Psychological Medicine (2012), **42**, 1409–1416. © Cambridge University Press 2011
doi:10.1017/S0033291711002662

ORIGINAL ARTICLE

Affective symptoms and intra-individual variability in the short-term course of cognitive functioning in bipolar disorder

C. A. Depp[1]*, G. N. Savla[1], L. A. Vergel de Dios[1], B. T. Mausbach[1] and B. W. Palmer[1,2]

[1] *Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA*
[2] *Veterans Medical Research Foundation, VA San Diego, La Jolla, CA, USA*

Background. Few studies have examined the short-term course of cognitive impairments in bipolar disorder (BD). Key questions are whether trajectories in symptoms covary with cognitive function and whether BD is associated with increased intra-individual variability in cognitive abilities.

Method. Forty-two out-patients with BD and 49 normal comparison (NC) subjects were administered a battery of neuropsychological tests at baseline, 6, 12 and 26 weeks, along with concurrent ratings of depressive and manic symptom severity. Mixed-effects regressions were used to model relationships between time, diagnosis and symptom severity on composite cognitive performance. Within-person variance in cognitive functioning across time was calculated for each subject.

Results. BD patients had significantly worse performance in cognitive ability across time points, but both groups showed significant improvement in cognitive performance over repeated assessments (consistent with expected practice effects). BD was associated with significantly greater intra-individual variability in cognitive ability than NCs; within-person variation was negatively related to baseline cognitive ability in BD but not NC subjects. Changes in affective symptoms over time did not predict changes in cognitive ability.

Conclusions. Moderate changes in affective symptoms did not covary with cognitive ability in BD. The finding of elevated intra-individual variability in BD may reduce capacity to estimate trajectories of cognitive ability in observational and treatment studies.

Received 6 May 2011; Revised 13 October 2011; Accepted 25 October 2011; First published online 9 December 2011

Key words: Bipolar disorder, cognitive functioning, depression, intra-individual variability, longitudinal statistical analysis, neuropsychology.

Introduction

Cognitive impairments are evident in 40–60% of people with bipolar disorder (BD) (Bearden *et al*. 2001; Burdick *et al*. 2006; Torres *et al*. 2007) and contribute substantially to functional disability among people with the illness (Jaeger *et al*. 2007; Bowie *et al*. 2010; Burdick *et al*. 2010). There is no single profile of cognitive deficits, but affected domains typically include sustained attention, working memory, verbal learning and memory and executive control (Torres *et al*. 2007; Arts *et al*. 2008). Although a hallmark feature of BD is its fluctuating course, there are few longitudinal studies that have examined within-person trajectories of cognitive abilities (Goodwin *et al*. 2008). Two key issues in regard to the short-term course of cognitive abilities are (1) the degree to which fluctuations in affective symptoms affect cognitive performance, and (2) whether performance on cognitive tests is more variable within subjects over time in BD compared to normative samples. Greater understanding of the predictors and pattern of cognitive abilities over time in BD is especially important in light of recently proposed clinical research on pharmacologic and non-pharmacologic treatments to ameliorate cognitive deficits associated with this illness (Burdick *et al*. 2007; Harvey *et al*. 2010).

The magnitude of the influence of affective symptoms on cognitive performance in BD is unclear, although many studies attempt to isolate the effects of state variation in affective symptoms from cognitive abilities by including only euthymic patients (Torres *et al*. 2007; Arts *et al*. 2008). Cross-sectional studies contrasting cognitive ability across depressed, manic and euthymic patients reveal greater cognitive deficits

* Address for correspondence: C. A. Depp, Ph.D., University of California, San Diego, Department of Psychiatry, 9500 Gilman Drive (0664), La Jolla, CA 92093-0664, USA.
(Email: cdepp@ucsd.edu)

in some states. For instance, Martinez-Aran *et al*. (2004) found that verbal fluency was more impaired among depressed patients than euthymic patients. Furthermore, in a sample of patients in euthymic, manic or depressed states, Malhi *et al*. (2007) found impairments in executive functioning in manic and depressed patients that were not present in euthymic patients; they also found that cognitive impairment correlated with functional disability only among symptomatic patients. However, in the few longitudinal studies that have assessed the impact of symptoms on change in cognitive abilities over 1- to 3-year periods (Depp *et al*. 2008; Arts *et al*. 2011; Chaves *et al*. 2011), the relationship between change in affective symptoms and cognitive trajectories was small or non-significant. Therefore, cross-sectional studies are incompatible with longitudinal studies on the issue of the impact of symptoms on cognitive abilities.

In addition to trajectories of cognitive functioning, a second dimension of the longitudinal course of cognitive abilities in BD is within-person instability or fluctuation in performance. Although longitudinal studies indicate that cognitive deficits persist over time in BD (Mur *et al*. 2008), greater within-person variability over time in mean levels of cognitive functioning has been seen in patients with BD relative to both healthy comparators and to patients with schizophrenia (Burdick *et al*. 2006; Depp *et al*. 2008). The clinical significance of instability in cognitive abilities in BD remains unclear, but related work has suggested that instability in cognitive functioning is evident in other neuropsychiatric illnesses, including people with attention deficit hyperactivity disorder, multiple sclerosis and age-associated cognitive decline (MacDonald *et al*. 2006). Greater within-person variation over serial cognitive measurements is linked with neurodegeneration, in particular reduced frontal grey matter density and white matter abnormalities (MacDonald *et al*. 2006). There has been some speculation that the cognitive deficits in BD arise from neurodegenerative processes (Krabbendam *et al*. 2000; Goodwin *et al*. 2008). Thus, within-person variation may represent a trait-like aspect of BD, possibly related to alterations in brain structure.

As the field moves toward clinical trials of pharmacologic and non-pharmacologic treatments of cognitive impairments, it is important to establish normative naturalistic trajectories of neuropsychological test performance over periods commensurate with the duration of typical clinical trials (Heaton *et al*. 2001). Nevertheless, a major limitation of research to date is that all but one of the studies have included only two serial cognitive assessments per person (the exception is Arts *et al*. 2011). No study to our knowledge has examined performance on repeated neuropsychological assessments among people with BD within the time frame proposed for interventions targeting cognitive abilities: according to the US Food and Drug Administration (FDA) and the National Institute of Mental Health (NIMH), the minimum duration is 6 months (Buchanan *et al*. 2005). In the present study, we assessed the short-term course of cognitive functioning in 42 adult out-patients with BD type I or II, in comparison to 49 demographically matched adults without psychiatric illnesses. Based on findings from cross-sectional studies and also those using single follow-up assessments, we hypothesized that slopes in depressive and manic symptoms would account for significant variation in slopes of cognitive ability, and that patients with BD would display greater within-person variation, adjusting for mean cognitive ability, in cognitive tests compared to the normal comparison (NC) subjects.

Method

Sample characteristics

The present data were collected as part of subjects' participation in a larger ongoing study of cognitive and other determinants and short-term course of capacity to consent to research among patients with BD. Participants included in the present analyses were 42 out-patients with a diagnosis of either BD I or II and 49 NC subjects. They were assessed at baseline, and at 6, 12 and 26 weeks of follow-up. The design and duration of the study was patterned after a typical randomized clinical trial, although no medications or other treatments were provided as a part of participating in the study. BD subjects were recruited from out-patient psychiatry clinics at the University of California, San Diego (UCSD) and the Veterans Affairs San Diego Healthcare System, and from local area assisted living (board and care) facilities (two subjects with BD resided in board and care homes). NC subjects were recruited through flyers posted in newspapers, online, or in the community. All subjects enrolled in the study provided written informed consent for participation, including a signed printed consent form reviewed and approved by the UCSD Human Research Protections Program and by the Institutional Review Board of the San Diego Veterans Administration Healthcare System.

DSM-IV BD I or II diagnosis, or absence of current or past psychiatric disorder for NCs, was established with the Structured Clinical Interview for the DSM-IV-TR (SCID; First *et al*. 2002). The SCID was conducted by trained research associates and confirmed in consensus meetings in which the SCID responses and other available information were presented to a

https://doi.org/10.1017/S0033291711002662 Published online by Cambridge University Press

board-certified psychiatrist. Participants in the parent study were included if they were fluent in English and were able to give written consent to participate in the study. Given that the present study addressed longitudinal change and within-person variability, we included participants with at least two of the four assessments, therefore excluding data from 25 patients with BD and 20 NC subjects in the larger dataset at the time these analyses were initiated (this was an ongoing study and some of the excluded participants had not yet undergone the follow-up evaluations). Potential participants were excluded from the parent study if they met DSM-IV criteria for substance dependence within 2 months preceding enrollment, had been diagnosed with dementia or other neurological/medical conditions known to affect cognitive functioning, or had medical problems that hindered their ability to complete the assessments. Three of the bipolar participants in the larger dataset were diagnosed with BD not otherwise specified (NOS) and were excluded from the analysis.

BD and NC participants completed an average of 3.4 assessments (S.D. = 0.7) with no differences between groups in the number of assessments completed [$F(1, 90) = 0.3$, $p = 0.585$]. Comparing participants who were included in the analyses to those excluded due to only one assessment, the 25 BD patients with only one assessment were significantly older than the 42 with two or more time points [$F(1, 68) = 6.6$, $p = 0.012$]. The 20 NCs who were excluded had a lower educational attainment than the 49 included in the present analyses [$F(1, 68) = 4.6$, $p = 0.046$]. No differences between included and excluded NCs or patients with BD were evident in sex, ethnicity, education or manic or depressive symptom severity.

Patients were assessed when in various states, as no restrictions were placed on symptom status at study entry. A total of 86% met criteria for BD I and 14% for BD II. Based on the SCID, 52% of patients were euthymic, 17% were depressed and 17% were exhibiting a hypomanic or mixed episode at the time of the baseline assessment. A total of 43% had a prior history of substance use disorders (none had current substance use disorders according to exclusion criteria). Seventy-one percent of the patients were taking a mood stabilizer, 43% an antipsychotic and 40% an antidepressant; 14 were on both antipsychotics and antidepressants, 26% on both antidepressants and mood stabilizers and 10% on all three. Seventeen participants were taking a second-generation antipsychotic and one a first-generation antipsychotic. Thirteen participants were taking a selective serotonin reuptake inhibitor (SSRI) or a serotonin–norepinephrine reuptake inhibitor (SNRI) and four were taking a tricyclic antidepressant.

Assessments

Demographic information and clinical history were collected from all participants at the baseline visit. Premorbid verbal ability was estimated with the Word Reading subtest of the Wide Range Achievement Test – 4th edition (WRAT-4; Wilkinson & Robertson, 2006).

Affective symptoms

The presence and severity of manic symptoms (among both BD and NC subjects) was measured with the Young Mania Rating Scale (YMRS; Young *et al*. (1978)). The presence and severity of depressive symptoms was assessed with the clinician-rated Montgomery–Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979). Symptoms were rated at all four study visits.

Cognitive ability

Participants' cognitive functioning was assessed at all four study visits with the same battery of neuropsychological tests. The neuropsychological battery comprised of the following tests: Digit Symbol-Coding and Letter-Number Sequencing from the Wechsler Adult Intelligence Scale – Third Edition [WAIS-IIII (Heaton *et al*. 2001) total correct]; a selected subtest from the Delis–Kaplan Executive Function System (D-KEFS; Delis *et al*. 2001), that is the Trail Making Test (total time for visual scanning, number sequencing, letter sequencing, number-letter switching, and motor speed), Verbal Fluency (total correct for letter and category fluency, and category switching), and Color-Word Interference (total time for the inhibition and inhibition/switching subtests); the Rey Tangled Lines Test (RTLT; total correct); and the Hopkins Verbal Learning Test – Revised [HVLT-R (Shapiro *et al*. 1999); total correct across three learning trials].

We calculated z scores for each of the tasks based on performance in the NC sample (when needed, scores were reflected such that high z values indicated better cognitive functioning across all tests). We then calculated a composite score from the z scores as our cognitive functioning indicator for our analysis. The composite was only calculated when scores from at least 10 of the 14 tests were available. None of the composite scores were excluded as a result of having less than 10 of the tests available. In addition, there were no differences between BD and NC samples in the number of available tests at baseline [$F(1, 90) = 0.120$, $p = 0.730$] or at the 6-week [$F(1, 90) = 1.4$, $p = 0.232$], 12-week [$F(1, 78) = 1.6$, $p = 0.205$] or 26-week follow-up [$F(1, 63) = 0.108$, $p = 0.744$]. The reliability of

https://doi.org/10.1017/S0033291711002662 Published online by Cambridge University Press

the composite score in BD and NC groups was high (Cronbach's $\alpha = 0.934$ and 0.914 respectively).

Statistical analyses

All variables were assessed for normality and transformations were conducted when necessary. Baseline demographic characteristics of NC *versus* BD groups were compared with ANOVAs for continuous variables and Pearson χ^2 for categorical variables. Demographic characteristics on which the two groups differed were used as covariates in subsequent analyses. We used linear mixed effects models to assess for linear trajectories in cognitive abilities over time. In our first model, we used random intercepts and slopes for participants, with time (in months) entered as our sole predictor. In this model, baseline values for time were coded as '0'. A second model added diagnostic group (to determine whether groups differed in cognitive performance across the study period) and a time × diagnostic group interaction to determine whether slopes in cognitive ability differed between diagnostic groups. A third model added time-varying values for MADRS and YMRS scores into the model to covary for their correlation with cognitive functioning across time. Finally, restricting the analysis to only the BD group, we conducted two final models to determine whether within-person change in MADRS or YMRS scores correlated with change in cognitive ability over time. As an example, the regression equation for the MADRS model was:

$$NP_{ij} = \pi_{oi} + \pi_{1i}(MADRS_{ij}) + e_{ij},$$

where NP refers to neuropsychological performance and MADRS refers to depression scores. In this model, each person's mean MADRS score was subtracted from each of their actual observations of these variables at the four assessments. The intercept, π_{oi}, is therefore participant i's neuropsychological performance score at their average level of depression. NP_{ij} is participant i's level of cognitive ability at assessment j. The slope, π_{1i}, is the change in participant i's cognitive ability for every one unit increase in depression. e_{ij} is the error term for person i. Each person therefore has unique regression parameters, representing their own relationship between depression and cognitive performance. We replicated this exact model using YMRS scores for our final model.

Intra-individual variability can be represented in terms of dispersion across tests with a testing session (i.e. dispersion) or variability relative to the within-person mean performance over time. We looked on intra-individual variability over time because we were focused on the predictors, pattern and variability of change within-person over time. We calculated the within-person standard deviation (W.S.D.) for each subject, and then conducted an ANCOVA comparing within-person variance between the NC and BD groups with mean cognitive ability as a covariate. Finally, we explored whether demographic characteristics, baseline cognitive ability and within-person variation in manic and depressive symptom severity predicted W.S.D. in cognitive ability.

Results

Sample characteristics (*Table 1*)

The NC and BD samples did not differ from one another in age, sex, ethnicity, education, or in estimated pre-morbid verbal ability (WRAT-4 reading scores). The BD group had lower mean cognitive composite scores and, as expected, more severe depressive and manic symptoms (MADRS and YMRS respectively). The level of severity of affective symptoms in the bipolar sample was, on average, in the mild range for both depression and manic symptom scores. In terms of the amount of variability over time in symptoms, the average patient experienced a range across time points of 7 points on the YMRS and 12.8 points on the MADRS. The mean age of onset of BD was 21.6 (S.D. = 9.06) years.

Trajectory and fluctuations in cognitive ability

Fig. S1 (see online Supplementary Material) displays the individual trajectories of performance on composite cognitive ability across the four time points in (*a*) BD and (*b*) NC samples. In the first analysis predicting composite cognitive ability, the single fixed effect visit (time) was positive and significant [estimate = 0.007, standard error (S.E.) = 0.001, $t = 6.6$, $p < 0.001$], indicating that cognitive performance improved over time. In the next analysis, diagnostic group and diagnostic group × visit were entered. Diagnostic group was associated with a significant main effect, with lower performance in the BD group by about $\frac{1}{2}$ S.D. (estimate = −0.479, S.E. = 0.122, $t = 3.9$, $p < 0.001$). However, the diagnosis by time interaction was not significant; that is, both groups improved in cognitive test performance over the repeated assessments (estimate = −0.002, S.E. = 0.002, $t = 1.0$, $p = 0.317$).

Impact of affective symptoms on cognitive ability

The addition of depressive symptoms (MADRS scores) and manic symptoms (YMRS scores) into the model revealed that neither MADRS score (estimate = 0.0001, S.E. = 0.002, $t = 0.052$, $p = 0.959$) nor YMRS score (estimate = −0.005, S.E. = 0.003, $t = 1.3$,

https://doi.org/10.1017/S0033291711002662 Published online by Cambridge University Press

Table 1. *Characteristics of normal comparison (NC) and bipolar disorder (BD) groups*

Variable	NC ($n=49$)	BD ($n=42$)	Test statistic F (df) or χ^2 (df)	p value
Age (years), mean (S.D.)	40.1 (14.6)	43.8 (11.6)	1.8 (1, 89)	0.181
Sex (% female)	57.1	40.5	2.5 (1)	0.113
Ethnicity (%)			0.682 (1)	0.409
Caucasian	63.3	71.4		
African American	4.1	9.5		
Hispanic/Latino	8.2	11.9		
Asian	12.2	0		
Other	12.2	7.2		
Education (years), mean (S.D.)	15.9 (2.2)	15.0 (2.4)	3.3 (1, 89)	0.071
MADRS score (baseline), mean (S.D.)	1.4 (2.3)	14.5 (11.1)	64.7 (1, 88)	<0.001
YMRS score (baseline), mean (S.D.)	1.3 (1.6)	8.6 (8.3)	35.0 (1, 87)	<0.001
WRAT-4, reading subtest, mean (S.D.)	97.6 (13.5)	98.3 (11.3)	0.6 (1, 79)	0.893
Cognitive composite score, mean (S.D.)				
Baseline	0.2 (0.5)	−0.34 (0.7)	14.8 (1, 89)	<0.001
6 weeks	0.2 (0.6)	−0.2 (0.7)	11.9	0.001
12 weeks	0.3 (0.6)	−0.2 (0.7)	15.1	<0.001
26 weeks	0.3 (0.5)	−0.2 (0.7)	11.9	0.001

MADRS, Montgomery–Asberg Depression Rating Scale; YMRS, Young Mania Rating Scale; WRAT-4, Wide Range Achievement Test 4; df, degrees of freedom; S.D., standard deviation.

$p=0.174$) was significant. In addition, the main effect of diagnosis remained significant after entering MADRS and YMRS scores into the model, indicating that between-group differences in affective symptoms did not account for the difference in cognitive ability between bipolar and NC groups. To examine the impact of change in affective symptoms and cognitive ability, we examined within-subject change in time-varying MADRS and YMRS scores in the bipolar group only. There was no significant relationship between changes in cognitive ability and changes in manic symptoms (estimate $=-0.007$, S.E. $=0.004$, $t=1.4$, $p=0.098$) or depressive symptoms (estimate $=-0.002$, S.E. $=0.002$, $t=0.7$, $p=0.452$). The correlations between cognitive ability and depressive symptoms (MADRS scores) were all non-significant: baseline: $r=0.052$, $p=0.746$, 6 weeks: $r=-0.090$, $p=0.561$, 12 weeks: $r=0.013$, $p=0.940$, and 26 weeks: $r=-0.027$, $p=0.888$. For manic symptoms (YMRS scores), the correlations were also non-significant: baseline: $r=0.008$, $p=0.958$, 6 weeks: $r=-0.060$, $p=0.699$, 12 weeks: $r=-0.049$, $p=0.665$, and 26 weeks: $r=-0.130$, $p=0.503$.

Intra-individual variability in cognitive functioning

The mean W.S.D. was significantly greater among BD *versus* NC subjects (W.S.D. $=0.13$, S.D. $=0.05$ *versus* W.S.D. $=0.10$, S.D. $=0.04$ respectively), adjusting for the mean cognitive composite score [$F(1,89)=5.1$, $p=0.026$]. However, it is notable that the mean composite score was differentially related to intra-individual variability across groups. In the BD sample, W.S.D. was significantly negatively correlated with baseline cognitive performance ($r=-0.308$, $p=0.047$). Among NCs, the relationship with baseline cognitive performance was significant and in the opposite direction from that in the bipolar sample ($r=0.355$, $p=0.012$). To determine whether these associations were significantly different across groups, we examined whether the interaction of diagnostic group with mean cognitive performance and diagnostic group was significant in predicting intra-individual variability. We found that the interaction term was significant [$F(1,89)=9.1$, $p=0.003$], indicating that cognitive ability was more negatively related to W.S.D. in the bipolar group. Finally, neither baseline manic nor depressive symptoms, nor within-person variation in depressive or manic symptoms, were related to W.S.D.

Discussion

The present study revealed several potentially notable aspects of the short-term course of cognitive abilities in BD. Consistent with cross-sectional studies contrasting BD and NC groups on cognitive ability, BD patients had worse overall cognitive performance than NC subjects across each of four assessments over a 6-month period. The effect size of the group mean differences was approximately 0.5 (Cohen's *d*), which

https://doi.org/10.1017/S0033291711002662 Published online by Cambridge University Press

is consistent with prior meta-analyses reporting comparisons on composite measures of cognitive ability (Torres *et al*. 2007; Arts *et al*. 2008). However, both groups showed slight improvements in mean performance over repeated assessments, as expected from normal practice effects associated with repeated exposure to the test materials; the magnitude of these practice effects did not differ between the NC *versus* BD groups. Contrary to our hypotheses, manic and depressive symptoms had little effect on the overall level of cognitive impairment or the discrepancy between NC and BD patients in cognitive ability. Additionally, change in symptoms was not associated with change in cognitive performance. Intra-individual variability in composite cognitive ability over time was greater in BD patients than among NCs. Taken together, our findings suggest that future research aimed at estimating and contrasting trends in cognitive ability in BD, such as in cognitive remediation trials, would need to account for practice effects (which are similar in magnitude to that in NCs) and intra-individual variability (which is elevated in BD). However, variation in affective symptoms among mildly symptomatic out-patients with BD is unlikely to substantially affect estimates of mean cognitive performance.

Our findings add to the small number of longitudinal studies suggesting relative independence of mood symptoms and cognitive abilities, at least over a 6-month period. One clear caveat to this lack of relationship is that patients with more acutely severe manic or depressive symptoms than present in this out-patient sample may indeed experience additional cognitive deficits, or at least have more difficulty engaging in cognitive testing and thereby showing performance, if not actual, cognitive decrements. Nonetheless, our findings suggest that fluctuations in symptoms at the level characteristic of ongoing out-patient status do not affect cognitive functioning. Given that many studies attempt to isolate cognitive deficits in BD by enrolling only euthymic patients (e.g. Torres *et al*. 2007; Arts *et al*. 2008), a concern for future clinical trials aimed at cognitive remediation is the extent to which affective symptoms might mask treatment effects (see Buchanan *et al*. 2005 for a discussion of this topic in schizophrenia). Based on the present findings, it seems unnecessary to exclude patients from such cognitive trials based on the presence of non-severe manic or depressive symptoms.

Nevertheless, BD patients did evidence significantly greater variability within-person variability in cognitive performance compared to the NCs. Greater variability has been found in prior studies in BD (with only two serial assessments over a longer period of time) in comparison to NC and schizophrenia samples (Burdick *et al*. 2006; Depp *et al*. 2008). There are two primary implications of greater intra-individual variability over time. First, if estimates of cognitive abilities are less reliable over time, the presence of this within-person variability reduces statistical power to detect treatment effects. A conundrum in interpreting these results is whether this variability reflects true variability in neuropsychological abilities (i.e. the underlying constructs) or instead is a reflection of fluctuations in test performance only (i.e. due to measurement error). Establishing a normative pattern of neuropsychological change is necessary to identify what constitutes a meaningful change in neuropsychological test performance among people with BD in the context of treatment (pharmacologic or psychosocial). Recent work suggests that the MATRICS (Measurement and Treatment Research to Improve Cognition in Schizophrenia) battery may be appropriate in terms of identifying deficits in BD relative to NCs (Burdick *et al*. 2011) (although some alterations to this battery for BD have been advised by a panel of international experts; Yatham *et al*. 2010), and so an important next step would seem to be to evaluate test–retest reliability in BD patients and to develop norms for change in this group (Harvey *et al*. 2005; Cysique *et al*. 2011). Moreover, whether variability in practice effects relates to patient characteristics would be important to evaluate; in our study the positive association between baseline cognitive ability and intra-individual variation within the NC group may have arisen from stronger practice effects among higher performing NCs, as has been found previously (Rapport *et al*. 1997).

Second, there is some basis for suspecting that intra-individual variability in cognitive ability in BD may be a trait-like aspect of BD related to the pathophysiology of the illness. Prior work has linked within-person variability with a variety of brain abnormalities, including changes in gray matter density and diminished connectivity in white matter tracts, with primary localization in frontal brain regions and association with executive control-type tasks (MacDonald *et al*. 2006). Without data on brain structure, we cannot comment on the mechanisms of intra-individual variability in BD; however, it was notable that intra-individual variability was more negatively related to baseline cognitive performance in the BD *versus* the NC group. Future research should examine the association with variability over time with executive control tasks, which may be relevant to recent hypotheses related to hypotheses about neurodegeneration in BD (Goodwin *et al*. 2008).

There are several limitations to this study. The present results are based on a modest-sized out-patient sample of BD patients, and the results may not

generalize to those in acute affective states, patients with active substance dependence and/or in-patients. Thus, in other more severely ill samples, the influence of symptoms on trajectories of cognitive abilities may be greater. Moreover, our study does not address whether more pronounced switches between states, for example from euthymic to severe depression, have an impact on cognitive impairment, or the potential impact of psychotic symptoms (both historical and current) on trends in cognitive ability. In future research it would also be useful to examine individual cognitive domains with a more comprehensive battery to determine whether some domains are associated with more variability or practice effects than others. Although our battery evaluated visual attention, we lacked a measure of sustained attention, which has been shown to be impaired in BD. Finally, there are a variety of methods for estimating net within-person variability (Ram & Gerstorf, 2009) that have differing advantages. Within-person S.D. does not remove the effects of trend, which is problematic given correlations between serial test performance (e.g. practice effects). Future studies with more frequent administration of cognitive tests may allow for better indicators of more 'pure' instability (e.g. the mean squared successive difference), enabling better understanding of the relevance of intra-individual variability to BD.

In summary, the course of global cognitive ability over a 6-month period in BD was found to be independent from affective symptom severity, associated with a similar level of practice effects as non-affected people, and more variable within patients over time than among NC subjects. For future clinical trials aimed at remediating cognitive impairment in BD, the development of standardized cognitive batteries used as outcome measures should take into account the likelihood of increase variability (e.g. using the reliable change index plus practice effect method to define meaningful cognitive changes). The sources and clinical significance of intra-individual variability in cognitive ability should be a focus of future research in BD.

Acknowledgments

This study was supported in part by NIMH Grants MH064722 (BWP), MH077225 (CAD), P30MH080002 and MH019934, and by the Department of Veterans Affairs.

Declaration of Interest

None.

Note

Supplementary material accompanies this paper on the Journal's website (http://journals.cambridge.org/psm).

References

Arts B, Jabben N, Krabbendam L, van Os J (2008). Meta-analyses of cognitive functioning in euthymic bipolar patients and their first-degree relatives. *Psychological Medicine* **38**, 771–785.

Arts B, Jabben N, Krabbendam L, van Os J (2011). A 2-year naturalistic study on cognitive functioning in bipolar disorder. *Acta Psychiatrica Scandinavica* **123**, 190–205.

Bearden CE, Hoffman KM, Cannon TD (2001). The neuropsychology and neuroanatomy of bipolar affective disorder: a critical review. *Bipolar Disorders* **3**, 106–150; discussion 151–153.

Bowie CR, Depp C, McGrath JA, Wolyniec P, Mausbach BT, Thornquist MH, Luke J, Patterson TL, Harvey PD, Pulver AE (2010). Prediction of real-world functional disability in chronic mental disorders: a comparison of schizophrenia and bipolar disorder. *American Journal of Psychiatry* **167**, 1116–1124.

Buchanan RW, Davis M, Goff D, Green MF, Keefe RS, Leon AC, Nuechterlein KH, Laughren T, Levin R, Stover E, Fenton W, Marder SR (2005). A summary of the FDA-NIMH-MATRICS workshop on clinical trial design for neurocognitive drugs for schizophrenia. *Schizophrenia Bulletin* **31**, 5–19.

Burdick KE, Braga RJ, Goldberg JF, Malhotra AK (2007). Cognitive dysfunction in bipolar disorder: future place of pharmacotherapy. *CNS Drugs* **21**, 971–981.

Burdick KE, Goldberg JF, Harrow M (2006). Neurocognition as a stable endophenotype in bipolar disorder and schizophrenia. *Journal of Nervous and Mental Disease* **194**, 255–260.

Burdick KE, Goldberg JF, Harrow M, Faull RN, Malhotra AK (2010). Neurocognitive dysfunction and psychosocial outcome in patients with bipolar I disorder at 15-year follow-up. *Acta Psychiatrica Scandinavica* **122**, 499–506.

Burdick KE, Goldberg TE, Cornblatt BA, Keefe RS, Gopin CB, Derosse P, Braga RJ, Malhotra AK (2011). The MATRICS consensus cognitive battery in patients with bipolar I disorder. *Neuropsychopharmacology* **36**, 1587–1592.

Chaves OC, Lombardo LE, Bearden CE, Woolsey MD, Martinez DM, Barrett JA, Miller AL, Velligan DI, Glahn DC (2011). Association of clinical symptoms and neurocognitive performance in bipolar disorder: a longitudinal study. *Bipolar Disorders* **13**, 118–123.

Cysique LA, Franklin Jr. D, Abramson I, Ellis RJ, Letendre S, Collier A, Clifford D, Gelman B, McArthur J, Morgello S, Simpson D, McCutchan JA, Grant I, Heaton RK (2011). Normative data and validation of a regression based summary score for assessing meaningful neuropsychological change. *Journal of Clinical and Experimental Neuropsychology* **33**, 505–522.

Delis D, Kaplan E, Kramer J (2001). *Delis-Kaplan Executive Function System* (*D-KEFS*). The Psychological Corporation: San Antonio, TX.

Depp CA, Savla GN, Moore DJ, Palmer BW, Stricker JL, Lebowitz BD, Jeste DV (2008). Short-term course of neuropsychological abilities in middle-aged and older adults with bipolar disorder. *Bipolar Disorders* **10**, 684–690.

First MB, Spitzer RL, Gibbon M, Williams JBW (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition* (*SCID-I/P*). Biometrics Research, New York State Psychiatric Institute: New York.

Goodwin GM, Martinez-Aran A, Glahn DC, Vieta E (2008). Cognitive impairment in bipolar disorder: neurodevelopment or neurodegeneration? An ECNP expert meeting report. *European Neuropsychopharmacology* **18**, 787–793.

Harvey PD, Palmer BW, Heaton RK, Mohamed S, Kennedy J, Brickman A (2005). Stability of cognitive performance in older patients with schizophrenia: an 8-week test-retest study. *American Journal of Psychiatry* **162**, 110–117.

Harvey PD, Wingo AP, Burdick KE, Baldessarini RJ (2010). Cognition and disability in bipolar disorder: lessons from schizophrenia research. *Bipolar Disorders* **12**, 364–375.

Heaton RK, Temkin N, Dikmen S, Avitable N, Taylor MJ, Marcotte TD, Grant I (2001). Detecting change: a comparison of three neuropsychological methods, using normal and clinical samples. *Archives of Clinical Neuropsychology* **16**, 75–91.

Jaeger J, Berns S, Loftus S, Gonzalez C, Czobor P (2007). Neurocognitive test performance predicts functional recovery from acute exacerbation leading to hospitalization in bipolar disorder. *Bipolar Disorders* **9**, 93–102.

Krabbendam L, Honig A, Wiersma J, Vuurman EF, Hofman PA, Derix MM, Nolen WA, Jolles J (2000). Cognitive dysfunctions and white matter lesions in patients with bipolar disorder in remission. *Acta Psychiatrica Scandinavica* **101**, 274–280.

MacDonald SW, Nyberg L, Backman L (2006). Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. *Trends in Neurosciences* **29**, 474–480.

Malhi GS, Ivanovski B, Hadzi-Pavlovic D, Mitchell PB, Vieta E, Sachdev P (2007). Neuropsychological deficits and functional impairment in bipolar depression, hypomania and euthymia. *Bipolar Disorders* **9**, 114–125.

Martinez-Aran A, Vieta E, Reinares M, Colom F, Torrent C, Sanchez-Moreno J, Benabarre A, Goikolea J, Comes M, Salamero M (2004). Cognitive function across manic or hypomanic, depressed, and euthymic states in bipolar disorder. *American Journal of Psychiatry* **161**, 262–270.

Montgomery SA, Asberg M (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry* **134**, 381–389.

Mur M, Portella MJ, Martinez-Aran A, Pifarre J, Vieta E (2008). Long-term stability of cognitive impairment in bipolar disorder: a 2-year follow-up study of lithium-treated euthymic bipolar patients. *Journal of Clinical Psychiatry* **69**, 712–719.

Ram N, Gerstorf D (2009). Time-structured and net intraindividual variability: tools for examining the development of dynamic characteristics and processes. *Psychology and Aging* **24**, 778–791.

Rapport L, Brines D, Axelrod B, Thiesen M (1997). Full scale IQ as mediator of practice effects: the rich get richer. *The Clinical Neuropsychologist* **11**, 375–380.

Shapiro AM, Benedict RH, Schretlen D, Brandt J (1999). Construct and concurrent validity of the Hopkins Verbal Learning Test – revised. *The Clinical Neuropsychologist* **13**, 348–358.

Torres IJ, Boudreau VG, Yatham LN (2007). Neuropsychological functioning in euthymic bipolar disorder: a meta-analysis. *Acta Psychiatrica Scandinavica* (Suppl.) **434**, 17–26.

Wilkinson GS, Robertson GJ (2006). *Wide Range Achievement Test 4, Professional Manual*. Psychological Assessment Resources: Lutz, FL.

Yatham LN, Torres IJ, Mahli GS, Frangou S, Glahn DC, Bearden CE, Burdick KE, Martinez-Aran A, Dittman S, Goldberg JF, Ozerdem A, Aydemir O, Chengappa KN (2010). The International Society for Bipolar Disorders – Battery for Assessment of Neurocognition (ISBD-BANC). *Bipolar Disorders* **12**, 351–363.

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