Supplementary methods

Meta-analysis

All analyses were conducted using the grey matter template included in AES-SDM using a voxel size of 1mm$^3$. We used a significance threshold of $p < .005$ with a peak Z value of $>1$ and cluster size of $>10$ voxels for all main analyses. We used a more conservative threshold of $p < .0005$ for meta-regressions as these are more susceptible to false positives, particularly when conducted on a small number of studies$^1$. These thresholds have been found to provide the optimal balance of sensitivity and specificity in previous research$^2$. As we had a large number of studies in each diagnostic group, we used a less conservative threshold of $p < .001$ for the comparison between MDD and BD. Regions where results were significant in both conditions were submitted to a conjunction analysis to identify regions that were robustly affected in both conditions, accounting for error in the estimation of p-values within individual meta-analyses$^3$. To protect against false positives in the conjunction analysis, the threshold was set at $p < .005$ for this analysis, without the correction described by Radua and others$^3$ in order to test against the conjunction null hypothesis (i.e. no significant difference in one or fewer disorders) rather than the global null (i.e. no significant difference in any disorder).
Supplementary results

Results of heterogeneity analysis in regions where there was significant heterogeneity but no main disorder effect – Bipolar Disorder

We found that studies with euthymic samples found greater volume reductions in the right caudate than those in other mood states (peak MNI = -4, 10, 4, Z = 2.21, p < .001, 120 voxels, Figure 2A). Although our overall analysis showed no differences between patients and controls here, we performed a subgroup analysis using only studies with euthymic patients and found a cluster of significant volume reductions in the caudate (peak MNI = -6, 14, 0, Z = -3.61, p < .001, 804 voxels), suggesting that this may be specific to the euthymic state.

Meta-regression analysis also revealed that the percentage of patients taking antipsychotic medication at the time of scanning was positively associated with effect size in the left fusiform gyrus and parahippocampal gyrus (peak MNI = -30, -28, -28, Z = -2.20, p < .001, 134 voxels), indicating that studies with higher numbers of patients receiving antipsychotic treatment found greater decreases in grey matter volume in this region (Figure 2B). Again, this region was not present in our main analysis but was present when we conducted a subgroup analysis using only studies with >50% of patients taking antipsychotic medication (peak MNI = -34, -28, -32, Z = 2.01, p = 0.001, 95 voxels). Meta-regression analyses with methodological variables showed that studies with higher strength scanners showed smaller volumes relative to controls in the right caudate (peak MNI = 10, 6, 6, Z = 3.19, p < .001, 597 voxels) and the right superior temporal gyrus (peak MNI = 62, -30, 2, Z = 2.57, p < .001, 269 voxels). Meta-regression analysis with spatial smoothing level revealed a negative relationship between smoothing level and effect size in the left inferior frontal gyrus (peak MNI = -44, 16, 4, Z = 4.14, p < .001, 101 voxels).

All other meta-regression results were non-significant, or did not overlap with regions of heterogeneity.
## Supplementary tables

### Major Depression Patients

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Mean Age</th>
<th>Sex (M/F)</th>
<th>Antidepressant %</th>
<th>Duration of Illness (y)</th>
<th>State</th>
<th>HAMD-17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe (2010)</td>
<td>21</td>
<td>48.1</td>
<td>11, 10</td>
<td>90</td>
<td>6</td>
<td>7 Depressed, 2 partially remitted, 12 fully remitted</td>
<td>9.2</td>
</tr>
<tr>
<td>Amico (2011)</td>
<td>33</td>
<td>32</td>
<td>19, 14</td>
<td>82</td>
<td>3.4</td>
<td>Depressed</td>
<td>18.6</td>
</tr>
<tr>
<td>Arnone (2012a)</td>
<td>39</td>
<td>36.3</td>
<td>12, 27</td>
<td>0</td>
<td>14.3</td>
<td>Depressed</td>
<td>21.8</td>
</tr>
<tr>
<td>Arnone (2012b)</td>
<td>25</td>
<td>34.5</td>
<td>5, 20</td>
<td>0</td>
<td>9.4</td>
<td>Remitted</td>
<td>0.3</td>
</tr>
<tr>
<td>Bergouignan (2009)</td>
<td>20</td>
<td>33.2</td>
<td>3, 17</td>
<td>100</td>
<td>8.5</td>
<td>Depressed</td>
<td>23.1</td>
</tr>
<tr>
<td>Chaney (2014)</td>
<td>37</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Depressed</td>
<td>-</td>
</tr>
<tr>
<td>Cheng (2010)</td>
<td>68</td>
<td>29.9</td>
<td>21, 47</td>
<td>0</td>
<td>11</td>
<td>Depressed</td>
<td>22.3</td>
</tr>
<tr>
<td>de Azevedo-Marques Périco (2011)</td>
<td>20</td>
<td>29.9</td>
<td>5, 15</td>
<td>55</td>
<td>0.7</td>
<td>-</td>
<td>19.6</td>
</tr>
<tr>
<td>Grieve (2013)</td>
<td>102</td>
<td>33.8</td>
<td>48, 54</td>
<td>0</td>
<td>11.3</td>
<td>Depressed</td>
<td>21</td>
</tr>
<tr>
<td>Guo (2014)</td>
<td>44</td>
<td>27.5</td>
<td>22, 22</td>
<td>0</td>
<td>1.6</td>
<td>Depressed</td>
<td>25.2</td>
</tr>
<tr>
<td>Inkster (2011a)</td>
<td>49</td>
<td>47.6</td>
<td>16, 33</td>
<td>-</td>
<td>14.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Inkster (2011b)</td>
<td>96</td>
<td>50.4</td>
<td>35, 61</td>
<td>-</td>
<td>14.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Jia (2010a)</td>
<td>36</td>
<td>34.7</td>
<td>20, 16</td>
<td>0</td>
<td>1.8</td>
<td>Depressed</td>
<td>22.3</td>
</tr>
<tr>
<td>Jia (2010b)</td>
<td>16</td>
<td>34.2</td>
<td>5, 11</td>
<td>0</td>
<td>6.7</td>
<td>Depressed</td>
<td>24.6</td>
</tr>
<tr>
<td>Kim (2008)</td>
<td>22</td>
<td>38.5</td>
<td>0, 22</td>
<td>45</td>
<td>17.4</td>
<td>Depressed</td>
<td>-</td>
</tr>
<tr>
<td>Klauser (2015)</td>
<td>56</td>
<td>34</td>
<td>16, 40</td>
<td>-</td>
<td>10.3</td>
<td>29 depressed, 27 remitted</td>
<td>-</td>
</tr>
<tr>
<td>Kong (2014)</td>
<td>28</td>
<td>34.4</td>
<td>11, 17</td>
<td>0</td>
<td>2.11</td>
<td>Depressed</td>
<td>21.7</td>
</tr>
<tr>
<td>Kroes (2011)</td>
<td>29</td>
<td>33.4</td>
<td>8, 21</td>
<td>76</td>
<td>-</td>
<td>Depressed</td>
<td>-</td>
</tr>
<tr>
<td>Lai (2010)</td>
<td>16</td>
<td>37.9</td>
<td>5, 11</td>
<td>0</td>
<td>0.4</td>
<td>Depressed</td>
<td>29.1</td>
</tr>
</tbody>
</table>

### Healthy Controls

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Mean Age</th>
<th>Sex (M/F)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe (2010)</td>
<td>42</td>
<td>48</td>
<td>22, 20</td>
<td></td>
</tr>
<tr>
<td>Amico (2011)</td>
<td>64</td>
<td>30.4</td>
<td>36, 28</td>
<td>1.5</td>
</tr>
<tr>
<td>Arnone (2012a)</td>
<td>66</td>
<td>32.1</td>
<td>20, 46</td>
<td>1.5</td>
</tr>
<tr>
<td>Arnone (2012b)</td>
<td>66</td>
<td>32.1</td>
<td>20, 46</td>
<td>1.5</td>
</tr>
<tr>
<td>Bergouignan (2009)</td>
<td>21</td>
<td>28.2</td>
<td>7, 14</td>
<td>1.5</td>
</tr>
<tr>
<td>Chaney (2014)</td>
<td>46</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>Cheng (2010)</td>
<td>68</td>
<td>30.5</td>
<td>21, 47</td>
<td>1.5</td>
</tr>
<tr>
<td>de Azevedo-Marques Périco (2011)</td>
<td>94</td>
<td>30.2</td>
<td>41, 53</td>
<td>3</td>
</tr>
<tr>
<td>Grieve (2013)</td>
<td>34</td>
<td>31.5</td>
<td>18, 16</td>
<td>3</td>
</tr>
<tr>
<td>Guo (2014)</td>
<td>44</td>
<td>29.4</td>
<td>20, 24</td>
<td>1.5</td>
</tr>
<tr>
<td>Inkster (2011a)</td>
<td>183</td>
<td>48</td>
<td>73, 110</td>
<td>1.5</td>
</tr>
<tr>
<td>Inkster (2011b)</td>
<td>183</td>
<td>48</td>
<td>73, 110</td>
<td>3</td>
</tr>
<tr>
<td>Jia (2010a)</td>
<td>52</td>
<td>37.1</td>
<td>24, 28</td>
<td>3</td>
</tr>
<tr>
<td>Jia (2010b)</td>
<td>52</td>
<td>37.1</td>
<td>24, 28</td>
<td>3</td>
</tr>
<tr>
<td>Kim (2008)</td>
<td>25</td>
<td>35.3</td>
<td>0, 25</td>
<td>1.5</td>
</tr>
<tr>
<td>Klauser (2015)</td>
<td>33</td>
<td>34.7</td>
<td>12, 21</td>
<td>1.5</td>
</tr>
<tr>
<td>Kong (2014)</td>
<td>28</td>
<td>32.1</td>
<td>14, 14</td>
<td>1.5</td>
</tr>
<tr>
<td>Kroes (2011)</td>
<td>29</td>
<td>32.5</td>
<td>13, 16</td>
<td>3</td>
</tr>
<tr>
<td>Lai (2010)</td>
<td>15</td>
<td>34.3</td>
<td>4, 11</td>
<td>3</td>
</tr>
</tbody>
</table>

### Methods

<table>
<thead>
<tr>
<th>Field strength (Tesla)</th>
<th>Smoothing (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>10</td>
</tr>
<tr>
<td>1.5</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Study</td>
<td>N</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Lai and Wu (2014)</td>
<td>38</td>
</tr>
<tr>
<td>Lee (2011)</td>
<td>47</td>
</tr>
<tr>
<td>Leung (2009)</td>
<td>17</td>
</tr>
<tr>
<td>Ma (2012a)</td>
<td>18</td>
</tr>
<tr>
<td>Ma (2012b)</td>
<td>17</td>
</tr>
<tr>
<td>MacGregor Legge (2015)</td>
<td>67</td>
</tr>
<tr>
<td>Machino (2014)</td>
<td>29</td>
</tr>
<tr>
<td>Modinos (2014)</td>
<td>23</td>
</tr>
<tr>
<td>Nakano (2014)</td>
<td>36</td>
</tr>
<tr>
<td>Peng (2011)</td>
<td>22</td>
</tr>
<tr>
<td>Peng (2014a)</td>
<td>18</td>
</tr>
<tr>
<td>Peng (2014b)</td>
<td>20</td>
</tr>
<tr>
<td>Redlich (2014)</td>
<td>58</td>
</tr>
<tr>
<td>Rodriguez-Cano (2014)</td>
<td>32</td>
</tr>
<tr>
<td>Salvadore (2011a)</td>
<td>58</td>
</tr>
<tr>
<td>Salvadore (2011b)</td>
<td>27</td>
</tr>
<tr>
<td>Scheuerecker (2010)</td>
<td>13</td>
</tr>
<tr>
<td>Serra-Blasco (2013a)</td>
<td>22</td>
</tr>
<tr>
<td>Serra-Blasco (2013b)</td>
<td>22</td>
</tr>
<tr>
<td>Serra-Blasco (2013c)</td>
<td>22</td>
</tr>
<tr>
<td>Shah (1998a)</td>
<td>20</td>
</tr>
<tr>
<td>Shah (1998b)</td>
<td>20</td>
</tr>
<tr>
<td>Soriano-Mas (2011)</td>
<td>70</td>
</tr>
<tr>
<td>Sprengelmeyer (2011)</td>
<td>17</td>
</tr>
</tbody>
</table>
### Supplementary Table 1. Characteristics of major depression studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size</th>
<th>Male (%)</th>
<th>Female (%)</th>
<th>Mean Age (y)</th>
<th>Depression Score</th>
<th>M</th>
<th>F</th>
<th>Years</th>
<th>HAMD-17 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stratmann (2014)</td>
<td>132</td>
<td>37.9</td>
<td>56, 76</td>
<td>94</td>
<td>7.8</td>
<td>3</td>
<td>8</td>
<td></td>
<td>20.5 132 37.8 58, 74</td>
</tr>
<tr>
<td>Treadway (2009)</td>
<td>19</td>
<td>35.2</td>
<td>9, 10</td>
<td>0</td>
<td>12.9</td>
<td>3</td>
<td>12</td>
<td></td>
<td>21.5 19 30.3 9, 10</td>
</tr>
<tr>
<td>Wagner (2011)</td>
<td>30</td>
<td>0</td>
<td>5, 25</td>
<td>-</td>
<td>6</td>
<td>1.5</td>
<td>12</td>
<td></td>
<td>20.1 30 0 5, 25</td>
</tr>
<tr>
<td>Wang (2012)</td>
<td>18</td>
<td>34</td>
<td>9, 9</td>
<td>0</td>
<td>0.4</td>
<td>3</td>
<td>6</td>
<td></td>
<td>25 18 35 9, 9</td>
</tr>
<tr>
<td>Wang (2014)</td>
<td>13</td>
<td>30.9</td>
<td>0, 13</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>4</td>
<td></td>
<td>- 10 29.8 0, 10</td>
</tr>
<tr>
<td>Yoshikawa (2006)</td>
<td>11</td>
<td>48.5</td>
<td>0, 11</td>
<td>0</td>
<td>-</td>
<td>1.5</td>
<td>8</td>
<td></td>
<td>- 29 48.6 0, 29</td>
</tr>
<tr>
<td>Zou (2010)</td>
<td>23</td>
<td>31.1</td>
<td>10, 13</td>
<td>0</td>
<td>0.6</td>
<td>3</td>
<td>10</td>
<td></td>
<td>24.4 23 36.6 10, 13</td>
</tr>
</tbody>
</table>

M = male, F = female, y = years, HAMD-17 = Hamilton Depression Rating Scale 17-item, † = unpublished voxel-based morphometry data from the sample reported in this study.
<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Age (M,F)</th>
<th>Lithium (%)</th>
<th>Antipsychotic medication (%)</th>
<th>Duration of illness (y)</th>
<th>Subtype</th>
<th>State</th>
<th>HAMD -17</th>
<th>YMRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adler (2005)45</td>
<td>32</td>
<td>31.2</td>
<td>19, 13</td>
<td>-</td>
<td>8.7</td>
<td>BD-I</td>
<td>5 manic, 2 depressed, 25 euthymic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Almeida (2009)46</td>
<td>27</td>
<td>31.9</td>
<td>10, 17</td>
<td>-</td>
<td>59.3</td>
<td>BD-I</td>
<td>17 euthymic, 10 depressed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambrosi (2013)47</td>
<td>20</td>
<td>42</td>
<td>5, 15</td>
<td>35</td>
<td>50.0</td>
<td>BD-II</td>
<td>12.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown (2011)48</td>
<td>15</td>
<td>46.2</td>
<td>7, 8</td>
<td>-</td>
<td>20.0</td>
<td>BD-I</td>
<td>19.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen (2007)50</td>
<td>24</td>
<td>38.2</td>
<td>6, 18</td>
<td>50</td>
<td>0.0</td>
<td>14.2</td>
<td>BD-I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen (2012)51</td>
<td>18</td>
<td>32</td>
<td>18, 0</td>
<td>83</td>
<td>16.7</td>
<td>4.2</td>
<td>BD-I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cui (2011)52</td>
<td>24</td>
<td>28.4</td>
<td>15, 9</td>
<td>-</td>
<td>6.1</td>
<td>BD-I</td>
<td>Manic 25.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Azevedo-Marques Périco</td>
<td>26</td>
<td>29.9</td>
<td>5, 15</td>
<td>23.1</td>
<td>42.0</td>
<td>0.5</td>
<td>BD-I</td>
<td>7.5</td>
<td>7.4</td>
</tr>
<tr>
<td>Doris (2004)53</td>
<td>11</td>
<td>40.5</td>
<td>6, 5</td>
<td>-</td>
<td>45.5</td>
<td>16.2</td>
<td>BD-I</td>
<td>8.3</td>
<td></td>
</tr>
<tr>
<td>Eker (2014)54</td>
<td>28</td>
<td>36.4</td>
<td>16, 12</td>
<td>23</td>
<td>17.0</td>
<td>16.3</td>
<td>BD-I</td>
<td>2.3</td>
<td>1</td>
</tr>
<tr>
<td>Emsell (2013)55</td>
<td>60</td>
<td>42</td>
<td>31, 29</td>
<td>77</td>
<td>56.7</td>
<td>13</td>
<td>BD-I</td>
<td>13.1</td>
<td></td>
</tr>
<tr>
<td>Ha (2009)56</td>
<td>23</td>
<td>35.2</td>
<td>8, 15</td>
<td>30</td>
<td>34.8</td>
<td>10.5</td>
<td>BD-II</td>
<td>10.5</td>
<td></td>
</tr>
<tr>
<td>Ha (2009a)56</td>
<td>23</td>
<td>35.6</td>
<td>8, 15</td>
<td>35</td>
<td>43.5</td>
<td>10.4</td>
<td>BD-I</td>
<td>4.5</td>
<td>1.2</td>
</tr>
<tr>
<td>Haldane (2008)57</td>
<td>44</td>
<td>42.7</td>
<td>20, 24</td>
<td>-</td>
<td>50.0</td>
<td>16.3</td>
<td>BD-I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kempton (Unpublished)58</td>
<td>26</td>
<td>42.1</td>
<td>9, 17</td>
<td>35</td>
<td>15.0</td>
<td>16.1</td>
<td>24 BD-I, 2 BD-II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li (2011)59</td>
<td>24</td>
<td>28.4</td>
<td>15, 9</td>
<td>83</td>
<td>-</td>
<td>6</td>
<td>BD-I</td>
<td>13.1</td>
<td></td>
</tr>
<tr>
<td>Lochhead (2004)50</td>
<td>11</td>
<td>38.2</td>
<td>6, 5</td>
<td>-</td>
<td>13.9</td>
<td>7 BD-I, 4 BD-II</td>
<td>11 depressed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McDonald (2005)52</td>
<td>37</td>
<td>40.7</td>
<td>15, 22</td>
<td>59</td>
<td>24.3</td>
<td>17.8</td>
<td>BD-I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McIntosh (2004a)53</td>
<td>26</td>
<td>40.5</td>
<td>14, 12</td>
<td>-</td>
<td>-</td>
<td>BD-I</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McIntosh (2004b)53</td>
<td>19</td>
<td>39.7</td>
<td>7, 12</td>
<td>-</td>
<td>-</td>
<td>BD-I</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>M</td>
<td>F</td>
<td>y</td>
<td>HAMD-17</td>
<td>BD-I</td>
<td>YMRS</td>
<td>Manic</td>
<td>Euthymic 1st</td>
<td>Depressed 1st</td>
</tr>
<tr>
<td>---------------------</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>---------</td>
<td>-----------------------</td>
<td>------</td>
<td>-------</td>
<td>--------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Molina (2011)</td>
<td>19</td>
<td>38.3</td>
<td>12, 7</td>
<td>84</td>
<td>5.3</td>
<td>12</td>
<td>BD-I</td>
<td>Euthymic</td>
<td>-</td>
</tr>
<tr>
<td>Narita (2011a)</td>
<td>17</td>
<td>41.4</td>
<td>9, 8</td>
<td>65</td>
<td>11.8</td>
<td>6.2</td>
<td>BD-II</td>
<td>6 euthymic, 9 depressed, 2 manic</td>
<td>-</td>
</tr>
<tr>
<td>Narita (2011b)</td>
<td>14</td>
<td>40.2</td>
<td>8, 6</td>
<td>76</td>
<td>21.4</td>
<td>8.6</td>
<td>BD-II</td>
<td>2 euthymic, 10 depressed, 2 manic</td>
<td>-</td>
</tr>
<tr>
<td>Nugent (2006a)</td>
<td>16</td>
<td>37</td>
<td>5, 11</td>
<td>0</td>
<td>0.0</td>
<td>17</td>
<td>-</td>
<td>Depressed</td>
<td>-</td>
</tr>
<tr>
<td>Nugent (2006b)</td>
<td>20</td>
<td>41</td>
<td>5, 15</td>
<td>31</td>
<td>2.8</td>
<td>23</td>
<td>-</td>
<td>Depressed</td>
<td>-</td>
</tr>
<tr>
<td>Redlich (2014)</td>
<td>58</td>
<td>37.5</td>
<td>21, 37</td>
<td>26</td>
<td>64.0</td>
<td>14.2</td>
<td>BD-I</td>
<td>Depressed</td>
<td>21</td>
</tr>
<tr>
<td>Rocha-Rego (2013a)</td>
<td>26</td>
<td>41.5</td>
<td>12, 14</td>
<td>38</td>
<td>30.8</td>
<td>15.8</td>
<td>BD-I</td>
<td>Euthymic</td>
<td>6.6</td>
</tr>
<tr>
<td>Rocha-Rego (2013b)</td>
<td>14</td>
<td>37.6</td>
<td>6, 8</td>
<td>-</td>
<td>0.0</td>
<td>18.8</td>
<td>BD-I</td>
<td>Euthymic</td>
<td>4.2</td>
</tr>
<tr>
<td>Scherk (2008)</td>
<td>35</td>
<td>43.3</td>
<td>18, 17</td>
<td>34</td>
<td>51.4</td>
<td>14.4</td>
<td>BD-I</td>
<td>Euthymic</td>
<td>2.5</td>
</tr>
<tr>
<td>Shepherd (2014)</td>
<td>30</td>
<td>39.1</td>
<td>12, 18</td>
<td>-</td>
<td>-</td>
<td>13.5</td>
<td>BD-I</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Stanfield (2009)</td>
<td>66</td>
<td>36.4</td>
<td>30, 36</td>
<td>-</td>
<td>47.0</td>
<td>15.4</td>
<td>BD-I</td>
<td>Euthymic</td>
<td>-</td>
</tr>
<tr>
<td>Tang (2014)</td>
<td>27</td>
<td>32</td>
<td>10, 17</td>
<td>33</td>
<td>-</td>
<td>4.2</td>
<td>-</td>
<td>Depressed</td>
<td>19.7</td>
</tr>
<tr>
<td>Yatham (2007)</td>
<td>15</td>
<td>36</td>
<td>6, 9</td>
<td>-</td>
<td>0.0</td>
<td>3.9</td>
<td>BD-I</td>
<td>Manic</td>
<td>-</td>
</tr>
<tr>
<td>Yüksel (2012)</td>
<td>27</td>
<td>32.9</td>
<td>17, 10</td>
<td>48</td>
<td>0.0</td>
<td>-</td>
<td>BD-I</td>
<td>18 manic, 5 mixed, 4 euthymic</td>
<td>7.2</td>
</tr>
</tbody>
</table>

**Supplementary Table 2.** Characteristics of bipolar disorder studies included in the meta-analysis. M = male, F = female, y = years, HAMD-17 = Hamilton Depression Rating Scale 17-item, BD-I = bipolar disorder type I, BD-II = bipolar disorder type II, YMRS = Young Mania Rating Scale
### Table S3. Clusters showing differences between major depression and controls that did not meet our criteria for robustness

<table>
<thead>
<tr>
<th>Peak MNI coordinate</th>
<th>Z</th>
<th>P</th>
<th>Voxels</th>
<th>Brodmann areas</th>
<th>Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>-66,-16,-12</td>
<td>2.64</td>
<td>0.002</td>
<td>19</td>
<td>21</td>
<td>Left middle temporal gyrus</td>
</tr>
<tr>
<td>-30,-72,44</td>
<td>2.51</td>
<td>0.003</td>
<td>18</td>
<td>7</td>
<td>Left inferior parietal lobule</td>
</tr>
<tr>
<td>-36,-34,-18</td>
<td>2.74</td>
<td>0.001</td>
<td>13</td>
<td>37</td>
<td>Left fusiform gyrus</td>
</tr>
<tr>
<td>-4,-44,-22</td>
<td>2.46</td>
<td>0.004</td>
<td>13</td>
<td>-</td>
<td>Cerebellar vermis</td>
</tr>
<tr>
<td>-10,-72,0</td>
<td>-1.04</td>
<td>&lt;0.001</td>
<td>207</td>
<td>18</td>
<td>Left lingual gyrus</td>
</tr>
<tr>
<td>-8,-52,18</td>
<td>-1.06</td>
<td>&lt;0.001</td>
<td>197</td>
<td>30</td>
<td>Left precuneus</td>
</tr>
<tr>
<td>-48,-50,-40</td>
<td>-1.01</td>
<td>&lt;0.001</td>
<td>129</td>
<td>-</td>
<td>Left cerebellum, crus I</td>
</tr>
</tbody>
</table>

### Table S4. Clusters showing significant between study heterogeneity in major depression

<table>
<thead>
<tr>
<th>Peak MNI coordinate</th>
<th>Z</th>
<th>P</th>
<th>Voxels</th>
<th>Brodmann areas</th>
<th>Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>-46,4,4</td>
<td>1.82</td>
<td>&lt;0.001</td>
<td>436</td>
<td>48</td>
<td>Left insula</td>
</tr>
<tr>
<td>2,34,-16</td>
<td>1.33</td>
<td>&lt;0.001</td>
<td>248</td>
<td>11</td>
<td>Left gyrus rectus, medial orbitofrontal cortex</td>
</tr>
<tr>
<td>4,-2,42</td>
<td>1.15</td>
<td>&lt;0.001</td>
<td>203</td>
<td>23, 24</td>
<td>Right midcingulate area</td>
</tr>
<tr>
<td>22,2,-18</td>
<td>1.33</td>
<td>&lt;0.001</td>
<td>207</td>
<td>34</td>
<td>Right amygdala</td>
</tr>
<tr>
<td>52,6,-20</td>
<td>1.23</td>
<td>&lt;0.001</td>
<td>173</td>
<td>21</td>
<td>Right temporal pole, middle temporal gyrus</td>
</tr>
<tr>
<td>-22,-16,-14</td>
<td>1.50</td>
<td>&lt;0.001</td>
<td>51</td>
<td>35</td>
<td>Left hippocampus</td>
</tr>
<tr>
<td>-26,20,60</td>
<td>1.18</td>
<td>&lt;0.001</td>
<td>42</td>
<td>8</td>
<td>Left middle frontal gyrus</td>
</tr>
</tbody>
</table>
### Table S5. Clusters showing differences between bipolar disorder and controls did not meet our criteria for robustness

<table>
<thead>
<tr>
<th>Peak MNI coordinate</th>
<th>Z</th>
<th>P</th>
<th>Voxels</th>
<th>Brodmann areas</th>
<th>Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>-62,-60,-10</td>
<td>2.71</td>
<td>0.002</td>
<td>20</td>
<td>37</td>
<td>Left inferior temporal gyrus</td>
</tr>
<tr>
<td>36,60,6</td>
<td>2.64</td>
<td>0.002</td>
<td>13</td>
<td>10</td>
<td>Right middle frontal gyrus</td>
</tr>
<tr>
<td>-14,-60,-44</td>
<td>-1.21</td>
<td>0.003</td>
<td>23</td>
<td>-</td>
<td>Left cerebellum, hemispheric lobule VIII</td>
</tr>
<tr>
<td>4,-14,-22</td>
<td>-1.33</td>
<td>0.002</td>
<td>16</td>
<td>-</td>
<td>Right pons</td>
</tr>
<tr>
<td>-14,-30,-34</td>
<td>-1.39</td>
<td>0.001</td>
<td>19</td>
<td>-</td>
<td>Middle cerebellar peduncles</td>
</tr>
<tr>
<td>4,-38,-18</td>
<td>-1.17</td>
<td>0.003</td>
<td>20</td>
<td>-</td>
<td>Cerebellar vermis</td>
</tr>
<tr>
<td>42,-56,6</td>
<td>-1.34</td>
<td>0.002</td>
<td>20</td>
<td>37</td>
<td>Right middle temporal gyrus</td>
</tr>
<tr>
<td>-10,46,-30</td>
<td>-1.38</td>
<td>0.001</td>
<td>11</td>
<td>11</td>
<td>Left orbitofrontal cortex</td>
</tr>
<tr>
<td>-32,-54,38</td>
<td>-1.22</td>
<td>0.003</td>
<td>11</td>
<td>40</td>
<td>Left inferior parietal gyrus</td>
</tr>
</tbody>
</table>

### Table S6. Clusters showing significant between study heterogeneity in bipolar disorder

<table>
<thead>
<tr>
<th>Peak MNI coordinate</th>
<th>Z</th>
<th>P</th>
<th>Voxels</th>
<th>Brodmann areas</th>
<th>Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>8, 12, 12</td>
<td>5.88</td>
<td>&lt;0.001</td>
<td>657</td>
<td>25</td>
<td>Right caudate nucleus</td>
</tr>
<tr>
<td>-46,38,-14</td>
<td>4.42</td>
<td>&lt;0.001</td>
<td>425</td>
<td>47</td>
<td>Left inferior frontal gyrus</td>
</tr>
<tr>
<td>30,-36,-18</td>
<td>3.297</td>
<td>0.001</td>
<td>169</td>
<td>37</td>
<td>Right fusiform gyrus</td>
</tr>
<tr>
<td>48,36,-12</td>
<td>3.94</td>
<td>&lt;0.001</td>
<td>119</td>
<td>47</td>
<td>Right inferior frontal gyrus</td>
</tr>
<tr>
<td>-58,-66,-12</td>
<td>3.33</td>
<td>&lt;0.001</td>
<td>61</td>
<td>37</td>
<td>Left inferior occipital gyrus</td>
</tr>
<tr>
<td>6,18,22</td>
<td>2.82</td>
<td>0.002</td>
<td>55</td>
<td>24</td>
<td>Right anterior cingulate cortex</td>
</tr>
<tr>
<td>50,16,2</td>
<td>2.77</td>
<td>0.002</td>
<td>33</td>
<td>45, 48</td>
<td>Right inferior frontal gyrus</td>
</tr>
<tr>
<td>62,-20,-4</td>
<td>2.72</td>
<td>0.003</td>
<td>30</td>
<td>21,22</td>
<td>Right superior temporal gyrus, middle temporal gyrus</td>
</tr>
<tr>
<td>6,-14,10</td>
<td>3.03</td>
<td>0.002</td>
<td>30</td>
<td>-</td>
<td>Right thalamus</td>
</tr>
<tr>
<td>-20,-22,-22</td>
<td>2.89</td>
<td>0.002</td>
<td>28</td>
<td>30</td>
<td>Left parahippocampal gyrus</td>
</tr>
<tr>
<td>-42,-46,46</td>
<td>2.60</td>
<td>0.003</td>
<td>24</td>
<td>40</td>
<td>Left inferior parietal gyrus</td>
</tr>
<tr>
<td>-52, 34, 24</td>
<td>3.39</td>
<td>0.001</td>
<td>18</td>
<td>45</td>
<td>Left inferior frontal gyrus, pars triangularis</td>
</tr>
</tbody>
</table>
Supplementary figures

**Figure S1.** Literature search

Results from 2 unpublished studies acquired from authors

14951 results from literature search

340 selected for full text review

14169 excluded based on title and abstract review (not relevant topic, duplicate search results, not VBM, ROI approach, did not use an adult sample, comorbid conditions)

256 excluded (ROI approach, sample overlap, no coordinates reported, did not use an adult sample, shared sample with another study, did not compare with a healthy group, comorbid neurological illness, did not use VBM analysis)

5 studies identified from review of reference lists

84 included in analyses
Figure S2. A) Results of mood state comparison in bipolar disorder. B) Results of meta-regression with antipsychotic medication load in bipolar disorder. C & D) Results of meta-regression with scanner field strength in major depression.
Supplementary references


69 Shepherd AM, Quidé Y, Laurens KR, O'Reilly N, Rowland JE, Mitchell PB et al. Shared intermediate phenotypes for schizophrenia and bipolar disorder: neuroanatomical


