Delineating ADHD and bipolar disorder: A comparison of clinical profiles in adult women

Glenn L. Kitsune, Jonna Kuntsi, Helen Costello, Sophia Frangou, Georgina M. Hosang, Gráinne McLoughlin, Philip Asherson

Keywords: Adults Attention/deeficit Hyperactivity-disorder Bipolar disorder Diagnostic errors Mood disorders

ARTICLE INFO

Article history:
Received 4 June 2015
Received in revised form
30 November 2015
Accepted 20 December 2015
Available online 23 December 2015

Abstract

Objective: Overlapping symptoms can make the diagnostic differentiation of attention-decit/hyperactivity disorder (ADHD) and bipolar disorder (BD) challenging in adults using current clinical assessments. This study sought to determine if current clinical measures delineate ADHD from BD in adults, comparing relative levels of ADHD, BD and emotional lability (EL) symptoms.

Methods: Sixty adult women with ADHD, BD or controls were compared on self-report and interview measures for ADHD symptoms, mania, depression, EL, and impairment.

Results: ADHD interview measures and self-ratings of ADHD symptoms best discriminated between ADHD and BD. Self-report measures of EL and depression showed non-specific enhancement in both clinical groups. BD-specific items may distinguish BD from ADHD if a retrospective time-frame is adopted.

Conclusions: Using measures which capture specific symptoms of ADHD and chronicity/episodicity of symptoms facilitates the delineation of ADHD from BD in adult women.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The diagnostic differentiation of attention-decit/hyperactivity disorder (ADHD) from bipolar disorder (BD) is important for the correct treatment and management of both conditions (Asherson et al., 2014; Atmaca et al., 2009; Galanter et al., 2005; Mosholder et al., 2009). Yet, similarities in symptoms such as restlessness, increased production of speech and distractibility in both conditions and evidence of persistent impulsive behaviours in euthymic BD (Najt et al., 2007; Peluso et al., 2007) make differentiation of the two conditions challenging (Galanter and Leibenluft, 2008; Kent and Craddock, 2003). The emerging evidence indicating high levels of emotional lability (EL) in ADHD (Barkley and Fischer, 2010; Skirrow et al., 2014, 2012; Surman et al., 2013), independent of comorbidity (Skirrow and Asherson, 2013), and the recognition of EL as an associated feature of ADHD (American Psychiatric Association, 2013), further complicate the diagnostic boundaries between ADHD and BD. In BD equivalent prevalence rates are observed in both men and women (Diflorio and Jones, 2010), while the ratio of males to females diagnosed with ADHD is 1:6:1 (Willcutt, 2012), with indications that ADHD persistence and patterns of comorbidity are similar in both genders (Biederman et al., 2011, 2012). However, it is acknowledged that there remains a lack of research into ADHD in females, particularly amongst adults (American Psychiatric Association, 2013).

Meta-analysis examining comorbidity of ADHD and BD in adults identified rates ranging from 5% to 47% (Wingo and Ghaemi, 2007), and studies of familial co-variation indicate that the disorders co-occur at a higher rate than in the general population, suggesting a potential familial relationship between them (Larsson et al., 2013; Skirrow et al., 2012). The existence of juvenile bipolar disorder, now reconceptualised as severe mood dysregulation in DSM-5, and its overlap with ADHD has been hotly debated (Kent and Craddock, 2003; Skirrow et al., 2012). Yet, despite clearer diagnostic conceptualisations in adults, there are few studies comparing the extent to which symptoms are similar or different between ADHD and BD, and address the challenges of delineation in adult populations. The few direct comparisons to date have used self-report measures of ADHD and depression symptoms, which may have limited scope in their potential to delineate the two disorders (Ibanez et al., 2012; Torralva et al., 2011). The
comparative degree and specificity of EL within each disorder is also an important question to clarify, as mood fluctuations are seen as a characteristic feature of BD, and could result in the misdiagnosis of adults with ADHD and high EL.

1.1. Aims of the study

The aims were to determine the potential of current clinical measures to delineate ADHD from BD in adults, comparing relative levels of ADHD, BD and EL symptoms across the two disorders. Based on previous studies in men with ADHD we hypothesised that EL frequently occurs in women with ADHD and for this reason cross-sectional measures of EL will not distinguish between ADHD and BD. We further hypothesise that women with ADHD will present with a significant number of ‘mania’ symptoms due to the overlap in symptom criteria. We propose that the key distinction to be made will be based on episodicity versus chronicity of the symptoms and which might not be easy to determine based on cross-sectional data alone.

2. Methods and materials

2.1. Sample

Participants with BD were recruited from a largely female sample that had previously participated in another research study (Hosang et al., 2012) and the Maudsley Psychosis Clinic. In ADHD, population studies have not reported gender differences in clinical-range symptoms (Das et al., 2012; de Zwaan et al., 2012), although there remains a relatively limited amount of data collected with adult female participants with ADHD. To address this need, and for purposes of sample matching with the BD group, we recruited an all-female sample in this study. Women with ADHD were therefore recruited from the National Adult ADHD Clinic at the Maudsley Hospital. Control participants were recruited from the Mindsearch volunteer database maintained by the Institute of Psychiatry, Psychology and Neuroscience, which comprises several thousand potential participants. Participants were randomly selected from all those meeting recruitment criteria for this study (described below). In total 60 adult women were recruited (20 with ADHD, 20 with BD and 20 control participants). The study received ethical approval by the Camberwell St Giles Research Ethics Committee (Ref: 11/LO/0438) and was conducted according to the Declaration of Helsinki. All participants provided informed consent.

2.2. Diagnosis and recruitment

Fifty-seven women with ADHD, 75 women with BD, and 120 control women matching requirements of age, gender and clinical diagnosis based upon DSM-IV criteria were approached to participate. The ADHD participants met current criteria for combined-type ADHD or inattentive-type ADHD with sufficient past reported symptoms of hyperactivity–impulsivity to have met combined-type criteria during childhood. Participants in the BD group had a diagnosis of Bipolar I Disorder (BD-I) with evidence of a past manic episode lasting one week or more. Eligibility to participate was continued until 20 participants were recruited for each group, as this was calculated to provide 80–90% power to detect a large effect size (0.8) (Table 1). Samples were age-matched at a group level during recruitment. ADHD participants were asked to stop stimulant medication 48-h before research assessments. For ethical reasons, BD participants were not asked to stop taking mood-stabilisers or any anti-psychotic medication they had been prescribed. All participants were asked to refrain from caffeinated drinks and nicotine for two hours prior to the assessment session.

2.3. Procedure

Participants attended a single research session to complete self-report measures and clinical interviews alongside other research evaluations. All participants completed the same set of assessments. For informant ratings, participants were given a questionnaire to take home in a stamped address envelope, for a family member or close friend to complete. Interview ratings were conducted by an experienced researcher (GK), trained by a consultant psychiatrist (PA) with experience of both ADHD and BD.

2.4. Measures

2.4.1. ADHD symptoms

Measures of ADHD symptoms were obtained using the 18-item Barkley Adult ADHD Rating Scale (BAARS-IV) (Barkley and Murphy, 2006), which consists of the DSM-IV items related to inattention and hyperactivity–impulsivity. Respondents indicated how frequently they experienced behaviours on a scale of 0–3 (never or

<table>
<thead>
<tr>
<th>Number approached</th>
<th>ADHD</th>
<th>BD</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recruitment</td>
<td>57</td>
<td>75</td>
<td>120</td>
</tr>
<tr>
<td>Un-contactable</td>
<td>17</td>
<td>26</td>
<td>45</td>
</tr>
<tr>
<td>Declined</td>
<td>4</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>Travel or childcare difficulties</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Did not attend or cancelled</td>
<td>1</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Exclusions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unsuitable diagnosis</td>
<td>3</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>ADHD with comorbid BD</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control with psychiatric disorder</td>
<td></td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Medical or neurological disorder</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Autism</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past ECT treatment</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participating in another research trial</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currently pregnant</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insufficient English language ability</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Sample</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

Legend: ADHD = attention-deficit/hyperactivity disorder; BD = Bipolar Disorder; ECT = Electroconvulsive therapy.
rarely, sometimes, often, very often) during the past 6 months. Total scores were calculated for each symptom dimension. The Barkley’s functional impairment scale (Barkley and Murphy, 2006) used the same scoring system and was included with the (BAARS-IV) to create a third impairment subscale, indexing functional impairments across several domains including occupational, daily responsibilities and social relationships. Both self-rated and informant-rated versions of the BAARS-IV were used to obtain measures of ADHD symptoms.

The Diagnostic Interview for ADHD in Adults (DIVA) (Kooij and Francken, 2007) was used to assess ADHD symptoms in participants. The DIVA, like the BAARS-IV, consists of 18 items used to define the DSM-IV symptom criteria for ADHD, but is a semi-structured interview conducted by a trained clinical investigator. Each item is scored “yes”, if the behavioural symptom is present often within the past 6 months. Outcomes were total current ADHD symptom score, and separate totals for inattentive and hyperactive-impulsive symptom domains.

2.4.2. Mania and depression symptoms

The Beck’s Depression Inventory II (BDI) (Beck et al., 1996) was included as a self-rated measure of depression symptoms. The scale has 21 questions, rated 0–3 based on the severity of symptoms, during the past two weeks. The test variable was total score. The self-report Altman Self-Rating Mania Scale (ASRM) (Altman et al., 1997) was used to measure mania symptoms in the past week. This is a 5-item measure scored 0–4 based on the strength of the behaviour. The total score was used as the test variable.

A second measure of mania symptoms was collected using the Young Mania Rating Scale (YMRS) (Young et al., 1978), completed by the investigator following clinical interview. This 11-item measure uses subjective report of mental phenomena and clinical observations to rate behaviours associated with mania in the past 48 h. Seven items are scored 0–4 based on severity and the remaining four items (irritability, rate or amount of speech, delusional/grandiose thought content, and severe aggressive or uncontrollable behaviour) are scored 0–8, as characteristic features of manic episodes. For this study, a change indicator asking “Is this how you normally feel?”, scored yes/no, was added to each item to distinguish between episodic symptoms which are characteristic of BD, and the more stable trait-like symptoms which are characteristic of ADHD. A further question asking “Has there ever been a time other than the last 48 h when you have felt….” was added to each of the 8 self-report items to count the number of symptoms experienced in the past, including worst ever episode, to determine the range of symptoms experienced by BD patients during episodes of mania. Outcomes for this measure were total score, number of present symptoms (excluding observer-rated items to make this comparable with the past symptoms scale) and number of past symptoms.

We also examined whether particular items on the YMRS, which related to specific symptom domains or loaded on previously identified factors, were able to delineate the two clinical groups. We compared two approaches. The first approach was based on diagnostic criteria for ADHD which grouped YMRS items on whether they overlapped with ADHD symptoms (increased motor activity/energy, increased rate of production of speech, and language/thought disorder including distractible thought processes and changing topics frequently); or did not overlap with ADHD (inappropriate elevated mood, increased/inappropriate sexual interest, delusions and grandiosity, and severe disruptive/aggressive or uncontrollable behaviour); or were associated features of ADHD that overlap with BD (difficulty sleeping and irritable mood). The second approach used three groupings previously identified by Hanwell and de Silva (2011) in a factor analysis of YMRS items, which were labelled: irritable mania (increased motor activity/energy, irritable moods, and severe disruptive/aggressive or uncontrollable behaviour); elevated mania (elevated mood, language/thought disorder, sexual interest and insight); and psychotic mania (increased motor activity/energy, motor activity, delusions and grandiosity, and appearance).

2.4.3. Emotional lability

The self-rated Affective Lability Scale Short Form (ALS-SF) (Oliver and Simons, 2004), comprising of 18 items scored 1–4 (very un-descriptive, rather un-descriptive, rather descriptive, very descriptive) was used as one of two measures of mood lability. The ALS-SF measures fluctuations from a normal mood to other emotional states from moment to moment during the past week, and has been shown to comprise of three domains of anxiety—depression, depression—elation and anger (Oliver and Simons, 2004). Total overall score and total score on each subscale were used as the test variables.

The second measure of emotional lability was the auxiliary subscale of the Centre for Neurologic Study-Lability Scale (CNS-LS) (Moore et al., 1997), adapted by removing two items related to impatience which have clear overlap with impulsive symptoms of ADHD. This created a self-rated 8-item measure focusing on negative emotions, such as getting easily frustrated, upset and angry occurring in the past month and past 5 years. Each item is scored on a scale of 0–4 (applies never, rarely, occasionally, frequently, most of the time), based on the frequency of each experience. Total scores for the past month and past 5 years were used as the test variables.

2.4.4. Intellectual ability

The Wechsler Abbreviated Scale of Intelligence Fourth Edition (WASI-IV) (Wechsler, 1999) was administered to all participants to derive an estimate of IQ.

2.4.5. Statistical analyses

Across the three samples, rating scale data were normally distributed for the BAARS-IV impairment scales (self-rated and informant), the ALS total score and the hyperactive-impulsive subscale of the DIVA. Otherwise, the most appropriate transformations were applied to the data (log or square-root). For the ALS and YMRS subscales, no available transformations normalised the data, so non-parametric Kruskal–Wallis tests were used. Group differences in normal and transformed-normal data were tested using univariate ANOVAs. Where appropriate, pairwise comparisons were conducted to discriminate which groups differed. On the YMRS, the number of symptoms past and present were compared using repeated-measures ANOVA to explore the interaction of group and symptom change over time. Additional post-hoc pairwise comparisons were used to investigate both group differences and differences between the number of past and presents symptoms within group. Analyses were carried out using STATA (Version 11) and SPSS (Version 21). Given the large number of sub-scales used in this study, and therefore high number statistical comparisons and associated risk of type-I error, all reported p-values were adjusted for multiple testing using family-wise Bonferroni corrections to maintain α = 0.05 for all 20 independent tests employed in the primary analysis and all subsequent post-hoc comparisons.

3. Results

The groups did not differ in mean age (Mean (SD): ADHD = 37.4 (7.65); BD = 40.3 (7.68); Control = 36.7 (4.28); F = 1.63, p = 0.21) or IQ (Mean (SD): ADHD = 104.5 (17.85); BD: 108 (12.50); Control = 112.35 (14.21); F = 1.37, p = 0.26). Mean scores on outcome
measures with adjusted p-values are shown in Table 2 and the standardised differences between groups for all measures are shown in Fig. 1.

### 3.1. ADHD symptoms

Group differences were present for: DIVA total symptom score \((F(2,57)=37.65, p<0.001)\), inattention \((F(2,57)=33.68, p<0.001)\) and hyperactive-impulsive subscales \((F(2,57)=25.65, p<0.001)\); self-rated BAARS inattention \((F(2,55)=40.51, p<0.001)\) and hyperactive-impulsive \((F(2,57)=18.39, p<0.001)\) subscales; and informant-rated BAARS inattention subscale \((F(2,48)=12.05, p=0.01)\). Post-hoc analysis indicated that both ADHD and BD groups had higher ADHD symptom scores than controls on self and informant reported ADHD rating scales. However, only the ADHD group had higher current ADHD symptom scores compared to controls when the DIVA interview was used as the measure of ADHD symptoms.

The ADHD group had significantly higher symptoms than the BD group for the DIVA and self-rated BAARS scores, but not for the informant-rated BAARS. To quantify the degree to which the DIVA and self-rated BAARS scores can distinguish between patients with ADHD and BD we calculated receiver operating characteristic (ROC) scores. To compare the BAARS with the DIVA we made binary variables from the BAARS scores for the absence (never, rarely or sometimes) or presence (often, very often) of each individual ADHD item. The results are summarised in Table 3 with optimal thresholds that balance sensitivity against specificity. There was very good sensitivity (90%) and specificity (95%) for the DIVA interview, particularly for the inattentive items when the symptoms threshold of 6 or more symptoms was applied. This compared to a much lower sensitivity of 65–70% using the BAARS, although specificity remained high (95–100%).

### 3.2. Impairment

Ratings of impairment showed significant group differences on both self-rated \((F(2,55)=41.55, p<0.001)\) and informant \((F(2,47)=10.92, p=0.003)\) scales. Both clinical groups reported elevated impairment compared to controls on both scales. The ADHD group had elevated scores compared to the BD group on the self-report measure, but not the informant measure.

### 3.3. Mania and depression

Group differences in self-reported current manic symptoms on the ASRM were not significant \((F(2,55)=1.71, p=1.00)\). Self-rated current depression symptoms on the DI showed group differences \((F(2,57)=13.79, p<0.001)\), with both ADHD and BD groups reporting higher scores than controls, but not differing compared to

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Mean and standard deviations for ADHD and BD symptom measures and emotional lability measures.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total n</td>
</tr>
<tr>
<td></td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>BAARS-IV self-rated</td>
<td></td>
</tr>
<tr>
<td>Inatt score</td>
<td>58</td>
</tr>
<tr>
<td>Hyp-Imp score</td>
<td>60</td>
</tr>
<tr>
<td>Impairment score</td>
<td>58</td>
</tr>
<tr>
<td>BAARS-IV informant</td>
<td></td>
</tr>
<tr>
<td>Inatt score</td>
<td>51</td>
</tr>
<tr>
<td>Hyp-Imp score</td>
<td>51</td>
</tr>
<tr>
<td>Impairment score</td>
<td>50</td>
</tr>
<tr>
<td>DIVA</td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>60</td>
</tr>
<tr>
<td>Inatt symptoms</td>
<td>60</td>
</tr>
<tr>
<td>Hyp-Imp symptoms</td>
<td>60</td>
</tr>
<tr>
<td>ASRM</td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>58</td>
</tr>
<tr>
<td>YMRS</td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>60</td>
</tr>
<tr>
<td>No. current symptom</td>
<td>60</td>
</tr>
<tr>
<td>No. past symptoms</td>
<td>60</td>
</tr>
<tr>
<td>State change (past 48 h)</td>
<td>60</td>
</tr>
<tr>
<td>DI</td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>60</td>
</tr>
<tr>
<td>ALS-SF</td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>60</td>
</tr>
<tr>
<td>Anxity-depression score</td>
<td>60</td>
</tr>
<tr>
<td>Elation-depression score</td>
<td>60</td>
</tr>
<tr>
<td>Anger score</td>
<td>58</td>
</tr>
<tr>
<td>CNS</td>
<td></td>
</tr>
<tr>
<td>Past month score</td>
<td>59</td>
</tr>
<tr>
<td>Past 5 year score</td>
<td>59</td>
</tr>
</tbody>
</table>

*Adj. p* family-wise Bonferroni corrected post-hoc p-values. "State change (past 48 h)" indexes manic-like symptom change in the past 48 h (appearing or disappearing). Other abbreviations: Inattention (Inatt), Hyperactive-Impulsive (Hyp-Imp), Barkley Adult ADHD Rating Scale-IV (BAARS-IV), Diagnostic Interview for ADHD in Adults (DIVA), Altman Self-Rating Mania Scale (ASRM), Young Mania Rating Scale (YMRS), Beck Depression Inventory (DI), Affective Lability Scale (ALS), Centre for Neurologic Study-Lability Scale (CNS-LS).

---

```
* p < 0.001.
  p < 0.01.
  p < 0.05.
```
each another (Table 2).

The YRMS interview showed a nominal group difference in total score for current symptoms (unadjusted \( p = 0.005 \)), which did not survive \( \alpha \)-correction (\( F(2,59) = 5.78, p = 0.10 \)). On the extension questions added for this study, the groups did not differ in the proportion of mania symptoms which had changed within the past 48 h (\( F(2,57) = 1.31, p = 1.00 \)). A repeated-measures ANOVA, comparing the number of past symptoms with the number of present symptoms, showed a significant main effect of group (\( F(2,57) = 6.88, p < 0.001 \), time period (\( F(1,57) = 15.56, p = 0.01 \)) and interaction of time period \( \times \) group (\( F(2,57) = 12.03, p < 0.001 \)). Overall, post-hoc tests indicated that the BD group had a higher number of mania symptoms in the past, but that the two clinical groups did not differ in the number of current symptoms. The ADHD group also showed more current mania symptoms than controls, although BD-control differences were not significant (Table 2).

We further examined if subsets of items from the YMRs were better able to discriminate ADHD from BD using the non-ADHD overlapping symptoms grouping (inappropriate elevated mood, increased/inappropriate sexual interest, delusions and grandiosity, and severe disruptive/aggressive or uncontrollable behaviour) and then only for past ‘worst episode’ symptoms and not current symptoms. Both clinical groups had elevated scores compared to controls on the ADHD overlapping symptom grouping (increased motor activity/energy, increased rate of production of speech, and language/thought disorder including distractible thought processes and changing topics frequently) for both current and past symptoms, but did not differ between themselves. The shared associated symptom grouping (sleep disturbance and irritability) indicated elevated scores for past symptoms in the BD group compared to controls, but no differences for other comparisons. The factor-based item groupings did not show any group differences for current symptoms. For past symptoms, on both elevated mania (elevated mood, language/thought disorder, sexual interest and insight) and psychotic mania (increased motor activity/energy, motor activity, delusions and grandiosity and appearance) item groupings the BD group had elevated scores compared to ADHD and controls. Additionally, on the psychotic mania cluster the ADHD group had higher scores than controls. For irritable mania items (increased motor activity/energy, irritable moods, and severe disruptive/aggressive or uncontrollable behaviour) both clinical groups scored higher than controls, but did not differ compared to one another.

### 3.4. Emotional lability (EL)

Group differences were detected for ALS total scores (\( F(2,59) = 11.86, p = 0.001 \)), the elation-depression subscale (\( H(2) = 17.60, p < 0.001 \)) and the anger subscale (\( H(2) = 17.20, p < 0.001 \)). Group differences for the anxiety-depression subscale did not survive \( \alpha \)-correction (\( H(2) = 9.34, p = 0.18 \)). Post-hoc pairwise comparisons did not distinguish ADHD and BD groups on either total scores or subscales (Table 2). The ADHD group had higher scores compared to controls on all scales of the ALS, with the BD group showing division of items only discriminated ADHD from BD using the non-ADHD overlapping symptoms grouping (inappropriate elevated mood, increased/inappropriate sexual interest, delusions and grandiosity, and severe disruptive/aggressive or uncontrollable behaviour) and then only for past ‘worst episode’ symptoms and not current symptoms. Both clinical groups had elevated scores compared to controls on the ADHD overlapping symptom grouping (increased motor activity/energy, increased rate of production of speech, and language/thought disorder including distractible thought processes and changing topics frequently) for both current and past symptoms, but did not differ between themselves. The shared associated symptom grouping (sleep disturbance and irritability) indicated elevated scores for past symptoms in the BD group compared to controls, but no differences for other comparisons. The factor-based item groupings did not show any group differences for current symptoms. For past symptoms, on both elevated mania (elevated mood, language/thought disorder, sexual interest and insight) and psychotic mania (increased motor activity/energy, motor activity, delusions and grandiosity and appearance) item groupings the BD group had elevated scores compared to ADHD and controls. Additionally, on the psychotic mania cluster the ADHD group had higher scores than controls. For irritable mania items (increased motor activity/energy, irritable moods, and severe disruptive/aggressive or uncontrollable behaviour) both clinical groups scored higher than controls, but did not differ compared to one another.

### Table 3

Receiver Operating Characteristics (ROC) scores showing sensitivity and specificity of BAARS and DIVA measures to ADHD diagnosis compared to BD diagnosis.

<table>
<thead>
<tr>
<th>ROC scores</th>
<th>AUC</th>
<th>Threshold</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAARS inattention</td>
<td>0.87</td>
<td>6/9</td>
<td>0.70</td>
<td>0.95</td>
</tr>
<tr>
<td>BAARS hyper-imp</td>
<td>0.83</td>
<td>6/9</td>
<td>0.45</td>
<td>0.90</td>
</tr>
<tr>
<td>BAARS total score</td>
<td>0.89</td>
<td>11/18</td>
<td>0.65</td>
<td>1.00</td>
</tr>
<tr>
<td>DIVA inattention</td>
<td>0.95</td>
<td>7/9</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td>DIVA hyper-imp</td>
<td>0.89</td>
<td>6/9</td>
<td>0.55</td>
<td>0.90</td>
</tr>
<tr>
<td>DIVA total score</td>
<td>0.97</td>
<td>11/18</td>
<td>0.90</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Abbreviations: Barkley Adult ADHD Rating Scale (BAARS), Diagnostic Interview for ADHD in Adults (DIVA), Area under the curve (AUC).

Standardised means total scores for each measure or subscale. ADHD symptoms are presented in the first part, mania and depression symptoms in the second and emotional lability in the third part of this graph. “Change (48h)” indexes manic-like symptom change in the past 48 hours (appearing or disappearing). “No. Sym Past” measures current manic-like symptoms, “No. Sym Past” measures symptoms experienced in the past. Other abbreviations: Inattention (Inatt), Hyperactive-Impulsive (Hyp-Imp), Anxiety-Depression (Anx-Dep), Elation-Depression (Ela-Dep), Barkley Adult ADHD Rating Scale-IV (BAARS-IV), Diagnostic Interview for ADHD in Adults (DIVA), Altman Self-Rating Mania Scale (ASRM), Young Mania Rating Scale (YMRS), Beck Depression Inventory (DI), Affective Lability Scale (ALS), Centre for Neurologic Study-Lability Scale (CNS-LS).

**Fig. 1.** Group differences on ADHD, mania, depression and emotional lability measures, as standardised scores with standard error.
higher scores than controls on total score and the elation-depression subscale.

The CNS-LS, indicated group differences on both time spans (last month: $F(2,56) = 11.35, p < 0.001$; last 5 years: $F(2,56) = 12.04, p = 0.001$). For CNS-LS ratings of EL in the last month, the ADHD group had elevated scores compared to both BD and controls, with the BD group not differing from controls (Table 2). For ratings based on worst ever in the last 5 years, both clinical groups had higher scores than controls, but were undifferentiated compared to each other.

### 3.5. Comorbidities

Primary analyses were rerun after excluding the four individuals which had a diagnosed comorbidity (ADHD: one depression, one OCD; BD: one BPD, one Anxiety Disorder). Overall, the results did not change with these participants excluded, except for two post-hoc pairwise comparisons, which then became non-significant after correcting for multiple testing. These were the ADHD-BD comparison on the CNS-LS (last month), where the adjusted $p$-value became non-significant (adj. $p = 0.11$, unadjusted $p = 0.04$) and the ADHD-control comparison for the number of present symptoms on the YMRS which weakened to a trend level (adj. $p = 0.07$, unadjusted $p = 0.008$).

### 4. Discussion

In this study we investigated the similarities and differences between female patients with typical ADHD, Bipolar I Disorder (who were not currently experiencing a manic episode) and healthy controls, using standard measures used in the diagnostic assessment of ADHD and BD. Using ratings for the current mental state, increased levels of ADHD and depression symptoms, emotional lability and functional impairment were seen in both the ADHD and BD groups compared to controls. The ADHD group generally showed higher levels of psychopathology than the BD group, particularly for current symptoms of emotional lability and mania. Using retrospective ratings for mania on the YMRS, which would measure past manic episodes, gave higher ratings in the BD than the ADHD group, although both groups had higher ratings than the controls. The DIVA interview was the best instrument for separating out the two clinical groups, with high sensitivity and specificity for ADHD. Overall, these findings show a significant level of residual symptoms and impairments in BD patients during non-manic periods, which was similar to the ADHD patients for depression and impairment, but did not reach the levels of ADHD, mania and emotional lability symptoms seen in the ADHD group.

#### 4.1. Distinguishing BD from ADHD

Making this distinction is important because BD patients often present with continued mood symptoms and functional impairments in between major affective episodes, raising the question of whether any observed psychopathology is due to persistence of BD or could be due to comorbid ADHD. Indeed this study showed considerable overlap between BD and ADHD using both rating scale and interview measures. Yet it was possible to distinguish between the ADHD and BD groups. We found that the interview measure for current ADHD symptoms provided very good sensitivity (around 90%) and specificity (around 95%) to identify ADHD in comparison with BD. In contrast, the self-reported ADHD measures showed enhancement of scores in both clinical groups, although self-ratings of ADHD inattention were moderately good at separating ADHD from BD. For BD the best discrimination came from the use of the YMRS mania interview, which was sensitive to
differences between ADHD and BD groups when using a retrospective ‘worst ever’ adaption included for this study. However, even using retrospective data, the ADHD group showed a significant level of symptoms on the YMRS, and for current symptoms the ADHD group had more mania symptoms than the BD group. These findings are similar to those reported in previous studies.

For the ADHD measures, the inattentive symptoms gave the best discrimination of ADHD from BD for both the self-rated and interview measure. This is similar to the results of Ibanez et al. (2012), who report higher self-rated inattention scores in ADHD compared to BD or control groups using the same self-rated scale of ADHD used in this study, although they did not observe a significant BD-control difference.

The Altman Self-Rating Mania Scale (ASRM) and the Beck Depression Inventory (DI) were not able to distinguish the two clinical disorders. High depression scores for both clinical groups highlight that symptoms of depression are commonly seen in ADHD. This replicates findings from Torralva et al. (2011) and Ibanez et al. (2012) of elevated depression scores on the DI in ADHD compared to BD.

Current symptom score and number of current symptoms on the YMRS mania interview showed that the ADHD group had higher levels of mania-like symptoms compared to both the controls and the BD group. These results, collected using the standard form of the measure, highlight the potential difficulties of delineating ADHD and BD using cross-sectional (present state) mania measures, and replicate findings for Ibanez et al. (2012), which reported higher mania symptom scores in the ADHD group compared to controls on this measure. In contrast, when the scale was applied to the number of past symptoms, the BD group had a greater number of mania symptoms than the ADHD or control groups, even though the ADHD continued to report higher past symptoms than controls. These findings in combination indicate two important points. Firstly, that the YMRS may only be effective at distinguishing ADHD and BD when comparing retrospective worst episode information, meaning that enquiring about historical manic episodes, and especially symptoms known not to overlap with ADHD such as grandiosity and increased sex drive (Skirrow et al., 2012), may assist clinicians in delineating ADHD and BD from BD. Secondly, that the YMRS should not be used exclusively to make a diagnosis of BD, as other conditions, such as ADHD, also overlap with ADHD such as grandiosity and increased sex drive. (Skirrow et al., 2012), may assist clinicians in delineating ADHD from BD.

Overall, these preliminary findings suggest that ADHD and BD might load separately on specific items within the YMRS, and therefore development of a subscale designed to delineate ADHD from euthymic BD in adults may be possible. However, as indicated by our findings, any measure will require a retrospective component to fully delineate ADHD from BD.

4.2. YMRS symptom clusters

We completed further exploratory analyses to investigate whether particular YMRS items might be specifically associated with either ADHD or BD. We compared a three-cluster DSM symptom model (ADHD overlap; no-overlap; mood and sleep problems) against an empirically-derived three-factor model (irritable mania; elevated mania; psychotic mania) identified by Hanwella and de Silva (2011). For current symptom scores, group differences were only found on the ADHD overlapping symptoms item grouping, consisting of increased motor activity/energy, increased rate of production of speech and language/thought disorder items, where both clinical groups scored significantly higher than controls. No other differences were found for the symptom-based or factor-based model. In line with our other analysis, this implies that the YMRS is poor at distinguishing ADHD and BD outside of a manic episode in its standard form.

For the number of past symptoms, based on worst ever episode, more differences emerged due to the higher number of symptoms reported by the BD group during past manic episodes. For the overlapping symptom grouping from the symptom-based model, and the irritable mania grouping from the factor-based model, both clinical groups had high scores compared to controls, indicating that these clusters both capture shared symptoms. However, only the elevated motor activity item was common between them (overlapping items: motor activity, speech rate, language/thought disorder; irritable mania: motor activity, irritability, disruptive/aggressive behaviour). Scores on item groupings for overlapping symptoms, mood, elevated mania and psychotic mania were all higher for the BD group, compared to both ADHD and control groups.

Overall, these findings illustrate the considerable overlap of symptoms in ADHD and BD. We found greater specificity for the ADHD symptoms elicited at interview to correctly identify the ADHD group, than for the traditional BD symptoms to correctly separate BD from ADHD. Thus the two disorders can usually be distinguished through a combination of detailed symptom review, elicited using a clinical interview, with a consideration of the time course and episodicity of the symptoms that are present. These data are consistent with other literature showing mood symptoms and EL are commonly seen in ADHD (Posner et al., 2014; Shaw et al., 2014; Skirrow and Asherson, 2013; Skirrow et al., 2014), and illustrate the importance of considering ADHD as a differential diagnosis in patients presenting with chronic (non-episodic and trait-like) mood symptoms, including symptoms of mania, depression and emotional lability, alongside other chronic, non-episodic, trait-like psychiatric conditions, including borderline personality disorder, cyclothymia and dysthymia. Therefore, a developmental account including age of onset and course (chronic versus episodic) should always be used to establish the diagnosis, and for an episode of mania/hypomania there must be evidence of a change from the premorbid mental state.

4.3. Chronicity and validity of symptom measurement

The interview measures provided better discrimination between ADHD and BD. One reason for this is likely to be that an interviewer is able to explore both the nature and time course of symptoms during an interview, to ensure any reported symptoms meet the question criteria. The DIVA measure provides several examples of behaviours associated with each symptom, allowing the interviewer to qualitatively explore each symptom before rating as present or absent. In contrast, the self-report measures only provide a question, but no examples and rely on the interpretation of an untrained person. This means that it is unknown if the items are being scored based on equivalent symptoms, as well as severity of symptoms, within each clinical group. In terms of the time course, the wording in the rating scales is also more ambiguous. For example, the DIVA interview items are scored when symptoms are present for at least six months or more. Although to a lesser extent this is also true of the self-report ADHD measures, the wording of questions is more ambiguous, stating that symptoms should be present during the last six months. These ratings could therefore reflect symptoms of any duration during this period, rather than the sustained trait-like symptoms that characterise ADHD.

The YMRS, on the other hand, is designed specifically to evaluate manic symptoms in a short-time window (past 48 h). Although our findings suggest that this measure would be effective at delineating ADHD from BD as a retrospective measure, or during a BD manic episode, it was not effective at delineating ADHD from
BD outside of a manic episode based on current symptoms alone. The YMRS therefore has discriminatory potential, and could be adapted either through development of a specific subscale using items which load selectively onto one of the clinical disorders, or by adapting the measure to compare the episodicity of symptoms; thereby making the distinction between chronic trait symptoms of ADHD from the episodic symptoms of BD. Our findings support arguments that chronicity versus episodicity is a key delineating factor between ADHD and BD in adulthood (Skirrow et al., 2012).

4.4. Emotional lability (EL)

EL is associated with both ADHD (Skirrow and Asherson, 2013) and euthymic BD (Judd et al., 2003). Our study supports the view of EL as a largely non-specific set of symptoms that are seen across different disorders, with high EL scores seen in both the ADHD and BD groups compared to controls. Indeed, EL occurred at higher rate in the ADHD patients, consistent with the emerging view of EL as an associated feature of ADHD. EL cannot therefore be relied upon to discriminate ADHD from BD. For this reason the current absence of EL from the DSM-5 ADHD criteria, but its inclusion as a characteristic feature of ADHD that supports the diagnosis of ADHD, remains a sensible decision (American Psychiatric Association, 2013). Health care professionals need to be reminded that the classification systems are not designed to capture all aspects of a clinical condition, but to provide an optimal algorithm that helps to separate one condition from another. In this regard, the DSM-5 ADHD items appear to be more specific to ADHD than the DSM-5 BD items are to BD (particularly if the definition of BD symptoms reflecting a change from the pre-morbid mental state is ignored).

4.5. Impairment ratings

ADHD participants showed more functional impairments than BD and controls, yet the BD group also showed higher levels of impairment compared to controls. This suggests that while impairment is known to be present in both disorders (Brassett-Harknett and Butler, 2007; Samalin et al., 2014), women with ADHD may be more severely impaired than women with BD who are not experiencing a manic episode. Studies into the individual disorders indicate that people with BD may show normal pre-morbid functioning (Reichenberg et al., 2002), while ADHD is associated with chronic functional impairment throughout lifespan (Brassett-Harknett and Butler, 2007; Lin et al., 2015). However, to date, there is insufficient cross-disorder evidence to determine if differences in psychosocial impairment between ADHD and BD represent a potentially useful marker for delineation.

4.6. Limitations and future directions

The samples are relatively small, consisting of selected patients with typical ADHD, typical BD-I and healthy controls, and focuses only on female participants. It is therefore not clear the extent to which these findings will generalise to more complex patients, of both genders, showing features of both ADHD and BD. ADHD is considered to reflect the extreme and impairing tail of a dimensional trait and symptoms commonly may also occur at sub-diagnostic levels (Hudziak et al., 1999; Simon et al., 2009). This means that BD patients are expected to display ADHD traits as part of a normal population distribution. However, we were unable to determine if the elevated ADHD symptoms in our BD sample represent a manifestation of BD or an independent subclinical expression of ADHD. The use of structured clinical interviews during research assessments, in addition to the symptom measures presented here, may have clarified matters further. In future, replication using prospective approaches, and structured clinical interviews, would be useful to determine if the development of a subscale based on item grouping in the YMRS has clinical utility for delineating ADHD from BD.

This study also focused on the most characteristic forms of ADHD and BD, in order to recruit homogeneous samples where the distinction in primary diagnosis was entirely clear. However, replication in clinical samples that are less highly selected and therefore closer to clinical realities would further enhance the usefulness of this research, as determining the potential for misdiagnosis between ADHD and BD under a broader range of conditions would be informative from a clinical perspective. For instance, delineation may be easier during manic episodes where BD symptoms are greatly pronounced, and may be more challenging in comparisons with BD-II and cyclothymia where reduced symptom severity will likely make the clinical differentiation of symptoms more difficult.

5. Conclusions

Overall, we show that ADHD is a chronic, impairing disorder, with a high degree of EL and hyperactivity which could be confused with symptoms of mania. Measures such as the DIVA interview which combine both a detailed disorder specific description of ADHD symptoms with a temporal component that captures the distinction between sustained traits and episodic symptoms that reflect a change in the pre-morbid mental state are best at discriminating ADHD from BD in adult women. We therefore conclude that interview measures combined with a developmental account of symptoms and impairments provide good discrimination compared to rating scale data, and should always be used as the primary diagnostic tool.

References


