Time Course of Recovery Showing Initial Prefrontal Cortex Changes at 16 weeks, Extending to Subcortical Changes by Three years in Pediatric Bipolar Disorder

Hongyu Yang\textsuperscript{a}, Lisa H. Lu\textsuperscript{a,c}, Minjie Wu\textsuperscript{a}, Michael Stevens\textsuperscript{d}, Ezra Wegbreit\textsuperscript{a}, Jacklynn Fitzgerald\textsuperscript{a,b}, Bryn Levitan\textsuperscript{a}, Stewart Shankman\textsuperscript{b} and Mani N. Pavuluri\textsuperscript{a,b,*}

\textsuperscript{a}Pediatric Brain Research and Intervention Center, Institute for Juvenile Research

Department of Psychiatry, University of Illinois at Chicago, Chicago

\textsuperscript{b}Department of Psychology, University of Illinois at Chicago, Chicago

\textsuperscript{c}Roosevelt University, Department of Psychology, Chicago

\textsuperscript{d}Olin Psychiatric Research Center, Connecticut

* Corresponding author:

Mani N. Pavuluri, MD PhD

Institute for Juvenile Research

University of Illinois at Chicago

1747 W Roosevelt Rd (M/C 747)

Chicago, IL 60608

Phone: (312) 413-0064

Fax: (312) 413-0063

Email: mpavuluri@psych.uic.edu
Abstract

Objective: Activation changes at the interface of affective and cognitive systems are examined over a three year period in pediatric bipolar disorder (PBD). Methods: Thirteen participants with PBD and 10 healthy controls (HC) matched on demographics and IQ were scanned at baseline, at 16 weeks, and after three years. All patients received pharmacotherapy based on a medication algorithm. A pediatric affective color matching paradigm was used to probe cognitive processing under emotional challenge. Results: At baseline, in response to emotional vs. neutral words, patients with PBD showed greater activation than HC in the right dorsal lateral prefrontal cortex (DLPFC) and amygdala, ventral lateral prefrontal cortex (VLPFC), bilateral anterior cingulate cortex (ACC), and ventral striatum. Increased activation in DLPFC in the PBD group normalized by 16 weeks. By three years, normalization was observed in VLPFC, ACC, amygdala, and striatum. Limitations: Small sample size renders the present findings preliminary. Conclusions: Greater activation in fronto-striatal and fronto-limbic circuits were observed in unmedicated patients with PBD. Present findings suggest the possibility that DLPFC is most malleable to pharmacological intervention with systematic pharmacotherapy leading to immediate response, which extended to amygdalostriatal and ventral cortical regions at three years. The seminal observation from this study is the prolonged length of recovery time in the normalization of subcortical activity along with their interfacing cortical regions. Findings from this proof of concept study need to be replicated in a larger sample.

Keywords: Pediatric bipolar disorder; Functional magnetic resonance imaging (fMRI); Emotion; treatment effects
1. Introduction

Pediatric bipolar disorder (PBD) is an early onset variant of the bipolar diathesis (Birmaher et al., 2009; Geller et al., 2002; Pavuluri et al., 2005), and patients are often burdened with persistent neurocognitive deficits, despite achieving symptomatic recovery in mood (Pavuluri et al., 2009a). The cognitive deficits in the domains of attention, response inhibition and executive function correlate with poor academic performance (Pavuluri et al., 2005). Emerging findings highlight the impairment at the interface of cognitive and affective brain circuitry that is responsible for combined deficits in executive function and mood stability in PBD (Leibenluft et al., 2007; Passarotti et al., 2010a, 2010b; Rich et al., 2005). One important question, however, is whether the neural circuitry abnormalities in PBD normalize over time due to normative brain development or treatment. Longitudinal studies of neural circuitry abnormalities in PBD are critical as significant changes occur in the brain throughout childhood and adolescence (Casey et al., 2005) and thus, any examination of brain “normalization” in these patients should take into account normal developmental processes. Additionally, longitudinal studies have the potential in aiding with prognosis as well as estimating the utility of the treatment algorithms.

To date, few longitudinal studies of brain function have been conducted in patients with PBD. Gogtay et al. (2007) examined structural brain changes in 9 patients over time and found that adolescents with bipolar disorder exhibited an increase in grey matter density in the left temporal cortex and a decrease in gray matter density in the anterior cingulate cortex. In another study, Bitter et al. (2011) found that over one year, those with PBD did not exhibit the normative increase in amygdala volume seen in healthy controls (HC). Despite these earlier clues from structural neuroimaging studies, little is known about the functional neuroanatomical changes that occur over time in PBD.
Normative development in adolescence has illustrated increased top-down regulation with maturation of fronto-limbic activity (Casey and Jones, 2010; Somerville et al., 2011; Somerville et al., 2010). Our recent studies have revealed that resting state limbic hyperconnectivity is associated with better cognitive and affective circuitry function (Wu et al., 2013) and that increased amygdala engagement in affective circuitry was associated with response to short-term pharmacotherapy (Wegbreit et al., 2011). Untreated patients with PBD have demonstrated hyperactive amygdala, insufficiently regulated by VLPFC (Pavuluri et al., 2007; Pavuluri et al., 2008; Pavuluri et al., 2010). Furthermore, the DLPFC and striatum serve multiple functions including the cognitive modulation of emotion (Badgaiyan, 2010; Hare et al., 2008), attention (Saint-Cyr, 2003) and response inhibition (Padmanabhan et al., 2011). Untreated patients with PBD showed greater activity in the striatum during response inhibition, which normalized with eight weeks of pharmacotherapy (Passarotti et al., 2011a; Pavuluri et al., 2011). It is not clear if the activity of these affective and cognitive fronto-limbic and fronto-striatal brain regions during the steep adolescent developmental curve will continue to improve, partially recover, or remain abnormal in PBD.

The current study utilized the pediatric color matching task that has previously yielded consistently reliable findings (Passarotti et al., 2010c; Pavuluri et al., 2008; Pavuluri et al., 2011; Pavuluri et al., 2010; Wegbreit et al., 2011) to probe the interface of affective VLPFC-amygdala circuitry and the cognitive DLPFC-striatal circuitry in PBD and HC. We followed participants over a three-year developmental period when the patients received systematic pharmacotherapy for optimal recovery. Based on the leads from conventional fMRI (Mayanil et al., 2011; Pavuluri et al., 2011), patients are predicted to show normalization, relative to HC, in cortico-subcortical
fMRI activity. Alternatively, PBD patients might show partial recovery among these circuits with residual impairment.

2. Methods

2.1. Participants
The sample consisted of 13 individuals with PBD (type I and II) and ten IQ and demographically matched HC (Table 1). All patients were unmedicated at baseline and received pharmacotherapy based on a standardized medication algorithm (Pavuluri et al., 2004). Participants were aged 10-18 years and were clinically assessed and scanned at baseline (13.4 ± 2.5 years), at 16 weeks (16 ± 14 weeks), and at three years follow-up (3.2 ± 1.1 years, Table 1). This study was approved by the institutional review board at the University of Illinois at Chicago. Verbal or written assent was obtained from all of the participants in addition to written consent from parents. Inclusion criteria for patients were a diagnosis of BD type I (mixed N = 2, manic N = 6) or type II (N = 5) according to the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV; American Psychiatric Association 1994), and a baseline score greater than 12 on the Young Mania Rating Scale (YMRS) (Young et al., 1978). Inclusion criteria for HC were no current and past DSM-IV diagnosis for axis I disorders or a family history of affective illness, and a score less than 12 on the YMRS at baseline. Exclusion criteria for both groups were substance abuse through urine toxicology screen; serious medical problems (e.g., epilepsy); IQ less than 70 as determined by the Wechsler Abbreviated Scale of Intelligence (WASI) (1999).

2.2. Clinical Assessment
Board-certified child psychiatrists completed the Washington University Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) (Geller et al., 1998) supplemented by the
episode characterization of bipolar disorder from the KSADS – Present and lifetime version (Kaufman et al., 2000). Diagnostic interviews were completed by doctoral-level clinicians with established inter-rater reliability (Cohen’s kappa = 0.94). Information from the interview and all other clinical measures were reviewed to make a consensus clinical diagnosis. The primary clinical measures were the YMRS and the Child Depression Rating Scale-Revised (CDRS-R) (Poznanski et al., 1984).

2.3. Pharmacotherapy

All PBD patients were medication free or were washed out of medication at baseline. Medications used in the treatment paradigm are summarized in electronic supplemental Table S1. Psychostimulants were used in six of the 13 participants, and non-stimulants were used in 2 participants with PBD with comorbid ADHD, at the end of three years. The average number of medications received by each patient was 2.0 ± 1.1 at 16 weeks and 2.8 ± 1.5 at three years.

2.4. Pediatric Affective Color Matching Task

The primary task for the study was the Pediatric Affective Color Matching Task. As described previously (Pavuluri et al., 2008), this paradigm had the advantage of (1) assessing the impact of an emotional challenge on the cognitive control required to match the color; (2) presenting words briefly (200 ms) to limit prefrontal processing of affective valence and semantic meaning; and (3) requiring a relatively simple cognitive judgment and response. Participants were asked to match the color of a word to one of two colored dots next to the word on a display screen (Pavuluri et al., 2008). The target words were negative (themes such as depression, disappointment, and rejection), positive (feelings of positivity, such as happiness, energy, accomplishment, and success), and neutral valence words. Words across affect conditions were
matched on frequency of use (Bradley, 1999; Gilhooly and Logie, 1980; Klein, 1964; Kucera et al., 1967) and were at an eight-year-old reading level.

In each three second trial, a word was presented for 200 ms, followed by a response period of 2800 ms, during which participants pressed a button (left or right) to match the color of the word to the correct color dot that appeared on either side of the word (Fig. S1). Participants were told to “respond as quickly as possible.” Participants were presented with separate blocks (30 s each) of positive, negative, and neutral word conditions (5 blocks of each condition) in a pseudo-random order, with each block consisting of ten trials. Each block was separated by a ten second fixation period (total of 15 fixation blocks), allowing the hemodynamic responses to return to resting level before the next block of trials. Trials were balanced for the location of correct responses (left or right dots) and the colors used. The total duration of the task was approximately ten minutes.

2.5. MRI Protocol

MRI studies were performed using a 3.0 Tesla whole body scanner (Signa, General Electric Medical System, Milwaukee, WI). Functional images were acquired using echo-planar imaging, which is sensitive to regional alterations in blood flow via blood oxygen level-dependent (BOLD) contrast effects. Parameters for functional scans were: TE=25 ms; flip angle=90°; field of view=20×20 cm; acquisition matrix=64×64; TR=2.5 s; 5 mm slice thickness with 1 mm gap. Anatomical images were acquired in the axial plane (three-dimensional spoiled gradient recalled, 1.5 mm thick contiguous axial slices) to coregister and normalize the functional data.

2.6. Clinical and Behavioral Data Analysis
Reaction time and accuracy for words (negative, positive, neutral), time (baseline, 16 weeks, three years), and groups (PBD, HC) were analyzed using repeated measures ANOVAs. Clinical and demographic data were evaluated by $t$ tests and $\chi^2$.

2.7. Image Processing and Data Analysis

The fMRI data were analyzed with FEAT (FMRI Expert Analysis Tool) Version 4.1.8, a tool of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Preprocessing included motion correction, slice-timing correction, spatial smoothing (a 8mm Gaussian filter at full-width-half-maximum), intensity normalization using a grand mean scaling factor, and high-pass temporal filtering (sigma=100.0s) to remove low frequency drift. Brain extraction and registration to high resolution and template (Montreal Neurological Institute, MNI152) images were carried out by BET (Brain Extraction Tool) (Smith, 2002) and FNIRT (FMRIB's Nonlinear Image Registration Tool) (Jenkinson et al., 2002; Jenkinson and Smith, 2001; Klein et al., 2009), respectively. Time-series statistical analysis was performed by FMRIB's Improved Linear Model (Woolrich et al., 2001), in which a general linear model (GLM) framework of the stimulus onset types as explanatory variables, convolved with a hemodynamic response function (HRF) with time derivatives, was used to fit the time series data of each voxel.

During first level analysis, the GLM model was set up with three regressors: negative, positive, and neutral words. First level analyses consisted of contrast images for the negative vs. neutral as well as positive vs. neutral words for each individual. In the higher level analysis, we conducted whole-brain repeated measures ANOVA with two groups (PBD and HC), two word conditions (negative vs. neutral words, positive vs. neutral words), and two time points. Separate analyses evaluated three combinations of two time points (baseline to three years, baseline to 16 weeks, and 16 weeks to three years). Z statistical images were thresholded using clusters determined by
Z > 2.3 and a corrected cluster significance threshold of \( p < 0.05 \) based on Gaussian random field (GRF) theory (Forman et al., 1995; Friston et al., 1993; Worsley et al., 1992). The relationship between brain responses and clinical measurements was assessed voxelwise in the whole brain by using YMRS and CDRS-R scores as covariates in regression analyses.

3. Results

3.1. Clinical and Demographic Characteristics

Demographic and clinical measurements are summarized in Table 1. No significant differences were found between PBD and HC groups for age, IQ, gender, race, and handedness. For the YMRS, there was a significant group by time interaction \( [F(2, 36) = 14.92, p < 0.01] \). Follow-up analyses indicated that the YMRS scores in the PBD group decreased from baseline to three years \( [t(12) = 8.90, p < 0.01] \), but they remained higher than the HC’s YMRS scores at both time points [PBD vs. HC at baseline: \( t(21) = 10.65, p < 0.01 \); PBC vs. HC at three years: \( t(21) = 4.11, p < 0.01 \)]. For the CDRS-R, there was also significant group by time interaction \( [F(2, 36) = 3.40, p < 0.05] \). Analogous to the YMRS, the CDRS-R scores in the PBD group tended to decrease from baseline to three years, but remained higher than the HC group, both at baseline \( [t(21) = 4.03, p < 0.01] \) and at the three year follow-up \( [t(21) = 3.96, p < 0.01] \).

3.2. Behavioral Data

Table S2 reports the average response time and accuracy for each group across time points. The ANOVAs revealed a main effect of group, reflecting slower response times for the PBD group compared to HC \( [F(3, 16) = 4.13, p < 0.05] \), as well as a main effect of time reflecting faster
response times at 3-year follow-up compared to baseline \(F(6, 13) = 3.45, p < 0.05\). No significant findings were found for accuracy at any time point (Table S2).

3.3. Imaging Data

Our focus was on the areas of the brain that responded differently between groups (PBD and HC) at two time points (baseline and three years), and whether this pattern was affected by the valence of words (negative, positive, neutral). The group by time by valence ANOVA revealed no three-way interactions, but did reveal a significant group by time (from baseline to three years) interaction in the right DLPFC [Broadmann Area (BA) 46], VLPFC [BA 47, right BA 45], bilateral anterior cingulate cortex (ACC), ventral striatum, and right amygdala (Figure 1A and Table 2). Follow-up analyses suggested that at baseline, the PBD group exhibited higher activation in the right DLPFC \(t(21) = 2.34, p = 0.03\) and bilateral ventral striatum \(t(21) = 2.65, p = 0.02\) (Fig. 1D). Interestingly, at three-year follow-up, the groups did not significantly differ on activation in these regions. We followed up above results to evaluate how things changed at 16 weeks. There was also a significant interaction of group by time (from baseline to 16 weeks) in the right DLPFC [Broadmann (BA) 46] and occipital cortex (Figure 1B and Table 2). From 16 weeks to three years, no interactions of group by time were found.

3.4. Correlation between activation and clinical measurements

Within subject activation difference between baseline and 3 years was correlated with treatment response, defined as change in YMRS and CDRS-R scores between these time periods. No significant relationship between neural change and treatment response was found in whole-brain correlational analyses.
4. Discussion

This is the first fMRI study to examine both the short (16 weeks) and long-term (three years) course of functional activity and response to treatment in PBD, revealing two distinct BOLD response patterns. Baseline pathophysiology in patients with PBD, relative to HC, is characterized by greater activation in the emotional and cognitive processing regions that included the right DLPFC, amygdala, VLPFC, bilateral ventral striatum, and ACC. After a 16 week short-term pharmacotherapy, heightened activation in the right DLPFC was reduced in the PBD group. After three years of treatment, subcortical regions including right amygdala and bilateral striatum showed normalization in the PBD group in addition to the ventral emotional and cognitive control regions that influence these subcortical areas i.e., right VLPFC and bilateral pregenual ACC. Behavioral performance improved in all participants over time, but patients demonstrated longer response time, relative to HC, at all times points.

4.1 Cortical and subcortical changes after pharmacotherapy in PBD

Bilateral striatum and right amygdala, intricately connected by amygdalostraiatal projections that extend to the ACC forming a limbic striatum (Zorrilla and Koob, 2012), instrumental in emotion processing and motivation (Berns et al., 2001; Hare et al., 2005; Jensen et al., 2003; Shohamy, 2011), took longer period to show normalization with treatment. Indeed, medications were shown to alter activity in subcortical regions unlike psychotherapy effects that are mostly limited to PFC changes in depression (Mayberg, 2002). Abnormally high subcortical activity is noted in untreated patients in both the pediatric and adult BD (Malhi et al., 2004; Pavuluri et al., 2007; Pavuluri et al., 2008; Pavuluri et al., 2009b) and the current study showed reversal in cortico-subcortical circuitry with systematic longer term therapy. It is critical to note that our previous studies of short term treatment showed change in amygdala activity (Passarotti et al., 2011a) or
improved connectivity within frontolimbic circuitry relative to HC (Wegbreit et al., 2011) but not normalization as shown for the first time through the current study.

As a crucial subregion of prefrontal cortex, the DLPFC is involved in several cognitive processes (i.e., shifting attention, working memory, and response inhibition) as well as affect regulation along with it’s interface with VLPFC (Chang et al., 2004; Fichtenholtz et al., 2004; Manes et al., 2002; Phillips et al., 2003). Using the same the pediatric affective color matching task, we were able to replicate the DLPFC normalization with treatment (Pavuluri et al., 2010). However, given that VLPFC and ACC regions took longer to recover along with subcortical regions, one can postulate that these critical affective and cognitive interface regions with dual functionality that are connected to amygdala and striatum directly and indirectly (Kelley et al., 1982), may need persistent medical treatment towards neuronal recovery in this complex illness.

4.2. Limitations

One limitation of the present study was the small sample size. Results should therefore be considered preliminary. However, the strengths of the study include comparison to developmental trajectory of HC who were matched on age, sex, race, handedness, and IQ, where all groups were scanned at three time points. Furthermore, the PBD patients received multiple psychiatric medications (Mean± SD, 3±1), although it is also a strength that they received systematic algorithm based long term pharmacotherapy over three years, from the unmedicated baseline point. This renders our findings generalizable to typical PBD patients that have the opportunity to receive optimal pharmacotherapy without the confound of medication effects at baseline. However, the exact time for subcortical recovery could have been longer than three
months and shorter than three years. Any adjuvant psychotherapy may have also hastened the recovery or worked through alternate brain regions, and pose new questions for future studies.

5. Conclusions

In sum, the present study demonstrated the long-term neural recovery in PBD along with the improvement in clinical symptoms over time. While changes in the DLPFC were evident after short-term treatment, changes in the amygdalo striatal and the interfacing ventral cortical activation were only apparent at the end of three years of pharmacotherapy. Present findings suggest the possibility that DLPFC is more immediately malleable to pharmacological intervention, whereas the subcortical and cortical regions of affective and cognitive interface do not normalize rapidly. Our findings support a longer period of pharmacotherapy in PBD to reverse the abnormal neural pathophysiology. This hypothesis needs to be evaluated with larger sample size involving more frequent followup in a longitudinal design.
Legends for Figures

**Figure S1.** Pediatric Affective Color Matching Paradigm.

**Figure 1.** Group differences in activation during affective color matching paradigm were moderated by time. Panel A) PBD differed from HC in activation of right DLPFC and bilateral ventral striatum between baseline and 3-year follow up. B) Interactions between groups and time were also explored for baseline to 16 weeks (left) and for 16 weeks to three years (right). Significant group differences between baseline to 16 weeks were found, but not between 16 weeks to three years. C) Group difference at each time point is shown to inform interactions. In red are regions where PBD > HC. In blue are regions where HC > PBD. D) Blood-oxygenation level dependent (BOLD) signal changes are plotted for the right DLPFC and bilateral ventral striatum to elucidate the nature of group differences across time. The interaction between time and group was significant \[F(4, 15) = 3.43, p = 0.04\]. The PBD group showed higher levels of activation at baseline relative to HC \[t(21) = 2.34, p = 0.03\] for right DLPFC, \[t(21) = 2.65, p = 0.02\] for bilateral ventral striatum], and over time, their activation normalized and became similar to that of the HC group.
Reference List
