Could glutamate spectroscopy differentiate bipolar depression from unipolar?

Matthew J Taylor

Department of Psychosis Studies, Institute of Psychiatry, King’s College London
SE5 8AF, UK

Corresponding author: Dr Taylor, Institute of Psychiatry, Dept Psychosis Studies
PO63, London SE5 8AF, UK.

email: matthew.j.taylor@kcl.ac.uk
Abstract

Background

Accurate differentiation of bipolar and unipolar depression is a key clinical challenge. A biological measure that could differentiate bipolar and unipolar depression might supplement clinical assessment. Magnetic Resonance Spectroscopy measurements of total glutamate and glutamine (Glx) in anterior cingulate cortex are one potential measure. The objective of this study was to assess the potential performance of this measure.

Methods

Meta-analysis of data from eleven studies where anterior cingulate Glx of depressed patients has been compared to that of healthy controls was performed. Effect sizes for bipolar and unipolar depression were calculated as Standardised Mean Differences. The best estimate of test classification performance on the basis of observed effects was calculated.

Results

People with unipolar depression had on average lower levels of Glx than healthy controls (effect size -1.05; 95% CI -1.53). People with bipolar depression tended towards higher Glx than healthy controls (effect size 0.40; 95% CI -0.85). This yielded a difference in Glx between unipolar and bipolar depression of effect size 1.46 (95% CI 0.80 to 2.11). Based on this difference, a test differentiating bipolar from unipolar depression by whether Glx was higher or
lower than the average in healthy population would have sensitivity .66 and specificity .85.

Limitations

There is an absence of studies directly comparing unipolar and bipolar depressed patients.

Conclusions

On available data, measurement of anterior cingulate Glx is a promising potential tool for differentiation of bipolar and unipolar depression. This potential effect requires direct validation within mixed clinical cohorts.

Keywords: Bipolar depression, Magnetic resonance spectroscopy, glutamate, glutamine, major depressive disorder
Introduction

Accurate identification of bipolar depression remains a major clinical challenge. Typically within bipolar disorder, episodes of depression predominate (Judd et al., 2003) and in practice many people with bipolar disorder are initially diagnosed and treated as unipolar depression (Hantouche et al., 1998; Hirschfeld et al., 2003; Smith et al., 2011). The adverse impact of missed diagnosis of bipolar disorder has been known for some time (Dunner, 2003). As growing evidence indicates that the optimal management of bipolar and unipolar depression differs substantially, the critical importance of early identification of bipolar depression becomes even clearer (Goodwin, 2012).

Episodes of major depression within bipolar disorder and unipolar depressive disorder show substantial overlap in presentation. Although some characteristics are more commonly associated with bipolar disorder, at present such differences remain probabilistic rather than being sufficient to guide differences in management by themselves (Mitchell et al., 2008). A number of screening tools have been developed in recent years to identify people most likely to receive a diagnosis of bipolar disorder (Angst et al., 2005; Ghaemi et al., 2005; Hirschfeld et al., 2000). However, even with expert clinical assessment, people who will go on to develop bipolar disorder but have as yet only experienced depressive episodes cannot be identified accurately.

A biological measure that could differentiate bipolar and unipolar depression would therefore be of interest as a supplement to clinical assessment. It might be of benefit clinically where historical information is unclear or contradictory, or as a
potential research tool in future studies, for example identifying those at highest risk of conversion to bipolar disorder. Glutamate neurotransmission has been of increasing interest in mood disorders including bipolar disorder in recent years (Sanacora et al., 2008). The imaging technique Magnetic Resonance Spectroscopy (MRS) allows for safe, non-invasive measurement of glutamate levels in the brain. In practice differentiation of glutamate from the related chemical glutamine is challenging with MRS, so the total combined glutamate and glutamine (Glx) level is often reported.

There is now a growing literature of glutamate MRS studies in mood disorders. An emerging pattern points to decreased Glx in unipolar depression and increased Glx in bipolar depression in some brain regions (Gigante et al., 2012; Yüksel and Öngür, 2010). Here we address quantitatively the question of whether this difference might be of diagnostic use, focusing on measures of Glx in anterior cingulate cortex, a region known to play an important role in bipolar disorder (Strakowski et al., 2012). We employ meta-analysis to quantify the differences described by the literature to date, and the extent of uncertainty around that estimate, and model the diagnostic test performance that would result from this difference.
Materials and Methods

Search strategy

Studies comparing levels of Glx between people with a current episode of major depression and a control population were sought. Studies where the MRS voxel encompassed medial prefrontal cortex including pregenual anterior cingulate cortex were included. Studies published to January 2014 were identified by searches of MEDLINE and EMBASE, reference lists of identified articles, and from personal reference collections.

Data extraction

Levels of Glx (mean and standard deviation) within anterior cingulate cortex for currently depressed participants and the control population were extracted. Levels reported either relative to creatine or tissue water were accepted. Where suitable data was not reported in the published article or could not be imputed from available data, further details were sought from the authors. No studies were excluded on the basis of spectral quality, but it was typical for studies to restrict their analysis to metabolite estimates meeting a set criterion e.g. Cramér-Rao SD less than 20% (Chen et al., 2014).

Meta-analysis was performed using a random effects model (DerSimonian-Laird), and standardised mean differences (Hedges’ adjusted $g$) employed for measures of effect given differences in technical approaches employed (Table 1). Statistical heterogeneity between studies was assessed by $I^2$ (Higgins et al., 2003) which expresses the percentage of total variation across studies that is
due to heterogeneity rather than chance. Meta-regression for effect of diagnosis was performed using a mixed-effects model (Viechtbauer, 2010). Differences in Glx levels between unipolar and bipolar depression were calculated by adjusted indirect comparison (Glenny et al., 2005).

Table 1 near here

A receiver operating curve (ROC) was calculated for the ability to differentiate bipolar from unipolar depression by Glx level given normal distributions for each disorder separated by the estimates of effect size established by meta-analysis. ROCs display the sensitivity and specificity of a test across all criterion levels that could be chosen. Total area under the curve (AUC) was calculated. An AUC of 1.0 indicates perfect performance, whereas an AUC of 0.5 indicates performance no better than chance alone. Statistical analyses were performed using R (version 3.0.2) with the meta (version 3.1-2) and metafor (version 1.9-2) packages (Schwarzer, 2013; Viechtbauer, 2010). 95% Confidence Intervals (CI) for estimates are presented where available.

Results

Eleven studies were identified as suitable for inclusion (Table 1). In these studies a total of 421 participants provided data. No studies were identified that had compared unipolar and bipolar depression directly. Seven studies compared anterior cingulate cortical Glx during episodes of unipolar depression to that of healthy controls. (In one study, one of 19 depressed participants had a diagnosis of bipolar disorder; (Auer et al., 2000)). Four studies compared Glx between
bipolar depression and healthy controls. Some additional studies of anterior cingulate glutamate spectroscopy were identified but could not be included as glutamate levels were reported but Glx was not reported and data could not be obtained from the authors.

Figure 1 near here

Glx differed significantly between the diagnostic groups ($p<.0001$). People with unipolar depression had on average lower levels of Glx than healthy controls. An effect size of -1.05 (95% CI -1.58 to -1.53) was found for this difference. There was moderately high statistical heterogeneity in estimates between studies ($I^2$ 64%). People with bipolar depression tended towards higher Glx than healthy controls. An effect size of 0.40 (95% CI -0.04 to 0.85) was found for this observation. The result was more consistent across studies, with lower statistical heterogeneity ($I^2$ 44%). The best estimate of the difference in Glx between unipolar and bipolar depression was of effect size 1.46 (95% CI 0.80 to 2.11).

The ROC analysis (Figure 1B) indicated that overall predicted test performance was good with AUC 0.85 (95% CI 0.72 to 0.93). The expected performance for differentiating bipolar from unipolar depression by whether Glx was higher or lower than the average in healthy population was found to be sensitivity .66 and specificity .85.
Discussion

Here we have addressed quantitatively the question of whether levels of Glx in anterior cingulate cortex measured using MRS have the potential to differentiate bipolar from unipolar depression. The published literature to date indicates that medium to large effect sizes are observed comparing either unipolar or bipolar depression to healthy controls. However, since the effects are in opposite directions, a very large effect size (1.46) is estimated for the difference between bipolar and unipolar depression.

The ROC expected performance of a test based on this measure has an AUC greater than some recent clinical measures, for example differentiation based on clinical features alone (Mitchell et al., 2008) has AUC .63 when employed in practice (Mitchell et al., 2011). The test would be expected to have similar performance to that reported by some clinical screening tools such as Bipolar Spectrum Diagnostic Scale (BSDS) (Ghaemi et al., 2005), Hypomania Checklist (HCL-32) (Angst et al., 2005) and the Mood Disorder Questionnaire (MDQ) (Hirschfeld et al., 2000). A recent application of two of these screening tools in primary care estimated AUC for HCL-32 of .81, and for BSDS .71 (Smith et al., 2011).

The neurobiological interpretation of the difference in Glx levels remains uncertain, since differences in Glx could be explained be alterations in either glutamate or glutamine levels. As technical advances begin to allow accurate differentiation of glutamate from glutamine, changes in glutamine:glutamate ratio appear to be more closely related to brain glutamatergic function. It is interesting
that during unipolar depression some studies have reported reduced glutamine (Walter et al., 2009) whereas in episodes of mania an increased ratio has been observed (Öngür et al., 2008).

The differences in anterior cingulate Glx reported may reflect altered glutamatergic signalling or neuronal function in that region. In keeping with this, recent functional imaging data points to differences between bipolar and unipolar depressed individuals in regions including anterior cingulate (Bertocci et al., 2012; Mourão-Miranda et al., 2012). It is of course possible that similar Glx differences by diagnosis may also be demonstrable in other brain regions, since a similar trend towards Glx reductions in unipolar depression and elevations in bipolar disorder has been noted in other grey matter regions (Yüksel and Öngür, 2010).

The origins of the heterogeneity in Glx effect observed remain unclear. Some studies have suggested the greatest alterations in MRS measures are associated with particular clinical characteristics such as anhedonia (Walter et al., 2009), or may emerge with disease progression over time (Portella et al., 2011). However, against this, one of the greatest magnitude effects in unipolar depression seen was in a young population (Mirza et al., 2004).

Medication is a logical potential confounding factor on any biological measure. While included studies vary by mediation status, it seems unlikely to explain the overall effect found here. It has proven challenging to demonstrate an direct effect of common medications such as the SSRIs on this measure (Taylor et al., 2010), and even for more novel glutamatergic agents such as ketamine, results
are mixed (Stone et al., 2012; Taylor et al., 2011). It has also been possible to replicate the finding in unmedicated samples (Hasler et al., 2007). It has been reported that pretreatment spectroscopic measures may predict in part ketamine treatment response, specifically that a surrogate of glutamine concentration may do so (Salvadore et al., 2012). Given that ketamine has promising treatment effects in both bipolar and unipolar depression (Berman et al., 2000; Zarate et al., 2012), it is therefore interesting to note that the differences in Glx do not rule out underlying similarities in e.g. glutaminergic function across disorders.

Small differences in voxel positioning between centres could increase heterogeneity if the Glx effect was highly localised. However, if the effect is more widespread as noted above this would decrease the magnitude of this factor. Interestingly one study that included both more dorsal and more ventral voxel positions reported a similar effect of diagnosis on Glx in both, whereas an effect on GABA levels was more localised (Hasler et al., 2007).

In usual clinical practice, no biological measure is likely to replace high quality clinical assessment. However if differing concentrations of Glx reflect core differences in neurobiology between the disorders, they might predate the first hypomanic or manic episode. Anterior cingulate Glx therefore warrants study in groups at high risk of progression to bipolar disorder as a possible adjunct to clinical assessment. As poor response to treatment of depression is associated with elevated risk of conversion to bipolar disorder (Li et al., 2012), future studies could investigate whether higher Glx levels were predictive of this.
On available data, measurement of anterior cingulate Glx by MRS is a promising potential tool for differentiation of bipolar and unipolar depression. This potential effect requires direct validation within mixed clinical cohorts. Future studies could investigate in particular whether this imaging measure adds discriminatory ability to clinical measures alone or as part of a multi-modal imaging classifier.
Legends

Figure 1

A) Forest plot of standardised mean differences (SMD) with 95% confidence intervals of differences in Glx concentration in anterior cingulate cortex between currently depressed individuals and healthy controls. Subgroup estimates presented for those with unipolar and bipolar depression. Data from Portella et al. 2011 presented separately for those with first presentation with depression (f) and chronic course (ch).

B) Receiver operating curve for differentiation of bipolar and unipolar depression by measurement of Glx levels in anterior cingulate cortex. Best estimate of test performance shown (solid line) with 95% confidence intervals (dashed lines) and the line of no effect (dotted oblique line). Estimated performance using mean of healthy control Glx levels as criterion is shown (dot).

Table 1

Characteristics of included studies.
A

<table>
<thead>
<tr>
<th>Study</th>
<th>SMD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dager 2004</td>
<td>0.41</td>
<td>[0.12; 0.93]</td>
</tr>
<tr>
<td>Frye 2007</td>
<td>0.88</td>
<td>[0.14; 1.61]</td>
</tr>
<tr>
<td>Patel 2008</td>
<td>0.62</td>
<td>[-0.12; 1.36]</td>
</tr>
<tr>
<td>Xu 2013</td>
<td>-0.27</td>
<td>[-0.99; 0.46]</td>
</tr>
<tr>
<td>Random effects model</td>
<td>0.40</td>
<td>[-0.04; 0.85]</td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2=43.7\%$, $tau^2=0.0905$, $p=0.1492$

B

Comprehensive analysis of sensitivity and specificity with various thresholds.
<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Diagnostic criteria</th>
<th>Mean age of patients (yr)</th>
<th>Symptom severity</th>
<th>Medication status</th>
<th>Magnet field strength (T)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dager</td>
<td>32 bipolar I or II, predominantly depressed or mixed state vs 26 matched controls</td>
<td>DSM-IV</td>
<td>30.3</td>
<td>HAMD (17 item) 17.7</td>
<td>Medication free</td>
<td>PEPSI 1.5</td>
<td>left anterior cingulate data included</td>
</tr>
<tr>
<td>Frye</td>
<td>26 bipolar I or II depression vs 12</td>
<td>DSM-IV</td>
<td>35.6</td>
<td>MADRS 27.5 +/- 6.2</td>
<td>5 receiving lithium</td>
<td>PRESS 1.5</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Sample Description</td>
<td>Diagnostic Tool</td>
<td>Score</td>
<td>Comparison Tool</td>
<td>Medication Status</td>
<td>Imaging Modality</td>
<td>Imaging Parameters</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>-------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Patel 2008</td>
<td>28 bipolar I, currently depressed vs 10 controls</td>
<td>K-SADS</td>
<td>15.5</td>
<td>CDRS-R</td>
<td>Medication free at time of scan</td>
<td>PRESS 1.5</td>
<td></td>
</tr>
<tr>
<td>Xu 2013</td>
<td>12 bipolar depressed vs 20 controls</td>
<td>DSM-IV</td>
<td>37</td>
<td>HDRS 19 ±</td>
<td>Medication free</td>
<td>2D MRSI 3</td>
<td>data from 12 manic patients not used</td>
</tr>
<tr>
<td>Auer 2000</td>
<td>19 depressed (one bipolar) vs 18 controls</td>
<td>ICD-10</td>
<td>50.2</td>
<td>14 severe, 5 medium</td>
<td>7 medication free or sedatives only by ICD-10 severity criteria</td>
<td>PRESS 1.5</td>
<td></td>
</tr>
<tr>
<td>Pfleiderer</td>
<td>17 patients vs 17</td>
<td>DSM-IV</td>
<td>61</td>
<td>MADRS</td>
<td>All but one</td>
<td>STEAM 1.5</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Study</td>
<td>Group A</td>
<td>Group B</td>
<td>Measure</td>
<td>Score</td>
<td>Additional Details</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>-------</td>
<td>--------------------</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>controls</td>
<td>37.7 +8.8</td>
<td>patient medication free</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>Mirza</td>
<td>13 MDD patients</td>
<td>DSM-IV</td>
<td>15.54</td>
<td>CDRS-R</td>
<td>Psychotropic PRESS 1.5</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>Hasler</td>
<td>20 MDD patients</td>
<td>DSM-IV</td>
<td>34</td>
<td>MADRS 27</td>
<td>Unmedicated PRESS-based J-editing 3</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>Horn</td>
<td>22 MDD patients</td>
<td>ICD-10</td>
<td>37.62</td>
<td>HAMD (21 item)</td>
<td>All patients PRESS 3</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>Portella</td>
<td>10 first episode 19 chronic MDD patients</td>
<td>DSM-IV-TR</td>
<td>44.5 (first episode)</td>
<td>HDRS 13</td>
<td>All patients PRESS 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>vs 13 controls</td>
<td></td>
<td>59.08 +8.8</td>
<td>naive</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>vs 20 controls</td>
<td></td>
<td>59.08 +8.8</td>
<td>naive</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>vs 20 controls</td>
<td></td>
<td>59.08 +8.8</td>
<td>naive</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Portella</td>
<td>15 controls</td>
<td>(chronic)</td>
<td>50.96</td>
<td>episode)</td>
<td>21.39 + remission</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MDD patients vs controls</td>
<td></td>
<td>50.96</td>
<td>episode)</td>
<td>21.39 + remission</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MDD patients vs controls</td>
<td></td>
<td>50.96</td>
<td>episode)</td>
<td>21.39 + remission</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MDD patients vs controls</td>
<td></td>
<td>50.96</td>
<td>episode)</td>
<td>21.39 + remission</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MDD patients vs controls</td>
<td></td>
<td>50.96</td>
<td>episode)</td>
<td>21.39 + remission</td>
<td></td>
</tr>
<tr>
<td>Chen</td>
<td>15 MDD patients vs 15 controls</td>
<td>DSM-IV</td>
<td>HDRS</td>
<td>No</td>
<td>PRESS</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------------------------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td>27.87</td>
<td></td>
<td>23.33 ±</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.5 (chronic) not used

4.36 medication for at least two weeks
References


Schwarzer, G., 2013. meta: Meta-analysis with r, ed.


