

The association between sleep and cognitive abnormalities in bipolar disorder

A. J. Bradley^{1,2}, K. N. Anderson³, P. Gallagher¹ and R. H. McAllister-Williams^{1,4}

Original Article

Cite this article: Bradley AJ, Anderson KN, Gallagher P, McAllister-Williams RH (2020). The association between sleep and cognitive abnormalities in bipolar disorder.

Psychological Medicine 50, 125–132. <https://doi.org/10.1017/S0033291718004038>

Received: 19 February 2018

Revised: 19 May 2018

Accepted: 13 December 2018

First published online: 16 January 2019

Key words:

Attention; bipolar disorder; cognition; cognitive dysfunction; psychomotor processing speed; sleep; sleep disturbance

Author for correspondence:

R. H. McAllister-Williams, E-mail: hamish.mcallister-williams@ncl.ac.uk

¹Institute of Neuroscience, Newcastle University, Wolfson Research Centre, Campus for Aging and Vitality, Westgate Road, Newcastle Upon Tyne, NE4 5LP, UK; ²Eli Lilly and Company Limited, Lilly House, Priestly Road, Basingstoke RG24 9NL, UK; ³Regional Sleep Service, Freeman Hospital, Freeman Road, High Heaton, Newcastle upon Tyne NE7 7DN, UK and ⁴Northumberland, Tyne and Wear NHS Foundation Trust, Newcastle, UK

Abstract

Background. Bipolar disorder (BD) is associated with attentional and processing abnormalities. Such abnormalities are also seen in healthy subjects with sleep disruption. We hypothesised cognitive abnormalities in BD patients would be worse in those with objectively verified sleep abnormalities.

Methods. Forty-six BD patients and 42 controls had comprehensive sleep/circadian rhythm assessment over 21 days alongside mood questionnaires. Cognitive function was assessed with a range of tasks including Psychomotor Vigilance Test (PVT), Attention Network Task (ANT) and Digit Symbol Substitution Test (DSST). BD participants with normal and abnormal sleep were compared with age- and sex-matched controls.

Results. BD patients had longer response times and made more lapses (responses >500 ms) than controls on the PVT (both $p < 0.001$). However, patients with normal sleep patterns did not differ from controls while those with sleep abnormalities did ($p < 0.001$). An identical pattern of effects were seen with the ANT response times, with the abnormality in bipolar abnormal sleepers related to the executive attentional network. Similarly, patients made fewer correct responses on the DSST compared with the controls ($p < 0.001$). Bipolar normal sleepers did not differ while those with abnormal sleep did ($p < 0.001$). All these differences were seen in bipolar abnormal sleepers who were euthymic ($p < 0.01$) and across the main abnormal sleep phenotypes.

Conclusions. We confirm impairment in attention and processing speed in BD. Rather than sleep abnormalities exacerbating such dysfunction, the impairments were confined to bipolar abnormal sleepers, consistent with sleep disturbance being the main driver of cognitive dysfunction.

Introduction

Bipolar disorder (BD) is associated with a range of well-described cognitive dysfunctions as detailed in an individual patient data meta-analysis of 2876 subjects, with deficits seen across all cognitive domains tested (Bourne *et al.*, 2013). While cognitive dysfunction is modestly associated with a history of psychotic symptoms and mania, it is seen across all clinical subgroups (Bora, 2018), including in patients when euthymic (Thompson *et al.*, 2005; Robinson *et al.*, 2006). There is little evidence of progression of dysfunction over time (Samame *et al.*, 2014). The dysfunction may relate to underlying attentional and processing speed abnormalities (Gallagher *et al.*, 2014). The attentional abnormalities include increased intra-individual variability (IIV) in response times in attentional tasks (Gallagher *et al.*, 2015). Sleep disruption, a common feature of all stages of BD, including euthymia (Harvey *et al.*, 2005), can also cause attentional abnormalities (Lim and Dinges, 2008). Again this is associated with an increased IIV, most commonly reported as an increase of ‘lapses’, or slow responses, in attentional tasks. Primary sleep disorders, such as sleep apnoea, also impact on attention and alertness (Bucks *et al.*, 2013) and are underdiagnosed in BD (Soreca *et al.*, 2015). Prior studies assessing cognition in patients with BD have rarely objectively assessed sleep alongside cognition and at best have relied upon self-report (Giglio *et al.*, 2010; Kanady *et al.*, 2017). There is a relative lack of correlation between objectively assessed and subjectively reported sleep abnormalities in BD (Harvey *et al.*, 2005; Bradley *et al.*, 2017). Therefore, there is an important lack of data describing the association between objectively measured sleep abnormalities and cognitive function in BD. Previous work from this group objectively assessed sleep and circadian rhythm in detail in a well-characterised cohort of BD patients (Bradley *et al.*, 2017). This revealed a range of sleep problems including obstructive sleep apnoea (OSA), circadian rhythm disturbance, hypersomnia and insomnia, with many patients having evidence of more than one of these disturbances. These objectively measured sleep disturbances were associated with impaired psychosocial functioning and quality of life. The aim of this study was to assess the relationship between

objectively assessed sleep and cognitive function in BD patients. Specifically this study has compared cognition in BD patients with and without abnormal sleep to age- and sex-matched controls. Our hypothesis was that BD patients would demonstrate evidence of cognitive dysfunction and that this would be worse in, though not exclusive to, those patients with objectively verified sleep abnormalities.

Method

Participants

The study was approved by the National Research Ethics Committee North East – Newcastle & North Tyneside. Outpatients with BD type I or II, in any mood state, were recruited from a research database, patient support groups and NHS services in the North East of England. Healthy controls, matched by age and gender, were recruited from Newcastle University, local volunteer databases and hospital staff and their families. All participants provided written informed consent before taking part in the research. Participants were aged 18–65 years and fluent in English. Exclusion criteria were: any significant medical or neurological disorder that might interfere with sleep or cognition; current alcohol or substance misuse disorder [defined by DSM IV criteria (American Psychiatric Association, 1994)]; current shift work and previous significant head injury. A BD diagnosis meeting DSM-IV criteria was confirmed using the Mini International Neuropsychological Interview (MINI) (Sheehan *et al.*, 1998). Patients with BD were excluded if they had any changes to their psychotropic medication in the previous 4 weeks. Exclusion criteria for controls were: personal or first degree relative history of a DSM IV Axis I disorder as determined by clinical history; prescribed psychotropic medications and any known sleep disorder. Additionally controls had to be psychiatrically well, confirmed by MINI interview; have a 17-item Hamilton Depression Rating Scale score (HAMD-17) (Hamilton, 1967) score <7; a Young Mania Rating Scale (YMRS) (Young *et al.*, 1978) score <5; a Pittsburgh Sleep Quality Index (PSQI) (Buysse *et al.*, 1989) score <5 and Epworth Sleepiness Scale (ESS) (Johns, 1991) score <10.

Overall study design

The study was cross-sectional, with participants assessed over a 3-week period.

Psychiatric symptoms and sleep assessments

Participants were assessed on days 1 and 21. A comprehensive battery of questionnaires and rating scales were used to assess psychiatric symptoms. These included the 17-item GRID-HAMD (Williams *et al.*, 2008), Beck Depression Inventory (BDI) (Beck *et al.*, 1961) and YMRS (Young *et al.*, 1978). All medications used by the patients in the BD group were recorded.

A single night of home partial polysomnography (Embla Systems, Bloomfield, New Jersey, USA) was used to screen for sleep apnoea. Respiratory events were scored according to the standard criteria of the American Association of Sleep Medicine (AASM) (Kushida *et al.*, 2005). An AHI of >5/h was considered abnormal and indicative of sleep apnoea. Severity was defined as mild (AHI 5–15), moderate (AHI 15–30) or severe (AHI >30). Participants also completed the Restless Legs Syndrome

(RLS) Rating Scale to assess for the presence and severity of RLS (Walters *et al.*, 2003). This scale includes 10 questions each scoring 0–4 with a score of 1–10 representing mild, 11–20 moderate, 21–30 severe and 31–40 very severe RLS.

For objective assessment of sleep/wake cycle (circadian rhythm), subjects wore a triaxial wrist accelerometer (GENEActiv; Activinsights, UK) on their non-dominant wrist for all 21 days of the study alongside completing a daily sleep log. The analysis of these data has been previously described (Bradley *et al.*, 2017). In brief, the accelerometer data were analysed using an open source R package sleep detection algorithm, GGIR, which has demonstrated a high sensitivity and specificity to detect periods of sleep (van Hees *et al.*, 2015). Sleep onset time, sleep offset time, TST, time in bed (TIB), sleep efficiency (defined as TST/TIB) and mean 24 h sleep duration (defined as nocturnal sleep plus daytime naps) were all derived. Correlation analysis was performed to check agreement between sleep logs and accelerometer-derived sleep variables. The relative amplitude between day and night activity was calculated during the least active 5 h (L5) and most active 10 h (M10) periods according to previously published methods (Van Someren *et al.*, 1999). Participants were then identified as normal sleepers (6–10 h sleep within 24 h with a regular sleep wake cycle), short sleepers (<6 h nocturnal sleep), long sleep [>10 h sleep within 24 h (Kaplan *et al.*, 2011)] and circadian rhythm disturbance, including a delayed sleep phase (habitual sleep onset after 2:00 am), advanced sleep phase, irregular sleep wake pattern (three to four periods of sleep but no consolidated overnight period) or non-24 h pattern in keeping with the definitions within the International Classification of Sleep Disorder – Third edition (ICSD-3) (American Academy of Sleep Medicine, 2014). Patient-reported sleep logs were used to assist in the interpretation of actigraphy data.

Cognitive assessment battery

Verbal IQ was assessed with the National Adult Reading Test (NART) (Nelson, 1982) and then a range of tasks were administered to assess psychomotor speed, attention and executive components of cognition. This comprised of the Psychomotor Vigilance Test (PVT) (Dinges and Powell, 1985), the Attention Network Test (ANT) (Fan *et al.*, 2002) and the Digit Symbol Substitution Test (DSST) (Wechsler, 1981). Participants completed two DSSTs one at the beginning and one at the end of the cognitive testing session to assess for any change in performance across the session. Reaction time (RT) data from the PVT and the ANT were fitted to an ex-Gaussian distribution (a mathematical convolution of the Gaussian normal and exponential distribution). The ex-Gaussian has been shown to approximate well to empirical RT distributions (Schmiedek *et al.*, 2007); this distribution model produces three parameters: μ and σ , the mean and standard deviation of the Gaussian component; and τ , which defines the exponential component and represents the 'slow-tail' of the distribution (Ratcliff, 1979). To derive these measures, ex-Gaussian probability density functions were fitted to the distribution of valid response times of each individual using the DISTRIB toolbox (Lacouture and Cousineau, 2008) in MATLAB® v.R2010b (The MathWorks Inc. Natick, MA, USA).

The Newcastle Spatial Working Memory (NSWM) Test (Pariante *et al.*, 2012), a test of verbal learning, TRAILS A and B and digit span test (Lezak *et al.*, 2004) were also performed.

Tests were performed at the same time of day for controls and BD patients at the end of the sleep assessment period.

Statistical analysis

All statistical analyses were performed using IBM SPSS statistical package version 22. Data are reported as means with standard deviations. Normality of distribution of data was tested using the Shapiro–Wilk test. Log10 or square root transformations were used where necessary to normalise the data. Parametric tests (e.g. *t* test) were used unless the data remained non-normally distributed despite transformation when equivalent non-parametric tests (e.g. Mann–Whitney *U* test) were used. In the comparisons between controls and BD patients, age and NART IQ score were examined as possible confounders. If there was a significant correlation of the outcome measure under investigation and either of these variables, then analysis of covariance (ANCOVA) was performed. Mood was not examined in this way due to this being significantly different between controls and BD patients. Rather sub-group analysis was performed taking advantage of those BD patients who were euthymic. A score of <8 on the BDI was used to define ‘euthymia’ (Keller, 2003). Note that all bipolar patients met euthymia criteria (YMRS < 10) with regards to manic symptoms. Overall, there was a strong correlation between HAMD-17 and BDI scores [$r_{(s)(45)} = 0.831, p < 0.001$]. The BDI was used in preference to the HAMD-17 due to it containing only one sleep variable, rather than the three within the HAMD-17. A significance threshold of $p < 0.05$ was used for all analyses.

Results

Participants

The participants and their sleep phenotypes have previously been described (Bradley *et al.*, 2017). In this current analysis, only participants with complete actigraphy and overnight sleep oximetry data sets were included. Eighty-two participants, 46 with BD (16 BD I and 30 BDII) and 36 controls, completed the study protocol. Table 1 shows the participant characteristics. Groups did not differ significantly with regard to age or gender but participants with BD had a greater body mass index (BMI) and scored more highly on mood rating scales. Twenty-one participants with BD scored ≥ 8 on the HAMD-17 (range 8–35), but none of the patients were considered clinically to be in a manic or hypomanic episode for the duration of the study (YMRS range 0–10). Of the 46 BD patients, 28 had objectively defined abnormal sleep including OSA ($n = 12$), circadian rhythm sleep disorder ($n = 12$), long sleep ($n = 14$), short sleep ($n = 4$) or a combination of these abnormalities. Sixteen had normal objective measures of sleep with the absence of sleep apnoea confirmed with an overnight sleep study. Two BD participants with normal accelerometry defined sleep patterns did not complete the overnight sleep apnoea study so were not included in the normal sleeping group. On this basis, the BD patients were divided into those with objectively normal ($n = 16$) and objectively abnormal sleep ($n = 28$). The characteristics of the various bipolar sub-groups are shown in Table 2. Thirty-six controls had objectively verified normal sleep without any level of OSA and were included in the analysis. Two BD participants did not complete the full battery of cognitive tests, one who became anxious and asked to stop and the other for undisclosed reasons. One aged 55 years was a normal sleeper and had a BDI score of 8. The other aged 25

years was a normal sleeper with a BDI of 28. Their data are included for the tests they completed.

Psychomotor Vigilance Test

The PVT was completed by 36 controls and 46 BD patients. One BD patient was omitted from the analysis as the number of lapses committed was more than three times the interquartile range and they were deemed an extreme outlier. As anticipated, BD patients differed from control participants having longer mean response times (RTs: 360 ± 50 v. 325 ± 32 ms; $p < 0.001$; $g = 0.81$, 95% CI 0.35–1.26) and making more ‘lapses’ (RT > 500 ms; 5.1 ± 6.4 v. 1.8 ± 2.8 ; $p < 0.001$; $g = 0.64$, 95% CI 0.19–1.09). There were no missed responses made by any participant in either group. The rate of ‘anticipations’ (defined as responding <100 ms after presentation of the target stimulus) was very low and did not differ between controls and BD patients (1.08 ± 1.4 v. 1.0 ± 1.3 respectively; $t = 0.275, p = 0.784$). As a result, anticipations were not analysed further. Mean RT and number of lapses was significantly greater ($p = 0.047$ and 0.050 , respectively) in the 26 euthymic BD patients compared with the controls, though the effect sizes were a little smaller [$g = 0.65$ (95% CI 0.13–1.17) and $g = 0.53$ (95% CI 0.02–1.05), respectively] than for the comparison including all patients. However, the BD patients with normal sleep did not differ from controls on either metric, while the BD patients with abnormal sleep did (mean RT: $p < 0.001$; $g = 1.18$, 95% CI 0.64–1.72; lapses: $p < 0.001$; $g = 0.91$, 95% CI 0.39–1.43). This was also the case for the BD patients with abnormal sleep but who were euthymic (mean RT: $p = 0.001$; $g = 1.26$, 95% CI 0.54–1.98; lapses: $p = 0.005$; $g = 1.04$, 95% CI 0.33–1.75). Mean RTs are shown in the various sub-groups in Fig. 1. The differences from controls were consistent across each of the three most common BD sleep phenotypes in this cohort ($p < 0.001$ for the difference in mean RT v. controls for all three – see Fig. 1). There were no significant correlations with age or NART IQ and so ANCOVA controlling for these variables was not performed.

Ex-Gaussian analysis of PVT RTs found a significantly greater μ ($p < 0.05$; $g = 0.52$, 95% CI 0.08–0.97) and τ ($p < 0.01$, $g = 0.65$, 95% CI 0.20–1.10), though not σ , in BD patients compared with controls. The increase in τ is consistent with increased IIV and the higher rate of lapses seen with the raw RT data. There were no significant differences from controls in either the euthymic BD sub-group nor the BD patients with normal sleep, in either μ , σ or τ . However, the BD patients with abnormal sleep had significantly greater μ ($p < 0.01$; $g = 0.79$, 95% CI 0.27–1.30), σ ($p < 0.005$; $g = 0.87$, 95% CI 0.34–1.39) and τ ($p < 0.001$; $g = 0.85$, 95% CI 0.33–1.37). In the sub-group of BD patients with abnormal sleep who were euthymic, σ and τ remained significantly greater than in the control participants with similar magnitude effect sizes. Ex-Gaussian distributions of the PVT RTs are shown in Fig. 2. There were no significant correlations of age or NART-IQ in the control group and therefore ANCOVA controlling for these variables was not performed.

Attention Network Task

The ANT was completed by 35 controls and 44 BD patients. One control was omitted from the analysis as their mean RT and conflict RT were more than three times the interquartile range and they were considered an extreme outlier. Mean ANT RTs showed

Table 1. General characteristics of study participants at baseline (day 1)

Characteristic	Controls (<i>n</i> = 36)	Bipolar disorder (<i>n</i> = 46)	Statistical test		
			χ^2 (df)	Mann-Whitney <i>U</i>	<i>p</i> Value
Female gender, <i>n</i> (%)	25 (69.4%)	31 (67.4%)	0.039 (1)		0.843
Age years: mean (s.d.)	42.8 (11.9)	46.8 (11.1)		670.0	0.140
(Range)	(19–64)	(23–64)			
NART (s.d.)	116.5 (8.0)	119.5 (7.2)		632.5	0.099
(Range)	(97–128)	(97–131)			
Years in full time education (s.d.)	15.3 (3.0)	15.8 (3.5)		755.5	0.495
(Range)	(11–22)	(10–24)			
BMI mean: kg/m ² (s.d.)	25.1 (4.4)	30.0 (6.7)		427.0	<0.001
(Range)	(19.5–36.4)	(21.0–52.0)			
BMI ≥ 30 kg/m ² , <i>n</i> (%)	6 (17.1%)	19 (41.3%)	5.438 (1)		0.02
Smoker, <i>n</i> (%)	3 (8.3%)	6 (13%)	0.459 (1)		0.724
Mean alcohol units per week	8.7 (7.7)	4.9 (9.8)		474.5	0.001
(range)	(0–30)	(0–50)			
HAMD-17 mean (s.d.)	0.3 (0.6)	9.1 (7.2)		42.0	<0.001
(Range)	(0–2)	(0–35)			
Remission (HAMD ≤ 7) <i>n</i> (%)	36 (100.0%)	25 (54.3%)			
BDI (s.d.)	0.7 (1.8)	12.2 (11.5)		116.5	<0.001
(Range)	(0–8)	(0–49)			
Remission (BDI ≤ 8) <i>n</i> (%)	36 (100%)	24 (52.2%)			
YMRS mean (s.d.)	0.1 (0.4)	0.9 (2.2)		728.0	0.128
(Range)	(0–2)	(0–10)			
PSQI Global score mean (s.d.)	2.3 (1.2)	8.6 (4.6)		125.0	<0.001
(Range)	(0–4)	(1–18)			
PSQI > 5 <i>n</i> (%)	0 (100%)	30 (65.2%)			
ESS mean (s.d.)	3.8 (2.4)	6.2 (4.9)		575.5	0.018
(Range)	(0–9)	(0–21)			
ESS ≥ 10	0 (0%)	9 (19.6%)			

Table 2. Comparison of characteristics of participants with bipolar disorder (BD) in different sub-groups

Characteristic	Euthymic BD patients (<i>n</i> = 24)	Non-euthymic BD patients (<i>n</i> = 22)	BD with normal sleep (<i>n</i> = 16)	BD with abnormal sleep (<i>n</i> = 28)	Euthymic BD with abnormal sleep (<i>n</i> = 11)
Female gender, <i>n</i> (%)	17 (70.8)	14 (63.6)	12 (75.0)	19 (67.9)	8 (72.7)
Age years: mean (s.d.)	47.1 (11.4)	46.1 (11.0)	44.1 (10.3)	49.2 (10.8)	50.2 (12.1)
NART (s.d.)	119.9 (5.9)	119.0 (8.5)	120.4 (7.1)	119.4 (6.7)	118.5 (4.2)
BMI mean: kg/m ² (s.d.)	29.8 (7.1)	30.1 (6.4)	28.3 (4.2)	31.0 (7.8)	32.2 (8.9)
HAMD-17 mean (s.d.)	4.5 (2.8)	14.1 (7.1)	5.8 (5.2)	10.1 (7.2)	5.3 (2.2)
BDI (s.d.)	3.7 (2.3)	21.4 (10.4)	6.7 (7.9)	13.8 (11.5)	3.6 (1.9)
YMRS mean (s.d.)	1.1 (2.6)	0.6 (1.8)	1.5 (3.1)	0.5 (1.6)	0.2 (0.4)
PSQI Global score mean (s.d.)	6.4 (3.7)	11.0 (4.4)	6.6 (4.2)	9.9 (4.6)	7.9 (4.3)
ESS mean (s.d.)	5.5 (4.4)	7.1 (5.3)	5.4 (3.6)	7.1 (5.4)	6.8 (5.7)

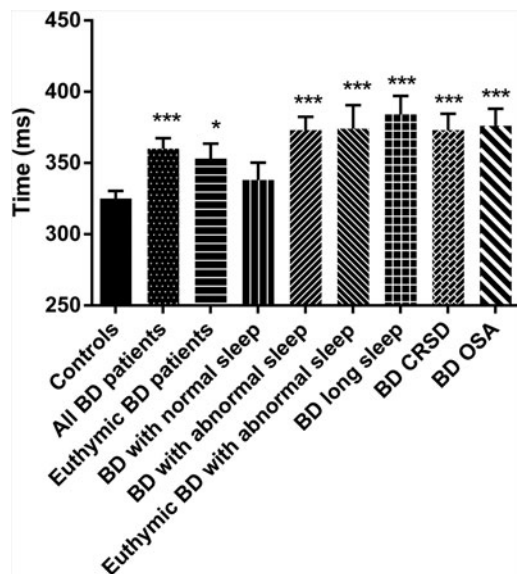


Fig. 1. PVT mean response times. Response times (mean with the error bars representing the SEM) on the PVT task for the healthy controls ($n = 36$) and BD patients. The total BD group is shown ($n = 45$) along with sub-sets of the BD patients including those who were euthymic ($n = 26$); those with normal sleep ($n = 16$); those with abnormal sleep (27); those with abnormal sleep who were euthymic ($n = 11$); those BD patients with long sleep ($n = 13$); those BD patients with CRSD ($n = 12$) and those BD patients with OSA ($n = 12$). * $p < 0.05$; *** $p < 0.001$ compared with controls. BD, bipolar disorder; CRSD, circadian rhythm sleep disorder; OSA, obstructive sleep apnoea; PVT, Psychomotor Vigilance Task; SEM, standard error of the mean.

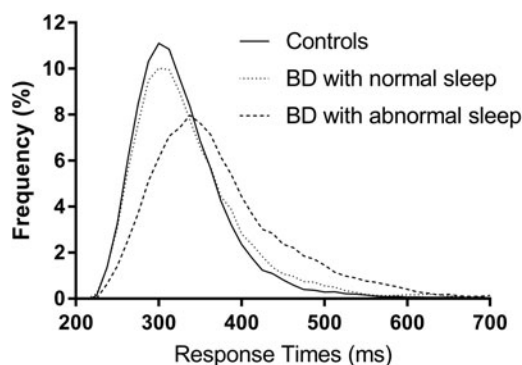


Fig. 2. PVT response time distributions. Response time distributions plotted for controls ($n = 36$), BD patients with normal sleep ($n = 16$) and BD patients with abnormal sleep ($n = 27$). Data plotted in Prism v7.01, 2016 (GraphPad Software Inc., La Jolla, CA, USA) with a smoothed curve based on a rolling average utilising four data points on each side of each data point. BD, bipolar disorder; PVT, Psychomotor Vigilance Task.

a similar pattern of effects to PVT RTs, with BD patients having significantly greater mean RTs than controls, euthymic and BD normal sleepers not differing from controls, but BD abnormal sleepers having significantly greater mean RTs than controls ($p = 0.001$; $g = 0.92$, 95% CI 0.40–1.45), including just those who were euthymic ($p = 0.002$; $g = 1.13$, 95% CI 0.41–1.85). Ex-Gaussian analysis also revealed an identical pattern of effects to that seen with the PVT RT.

The main purpose of the ANT is to assess the processing efficiency of the alerting, orientating and executive attentional networks (Fan *et al.*, 2002). BD patients differed from control participants on the orientating ($p < 0.05$; $g = 0.49$; 95% CI 0.04–0.95) and conflict ($p = 0.005$; $g = 0.68$; 95% CI 0.22–1.14) RTs, the latter representing the executive network. In the comparison between

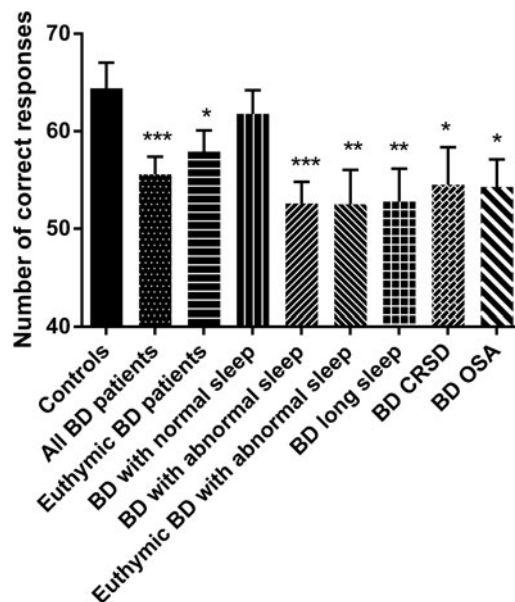


Fig. 3. DSST performance. Number of correct responses made in 90 s (mean with the error bars representing the SEM) for the healthy controls ($n = 36$) and BD patients. The total BD group is shown ($n = 46$) along with sub-sets of the BD patients including those who were euthymic ($n = 26$); those with normal sleep ($n = 16$); those with abnormal sleep (28) and those with abnormal sleep who were euthymic ($n = 11$); those BD patients with long sleep ($n = 14$); those BD patients with CRSD ($n = 12$) and those BD patients with OSA ($n = 12$). * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ compared with controls. BD, bipolar disorder; CRSD, circadian rhythm sleep disorder; DSST, Digit Symbol Substitution Test; OSA, obstructive sleep apnoea; SEM, standard error of the mean.

controls and euthymic BD patients and BD patients with normal sleep, the orientating RTs remained significantly different, though the conflict RTs were not significantly different. However, the BD patients with abnormal sleep differed from controls on the conflict RT with a large effect size ($p < 0.001$; $g = 1.14$; 95% CI 0.61–1.68) but with no difference in alerting or orientating RTs. The euthymic BD abnormal sleepers also differed from controls on the conflict RT ($p < 0.01$; $g = 1.15$; 95% CI 0.43–1.87) but not the other two RTs. This finding suggests that abnormal BD sleepers have an impaired executive attentional network compared with controls and that is not dependent on mood state. This impairment in executive attentional network was also evident in all three BD sleep phenotypes ($p < 0.01$ for all phenotypes).

There was no correlation between age or NART IQ and mean ANT RT, conflict or orientating RT and so no ANCOVA was performed.

Digit Symbol Substitution Test

The DSST was completed at the beginning and end of the cognitive test battery to check for evidence of fatigue. Performance on the second occasion was statistically significantly better, though with only a small effect size. Only data from the first DSST to be performed is reported here, though the findings are identical if using the second DSST data. The findings are shown in Fig. 3. As can be seen, the pattern of effects in the various sub-groups is very similar to that seen for the mean PVT RT shown in Fig. 1. As expected, BD patients made fewer correct responses than the controls ($p < 0.001$; $g = -0.77$; 95% CI -1.22 to -0.32). This was also the case in the euthymic BD patients ($p < 0.05$;

$g = -0.60$; 95% CI -1.11 to -0.08). However, the BD patients with normal sleep did not differ from controls, while those with abnormal sleep did ($p < 0.001$; $g = -0.98$; 95% CI -1.48 to -0.44), including just those who were euthymic ($p < 0.01$; $g = -0.91$; 95% CI -1.61 to -0.21). The sub-groups of BD patients with the three main abnormal sleep phenotypes seen also all different from controls (all $p < 0.05$). There was a significant Pearson's correlations between age and DSST score in controls [$r_{(36)} = 0.689$, $p < 0.001$] and patients [$r_{(46)} = 0.472$, $p = 0.003$]. Age was therefore included as a covariate in ANCOVA analysis of all of the comparisons between BD patient sub-groups and controls. All differences remained significant.

TRAILS A and B

One control participant was not included in this analysis as their TRAILS B – A score was greater than three times the interquartile range and was therefore deemed extreme outlier. There were no significant differences between controls and BD patients in performance on the TRAILS A or B, or the B minus A score. As a result, no sub-group analysis was performed.

Digit span test

There were no significant differences between controls and BD patients in performance on the digit span total scores. Correlation analysis however found a significant Pearson's correlation between NART-IQ and digit span score in controls [$r_{(36)} = 0.484$, $p = 0.003$] and BD patients [$r_{(46)} = 0.302$, $p = 0.042$]. ANCOVA was therefore performed to control for the effects of IQ. This revealed a significant difference between controls and patients [$F_{(1,78)} = 5.833$, $p = 0.018$], including just those who were euthymic [$F_{(1,58)} = 4.159$, $p = 0.046$] but not those with normal sleep [$F_{(1,50)} = 3.355$, $p = 0.073$]. However, patients with abnormal sleep did differ from controls [$F_{(1,60)} = 5.165$, $p = 0.027$] although in this instance this was not seen in the abnormal sleepers who were euthymic [$F_{(1,43)} = 2.717$, $p = 0.107$].

Newcastle Spatial Working Memory Test

BD patients committed significantly more between search errors on the NSWM test than controls ($p < 0.05$; $g = 0.023$, 95% CI 0.07 – 0.99). There was however a moderate correlation between age and between search errors in controls [$r_{(33)} = 0.562$, $p = 0.001$] and a weak correlation in BD patients [$r_{(43)} = 0.300$, $p = 0.051$]. After controlling for age with ANCOVA, the differences between controls and BD patients were no longer significant. As a result, no further analysis was performed.

Verbal learning test

BD patients recalled significantly fewer words on the immediate recall of the verbal learning test ($p = 0.002$; $g = -0.60$, 95% CI -1.05 to -0.16). There was a significant Pearson's correlation between NART-IQ and immediate recall score in BD patients [$r_{(46)} = 0.413$, $p = 0.004$] and a trend towards a significant correlation in controls after winsorising a significant outlier [$r_{(36)} = 0.314$, $p = 0.066$]. ANCOVA covarying for NART-IQ found a significant difference between BD patients and controls [$F_{(1,78)} = 13.999$, $p < 0.001$], which was also seen for the euthymic [$F_{(1,58)} = 4.551$, $p = 0.037$] but not normal sleeping [$F_{(1,48)} = 3.681$, $p = 0.061$] sub-groups. BD abnormal sleepers, however, did differ

from controls [$F_{(1,60)} = 12.806$, $p = 0.001$] though not just those who were in euthymia [$F_{(1,43)} = 1.982$, $p = 0.166$].

Discussion

This is the most comprehensive study to date to assess objectively defined abnormal sleep/circadian rhythm and cognition in BD patients compared with controls. In summary, across a wide range of cognitive assessments, there were the expected significant differences between BD patients as a group compared with age- and sex-matched controls. However, these differences were almost entirely associated with those BD patients with objectively defined abnormal sleep. BD patients with normal sleep did not differ from controls on any cognitive measure with the exception of the ANT orientating network RTs. Given that no corrections were made for multiple comparisons, this could be a type I error. Conversely, BD patients with abnormal sleep differed from controls having longer PVT and mean ANT RTs, greater RT IIV, impaired executive attentional network function, fewer correct responses on the DSST, poorer digit span performance and worse verbal memory recall. While BD patients in our sample with abnormal sleep had lower mood than those with normal sleep (Bradley *et al.*, 2017), BD patients with abnormal sleep but who were euthymic, also differed from the controls on all of these cognitive measures except digit span and verbal memory. Given the relatively small sample size of BD patients with abnormal sleep but who were in euthymia, the lack of statistical significance in the difference in digit span and verbal memory may well be due to a lack of statistical power. While we therefore confirmed cognitive dysfunction in BD, especially in measures of attention, rather than sleep abnormalities exacerbating such dysfunction, the cognitive abnormalities were entirely confined to BD patients with sleep abnormalities. As such, these findings are therefore consistent with sleep disturbance being the main driver of the cognitive abnormalities seen in BD.

Limitations

Weaknesses of the study include a lack of the gold standard measure of sleep using video polysomnography. However, this was a field study and it was felt that studying patients with partial polysomnography in their own homes would increase compliance. The sample was opportunistic with a potential bias for over-representation of patients with BD and sleep disorders, while controls were screened out if they suffered significant sleep problems. A larger, more representative, sample size would have allowed greater sub-group analysis of the differential effect of the different patterns of sleep and circadian rhythm. Future studies might include those with other psychiatric disorders or control subjects with primary sleep disorders to study the impact upon cognition.

Actigraphy assesses physical activity and hence is a surrogate marker of sleep. However, it has been the most widely used and published technique to assess sleep/wake patterns and circadian rhythm for many years (Melo *et al.*, 2016). The American Academy of Sleep Medicine also recognise actigraphy as sufficient, alongside clinical evaluation, to make the diagnosis of circadian rhythm disorders (American Academy of Sleep Medicine, 2014). In BD, sleep variables derived from actigraphy have been shown to highly correlate with gold standard polysomnography (Kaplan *et al.*, 2012). As a result, while the use of actigraphy is a limitation, we believe that it is a legitimate and pragmatic method for assessing sleep.

The patients included in this study were not medication free and it is not possible to exclude the possibility that this influenced our findings. However, there was no difference in the rates of usage of different medications between the different bipolar subgroups with the exception of a significantly higher rate of hypnotic use in long sleepers (42.9%) compared with normal sleepers (6.3%) (Fisher's exact test $p = 0.031$). This finding and details of the medication used is provided in our previous publication (Bradley *et al.*, 2017).

No correction was made for multiple statistical testing. However, it seems unlikely that our findings are type I errors for two main reasons. Firstly, there was consistency in the findings both between measures of the same cognitive domain (attention – PVT and ANT) and across domains. Secondly, if a conservative Bonferroni correction had been used, the main bulk of our findings would have remained significant, including the difference between bipolar patients with abnormal sleep and controls (PVT RTs and lapses, ANT RT, DSST, verbal learning) and between euthymic bipolar patients with abnormal sleep and controls (PVT RTs and lapses, ANT RT). The findings that would not have remained significant are the DSST difference between controls and euthymic bipolar abnormal sleepers and controls and the findings with the digit span test. We did not apply a Bonferroni correction since we were not interested in whether patients and controls differed cognitively in different domains, but if they differed with a particular *a priori* focus on attention and psychomotor processing speed. Applying corrections for multiple comparisons would simply have increased the risk of type II errors. Rather we rely upon describing the statistical tests performed and the pattern of effects seen (Perneger, 1998). For all contrasts, we include estimates of effect size with accompanying 95% CI which are more informative to future research than the point estimates of significance which we agree are linked to sample size and therefore limited in their utility.

The study also had a number of strengths. These include a well-characterised group and comprehensive measures of cognition, sleep and circadian rhythm across 21 days. The control group were carefully matched and a wide range of sleep disturbances was included to reflect real-life clinical practice where many BD patients have a variety of reasons for disturbed sleep. We were fortunate enough to have a reasonable number of both normal and abnormal sleepers, and of euthymic patients in the latter group. This allowed for an examination of our hypotheses without the need for spuriously covarying for mood in comparisons between two groups (patients and controls) that differed on their mood ratings.

In conclusion, we have found that many patients with BD have disrupted sleep and that this has a significant impact upon tests of cognition including in those patients in remission. This suggests that objective measures of sleep disturbance need to be taken into account in future studies of cognition in BD and that sleep may be a potential target for treatment of cognitive disturbance in BD.

Acknowledgements and Financial support. We would like to thank all of the participants for taking part in the study. The study was sponsored by Northumberland, Tyne and Wear NHS Foundation Trust. Some funding to cover running costs was received from Eli Lilly as an unrestricted investigator initiated and led award. The design, conduct and analysis of the study were entirely under the control of the authors.

Conflict of interest. None.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and

institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

- American Academy of Sleep Medicine** (2014) *International Classification of Sleep Disorders*. Darien, Illinois: American Academy of Sleep Medicine.
- American Psychiatric Association** (1994) *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV)*. Washington, DC: American Psychiatric Association.
- Beck AT, Ward CH, Mendelson M, Mock J and Erbaugh J** (1961) An inventory for measuring depression. *Archives of General Psychiatry* **4**, 561–571. Available at <http://www.sciencedirect.com/science/article/B6WVB-4F9MR49-1SB/2/ae36a31f5817e3647fa0ff86fab4a2ea>.
- Bora E** (2018) Neurocognitive features in clinical subgroups of bipolar disorder: a meta-analysis. *Journal of Affective Disorders* **229**, 125–134. Available from: PM:29306692.
- Bourne C, Aydemir O, Balanza-Martinez V, Bora E, Brissos S, Cavanagh JT, Clark L, Cubukcuoglu Z, Dias VV, Dittmann S, Ferrier IN, Fleck DE, Frangou S, Gallagher P, Jones L, Kieseppa T, Martinez-Aran A, Melle I, Moore PB, Mur M, Pfennig A, Raust A, Senturk V, Simonsen C, Smith DJ, Bio DS, Soeiro-de-Souza MG, Stoddart SD, Sundet K, Szoke A, Thompson JM, Torrent C, Zalla T, Craddock N, Andreassen OA, Leboyer M, Vieta E, Bauer M, Vorhusen PD, Tzagarakis C, Rogers RD, Geddes JR and Goodwin GM** (2013) Neuropsychological testing of cognitive impairment in euthymic bipolar disorder: an individual patient data meta-analysis. *Acta Psychiatrica Scandinavica* **128**, 149–162. Available from: PM:23617548.
- Bradley AJ, Webb-Mitchell R, Hazu A, Slater N, Middleton B, Gallagher P, McAllister-Williams H and Anderson KN** (2017) Sleep and circadian rhythm disturbance in bipolar disorder. *Psychological Medicine* **47**, 1678–1689. Available from: PM:28173893.
- Bucks RS, Olaithe M and Eastwood P** (2013) Neurocognitive function in obstructive sleep apnoea: a meta-review. *Respirology* **18**, 61–70. Available from: PM:22913604.
- Buysse DJ, Reynolds III CF, Monk TH, Berman SR and Kupfer DJ** (1989) The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research* **28**, 193–213. Available from: PM:2748771.
- Dinges DF and Powell JW** (1985) Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods* **17**, 652–655.
- Fan J, McCandliss BD, Sommer T, Raz A and Posner MI** (2002) Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience* **14**, 340–347.
- Gallagher P, Gray JM, Watson S, Young AH and Ferrier IN** (2014) Neurocognitive functioning in bipolar depression: a component structure analysis. *Psychological Medicine* **44**, 961–974. Available from: PM:23800475.
- Gallagher P, Nilsson J, Finkelmeyer A, Goshawk M, Macritchie KA, Lloyd AJ, Thompson JM, Porter RJ, Young AH, Ferrier IN, McAllister-Williams RH and Watson S** (2015) Neurocognitive intra-individual variability in mood disorders: effects on attentional response time distributions. *Psychological Medicine* **45**, 2985–2997. Available from: PM:26073667.
- Giglio LM, Magalhaes PV, Kapczinski NS, Walz JC and Kapczinski F** (2010) Functional impact of biological rhythm disturbance in bipolar disorder. *Journal of Psychiatric Research* **44**, 220–223. Available from: PM:19758600.
- Hamilton M** (1967) Development of a rating scale for primary depressive illness. *British Journal of Social & Clinical Psychology* **6**, 278–296.
- Harvey AG, Schmidt DA, Scarna A, Semler CN and Goodwin GM** (2005) Sleep-related functioning in euthymic patients with bipolar disorder, patients with insomnia, and subjects without sleep problems. *American Journal of Psychiatry* **162**, 50–57.
- Johns MW** (1991) A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* **14**, 540–545. Available from: PM:1798888.
- Kanady JC, Soehner AM, Klein AB and Harvey AG** (2017) The association between insomnia-related sleep disruptions and cognitive dysfunction

- during the inter-episode phase of bipolar disorder. *Journal of Psychiatric Research* **88**, 80–88, Available from: PM:28088728.
- Kaplan KA, Gruber J, Eidelman P, Talbot LS and Harvey AG** (2011) Hypersomnia in inter-episode bipolar disorder: does it have prognostic significance?. *Journal of Affective Disorders* **132**, 438–444, Available from: PM:21489637.
- Kaplan KA, Talbot LS, Gruber J and Harvey AG** (2012) Evaluating sleep in bipolar disorder: comparison between actigraphy, polysomnography, and sleep diary. *Bipolar Disorder* **14**, 870–879, Available from: PM:23167935.
- Keller MB** (2003) Past, present, and future directions for defining optimal treatment outcome in depression: remission and beyond. *Journal of the American Medical Association* **289**, 3152–3160, Available from: PM:12813121.
- Kushida CA, Littner MR, Morgenthaler T, Alessi CA, Bailey D, Coleman Jr J, Friedman L, Hirshkowitz M, Kapen S, Kramer M, Lee-Chiong T, Loube DL, Owens J, Pancer JP and Wise M** (2005) Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep*, **28**, 499–521, Available from: PM:16171294.
- Lacouture Y and Cousineau D** (2008) How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. *Methods for Psychology* **4**, 35–45.
- Lezak MD, Howieson DB and Loring DW** (2004) *Neuropsychological Assessment*. Oxford, UK: Oxford University Press.
- Lim J and Dinges DF** (2008) Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences* **1129**, 305–322, Available from: PM:18591490.
- Melo MC, Garcia RF, Linhares Neto VB, Sa MB, de Mesquita LM, de Araujo CF and de Bruin VM** (2016) Sleep and circadian alterations in people at risk for bipolar disorder: a systematic review. *Journal of Psychiatric Research* **83**, 211–219, Available from: PM:27661417.
- Nelson HE** (1982) *The National Adult Reading Test (NART): Test Manual*. Windsor, England: NFER Publishing Co.
- Pariante CM, Alhaj HA, Arulnathan VE, Gallager P, Hanson A, Massey AE and McAllister-Williams RH** (2012) Central glucocorticoid receptor-mediated effects of the antidepressant citalopram in humans: a study using EEG and cognitive testing. *Psychoneuroendocrinology* **37**, 618–628.
- Perneger TV** (1998) What's wrong with Bonferroni adjustments. *British Medical Journal* **316**, 1236–1238, Available from: PM:9553006.
- Ratcliff R** (1979) Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin* **86**, 446–461, Available from: PM:451109.
- Robinson LJ, Thompson JM, Gallagher P, Goswami U, Young AH, Ferrier IN and Moore PB** (2006) A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *Journal of Affective Disorders* **93**, 105–115, Available from: PM:16677713.
- Samame C, Martino DJ and Strejilevich SA** (2014) Longitudinal course of cognitive deficits in bipolar disorder: a meta-analytic study. *Journal of Affective Disorders* **164**, 130–138, Available from: PM:24856566.
- Schmiedek F, Oberauer K, Wilhelm O, Suss HM and Wittmann WW** (2007) Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General* **136**, 414–429, Available from: PM:17696691.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R and Dunbar GC** (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. *Journal of Clinical Psychiatry* **59**(suppl. 20), 22–33, Available from: PM:9881538.
- Soreca I, Buttenfield JA, Hall MH and Kupfer DJ** (2015) Screening for obstructive sleep apnea in patients with bipolar I disorder: comparison between subjective and objective measures. *Bipolar Disorder* **17**, 345–348, Available from: PM:25187351.
- Thompson JM, Gallagher P, Hughes JH, Watson S, Gray JM, Ferrier IN and Young AH** (2005) Neurocognitive impairment in euthymic patients with bipolar affective disorder. *British Journal of Psychiatry* **186**, 32–40.
- van Hees VT, Sabia S, Anderson KN, Denton SJ, Oliver J, Catt M, Abell JG, Kivimaki M, Trenell MI and Singh-Manoux A** (2015) A novel, open access method to assess sleep duration using a wrist-worn accelerometer. *PLoS ONE* **10**, e0142533, Available from: PM:26569414.
- Van Someren EJ, Swaab DF, Colenda CC, Cohen W, McCall WV and Rosenquist PB** (1999) Bright light therapy: improved sensitivity to its effects on rest-activity rhythms in Alzheimer patients by application of non-parametric methods. *Chronobiology International* **16**, 505–518, Available from: PM:10442243.
- Walters AS, LeBrocq C, Dhar A, Hening W, Rosen R, Allen RP and Trenkwalder C** (2003) Validation of the International Restless Legs Syndrome Study Group rating scale for restless legs syndrome. *Sleep Medicine* **4**, 121–132, Available from: PM:14592342.
- Wechsler D** (1981) *WAIS-R Manual, Wechsler Adult Intelligence Scale-Revised*. Cleveland, OH: Psychological Corp.
- Williams JB, Kobak KA, Bech P, Engelhardt N, Evans K, Lipsitz J, Olin J, Pearson J and Kalali A** (2008) The GRID-HAMD: standardization of the Hamilton Depression Rating Scale. *International Clinical Psychopharmacology* **23**, 120–129.
- Young RC, Biggs JT, Ziegler VE and Meyer DA** (1978) A rating scale for mania: reliability, validity and sensitivity. *British Journal of Psychiatry* **133**, 429–435, Available from: PM:728692.