Possible structural abnormality of the brainstem in unipolar depressive illness: a transcranial ultrasound and diffusion tensor magnetic resonance imaging study

J D Steele, M E Bastin, J M Wardlaw, K P Ebmeier

Background: Most empirically derived antidepressants increase monoamine levels. The nuclei of cells synthesising these monoamines are located in the brainstem, and projection tracts such as the medial forebrain bundle reach virtually all other brain areas. Two studies of unipolar depressive illness using transcranial ultrasound have reported reduced echogenicity of the brainstem midline in unipolar depressed patients. This may be consistent with disruption of white matter tracts, including the medial forebrain bundle, and it has been suggested that the effect of such disruption could be reversed by antidepressants.

Objective: To replicate these findings in a group of unipolar depressed patients and controls.

Methods: Fifteen unipolar depressed patients and 15 controls were studied using transcranial ultrasound imaging and diffusion tensor magnetic resonance imaging (DT-MRI).

Results: No difference in echogenicity of the brainstem midline of unipolar depressed patients was found. A possible trend (Cohen’s d = 0.39) in the direction of previous studies was found. Although the echogenicity of the brainstem midline of the control group was found to be similar to previous reports, there was no reduction in the patient group. Additionally, no structural abnormality of the brainstem was identified using DT-MRI.

Conclusions: While these data do not replicate the findings of previous studies reporting a significant reduction in the echogenicity of the brainstem midline in unipolar depressed patients, the ultrasound investigation indicated that there may be a trend in this direction. Given the importance of identifying the causes of depressive illness, it is important that other groups attempt similar studies.

Virtually all empirically derived antidepressants are known to increase monoamine levels. This observation led to the well known monoamine hypothesis of mood disorder. The nuclei of cells synthesising serotonin, noradrenaline, and dopamine are located in the brainstem, and axonal projections from these nuclei to other brain areas form white matter tracts such as the medial forebrain bundle. The anterior cingulate has often been reported as having abnormal function and structure in depressive illness.

Studies on primates indicate that there is a topographical projection from the cingulate to the pons. The anterior cingulate projects to the midline pons and the posterior cingulate to the lateral pons. Although virtually all regions of the anterior cingulate have been reported as abnormal in depressive illness, the subgenual region has been a particular focus of reports. In primates, the subgenual anterior cingulate projections to the brainstem are restricted to the posterior midline periaqueductal grey matter and raphe serotonergic nuclei.

A series of studies using transcranial ultrasound has reported a structural abnormality of the midbrain midline of patients with unipolar depressive illness compared to controls. The same structural abnormality has also been reported when depressed patients have been compared to non-depressed patients, both having a variety of neurological diseases, for example, Parkinson’s disease, and Huntington’s disease, but not multiple sclerosis. The structural abnormality was reported to occur in unipolar depressed patients, was unrelated to severity of current illness, and was absent in patients with bipolar disorder and schizophrenia. In all cases the abnormality was reduced echogenicity of the midline. It should be noted that Becker and colleagues referred to the abnormality as occurring within the raphe, and not specifically the medial forebrain bundle, due to many other fibre tracts being present in this region.

These ultrasound investigations have been supplemented by $T_2$ weighted MRI studies. Increased intensity of the midline has been reported for unipolar depressed patients when compared to bipolar patients and controls in a retrospective study using $T_2$ weighted MRI. Additionally, a prospective study of depressed patients with Parkinson’s disease reported a shift of signal in the midbrain using a semi-quantitative assessment of relaxation time. It has been suggested that reduced echogenicity of the midline and increased $T_2$ weighted signal intensity may be consistent with disruption of myelinated fibre tracts running along the brainstem, which include the monoaminergic tracts. There are few postmortem studies of the brainstem of patients who suffered from major depression. Nevertheless, one small study has reported that patients with unipolar depression have a distinct disruption of the mesencephalic fibre tracts which may be consistent with both the ultrasound and MRI findings.

It has been suggested that such an abnormality could result in a reduction of brain monoamines and be reversed by antidepressants. Clearly then, the collective replicated ultrasound, MRI, and histopathological studies are an important work with major implications for understanding the potential causes of unipolar depressive illness.

Abbreviations: BDI, Beck Depression Inventory; DT-MRI, diffusion tensor magnetic resonance imaging; DW, diffusion weighted; EP, echo planar; FA, fractional anisotropy; FDR, false discovery rate; FSE, fast spin echo; MNI, Montreal Neurological Institute
A literature search did not reveal any reports of attempted replication by independent groups. Therefore, the aim of this study was to attempt replication of these findings using transcranial ultrasound imaging and to explore the use of a new technique, diffusion tensor magnetic resonance imaging (DT-MRI). The latter may be the imaging modality of choice, since it obtains information on the structural integrity of white matter tracts which can be analysed in an objective automated manner. In contrast to ultrasound imaging, it does not rely on the clinical skill of an operator trying to record a subtle abnormality in a low signal/noise image, and removes the possibility of knowledge of the patient’s condition introducing bias. Although ultrasound images may be analysed independently and blindly after acquisition, the acquisition process itself (and hence images) could be influenced by the operator being aware of the patient’s appearance and guessing their clinical condition.

A power analysis was done for the planned ultrasound investigation. For the first pilot study, mean (standard deviation, SD) echogenicity scores of 1.3 (0.47) and 2.8 (0.64) for the patient and control groups, respectively, were reported. For the later study, echogenicity scores of 1.4 (0.6) and 2.8 (0.5) were reported. These correspond to Cohen’s d effect sizes of 2.68 and 2.54, respectively, which are extremely large. Based on an effect size of 2.5, of 0.05, 15 patients, and 15 controls, the power of this current study was calculated to be 100%.

**METHODS**

**Subjects**

Permission for the study was obtained from the local ethics committee and written informed consent obtained from each subject. A total of 30 right handed subjects participated in the study. One subject was unable to tolerate scanning, resulting in data from 15 patients and 14 controls (table 1). The power of the ultrasound investigation was recalculated for 29 subjects and still found to be 100%. The subjects also took part in an fMRI study which has been reported separately.

In addition to standard contraindications to MRI, subjects were excluded if they had a history of significant head injury or structural brain abnormality (including vascular disease), substance (including alcohol) misuse, a physical disorder known in some cases to be associated with abnormal brain structure (for example, epilepsy), or were receiving medication which might alter brain structure (for example, steroids). The upper age limit was not restricted to 65 years. All patients had an unequivocal diagnosis of recurrent unipolar major depressive illness in the opinion of both the patient’s consultant and the author involved in recruitment (JDS). Patients were excluded if the diagnosis of the current or any previous illness was in doubt (for example, a possible previous manic episode). Most patients were receiving a variety of medications (table 2) at doses similar to those described in a previous study. Medication had remained at a constant level for at least 2 weeks prior to scanning. Patients were recruited from inpatients and outpatients at the Royal Edinburgh Hospital. Controls were acquired from work colleagues, friends, and relatives. Potential controls were excluded if they admitted to a history of any previous psychiatric illness.

Details of the subjects are given in table 1. The control group was matched on the basis of age, National Adult Reading Test score, and percentage of male subjects. All subjects completed a Beck Depression Inventory (BDI) as a screening test. Patients satisfied DSM-IV criteria for major depressive disorder and a Hamilton depression rating was obtained as an index of depression severity. The ratings and scans were obtained on the same day and time of day (late morning) because of diurnal variation in mood which was present in some patients.

**Ultrasound image acquisition and analysis**

Ultrasound images were obtained from subjects using a phased array system equipped with a Siemens Elegra 2 MHz transducer (Siemens, Erlangen, Germany). All images were acquired and rated by the same author (JMW), a consultant neuroradiologist with extensive experience of transcranial, including brainstem ultrasound, scanning. Images were obtained blind to subject details. Additionally, subjects were told not to speak or otherwise communicate with JMW before or during scanning. Depressed patients typically look unwell in their general demeanour, and so it is difficult to achieve full and effective blinding. This was considered relevant, since moving the ultrasound probe by a small amount can cause a midline echo to disappear, and the probe is unstable since it is hand held. Therefore, as a check on the effectiveness of blinding, JMW was asked upon completion of each scan session to guess whether the subject was a patient or a control.

Images of the midbrain with red nucleus and rostral pontine brainstem were obtained in an axial scanning plane through a preauricular acoustic bone window. Acquisition was standardised such that the contralateral skull surface was just visible, resulting in a typical penetration depth of 14 cm. The dynamic range was 28 dB with consequent high tissue contrast. Typically, around seven images were recorded from each subject which always included views obtained from both head sides. The objective was to record the clearest images of the midline echo from the lower midbrain.

| Table 1 Mean clinical characteristics of patient and control groups |
|-----------------------------|-----------------------------|
| Patients | Controls | p  |
| Age | 45.9 | 43.0 | NS |
| Females | 11 | 7 | NS |
| NART | 12.4 | 8.5 | NS |
| BDI | 36.9 | 1.1 | <0.001 |
| Hamilton | 27.3 | - | - |

Although there were more females in the patient group, this did not reach significance at a conventional 5% level. BDI, Beck Depression Inventory; NART, National Adult Reading Test; NS, not significant.

<table>
<thead>
<tr>
<th>Table 2 Patient medication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total daily dose/mg</td>
</tr>
<tr>
<td>Sertraline</td>
</tr>
<tr>
<td>Fluoxetine</td>
</tr>
<tr>
<td>Phenoxyzine</td>
</tr>
<tr>
<td>Lithium carbonate</td>
</tr>
<tr>
<td>L-Tryptophan</td>
</tr>
<tr>
<td>Venlafaxine</td>
</tr>
<tr>
<td>Mirtazapine</td>
</tr>
<tr>
<td>Modafinil</td>
</tr>
<tr>
<td>Clonazepam</td>
</tr>
<tr>
<td>Imipramine</td>
</tr>
<tr>
<td>Amitriptyline</td>
</tr>
<tr>
<td>Fluphenazine</td>
</tr>
<tr>
<td>Quetiapine</td>
</tr>
<tr>
<td>Sodium valproate</td>
</tr>
<tr>
<td>Diazepam</td>
</tr>
<tr>
<td>Zolpidem</td>
</tr>
<tr>
<td>No medication</td>
</tr>
</tbody>
</table>

Each dose reflects a daily dose of medication for a single patient. No patient was receiving more than one antidepressant simultaneously. Patients receiving antidepressants were also prescribed lithium carbonate as a standard augmentation strategy.
Digitised images were stored for later analysis in the standard Siemens Elegra format.

For analysis, the anonymised stored images were rated (by JMW) according to the semi-quantitative method described by Becker and colleagues.10 This is a four point scale of echogenicity of the brainstem raphe compared to the red nucleus: 1, raphe not visible/soiechogenic compared to adjacent brain parenchyma; 2, decreased echogenicity of the raphe as compared to the echogenicity of the red nucleus; 3, normal/identical echogenicity compared to the red nucleus; and 4, increased echogenicity and/or widened raphe structure. As with previous studies, the null hypothesis of no difference of midline echogenicity in patients compared to controls was investigated with a U test.

Additionally, for each subject a single scan was identified at the time of rating which showed the brainstem midline most clearly. Quantitative image analysis was done on these scans avoiding the need for subjective rating. A template image of the brainstem (fig 1) was constructed using a transverse slice through the lower midbrain of a high resolution $T_1$ weighted MRI scan (http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml) which conforms to Montreal Neurological Institute (MNI) space. Landmarks were defined on the template which could also be identified on the ultrasound scans. Coregistration of the ultrasound images to the template was done (blind to subject details) using the Matlab image processing toolbox (The Mathworks, Natick, MA), which includes routines for landmark based coregistration.

An affine transformation was used to coregister the images. The echo intensity of each image was normalised to the average intensity of the template image excluding a wide midline region (fig 2). The normalisation region included, but was not limited to, the red nucleus. A one group, two tailed $t$ test was used to test the null hypothesis of no difference in echo intensity at each pixel compared to the average intensity of the normalisation region. Correction for multiple testing was done using the false discovery rate (FDR) method with a conventional threshold of $p<0.05$. In a similar manner, a two group independent $t$ test was calculated to test the two tailed null hypothesis of no difference in average midline echo intensity of patients compared to controls.

**DT-MRI acquisition and analysis**

All MRI data were obtained using a GE Signa LX 1.5 T (General Electric, Milwaukee, WI) research dedicated scanner, equipped with a self shielding gradient set (22 mT/m maximum gradient strength) and manufacturer supplied “birdcage” quadrature head coil. Each subject underwent an axial $T_2$ weighted fast spin echo (FSE) sequence to identify silent brain pathology and a $T_1$ weighted volume scan. This was followed by a DT-MRI protocol specifically designed to image the brainstem.

For the DT-MRI acquisition, diffusion weighted (DW) images were acquired from 11 slice locations covering the region from the pons to the body of the lateral ventricles using a single shot spin echo planar (EP) imaging sequence. The centre slice was aligned with the midbrain-pontine junction. Sets of axial DW-EP images ($b=0$ and 1000 s/mm$^2$) were collected with diffusion gradients applied sequentially along six non-collinear directions.28 Seven acquisitions consisting of a baseline $T_2$ weighted EP image and six DW-EP images, a total of 49 images, were collected per slice position. The acquisition parameters for the DW-EP imaging sequence were 11 contiguous axial slices of 4 mm thickness, a field of view of 180 x 180 mm, an acquisition matrix of 96 x 96 (zero filled to 256 x 256), a TR of 8.0 s and a TE of 98.6 ms.

From the eddy current corrected MRI data, the apparent diffusion tensor of water (D) was calculated in each voxel from the signal intensities in the component EP images. Maps of the $T_2$ weighted signal intensity and fractional anisotropy (FA) were generated on a voxel by voxel basis from the sorted eigenvalues of D and converted into Analyze (Mayo Foundation, Rochester, MN) format.

Statistical pre-processing and analysis of the DT-MRI images used SPM99 (http://www.fil.ion.ac.uk.spm) and was therefore automated. For pre-processing, $T_2$ weighted EP images were spatially normalised to the SPM template using an affine transformation which generated both the spatially normalised image and the corresponding affine transformation matrix. As a manual check on the automated spatial normalisation procedure, coordinates corresponding to anatomical landmarks in each $T_2$ weighted EP image were compared to the corresponding coordinates in the high resolution $T_1$ weighted image; no problems with normalisation were found. The $T_2$ weighted EP image normalisation parameters were then applied to the FA images, which were smoothed with an 8 mm isotropic Gaussian filter. Visual inspection of the FA images did not reveal any obvious artefacts in the brainstem. Since signal detection is in part dependent on optimal smoothing, which cannot be specified accurately a priori, DT-MRI data were analysed in three different ways.

Firstly, the affine spatial transformation matrix for each subject was inverted to allow specification of a single stereotactic (MNI) coordinate and calculation of the corresponding coordinate in the original, non-normalised, non-smoothed FA images for each subject. This method of sampling the original data avoided smoothing and potential
additional loss of information due to interpolation which is part of spatial normalisation. MNI coordinates corresponding to a cubic volume of interest centred in the inferior midline of the midbrain were then defined and a t test calculated at each MNI coordinate for the FA images. Correction for multiple testing was done using the FDR technique.\(^{27}\) Secondly, the spatially normalised and smoothed FA images were analysed with a 5 mm diameter spherical volume of interest centred on the inferior midline of the midbrain. The MarsBar (http://marsbar.sourceforge.net/) SPM99 toolbox was used to test the null hypothesis of no difference between the patient and control groups for the average anisotropy indices within the volume of interest. Thirdly, a voxel based SPM99 analysis of the smoothed and spatially normalised FA images was done using a small volume correction with the same 5 mm diameter volume of interest defined in the inferior midline of the midbrain.

**RESULTS**

**Ultrasound imaging**

A \(\chi^2\) test indicated that the correct diagnosis of the subject was not guessed more often than would have happened by chance alone, \(p = 0.53\) (table 3). Table 4 shows the results of rating the anonymised recorded scans. The mean (SD) echogenicity scores were calculated as 2.07 (1.39) and 2.57 (1.16) for the patient and control groups, respectively. This difference is not significant (U test, \(p = 0.37\)). The score for the control group is very similar to Becker and colleagues’ previous reports: 2.80 (0.5 and 0.64). In contrast, the patient score is different from previous reports: 1.3 and 1.4 (0.47 and 0.60). Cohen’s \(d\) is 0.39, which is of small to medium size, and in the direction of previous studies. For an effect size of 0.39, an a priori power analysis was calculated using GPower,\(^{19}\) assuming an asymptotic relative efficiency of 0.955 for a U test. For an \(\alpha\) of 0.05, a study would have a power of 70% if it included a total of 173 subjects (two tailed hypothesis) or 132 subjects (one tailed hypothesis).

Figure 3 shows the average spatially coregistered and echo intensity normalised images for the patient and control groups. The hyperechogenic basal cisterns (BC) and aqueductal region (AQ) are visible as are hyperechogenic red nuclei (RN). The midline (M) structure also appears to be present in both the average images. Figure 4 shows the result of testing the two tailed null hypothesis of no difference between the average echo intensity of the normalisation region (fig 2) and any voxel in the image. The BC and AQ regions are significantly increased in intensity as are the RN on the sides of the brainstem closest to the probe. There are two small regions of significantly increased intensity in the midline (M) of the control group but not the patient group, but far larger are the regions of significantly decreased (D) echo intensity lateral to the midline (fig 4). Consistent with the results of the semi-quantitative analysis, when the patient and control groups were directly compared using an independent two group \(t\) test, no significant differences were found.

**DT-MRI**

No significant differences in FA between patient and control groups were identified in the midbrain using the three methods stated. As an exploratory post hoc analysis, the smoothing was varied from 0 to 10 mm, and the analyses repeated. No significant differences were identified at any of the smoothing levels. Table 5 shows the results of a voxel based analysis comparing the two groups without specifying the midbrain region of interest. Two regions of decreased FA (patients compared to controls) were identified outside the brainstem. The first is located in the right lateral temporal lobe and may lie within a region of reduced grey matter reported for patients with treatment resistant depressive illness.\(^{32}\) Inspection of the FA scans from each subject did not identify an obvious artefact in this region. The other significant region is located within the left uncinate fasciculus and might be related to previous reports of structural and functional abnormality of the left prefrontal and temporal lobes in depressive illness.\(^{31}\) However, this region does not remain significant after correction for multiple testing using the FDR method. No regions of increased FA were identified.

**DISCUSSION**

As noted earlier, previous studies by Becker and colleagues reported extremely large effect sizes. Based on this, including 29 subjects (as here) should have had virtual certainty of rejecting the null hypothesis at \(p < 0.05\). However, no significant difference between patient and control groups was found. With regard to blinding, depressed patients typically look unwell in their general demeanour, and so it is difficult to achieve full and effective blinding. It is unclear how effective the blinding was in previous studies, since a check on blinding was not described. The analysis of the coregistered ultrasound images from each group identified significant features within the brainstem, including

---

**Table 3**

<table>
<thead>
<tr>
<th>Guess</th>
<th>Patient</th>
<th>Control</th>
<th>DK</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Control</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>14</td>
</tr>
</tbody>
</table>

Although JMW did guess the correct identity of a patient more often than not, this did not reach significance at a conventional 5% level. DK, don’t know.

**Table 4**

<table>
<thead>
<tr>
<th></th>
<th>Score 1</th>
<th>Score 2</th>
<th>Score 3</th>
<th>Score 4</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>2.57 (1.16)</td>
</tr>
<tr>
<td>Patients</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>2.07 (1.39)</td>
</tr>
</tbody>
</table>

The difference between patient and control groups is not significant.

**Figure 3** Average echo intensity normalised images for control group (left) and patient group (right). Anterior is left and probe scanned from top. Basal cistern (BC), aqueduct (AQ), midline of midbrain (M), and red nucleus (RN) are indicated.
statistically significant red nuclei echoes. Consequently, it is unlikely that the images were of insufficient quality. This interpretation is additionally supported by the semi-quantitative analysis. Becker and colleagues reported an echogenicity score for the control group similar to that found here. The main difference between this and previous studies, is that the patient score was not found to be approximately half the control score. An echogenic midline was found for a much higher percentage of patients than in the studies of Becker and colleagues.

Inability to replicate previous results might be due to a difference between patient samples. However, the average Hamilton ratings of illness severity in previous studies were 22.8 (6.6) and 17.5 (8.4) which appears to be less than the 27.3 (10.7) in this study. Nevertheless, Becker and colleagues did not find any correlation between illness severity and echogenicity of the brainstem. Consequently, it is unlikely that imaging more severely unwell patients explains the discrepancy. However, if not due to illness severity, the discrepancy might be due to comorbidity. Patients with comorbid substance (for example, alcohol) misuse were excluded from our study. Alcohol misuse is very common in depressed patients, and is generally recognised to be a cause of widespread structural brain changes. It is unclear if similar patients were excluded in the previous studies. In this study, consistent with the results of the ultrasound investigation, no difference in FA of the brainstem of patients was found.

In summary, we were not able to replicate Becker and colleagues’ reports of a significant reduction in the echogenicity of the brainstem midline in unipolar depressed patients. Nevertheless, the ultrasound investigation indicated that there may be a trend in the direction of previous studies, suggesting that the magnitude of the effect is less than previously suggested, at least for the subject inclusion criteria adopted for this study. Given the implications of Becker and colleagues’ reports with regard to identifying the causes of depressive illness, it is important that other groups attempt similar studies. The power estimates detailed earlier may be useful in this regard.

ACKNOWLEDGEMENTS
The authors thank all Royal Edinburgh Hospital consultants who referred their patients and in particular Dr Diana Morrison. Mr Tom Anderson, Medical Physics, Edinburgh University, assisted with transfer of ultrasound images. Dr Georg Becker, Neurologische Universitätsklinik, Wurzburg, Germany, very helpfully commented on pilot ultrasound images of the midbrain obtained from two subjects during initial study planning. All MRI data were acquired at the SHEFC Brain Imaging Research Centre for Scotland (http://www.dcn.ed.ac.uk/bic).

Table 5 Regions of significantly reduced (p<0.001, uncorrected) FA of patient group compared to controls

<table>
<thead>
<tr>
<th>Region</th>
<th>MNI coordinate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right lateral temporal cortex</td>
<td>(64, -34, -26)*</td>
</tr>
<tr>
<td>Left uncinate fasciculus</td>
<td>(-24, 10, -14)</td>
</tr>
</tbody>
</table>

No regions of increased FA were identified.

<p>0.05, after correction for multiple testing using FDR method.

REFERENCES

31 Basser PJ. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. NMR Biomed 1995;8:333–44.
Possible structural abnormality of the brainstem in unipolar depressive illness: a transcranial ultrasound and diffusion tensor magnetic resonance imaging study

J D Steele, M E Bastin, J M Wardlaw and K P Ebmeier

J Neurol Neurosurg Psychiatry 2005 76: 1510-1515
doi: 10.1136/jnnp.2004.057612