Neurocognitive endophenotypes of obsessive-compulsive disorder

Lara Menzies, Sophie Achard, Samuel R. Chamberlain, Naomi Fineberg, Chi-Hua Chen, Natalia del Campo, Barbara J. Sahakian, Trevor W. Robbins and Ed Bullmore

1Brain Mapping Unit, University of Cambridge, 2Department of Psychiatry, Queen Elizabeth II Hospital, Welwyn Garden City, 3Department of Experimental Psychology, Behavioural and Clinical Neurosciences Institute, University of Cambridge, Cambridge, CB2 3EB, 4Department of Psychiatry, Addenbrooke’s Hospital, Cambridge, CB2 2QQ and 5Clinical Unit Cambridge, Addenbrooke’s Centre for Clinical Investigations, Clinical Pharmacology & Discovery Medicine, GlaxoSmithKline, Cambridge CB2 2QQ, UK

Correspondence to: Professor Ed Bullmore, Brain Mapping Unit, Department of Psychiatry, University of Cambridge, Addenbrooke’s Hospital, Cambridge, CB2 2QQ, UK
E-mail: etb23@cam.ac.uk

Endophenotypes (intermediate phenotypes) are objective, heritable, quantitative traits hypothesized to represent genetic risk for polygenic disorders at more biologically tractable levels than distal behavioural and clinical phenotypes. It is theorized that endophenotype models of disease will help to clarify both diagnostic classification and aetiological understanding of complex brain disorders such as obsessive-compulsive disorder (OCD). To investigate endophenotypes in OCD, we measured brain structure using magnetic resonance imaging (MRI), and behavioural performance on a response inhibition task (Stop-Signal) in 31 OCD patients, 31 of their unaffected first-degree relatives, and 31 unrelated matched controls. Both patients and relatives had delayed response inhibition on the Stop-Signal task compared with healthy controls. We used a multivoxel analysis method (partial least squares) to identify large-scale brain systems in which anatomical variation was associated with variation in performance on the response inhibition task. Behavioural impairment on the Stop-Signal task, occurring predominantly in patients and relatives, was significantly associated with reduced grey matter in orbitofrontal and right inferior frontal regions and increased grey matter in cingulate, parietal and striatal regions. A novel permutation test indicated significant familial effects on variation of the MRI markers of inhibitory processing, supporting the candidacy of these brain structural systems as endophenotypes of OCD.

In summary, structural variation in large-scale brain systems related to motor inhibitory control may mediate genetic risk for OCD, representing the first evidence for a neurocognitive endophenotype of OCD.

Keywords: neuroimaging; inhibition; obsessive-compulsive; multivoxel; familial

Abbreviations: ANOVA = analysis of variance; DLPFC = dorsolateral prefrontal cortex; LSD = Least Significant Difference; MADRS = Montgomery Asberg Depression Rating Scale; MNI = Montreal Neurological Institute; MRI = magnetic resonance imaging; NART = National Adult Reading Test; OCD = obsessive-compulsive disorder; OCI-R = Obsessive Compulsive Inventory - Revised; OFC = orbitofrontal cortex; PLS = partial least squares; QTL = quantitative trait locus; RIFG = right inferior frontal gyrus; SPECT = Single Photon Emission Computed Tomography; SSRT = stop-signal reaction time; YBOCS = Yale Brown Obsessive Compulsive Scale


Introduction

An endophenotype was originally defined in the 1960s as a ‘measurable component unseen by the unaided eye on the pathway between disease (phenotype) and distal genotype’ (John and Lewis, 1966; Gottesman and Shields, 1973). It is a heritable quantitative trait associated with increased genetic risk for a disorder and therefore present in both patients and their clinically unaffected relatives (Gottesman and Gould, 2003; Bearden and Freimer, 2006). Interest in endophenotypes (or intermediate phenotypes) has been stimulated by the difficulties encountered in establishing specific genetic causes for complex disorders by classical linkage or association designs. Although most major psychiatric syndromes are highly heritable, several decades of effort to discover causative genes has yielded disappointing results. It is argued that traditional clinical phenotypes, such as a
of the disorder may reflect a loss of normal inhibitory processes (Chamberlain et al., 2005).

According to theoretical models of OCD, symptoms and associated cognitive impairments emerge from disordered structure and function of large-scale neurocognitive networks, specifically ‘limbic’ or ‘affective’ cortico-striato-thalamic circuits including orbitofrontal cortex (OFC) (Graybiel and Rauch, 2000; Saxena et al., 2001). These circuits, first identified by anatomical studies in primates (Alexander et al., 1986; Lawrence et al., 1998), have been implicated in the pathophysiology of OCD by human imaging and lesion-based studies (Rapoport and Wise, 1988; Saxena, 2003). The most consistent finding from structural MRI measurements of selected regions-of-interest has been reduced grey matter volume of OFC in patients with OCD; there is less consistent evidence in support of volume changes in caudate nucleus, medial temporal lobe structures, and anterior cingulate cortex (Scarone et al., 1992; Robinson et al., 1995; Aylward et al., 1996; Jenike et al., 1996; Rosenberg and Keshavan, 1998; Szeszko et al., 1999; Gilbert et al., 2000; Kwon et al., 2003; Choi et al., 2004, 2006; Kang et al., 2004; Szeszko et al., 2004; Atmaca et al., 2007); see Supplementary Fig. 1 for detail.

More recently, studies have used computational techniques, such as voxel-based morphometry (Ashburner and Friston, 2000), to map local structural differences in case-control designs without restriction a priori to selected regions (Kim et al., 2001; Pujol et al., 2004; Valente et al., 2005). However, the few studies adopting this approach to date have produced inconsistent results (see Supplementary Fig. 2). There are various possible reasons for the limited replicability of imaging studies of OCD, including clinical heterogeneity or co-morbidity of patient samples, medication effects and small sample sizes in the context of conservative significance thresholds mandated by the large number of voxels to be tested in a whole brain approach. Another possibility is that regional or voxel level analysis, focused on local changes in brain structure, may be less than optimally powerful to detect case-control differences in brain structure which are theoretically expected at the more distributed level of large-scale neurocognitive systems.

Two prior imaging studies have used theoretically more appropriate methods designed to characterize brain abnormalities associated with OCD at a systems level. Soriano-Mas et al. (2007) described a whole brain profile of anatomical abnormality in patients with OCD compared to healthy controls by testing the sum of t-statistics over all voxels against a chi-square distribution (Worsley et al., 1995, 1997). Harrison et al. (2006) used a non-parametric test in conjunction with canonical variates analysis [implemented using NPAIRS software (Strother et al., 2002)] to identify a cortico-striatal system of abnormal brain activation in patients with OCD performing the Stroop task during PET scanning. This study also demonstrated relatively greater power of this multivariate method.
to detect brain functional abnormalities when compared to
the results of a more traditional mass or multiple univariate
approach to analysis, entailing a significance test at each
individual voxel (Harrison et al., 2006). However, no prior
MRI studies have provided evidence for heritability or
familiality of brain structural abnormalities in OCD, as is
required to support the candidacy of neurocognitive
systems as endophenotypes of OCD.

In this context, we were motivated to address four key
hypotheses: (i) that motor response inhibition is abnormal
in patients with OCD and their first-degree relatives;
(ii) that variation in motor inhibition is associated with
structural variation in large-scale brain systems identified by
a multivoxel analysis of MRI data; (iii) that patients
with OCD and their first-degree relatives have abnormal
grey matter density in these motor inhibition systems,
likely including orbitofronto-striatal regions previously
implicated in OCD and (iv) that there are familial effects
on variation in inhibitory function and associated brain
systems.

Fig. 1 Motor inhibitory behaviour (stop-signal reaction time; SSRT) and associated brain scores (summary measures of grey matter
correlation with SSRT over the whole brain). (A) Scatter plot showing the relationship between brain score (x-axis) and log-transformed
SSRT (y-axis); Pearson’s $r = 0.82$, $N = 93$, $P < 0.001$. (B) Boxplots of brain score by group showing significantly larger scores for both patients
and relatives compared to controls. For each boxplot, thick bar indicates median; box and whiskers represent interquartile range and
range, respectively.

Fig. 2 Brain maps illustrating regions where grey matter density was most strongly correlated with latency of motor inhibitory response
(SSRT). Red/yellow regions indicate areas in which increased grey matter density is associated with prolonged SSRT (impaired response
inhibition); blue regions indicate areas where decreased grey matter density is associated with prolonged SSRT. Colour bar indicates
strength of correlation between SSRT and grey matter density for each voxel; R and L markers indicate side of the brain, numbers denote
the z dimension of each slice in MNI space.
To test these predictions, we estimated the stop-signal reaction time (SSRT) using a well-validated stop-signal task (Logan et al., 1984), and acquired structural MRI data, in patients with OCD, their unaffected first-degree relatives, and healthy volunteers. The available evidence suggests structural abnormalities at a systems level in OCD, and prior functional imaging studies of motor inhibition indicate this processes is subserved by an ‘inhibitory neurocognitive network’ (Rubia et al., 2001b). Therefore, we used the statistical method of partial least squares (PLS) (McIntosh et al., 1996; McIntosh and Lobaugh, 2004), in an innovative application to structural neuroimaging data, to identify grey matter systems (comprising multiple voxels) maximally correlated with variation in SSRT. Finally, we assessed the familiality of cognitive and MRI markers by a permutation test of their variation within proband-relative pairs, each pair comprising a patient with OCD and their clinically unaffected first-degree relative.

Materials and Methods
Participants and clinical assessments
The sample comprised 31 patients with a diagnosis of OCD, 31 unaffected first-degree relatives of a patient with OCD, and 31 unrelated healthy volunteers (including 30 complete pairs of a proband and their first-degree relative). Patients were recruited from an outpatient service by a consultant psychiatrist (NF) and satisfied DSM-IV criteria (American Psychiatric Association, 1994) for a diagnosis of OCD. Patients with symptoms of excessive washing or checking, in the absence of hoarding or motor tics, were selected by clinical interview, using a well-validated screening instrument; the YBOCS Symptom Checklist (Goodman et al., 1989). This careful clinical screening process was adopted to minimize co-morbidity and maximize symptomatic homogeneity in the patient sample selected for subsequent assessment by cognitive testing and MRI, since there is evidence that OCD subgroups with different symptom profiles may have different underlying profiles of brain abnormality (Mataix-Cols et al., 2004).

Eligible patients gave consent for a first-degree relative to be contacted (preferably a similarly aged sibling; alternatively a parent or child). Unrelated healthy volunteers were recruited by local community advertisements. Participants were excluded if they had an axis I psychiatric disorder (apart from OCD in the patients), serious head injury, substance abuse, epilepsy or MRI contraindications. The Mini International Neuropsychiatric Inventory (MINI) (Sheehan et al., 1998) was used to screen for axis I psychiatric disorders; the Montgomery–Asberg depression rating scale (MADRS) (Montgomery and Asberg, 1979) to measure current depressive symptoms; and two instruments, the clinician-rated yale-brown obsessive compulsive scale (YBOCS) (Goodman et al., 1989) and the self-rated obsessive compulsive inventory-revised (OCI-R) (Foa et al., 2002), were used to measure obsessive-compulsive symptom severity. Verbal IQ was estimated using the National Adult Reading Test (NART) (Nelson, 1982).

Patients were clinically medicated as follows: 21 were prescribed selective serotonin reuptake inhibitors; 1 was prescribed clomipramine and 1 was prescribed quetiapine. Eight patients were unmedicated. Relatives and healthy volunteers were not taking psychotropic medication.

All participants gave written informed consent and the study was approved by the Addenbrooke’s NHS Trust Local Research Ethics Committee (Cambridge, UK). Behavioural performance on this motor inhibition task was previously reported for a subset of the individuals in this sample (Chamberlain et al., 2007). Here we report for the first time both behavioural and MRI data on the full sample.

Behavioural testing
We used a computerized version of the stop-signal task to assess inhibition of prepotent motor responses (Logan et al., 1984); for a full description see Aron et al. (2003a). Briefly, participants watched a computer screen on which a series of five blocks of 64 arrows per block were visually presented. Arrows pointed either to the right (50%) or the left (50%) and subjects responded accordingly by pressing the appropriate button; the order of presentation of right- and left-pointing arrows was randomized. In a randomly assigned proportion (25%) of trials, an audible stop-signal was heard after presentation of the arrow and subjects were instructed to inhibit their motor response to these trials. The interstimulus interval (ISI) and the stop-signal delay were varied according to the subject’s performance such that subjects were able successfully to inhibit their responses to 50% of the stop trials. From these behavioural data, the stop-signal reaction time (SSRT, ms), i.e. the processing time required to inhibit a prepotent motor response, was calculated for each subject. SSRT data were normalized by log transformation before statistical analysis in SPSS v11 for Windows.

MRI data acquisition
Structural MRI data were obtained using a GE Signa system (General Electric, Milwaukee, USA) operating at 1.5T in the Magnetic Resonance Imaging and Spectroscopy Unit, Addenbrooke’s Hospital, Cambridge, UK. Axial 3D T1-weighted images were acquired using a spoiled gradient recall (SPGR) sequence and the following parameters: number of slices = 124, slice thickness = 2 mm, TR = 33 ms, TE = 3 ms, field of view = 24 cm, flip angle = 40°, matrix size = 256 × 256, voxel dimensions = 0.94 mm × 0.94 mm; scanning time = 20 min. In addition, axial dual-echo, fast spin echo (T2- and PD-weighted) images were acquired with number of slices = 40, slice thickness = 4 mm, TR = 5625 ms, TE = 20 and 102 ms with an 8-echo train length, field of view = 24 cm, matrix size = 256 × 256, voxel dimensions = 0.94 mm × 0.94 mm; scanning time = 10 min. Total scanning time (including a localizer scan and a diffusion tensor imaging sequence not reported here) amounted to 40 min.

MRI data analysis: pre-processing
First, non-brain tissue was removed using an automated brain extraction procedure (Smith, 2002). The T1-, T2- and PD-weighted images were then segmented using a multichannel tissue classification algorithm and probabilistic maps of grey matter, white matter, CSF and dural tissues were created by estimating the partial volume coefficient for each voxel, which represents the probability of each voxel belonging to one of four tissue classes (Zhang et al., 2001). The segmented grey matter partial-volume maps were registered in Montreal Neurological...
Institute (MNI) standard space by an affine transformation using a segmented grey matter template in FSL (Sheehan et al., 1998; Jenkinson and Smith, 2001; Jenkinson et al., 2002). The registered data were spatially smoothed by a 2D Gaussian kernel with full width at half maximum (FWHM) = 3 mm. All preprocessing was performed using FSL software (http://www.fmrib.ox.ac.uk/fsl).

MRI data analysis: partial least squares

To identify grey matter systems optimally correlated with SSRT, we used the statistical technique of partial least squares; PLS (McIntosh et al., 1996). For implementation, we used PLSGUI software (http://www.rotman-baycrest.on.ca/pls/) running in MatLab.

Briefly, over all participants, we estimated the correlations between log-transformed SSRT scores and normalized grey matter density at each voxel in the registered images where the probability of grey matter \( P(GM) \) was greater than 0.1 (thus excluding from consideration all voxels representing predominantly white matter or CSF). The normalization of grey matter density involved dividing each voxel’s density estimate by the mean density of grey matter over all voxels in the brain (thus correcting individual voxel values for between-subject differences in global grey matter volume). The overall strength of correlation between grey matter density and SSRT was summarized by the scalar \( B(\cdot) = \sum_i r_i^2 / \sum_i \) where \( r_i \) is the correlation at the \( i \)th voxel and the sum is over all \( i = 1, 2, 3, \ldots, V \) voxels with \( P(GM) > 0.1 \). The brain score for each participant was calculated as the sum of grey matter probabilities multiplied by the local weighted correlations with SSRT: i.e. \( B(j) = \sum_i P(GM)_i r_i / d \), where \( B(j) \) is the brain score for the \( j \)th participant, \( P(GM)_i \) is the probability of grey matter at the \( i \)th voxel for the \( j \)th participant, and the sum is over all \( V \) voxels for each participant.

The association between grey matter probability and SSRT was tested for statistical significance by a permutation test of \( d \). The ordering of SSRT scores was randomly permuted before recalculation of the correlations with grey matter at each voxel, leading to an estimate of \( d \) under the null hypothesis. This process was repeated 1000 times to sample the permutation distribution of \( d \) and the observed value was compared to the 950th value of the ranked permutation distribution for a test with one-tailed probability of type 1 error, \( P = 0.05 \). Brain scores were also compared between groups by analysis of variance (ANOVA) and post hoc testing (SPSS v11 for Windows).

The anatomical configuration of brain systems strongly associated with SSRT was visualized by thresholding the correlations at each voxel with an arbitrary threshold, \( |r_i| > 0.14 \) and a minimum cluster size of 400 voxels. This cluster size threshold was chosen for illustrative purposes, to best demonstrate the large-scale anatomical covariation with SSRT scores. The choice of visualization thresholds makes no difference to the statistical significance of \( d \) (the overall strength of correlation between grey matter density and SSRT) or the calculation of brain scores for each participant.

We also used this thresholded set of correlated voxels as a ‘mask’, applied to the preprocessed maps of grey matter probability, to estimate the grey matter density represented by the thresholded system and its component regions in each participant. These measures of regional grey matter density were also compared between groups by ANOVA and post hoc testing.

Assessment of familiality

We used two complementary techniques to assess the familiality of SSRT scores, or related grey matter systems (brain score and grey matter density), in the OCD patients and their first-degree relatives. First, we calculated the variance of the within-pair difference in SSRT (or brain scores or grey matter density): \( \sigma_{\text{proband–relative pair}} = \sum (u_i - \bar{u})^2 / N \) where \( u_i \) is the observed within-pair difference of the measure for the \( i \)th pair of participants, \( \bar{u} \) is the mean within-pair difference and \( N \) is the total number of pairs (\( N = 30 \)). Then we randomly reassigned the observations to new pairs, so that each patient was now paired with a clinically unaffected individual to whom they were not personally related. We recalculated the variance of the within-pair difference in phenotype after each random permutation and repeated this process 100 000 times to sample the permutation distribution of \( \sigma_{\text{proband–relative pair}} \) under the null hypothesis that the observed variance in within-pair differences was not determined by the familial relatedness of the observed pairs. On the alternative hypothesis that the observed variance would be small, we compared it to the 5000th value of the permutation distribution for a test with one-tailed \( P < 0.05 \).

Second, as another exploration of the similarity between patients and their relatives on behavioural and brain-based measures, we examined the strength and significance of the within-pair correlation of SSRT, brain score and grey matter densities between patients and their own relatives. We used within-pair correlation rather than intraclass correlation because there was a natural ordering (patient or relative) within each pair.

Results

Demographic and clinical data

The three groups were well matched for age, verbal IQ, gender and handedness (Table 1). As expected, there were significant differences between groups on both measures of obsessive-compulsive symptom severity (YBOCS: ANOVA, \( F_{2,90} = 337, P < 0.001 \); OCIR: ANOVA, \( F_{2,86} = 58.0, P < 0.001 \)). Post hoc least significant difference (LSD) tests confirmed that patients scored significantly higher on both instruments than either healthy volunteers (YBOCS: \( df = 60, P < 0.001 \); OCIR: \( df = 57, P < 0.001 \)) or relatives (YBOCS: \( df = 60, P < 0.001 \); OCIR: \( df = 56, P < 0.001 \)) whereas relatives did not differ significantly from healthy volunteers on either instrument (YBOCS: \( df = 60, P = 0.08 \); OCIR: \( df = 59, P = 0.54 \)). Although mean depressive symptom severity scores measured using MADRS were below the threshold for a diagnosis of depressive disorder in all three groups, patients with OCD had higher scores than relatives and healthy volunteers (ANOVA, \( F_{2,90} = 14.2, P < 0.001 \)). LSD tests, patients compared with healthy volunteers: \( df = 60, P < 0.001 \); patients compared with relatives: \( df = 60, P < 0.001 \). Relatives did not differ from healthy volunteers (\( df = 60, P = 0.36 \) (Table 1).

Stop-signal task performance

There was a significant difference between groups in mean SSRT (ANOVA, \( F_{2,90} = 9.07, P < 0.001 \) (Table 1). Post hoc analysis demonstrated that both patients (\( df = 60, P = 0.001 \)
Table 1: Demographic, clinical and behavioural data for patients with OCD, their first-degree relatives and healthy, unrelated volunteers

<table>
<thead>
<tr>
<th>Variable</th>
<th>OCD patients (N = 31)</th>
<th>First-degree relatives (N = 31)</th>
<th>Healthy unrelated volunteers (N = 31)</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Demographic data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>32.5</td>
<td>11.1</td>
<td>36.7</td>
<td>13.4</td>
</tr>
<tr>
<td>Handedness (right:left)</td>
<td>27:3</td>
<td></td>
<td>26:4</td>
<td></td>
</tr>
<tr>
<td>Gender (male:female)</td>
<td>9:22</td>
<td>9:22</td>
<td>11:20</td>
<td></td>
</tr>
<tr>
<td>NART</td>
<td>113.3</td>
<td>70</td>
<td>114.4</td>
<td>8.7</td>
</tr>
<tr>
<td>Clinical data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YBOCS</td>
<td>21.9</td>
<td>5.5</td>
<td>1.7</td>
<td>3.3</td>
</tr>
<tr>
<td>OCI-R</td>
<td>36.5</td>
<td>18.3</td>
<td>71</td>
<td>10.8</td>
</tr>
<tr>
<td>MADRS</td>
<td>5.6</td>
<td>6.3</td>
<td>1.5</td>
<td>2.6</td>
</tr>
<tr>
<td>Age of onset of symptoms (years)</td>
<td>15.6</td>
<td>8.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of illness (years)</td>
<td>16.4</td>
<td>11.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSRT (log transformed)</td>
<td>5.36</td>
<td>0.34</td>
<td>5.42</td>
<td>0.45</td>
</tr>
<tr>
<td>Median go reaction time (ms)</td>
<td>457</td>
<td>103</td>
<td>406</td>
<td>124</td>
</tr>
</tbody>
</table>

Abbreviations: OCD; obsessive-compulsive disorder, SD; standard deviation.

a df = 2, 89 for NART and df = 2, 86 for OCI-R due to data unavailability.
b Fisher’s exact test, total N = 91 due to data unavailability.
c $\chi^2$ test, df = 2.

and relatives (df = 60, $P < 0.001$) had significantly greater SSRT than healthy volunteers. There was no significant difference in SSRT between patients and relatives (df = 60, $P = 0.50$). Of note, these between-group differences in response inhibition were not accompanied by significant non-specific latency differences in responding to uninhibited trials (median ‘go’ reaction time: ANOVA, $F_{2,90} = 2.76$, $P = 0.07$). Additionally, there was no significant correlation between age and SSRT in relatives ($r = 0.18$, $N = 31$, $P = 0.33$), excluding the possibility that younger relatives perform worse on the task and might therefore represent individuals with an OCD prodrome.

**Grey matter systems correlated with SSRT**

There was a significant correlation between SSRT and grey matter probability ($d = 44.9$, permutation test, $P = 0.05$) and individual brain scores were strongly positively correlated with SSRT ($r = 0.82$, $N = 93$, $P < 0.001$) (Fig. 1A). As expected, brain scores were significantly different between groups (ANOVA, $F_{2,90} = 4.18$, $P = 0.018$) (Fig. 1B) and post hoc analysis demonstrated that this was due to significantly greater brain scores in patients compared to healthy volunteers (df = 60, $P = 0.021$) and in relatives compared to healthy volunteers (df = 60, $P = 0.010$); there was no significant difference in brain scores between patients and relatives (df = 60, $P = 0.78$).

An anatomical map of voxels strongly correlated with SSRT over all participants highlighted two extensive systems which were, respectively, positively and negatively correlated with latency of inhibitory processing (Fig. 2). In the positively correlated system, longer SSRT was associated with increased grey matter probability. This predominantly parieto-cingulo-striatal system comprised middle and posterior cingulate cortices (approximate Brodmann areas [BA] 23, 24, 31), bilateral putamen/caudate and amygdala, bilateral parietal cortical areas (BA 39, 40) and bilateral cerebellum. In the negatively correlated system, longer SSRT was associated with decreased grey matter probability. This predominantly frontal system comprised bilateral middle and medial orbitofrontal cortex (BA 11, 47), inferior frontal gyri (BA 44, 45), superior frontal and premotor cortices (BA 6, 8, 9), anterior cingulate cortex (BA 32) and bilateral temporal cortical areas (BA 21, 22, 37, 42) (Table 2).

To explore these results further, we extracted grey matter values for the systems in Fig. 2 that were correlated with SSRT. We confirmed that grey matter density in the parieto-cingulo-striatal system was positively correlated with SSRT ($r = 0.70$, $N = 93$, $P < 0.001$), grey matter density in the frontal system was negatively correlated with SSRT ($r = -0.78$, $N = 93$, $P < 0.001$), and grey matter density in the two systems was negatively correlated with each other ($r = -0.78$, $N = 93$, $P < 0.001$), indicating that individuals with prolonged SSRT and larger positive brain scores (typically patients and relatives) tended to have both increased grey matter in the parieto-cingulo-striatal system and reduced grey matter in the frontal system.

There were significant between-group differences in grey matter probability measured in both the parieto-cingulo-striatal system (ANOVA, $F_{2,90} = 4.27$, $P = 0.017$) and frontal system (ANOVA, $F_{2,90} = 3.36$, $P = 0.039$). When we explored
group differences separately for each anatomically distinct region in each system, we found that left inferior parietal and dorsal occipital regions demonstrated the greatest between-group difference in grey matter probability among all regions positively correlated with SRRT (ANOVA, $F_{2,90} = 7.48$, $P = 0.001$); whereas bilateral orbito-frontal cortex demonstrated the greatest between-group difference among all regions negatively correlated with SSRT (ANOVA, $F_{2,90} = 4.92$, $P = 0.009$) (Table 2, Fig. 3). There were no regions in which there were significant differences in grey matter between patients and relatives (LSD; $df = 60$, $P > 0.1$).

### Familiarity of cognitive and neuroimaging phenotypes

By a permutation test of the variance of within-pair differences in brain scores, we were able to show that the observed within-pair variance was small ($\sigma$$_{proband–relative pair} = 127$) compared to the distribution of variances in a

---

**Table 2** Anatomical details for brain regions where grey matter density was positively or negatively correlated with stop–signal reaction time (SSRT)

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>Size (voxels)</th>
<th>Peak correlation ($r$)</th>
<th>MNI coordinates (mm)</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$X$</td>
<td>$Y$</td>
</tr>
<tr>
<td><strong>Red regions (positive correlation between grey matter density per voxel and SSRT)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4172</td>
<td>0.408</td>
<td>0</td>
<td>-20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>-50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-2</td>
<td>-56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>-58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-8</td>
<td>-72</td>
</tr>
<tr>
<td>2</td>
<td>1720</td>
<td>0.373</td>
<td>-26</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-12</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-28</td>
<td>-6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-30</td>
<td>-12</td>
</tr>
<tr>
<td>3</td>
<td>1016</td>
<td>0.360</td>
<td>-22</td>
<td>-92</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-30</td>
<td>-70</td>
</tr>
<tr>
<td>4</td>
<td>1176</td>
<td>0.344</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>26</td>
<td>-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>44</td>
<td>-4</td>
</tr>
<tr>
<td>5</td>
<td>486</td>
<td>0.344</td>
<td>28</td>
<td>-84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>42</td>
<td>-78</td>
</tr>
<tr>
<td>6</td>
<td>3882</td>
<td>0.340</td>
<td>-34</td>
<td>-56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td>-82</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-4</td>
<td>-76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>-52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-14</td>
<td>-60</td>
</tr>
<tr>
<td><strong>Blue regions (negative correlation between grey matter density per voxel and SSRT)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1398</td>
<td>-0.407</td>
<td>40</td>
<td>-60</td>
</tr>
<tr>
<td>8</td>
<td>1665</td>
<td>-0.371</td>
<td>58</td>
<td>-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>60</td>
<td>-18</td>
</tr>
<tr>
<td>9</td>
<td>2410</td>
<td>-0.357</td>
<td>-26</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-52</td>
<td>34</td>
</tr>
<tr>
<td>10</td>
<td>1331</td>
<td>-0.354</td>
<td>52</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>58</td>
<td>-14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>46</td>
<td>-10</td>
</tr>
<tr>
<td>11</td>
<td>1188</td>
<td>-0.328</td>
<td>-12</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-6</td>
<td>22</td>
</tr>
<tr>
<td>12</td>
<td>1900</td>
<td>-0.311</td>
<td>-40</td>
<td>22</td>
</tr>
<tr>
<td>13</td>
<td>1715</td>
<td>-0.309</td>
<td>-52</td>
<td>-26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-38</td>
<td>-22</td>
</tr>
<tr>
<td>14</td>
<td>873</td>
<td>-0.292</td>
<td>-46</td>
<td>-52</td>
</tr>
</tbody>
</table>

Abbreviations: BA = Brodmann’s area; L = left; R = right.
sample of 100,000 randomly permuted pairs of patients and first-degree relatives (permutation test, \( P = 0.014 \)) (Fig. 4A). This observation is compatible with the (alternative) hypothesis that variance of true proband-relative pair differences is smaller because the phenotype in question, e.g. brain scores, is determined by familial factors in common to both the patient and the first-degree relative in each pair.

We applied the same permutation test to analyses of the variance of within-pair differences in grey matter density of the parieto-cingulo-striatal system positively correlated with SSRT and in grey matter density of the frontal system negatively correlated with SSRT. In both cases, we found that the observed within-pair variance was significantly small compared to the appropriate permutation distributions (permutation tests, \( P = 0.009 \) and \( P = 0.015 \), respectively), again implying that familial factors shared between patients and relatives determined grey matter volume in these systems. Interestingly, applying the same approach to analysis of the SSRT scores, provided no evidence for significantly reduced variance of within-pair differences on this cognitive measure (permutation test, \( P = 0.25 \)).

A broadly consistent pattern of results was obtained by analysis of correlation between patients and their relatives: these were significantly different from zero for brain score (\( r = 0.41, N = 30, P = 0.023 \)), mean grey matter probability in the parieto-cingulo-striatal system (\( r = 0.46, N = 30, P = 0.012 \)) and mean grey matter probability in the frontal system (\( r = 0.42, N = 30, P = 0.022 \)). In contrast, there was no significant within-pair correlation for the SSRT score.

Fig. 3 Brain maps highlighting regions of most significant group difference in grey matter density between OCD patients and first-degree relatives compared to healthy volunteers. (A) Most significant red regions were in left occipital and inferior parietal areas (BA 19, 40) (Table 2; cluster 3); one way ANOVA (\( F_{2,90} = 7.48, P = 0.001 \), post hoc tests; patients versus healthy volunteers; \( P = 0.002 \), relatives versus healthy volunteers; \( P = 0.001 \), patients versus relatives; \( P = 0.85 \)). (B) Most significant blue regions were in bilateral orbitofrontal and left inferior frontal gyral regions (Table 2; cluster 9) (BA II, 44, 45, 47); one way ANOVA (\( F_{2,90} = 4.92, P = 0.009 \), post hoc tests; patients versus healthy volunteers; \( P = 0.013 \), relatives versus healthy volunteers; \( P = 0.005 \), patients versus relatives; \( P = 0.76 \)). R marker indicates right side of the brain; \( x \), \( y \), and \( z \) indicate planes of brain maps; cross-hairs indicate point of peak correlation with the behavioural measure (SSRT).

Fig. 4 Estimating familiality effects on MRI and behavioural variation in patients and their own first-degree relatives (\( N = 30 \) per group) (A) Histogram showing distribution of variance of within-pair difference in brain scores for randomly permuted pairs of patients and relatives, compared with the observed variance of within-pair difference in brain score for patients and their own relative (arrow). (B) Scatter plots exploring (within-pair) correlation between patients and their relatives for SSRT (top panel) and brain score (bottom panel).
(r = 0.16, N = 30, P = 0.41) (Fig. 4B). Taken together with the results on variance of within-pair differences in MRI and cognitive phenotypes, these data suggest that compared with a behavioural measure of response inhibition, the MRI systems correlated with response inhibitory processing are more strongly determined by familial factors shared between true proband–relative pairs.

**Discussion**

These data provide empirical support for each of the four hypotheses motivating this study. We have confirmed that response inhibition, indexed by SSRT, is abnormal in patients with OCD and their first-degree relatives. We have identified extensive brain systems where grey matter density is (positively or negatively) correlated with variability in stop-signal task performance; and we have shown that patients with OCD and their relatives have structural abnormalities in these systems compared to healthy volunteers. Finally, we have exploited our proband–relative pair design to assess the familiality of variation in cognitive and associated MRI phenotypes and shown that variation in brain systems correlated with inhibitory function is likely determined by familial factors in common between patients and their first-degree relative. In short, we have combined structural neuroimaging and cognitive testing to identify for the first time a neurocognitive endophenotype of OCD.

**Inhibition and OCD**

The classical clinical symptoms of OCD are persistent, obsessional thoughts attended by an inability to inhibit compulsive behaviour repetition. It therefore seems almost self-evident that inhibitory processes might be abnormal in OCD and there is empirical evidence in support of this hypothesis. Patients with OCD are impaired across a range of tests of inhibitory function including motor inhibition tasks, e.g. go/no-go and stop-signal tasks (Bannon et al., 2002; Chamberlain et al., 2005, 2006; Penades et al., 2006); attentional set shifting tasks, such as the object alternation task (Abbruzzese et al., 1997; Aycicegi et al., 2003); the intra-dimensional/extra-dimensional task (Veale et al., 1996; Watkins et al., 2005; Chamberlain et al., 2006) which requires inhibition of a previously successful response strategy in response to changing criteria for task performance; and the Stroop task, a putative test of cognitive inhibition (van den Heuvel et al., 2005b; Penades et al., 2006). Our data, indicating that patients are impaired on a motor inhibition task, are consistent with this literature.

**Brain systems associated with motor inhibition**

Prior data on human brain systems underlying motor inhibition have been provided by lesion studies and neuroimaging. In a structural MRI study of patients following focal but variably located brain injuries, Aron et al. (2003b) found that grey matter volume deficit in the right inferior frontal gyrus (RIFG) was specifically predictive of prolonged SSRT; for a review of further evidence implicating the RIFG in motor inhibition see Aron et al. (2004). Functional neuroimaging studies of motor inhibition have generally identified a more extensive but predominantly right-sided system of regions including orbitofrontal, dorsolateral and medial frontal, temporal and parietal cortices, the cerebellum and the basal ganglia (Godefroy et al., 1996; Humberstone et al., 1997; Garavan et al., 1999; Rubia et al., 1999, 2000, 2001a, b and c; Horn et al., 2003).

In keeping with the focus on the RIFG in previous literature, we also found evidence that reduced grey matter density in this region was associated with prolonged SSRT. However, consistent with the functional neuroimaging data suggesting involvement of a network of regions in inhibition, we found that brain areas negatively correlated with latency of inhibitory processing in our data were not restricted to the RIFG but included regions such as bilateral orbitofrontal cortex, right premotor and anterior cingulate cortex, left dorsolateral prefrontal cortex and bilateral temporal cortex. We also found a number of regions in cingulate cortex, parietal and dorsal occipital cortex, and basal ganglia where SSRT was positively correlated with grey matter density, i.e. impaired inhibitory function was predicted by increased grey matter density.

**Brain systems implicated in OCD**

‘Affective’ fronto-striatal circuits including the orbitofrontal cortex, the striatum and anterior cingulate have been invoked theoretically to account for OCD. As already discussed (Introduction, Supplementary Figs 1 and 2), there is some inconsistent evidence in support of this hypothesis from structural MRI studies published to date. However, there is additional evidence for orbitofrontal dysfunction in OCD from positron emission tomography (PET) studies reporting abnormal resting or task-related orbitofrontal metabolism in OCD (Baxter et al., 1987, 1988; Nordahl et al., 1989; Swedo et al., 1989; McGuire et al., 1994; Rauch et al., 1994). Functional MRI studies investigating executive function in OCD have also identified fronto-striatal abnormalities in patients (Maltby et al., 2005; van den Heuvel et al., 2005a; Remijnse et al., 2006; Rauch et al., 2007); and there is evidence that affective fronto-striatal circuits are abnormally activated in patients during symptom provocation (Breiter et al., 1996; Adler et al., 2000; Phillips et al., 2000; Mataix-Cols et al., 2004; Nakao et al., 2005; Schienle et al., 2005). Moreover, fMRI studies have often also shown changes in activation of theoretically unanticipated regions such as the dorsolateral prefrontal cortex (DLPFC) and parietal cortex. For example, van den Heuvel et al. (2005a) found decreased DLPFC activation in
OCD patients compared with controls, Maltby et al. (2005) found hyperactivity in the anterior and posterior cingulate and lateral prefrontal cortex during unsuccessful stopping in a go/no-go task, and Schienle et al. (2005) found increased activation in DLPFC and parietal areas. There is corroborative evidence for parietal cortical abnormalities in OCD from PET (Nordahl et al., 1989; McGuire et al., 1994; Rauch et al., 1994) and SPECT (Lucey et al., 1995) activation studies.

Thus our findings of extensive grey matter abnormality in orbitofrontal cortex, ventral and dorsal prefrontal cortex, cingulate cortex, parietal cortex, striatum and cerebellum include many of the regions anticipated by an orbitofronto-striatal model; but also include other regions (such as parietal cortex or cerebellum), which have been reported in the functional neuroimaging literature and by some of the prior structural MRI studies (Pujol et al., 2004; Valente et al., 2005; Soriano-Mas et al., 2007), yet are not so readily accommodated by an exclusively orbitofronto-striatal model. As the neuroimaging evidence base grows and becomes more replicable, we predict that this will drive development of systems-level theory beyond the model of abnormality in a single cortico-striatal circuit.

**Candidacy of cognitive and MRI endophenotypes of OCD**

There is no universally accepted set of criteria to judge the validity of a candidate endophenotype. However, Gottesman proposed that endophenotypes are quantitative heritable traits that are abnormal in both probands and their relatives (Gottesman and Gould, 2003). How well do our data on behavioural and MRI markers of inhibitory processing satisfy these criteria?

Both behavioural and MRI markers were quantitatively abnormal on average in both patients and relatives compared to healthy volunteers. However, the demonstration of strict sense heritability is impossible in the absence of a twin design controlling for shared environmental influences on trait variation in genetically related individuals. We have therefore adopted the logistically more feasible approach of assessing familial (rather than strictly heritable) effects on trait variation in a proband–relative pair design. We have used an innovative permutation test of the within-pair variance in trait differences, and within-pair correlations, to quantify familiality of variation in discordant proband–relative pairs, finding evidence for significant familial effects on variation in the MRI systems associated with inhibitory processing, but not on the behaviourally derived SSRT measure. We conclude that the MRI markers of inhibitory processing more completely satisfy Gottesman’s criteria (by virtue of their greater familiality), perhaps reflecting the fact that structural variation in brain systems is more proximal to genetic effects than variation in task performance. This result strengthens the rationale for searching for neurocognitive endophenotypes in other complex behavioural disorders such as schizophrenia and bipolar disorder.

**Utility and specificity of endophenotypes of OCD**

Using probands and first-degree relatives to identify MRI endophenotypes has the immediate advantage of discounting any non-familial explanations (such as exposure to psychotropic medication in the probands) for abnormal patterns of brain structure. Endophenotypes could also be used to refine diagnostic subclassification of patients (based on the extent of their expression of endophenotypic abnormality), or to highlight additional abnormalities occurring only in patients (not relatives), although these were not objectives of the current study. However, it is interesting also to consider how our results could be exploited in future to identify specific genes determining variation in brain systems important for motor inhibition. In principle, the grey matter density of the motor inhibitory system could be used as a quantitative trait in a genome-wide search for associated polymorphisms by quantitative trait locus (QTL) analysis. Although we currently lack sufficient experience of genome-wide QTL mapping based on human imaging, this has been successfully used to identify genetic markers associated with imaging measurements of cortical and subcortical grey matter volumes in inbred strains of mice (Beatty and Laughlin, 2006).

Another question concerns the diagnostic specificity of an inhibitory endophenotype for OCD. Since the present study focused on patients with predominantly classical washing/checking symptoms, these results may not generalize to other OCD subgroups. On the other hand, recent evidence has accumulated to suggest that a deficit in response inhibition may also be an endophenotype for attention-deficit/hyperactivity disorder (ADHD) (Aron and Poldrack, 2005; Crosbie and Schachar, 2001). Behavioural deficits in response inhibition have been identified in ADHD (Casey et al., 1997; Vaidya et al., 1998; Chamberlain and Sahakian, 2007) and previously related to abnormal activation of right inferior prefrontal cortex and left caudate during the stop-signal task (Rubia et al., 1999). Clinically unaffected first-degree relatives of ADHD probands have also shown deficits on motor response inhibition in the go/no-go task (Slaats-Willemse et al., 2003). Future studies of ADHD patients and their relatives would establish if impairments in response inhibition are underpinned by anatomical variation in the same brain systems that we have identified in OCD; and test that behavioural and/or imaging markers of impaired inhibitory function satisfy the Gottesman criteria as endophenotypes of ADHD. Without such data, it is speculative but intriguing to consider that the same neurocognitive endophenotype might be related to important dimensions of these two traditionally distinct clinical syndromes.
Methodological considerations

An innovative aspect of this study was the use of the PLS method, a technique previously employed mainly in the analysis of functional neuroimaging data (McIntosh and Lobaugh, 2004), to find structural brain systems optimally correlated with a behavioural variable. PLS was attractive for our purpose because there were prior theoretical reasons to expect that inhibitory deficits in OCD might be related to structural abnormalities at a systems level, rather than in a discrete brain region, and PLS is designed to optimize correlation between one or more exogenous (behavioural) variables and a set of correlated image voxels, without specifying a priori which voxels are likely to be components of the behaviourally correlated system. Characterization of structure–function relationships at systems level has the considerable merit of minimizing the number of significance tests required. To search for significant structure–function associations at voxel level would entail approximately 150,000 tests, with concomitantly severe thresholds for significance to mitigate the multiple comparisons problem; whereas testing for a systems level association in PLS required only one significance test, which could be thresholded conventionally at $P \leq 0.05$.

In principle, partial least squares can be used as a multivariate analysis method to explore the relationships between multiple behavioural variables and multiple imaging variables. Here we have used it to test for association between a single behavioural variable and grey matter density at multiple voxels. To distinguish this application from the more general multivariate case, where PLS is used to find associations between multiple behavioural variables and imaging measures at multiple voxels, we have referred to our application as a multivoxel analysis because the (non-parametric) test for significant association is based on behaviour correlations summed across all voxels in the brain. Thus PLS with a single behavioural variable is conceptually close to the analysis of OCD imaging data by Soriano-Mas et al. (2007), based on the method proposed by Worsley et al. (1995, 1997). The main differences are that Soriano-Mas et al. (2007) tested a multivoxel measure of between-group difference, whereas we have tested a multivoxel measure of continuous covariation with cognitive function across diagnostic groups. More technically, the PLS approach has the relative merit of an entirely non-parametric (resampling-based) approach to significance testing which confers greater flexibility in choice of test statistics and greater robustness against violation of the conditions required for validity of parametric tests (Worsley et al., 1995, 1997).

Conclusions

To the best of our knowledge, this is the first example of a potentially powerful experimental and data analytic strategy to identify cognitive and related brain structural endophenotypes of heritable but genetically complex neuropsychiatric disorders. In a sample of OCD patients, their first-degree relatives and unrelated healthy volunteers, we have found substantial evidence that variation in motor inhibitory control is correlated with grey matter density changes in an extensive system comprising orbitofrontal, cingulate and parietal cortical areas as well as striatal and other subcortical regions. We have also tested rigorously by Gottesman’s criteria the candidacy of these inhibition-related brain systems as the first neurocognitive endophenotype for obsessive-compulsive disorder.

Supplementary material

Supplementary material is available at Brain online.

Acknowledgements

This work was funded by the National Alliance for Research on Schizophrenia and Depression (Distinguished Investigator Award to E.B.), the Wellcome Trust (T.W.R., B.J.S.), the Harnett Fund (University of Cambridge; MB/PhD studentship to L.M.), the Medical Research Council (MB/PhD studentship to S.R.C.) and the National Institutes of Mental Health and Biomedical Imaging & Bioengineering (Human Brain Project; E.B.). The Behavioural and Clinical Neuroscience Institute is supported by a joint award from the Medical Research Council and Wellcome Trust. Mr Chamberlain and Dr Robbins and Dr Sahakian consult for Cambridge Cognition. All other authors report no competing interests. We thank Drs Randy McIntosh and Nancy Lobaugh for expert support in using their PLSGUI software. Funding to pay the Open Access publication charges for this article was provided by the Wellcome Trust.

References


Neurocognitive endophenotypes of OCD


