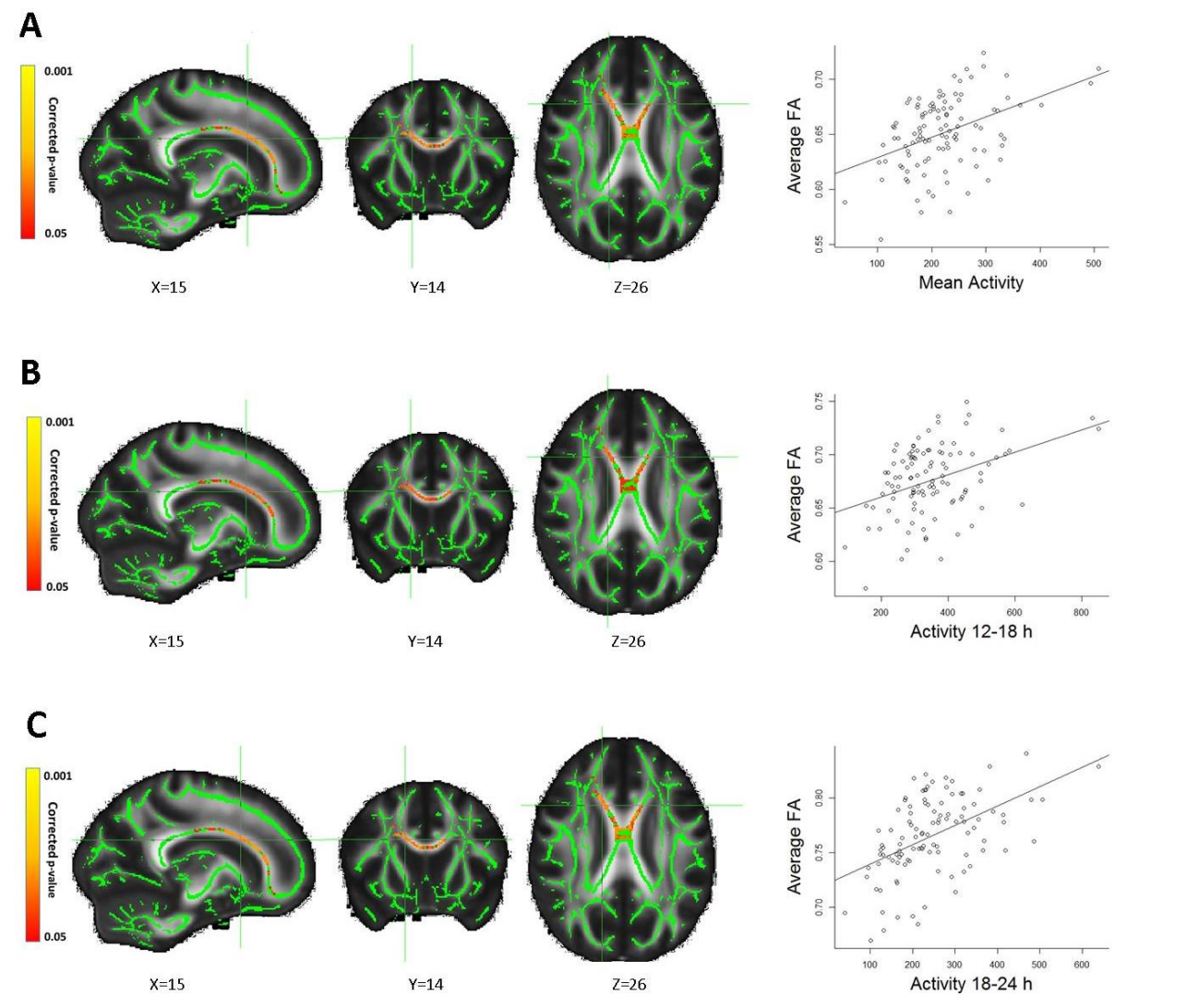
Xie, L., Kang, H., Xu, Q., Chen, M.J., Liao, Y., Thiyagarajan, M., O'Donnell, J., Christensen, D.J., Nicholson, C., Iliff, J.J., Takano, T., Deane, R., Nedergaard, M., 2013. Sleep drives metabolite clearance from the adult brain. Science 342, 373–7. doi:10.1126/science.1241224

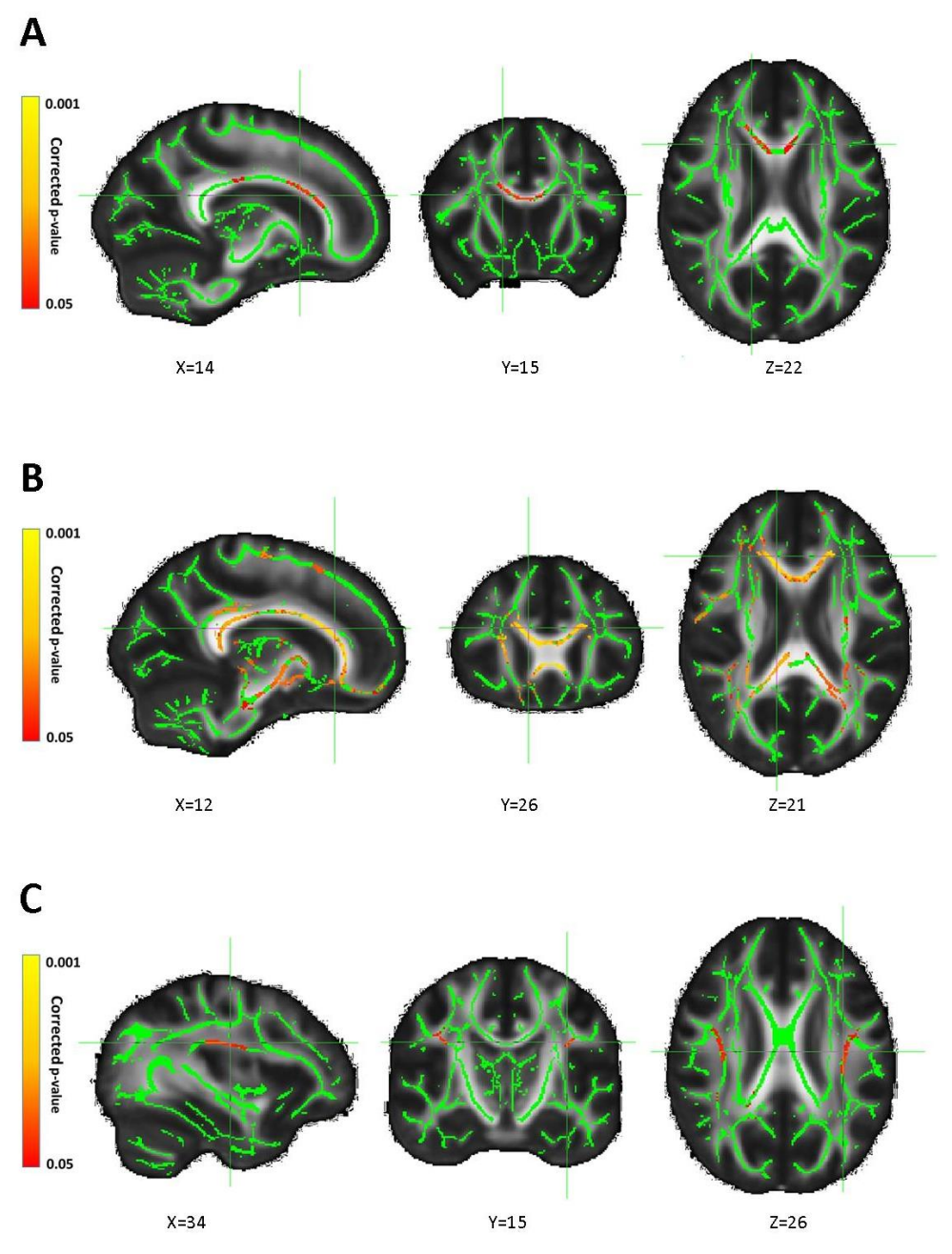
Yaffe, K., Nasrallah, I., Hoang, T.D., Lauderdale, D.S., Knutson, K.L., Carnethon, M.R., Launer, L.J., Lewis, C.E., Sidney, S., 2016. Sleep Duration and White Matter Quality in Middle-Aged Adults 39, 1743–1747.

Yip, S.W., Chandler, R. a, Rogers, R.D., Mackay, C.E., Goodwin, G.M., 2013. White matter alterations in antipsychotic- and mood stabilizer-naïve individuals with bipolar II/NOS disorder. NeuroImage. Clin. 3, 271–8. doi:10.1016/j.nicl.2013.08.005

Zhu, N., Jacobs, D.R.J., Schreiner, P.J., Launer, L.J., Whitmer, R. a, Sidney, S., Demerath, E., Thomas, W., Bouchard, C., He, K., Erus, G., Battapady, H., Bryan, R.N., 2015. Cardiorespiratory fitness and brain volume and white matter integrity: The CARDIA Study. Neurology 84, 2347–2353. doi:10.1212/WNL.0000000000001658


0.001
Corrected p-value
0.05
X=15
Y=-14
Z=26


A
0.001
Corrected p-value
0.05
X=15
Y=14
Z=26
Average FA
Mean Activity
B
0.001
Corrected p-value
0.05
X=15
Y=14
Z=26
Average FA
Activity 12-18 h
C
0.001
Corrected p-value
0.05
X=15
Y=14
Z=26
Average FA
Activity 18-24 h


A
0.001
Corrected p-value
0.05
X=14
Y=15
Z=22
B
0.001
Corrected p-value
0.05
X=12
Y=26
Z=21
C
0.001
Corrected p-value
0.05
X=34
Y=15
Z=26

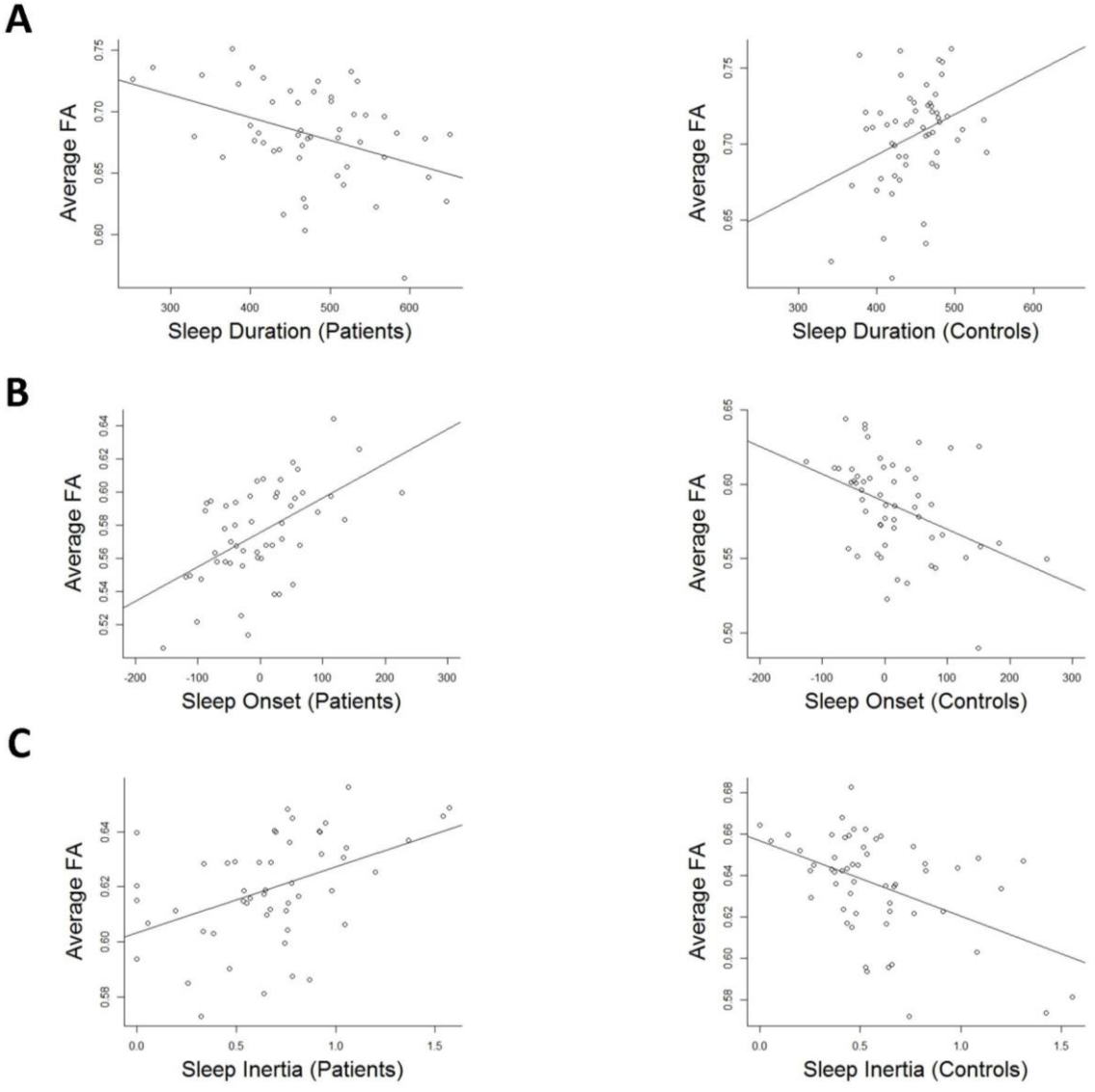

A
Average FA
Sleep Duration (Patients)
Average FA
Sleep Duration (Controls)
B
Average FA
Sleep Onset (Patients)
Average FA
Sleep Onset (Controls)
C
Average FA
Sleep Inertia (Patients)
Average FA
Sleep Inertia (Controls)

Table 1 – Sample characteristics

	Patients	**Controls**	***F* / χ^2**	**p-value**
Age M (sd)	49.5 (11.4)	45.5 (15.8)	2.18	0.14
Gender M/F (% male)	28/23 (54.9%)	25/30 (45.5%)	0.44	0.22
Handedness R/L/B (% right)	42/6/3 (82.4%)	44/11/0 (80.0%)	4.37	0.11
IQ M (sd)	99.6 (15.7)	105.7 (14.4)	4.18	0.04*
Level of education n (%)			6.32	0.28
1. Low education	5 (9.3%)	4 (7.8%)		
2. Intermediate secondary education	4 (7.4%)	9 (17.6%)		
3. Intermediate professional education	7 (13.0%)	10 (19.6%)		
4. High prepatory vocational / pre-university	13 (24.1%)	6 (11.8%)		
5. Bachelor degree	11 (20.4%)	13 (25.5%)		
6. Master or PhD degree	14 (25.9%)	9 (17.6%)		
Inventory of Depressive Symptoms M (sd)	14.3 (10.1)	5.2 (3.9)	35.60	<0.001*
Altman Self-Rating Mania Scale M (sd)	1.9 (1.8)	1.7 (2.3)	0.41	0.53
Body Mass Index M (sd)	27.4 (4.3)	25.2 (3.3)	8.45	0.01*
Nr of workdays M (sd)	4.2 (3.5)	5.7 (3.7)	4.11	0.05*
Illness duration/ years M (sd)	16.5 (14.2)	-		
Number of episodes M (sd)	10.0 (10.8)	-		
Age at onset M (sd)	31.7 (12.2)	-		
History of psychotic symptoms N (%)	35 (70.0%)	-		

* Significant group difference at $p<0.05$

Table 2 - Cluster size (number of voxels), p-value, t-value, MNI coordinates of center of gravity and anatomical location of significant differences in FA between patients and controls

Analysis	**Cluster size**	**Corrected p-value**	**T-value**	**MNI coordinates (x; y; z)**	**Anatomical region**
Significant group difference in FA between patients and controls	44232	<0.001	3.08	2.37; -21.4; 16.8	Genu of corpus callosum Body of corpus callosum Splenium of corpus callosum R cerebral peduncle L cerebral peduncle R posterior limb of internal capsule L posterior limb of internal capsule L retrolenticular part of internal capsule R anterior corona radiata L anterior corona radiata R superior corona radiata Lsuperior corona radiata Right posterior corona radiata L posterior corona radiata R posterior thalamic radiation (include optic radiation) L posterior thalamic radiation (include optic radiation) R external capsule L external capsule L cingulum (cingulate gyrus) R superior longitudinal fasciculus L superior longitudinal fasciculus

Table 3 – Difference between sleep and activity measures for patients and controls

	Patients	**Controls**	**p-value**
Sleep duration M (sd)	475.2 (86.3)	447.2 (40.4)	0.01
Sleep onset M (sd)	-0.50 (75.3)	17.6 (73.0)	0.22
Sleep onset latency M (sd)	5.0 (5.7)	4.9 (5.2)	0.89
Sleep offset M (sd)	474.7 (88.5)	464.8 (74.8)	0.33
Sleep intertia M (sd)	6.8 (7.2)	5.4 (6.3)	0.37
Wake after sleep onset M (sd)	59.6 (27.0)	48.5 (19.5)	0.03
Sleep efficiency M (sd)	84.8 (8.1)	87.3 (4.9)	0.12
Mean Activity M (sd)	204.7 (89.3)	234.4 (54.5)	0.08
Activity 0 to 6 h M (sd)	42.1 (36.6)	45.0 (41.5)	0.82
Activity 6 to 12 h M (sd)	227.3 (127.8)	260.8 (91.6)	0.12
Activity 12 to 18 h M (sd)	333.4 (143.1)	361.8 (93.9)	0.34
Activity 18 to 24 h M (sd)	215.6 (109.6)	270.4 (84.9)	0.02

Means (M) and standard deviations (sd) on sleep and activity measures. Models are corrected for age and gender. Sleep duration, sleep onset latency, sleep inertia and WASO are measured in minutes. Sleep onset and sleep offset describe minutes prior or past midnight. Sleep efficiency is a percentage ranging from 1-100. All activity variables are activity counts averaged per minute.

Table 4 - Cluster size (number of voxels), p-value, t-value, MNI coordinates of center of gravity and anatomical location of significant associations between sleep and activity parameters with FA

Analysis	Cluster size	Corrected p-value	T-value	MNI coordinates (x; y; z)	Anatomical region
*Significant Group*Sleep duration interaction with FA*	651	0.04	3.74	10.2; 20.5; 19.5	Genu of corpus callosum
	16	0.05	2.34	-6.25; 12.7; 22	Body of corpus callosum
*Significant Group*Sleep onset interaction with FA*	292	0.039	4.36	37.6; -11.9; 29.4	R superior longitudinal fasciculus
	266	0.038	4.85	-37; -15.6; 27.2	L superior longitudinal fasciculus
	134	0.047	3.39	2.29; 24.9; 13.2	Genu of corpus callosum
	75	0.046	3.78	18.1; -51.6; 30.4	Splenium of corpus callosum R posterior corona radiata
	45	0.047	3.63	12.2; 37.9; 21	Splenium of corpus callosum
	33	0.047	4.96	38; -40.5; 24.8	R superior longitudinal fasciculus
*Significant Group*Sleep inertia interaction with FA*	21826	0.006	3.46	12.2; -13.3; 17.2	Genu of corpus callosum Body of corpus callosum Splenium of corpus callosum
	227	0.047	3.10	-11.6; 8.42; -7.77	Cerebral peduncle Posterior limb of left internal capsule
	202	0.049	2.19	8.01; -26; -28.8	R corticospinal tract Pontine crossing tract R cerebral peduncle
	162	0.046	3.96	-7.18; -28.3; -30.7	L corticospinal tract Pontine crossing tract L cerebral peduncle
	104	0.049	2.63	-29.4; -23.4; -1.7	Retrolenticular part of left internal capsule Sagittal stratum Fornix (cres) / Stria terminalis
	78	0.049	2.58	-37.9; -47.6; 1.23	Posterior thalamic radiation Sagittal stratum
	38	0.05	3.57	-35.1; -58.4; -5.16	Posterior thalamic radiation Sagittal stratum
	15	0.05	1.93	-40.6; -25.7; -4.6	Sagittal stratum
	13	0.05	2.39	-32; -203;	L external capsule

				0.769	
Significant correlation of mean activity with FA	3153	0.019	2.84	9.01; 19.2; 18.7	Genu of corpus callosum Body of corpus callosum R anterior corona radiata
Significant correlation of activity from 12h to 18h with FA	2466	0.028	3.24	7.37; 17; 19.7	Genu of corpus callosum Body of corpus callosum R anterior corona radiata
Significant correlation of activity from 18h to 24h with FA	333	0.044	3.46	3.64; 25.7; 10.9	Genu of corpus callosum Body of corpus callosum

Table 5 – Association between average FA and sleep parameters, controlled for age and gender

	β	**R^2**	**p-value**	**N**
Sleep duration	-0.39	0.32	0.004	51
Sleep onset	0.49	0.37	<0.001	51
Sleep inertia	0.34	0.33	0.008	51

Table 6 – Correlation (*r*) between sleep parameters and psychotropic medication, and average FA of the significant clusters and psychotropic medication. P < 0.05 indicated with one asterisk, p<0.001 indicated with double asterisk.

	Lithium	*Other mood stabilizers*	*Benzodiazepines*	*Antipsychotics*
Sleep duration	0.10	0.12	0.07	0.47*
FA sleep duration	0.34 *	-0.33 *	-0.03	-0.31*
Sleep onset	-0.10	-0.04	0.13	-0.40**
FA sleep onset	0.28*	-0.15	0.20	-0.26
Sleep inertia	0.17	-0.14	0.01	-0.23
FA sleep inertia	0.38**	-0.34*	0.05	-0.17

Suppl. Table 1- Association between sleep onset and average FA while controlling for history of psychotic symptoms or number of episodes

	Beta	**p-value**	**R^2**	N
Model + psychotic symptoms	0.41	0.002	0.35	50
Model + number of episodes	0.35	0.02	0.40	41

Suppl. Table 2 - Cluster size (number of voxels), p-value, t-value, MNI coordinates of center of gravity and anatomical location of significant associations between interdaily stability and sleepduration (latter analysis restricted to patients using antipsychotics) with FA

Analysis	Cluster size	Corrected p-value	T-value	MNI coordinates (x; y; z)	White matter tracts
Significant correlation of interdaily stability with FA	9935	0.018	3.12	6.22; -25; 28.6	Genu of corpus callosum Body of corpus callosum Splenium of corpus callosum Right anterior corona radiata Left anterior corona radiata Right superior corona radiata Left superior corona radiata Right posterior corona radiata Left posterior corona radiata Right posterior thalamic radiation (include optic radiation) Left cingulum (cingulate gyrus) Right superior longitudinal fasciculus
	989	0.043	3.07	-40.2; -5.46; 27.3	Left superior longitudinal fasciculus
	480	0.044	3.83	38.1; -22.4; -4.51	Retrolenticular part of right internal capsule Sagittal stratum Right external capsule Fornix (cres) / Stria terminalis
	311	0.043	3.81	48.4; -32.8; -10.6	Sagittal stratum
*Significant Group*Sleep duration interaction with FA in patient using antipsychotics and controls*	53	0.048	3.95	-16.9; 27.2; 21.7	Genu of corpus callosum Left anterior corona radiata

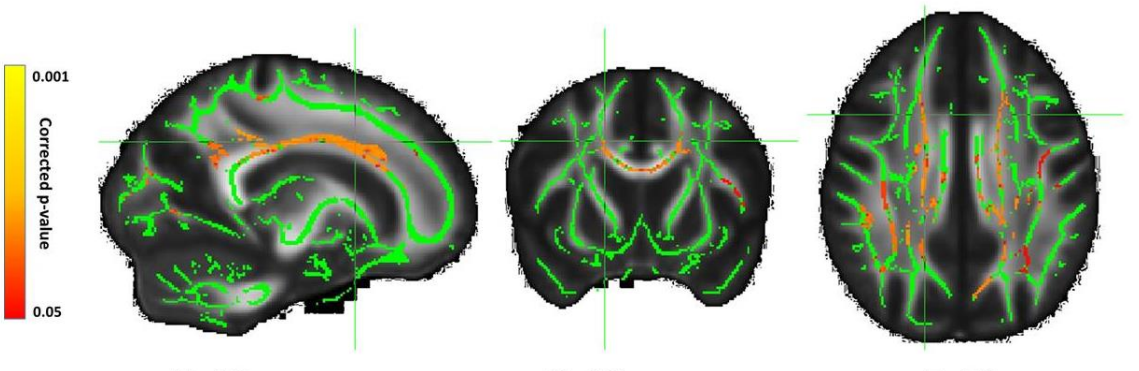
Abstract

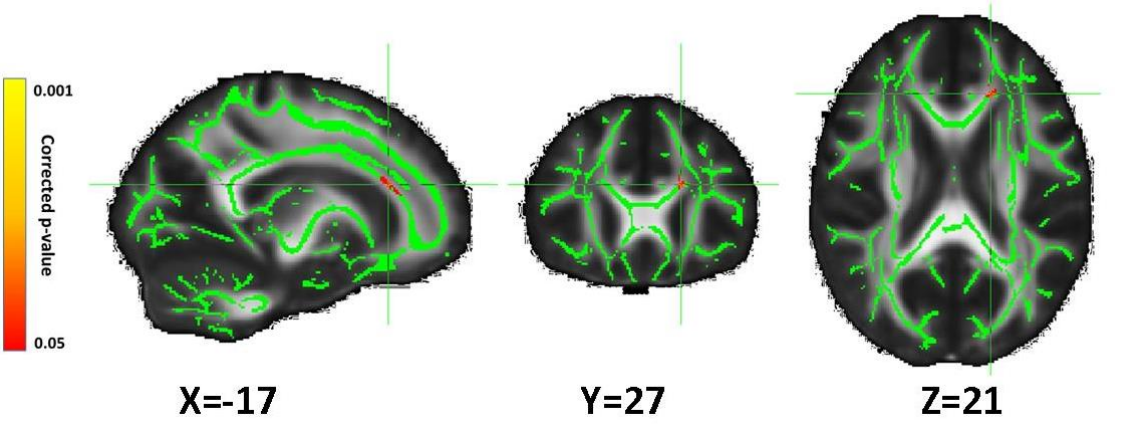
We investigate how the sleep disruptions and irregular physical activity levels that are prominent features of bipolar disorder (BD) relate to white matter microstructure in patients and controls. Diffusion tension imaging (DTI) and 14-day actigraphy recordings were obtained in 51 BD I patients and 55 age-and-gender-matched healthy controls. Tract-based spatial statistics (TBSS) was used for voxelwise analysis of the association between fractional anisotropy (FA) and sleep and activity characteristics in the overall sample. Next, we investigated whether the relation between sleep and activity and DTI measures differed for patients and controls. Physical activity was related to increased integrity of white matter microstructure regardless of bipolar diagnosis. The relationship between sleep and white matter microstructure was more equivocal; we found an expected association between higher FA and effective sleep in controls but opposite patterns in bipolar patients. Confounding factors such as antipsychotic medication use are a likely explanation for these contrasting findings and highlight the need for further study of medication-related effects on white matter integrity.

Keywords: bipolar disorder, actigraphy, DTI, sleep disturbances, physical activity

Highlights:

- Physical activity was related to increased fractional anisotropy (FA)
- Controls showed positive associations between higher FA and effective sleep
- In patients sleep was inversely related to FA
- The relation between sleep and FA in patients was confounded by antipsychotics


0.001
Corrected p-value
0.05
X=18
Y=13
Z=32


0.001
Corrected p-value
0.05
X=-17
Y=27
Z=21