Social cognitive deficits and their neural correlates in progressive supranuclear palsy

Boyd C. P. Ghosh,1,2 Andrew J. Calder,3 Polly V. Peers,3 Andrew D. Lawrence,4 Julio Acosta-Cabronero,2 João M. Pereira,2 John R. Hodges5 and James B. Rowe2,3,6

1 Wessex Neuroscience Centre, Southampton SO16 6YD, UK
2 Department of Clinical Neurosciences, Cambridge University, Cambridge CB2 0QQ, UK
3 MRC Cognition and Brain Sciences Unit, Cambridge CB2 7EY, UK
4 School of Psychology, Cardiff University, Cardiff CF10 3AT, UK
5 Neuroscience Research Australia, Sydney, New South Wales 2031, Australia
6 MRC Behavioural and Clinical Neuroscience Institute, Cambridge CB2 2EB, UK

Correspondence to: Dr B Ghosh, Wessex Neuroscience Centre, Mailpoint 101, Southampton University Hospitals NHS Trust, Tremona Road, Southampton SO16 6YD, UK
E-mail: boydghosh@doctors.org.uk

Although progressive supranuclear palsy is defined by its akinetic rigidity, vertical supranuclear gaze palsy and falls, cognitive impairments are an important determinant of patients’ and carers’ quality of life. Here, we investigate whether there is a broad deficit of modality-independent social cognition in progressive supranuclear palsy and explore the neural correlates for these. We recruited 23 patients with progressive supranuclear palsy (using clinical diagnostic criteria, nine with subsequent pathological confirmation) and 22 age- and education-matched controls. Participants performed an auditory (voice) emotion recognition test, and a visual and auditory theory of mind test. Twenty-two patients and 20 controls underwent structural magnetic resonance imaging to analyse neural correlates of social cognition deficits using voxel-based morphometry. Patients were impaired on the voice emotion recognition and theory of mind tests but not auditory and visual control conditions. Grey matter atrophy in patients correlated with both voice emotion recognition and theory of mind deficits in the right inferior frontal gyrus, a region associated with prosodic auditory emotion recognition. Theory of mind deficits also correlated with atrophy of the anterior rostral medial frontal cortex, a region associated with theory of mind in health. We conclude that patients with progressive supranuclear palsy have a multimodal deficit in social cognition. This deficit is due, in part, to progressive atrophy in a network of frontal cortical regions linked to the integration of socially relevant stimuli and interpretation of their social meaning. This impairment of social cognition is important to consider for those managing and caring for patients with progressive supranuclear palsy.

Keywords: progressive supranuclear palsy; voxel-based morphometry; social cognition; theory of mind; emotion perception

Abbreviations: PSPRS = progressive supranuclear palsy rating scale; TASIT = the awareness of social inference test; UPDRS = Unified Parkinson Disease Rating Scale
Introduction

When Richardson et al. (1963) first described progressive supranuclear palsy, they described a set of characteristic features that remain the principal diagnostic criteria: a vertical supranuclear gaze palsy and postural instability with falls (Litvan et al., 2003). Generally, patients also have akinetic rigidity and progressive dysarthria and dysphagia. However, significant cognitive problems are common, including apathy and a dysexecutive syndrome (Grafman et al., 1995; Bak and Hodges, 1998; Millar et al., 2006; Brown et al., 2010) and psychiatric comorbidity including depression (Schrag et al., 2003; Herting et al., 2007; Bak et al., 2010) or anxiety (Litvan et al., 1996b; Aarsland et al., 2001; Borroni et al., 2008). Half of patients also report social impairments as a negative influence on their quality of life (Schrag et al., 2003).

Recently, we reported deficits in basic emotion recognition in progressive supranuclear palsy (Ghosh et al., 2009). We proposed that these deficits were part of a wider deficit in social cognition that may contribute to the common difficulties patients experienced in social interactions (Schrødt et al., 2003). However, social cognition is not a unitary phenomenon. At one level, it requires the ability to recognize basic emotions (e.g. happy, sad, angry, disgust and surprise) ‘emotion knowledge’ (Izard, 1971). It also encompasses an appreciation of more complicated emotions (e.g. sarcasm and humour) and higher order social inferences. These require greater understanding and representation of other people’s mental states in a given context, known as ‘Theory of Mind’ (Premack and Woodruff, 1978).

Several neurological disorders affect social cognition, including emotion recognition and theory of mind. In addition to progressive supranuclear palsy (Ghosh et al., 2009), these include behavioural variant frontotemporal degeneration (Keane et al., 2002; Rankin et al., 2005), Parkinson’s disease (Sprengelmeyer et al., 2003; Lawrence et al., 2007; Gray and Tickle-Degnen, 2010) and Huntington’s disease (Sprengelmeyer et al., 1996; Calder et al., 2010). Basic emotion recognition deficits in these degenerative disorders have been investigated using static pictures of faces showing basic emotions (Keane et al., 2002; Rosen et al., 2002b, 2004; Diehl-Schmid et al., 2007; Kipps et al., 2009a; Rankin et al., 2009); sounds expressing emotion (Keane et al., 2002; Rankin et al., 2009); and videos of emotions (Werner et al., 2007; Kipps et al., 2009b; Rankin et al., 2009). Theory of mind abilities in these clinical populations have been assessed using both static and dynamic tests such as the Awareness of Social Inference Test (TASIT) (Kipps et al., 2009b; Rankin et al., 2009); the Mind in the Eyes Test (Baron-Cohen et al., 2001; Gregory et al., 2002; Stone et al., 2003; Torralva et al., 2007; Hirao et al., 2008); cartoons of social situations (Snowden et al., 2003; Lough et al., 2006; Rankin et al., 2009); and faux pas stories (Stone et al., 1998, 2003; Gregory et al., 2002; Torralva et al., 2007; Roca et al., 2010).

The critical neural structures supporting these social cognitive functions have been studied with structure–function correlations such as voxel-based morphometry of MRI (Rosen et al., 2002b; Henley et al., 2008; Kipps et al., 2009b; Rankin et al., 2009; lesion studies (Adolphs et al., 1996, 2000, 2002; Channon and Crawford, 2000; Roca et al., 2010; Shamay-Tsoory et al., 2009) and functional MRI (Hennenlotter et al., 2004; Brunet-Gouet and Decety, 2006; Carrington and Bailey, 2009). These complimentary methods reveal a partial convergence onto a frontotemporal network for social cognition.

A functional anatomical distinction appears to exist between unimodal and supramodal representations of emotion. For example, discrete emotion associations have been found between the insula and disgust (Phillips et al., 1997; Calder et al., 2000; Kipps et al., 2007) and between the amygdala and fear (Adolphs et al., 1994; Calder et al., 1996, 2001). Other simple associations are less consistent, for example, between anger and the medial temporal gyrus (Sprengelmeyer et al., 1998) or happiness and the middle temporal gyrus and amygdala (Johnstone et al., 2006; Kipps et al., 2007).

For the supra- or cross-modal recognition of emotion stimuli, several critical regions have been proposed, including the inferior frontal gyrus (Sprengelmeyer et al., 1998; Nakamura et al., 1999; Schirmer and Kotz, 2006; Beaucousin et al., 2007; Philippi et al., 2009; Leitman et al., 2010), the orbitofrontal cortex (Keane et al., 2002; Hornak et al., 2003; Paulmann et al., 2010) and the medial frontal cortex (Peelen et al., 2010). Other regions have been associated with the coding of multiple emotions but within a single sensory domain, such as the auditory domain (Murphy et al., 2003; Wildgruber et al., 2005; Ethofer et al., 2006; Schirmer and Kotz, 2006).

For higher order aspects of social cognition, including empathy and theory of mind, a frontotemporal cortical network has been implicated that includes regions affected by progressive supranuclear palsy pathology. For example, the inferior frontal gyrus together with components of the emotional network have been linked to emotional empathy—the ability to ‘feel’ the emotion of others. This may be mediated by its putative role in the mirror neuron system (Schulte-Ruther et al., 2007; Bastiaansen et al., 2009; Shamay-Tsoory et al., 2009). Such ‘emotional empathy’ can be contrasted with ‘cognitive empathy’, the determination of what others think, feel or intend without necessarily sharing the ‘feeling’.

On the basis of this prior knowledge of the functional anatomy of social cognition, one can use a region of interest approach to increase sensitivity and localization of pathological changes correlated with empathy. It can be difficult to isolate empathic aspects of the task and control for the mode of testing (for example, videos versus cartoons). However, there is a consensus that cognitive empathy is associated with the medial prefrontal cortex, orbitofrontal cortex, posterior superior temporal sulcus, the temporoparietal junction and striatum (Gregory et al., 2002; Frith and Frith, 2003; Saxe and Kanwisher, 2003; Snowden et al., 2003; Amodio and Frith, 2006; Gilbert et al., 2006; Hirao et al., 2008; Carrington and Bailey, 2009; Kipps et al., 2009b; Rankin et al., 2009; Shamay-Tsoory et al., 2009). Of these, the correlation with the medial prefrontal cortex is most consistent (Carrington and Bailey, 2009). Therefore, it is relevant that with progressive supranuclear palsy, the foci of cerebral atrophy and highest pathological burden overlap with these regions associated with social cognition (Brenneis et al., 2004; Price et al., 2004; Padovani et al., 2006;
Paviour et al., 2006b; Nilsson et al., 2007; Josephs et al., 2008; Nicoletti et al., 2008; Rizzo et al., 2008; Schofield et al., 2011), especially in the frontal areas.

Based on these associations between structure, function and pathology, we formed three specific hypotheses. These were tested in a cohort of patients with progressive supranuclear palsy using voxel-based morphometry of MRI in conjunction with neuropsychological assessment of social cognition. Our hypotheses were that: (i) progressive supranuclear palsy causes a multimodal deficit in emotion recognition, such that patients have deficits in basic emotion recognition in auditory and visual modalities. Demonstration of deficits in an auditory test would additionally overcome a potential confound in previous visual emotion studies, arising from disordered eye movements; (ii) progressive supranuclear palsy impairs higher order social cognition, including theory of mind and (iii) greater social cognition dysfunction in progressive supranuclear palsy correlates with more severe regional atrophy of the inferior frontal gyrus and medial frontal cortex.

Materials and methods

Participants

Twenty-three patients were recruited prospectively from a specialist neurological clinic for patients with progressive supranuclear palsy and related disorders between 2007 and 2009. Clinical diagnostic criteria (Litvan et al., 2003) were used by an experienced neurologist. To date, nine patients have subsequently undergone post-mortem examination; all nine had progressive supranuclear palsy. Patients were excluded if they had another significant neurological illness. Twenty-two age- and education-matched controls were recruited from the panel of volunteers at the Medical Research Council’s Cognition and Brain Sciences Unit or from spouses of patients. Control participants had normal hearing and corrected vision and did not have significant neurological or psychiatric comorbidity. Not all participants were able to complete all tests. Three patients had either poor visual acuity or intercurrent illness and therefore could not complete all the neuropsychological test sessions. Two controls and one patient were unable to complete the volumetric MRI scan.

Patients and controls underwent the same testing protocol. Examination included the motor section of the Unified Parkinson Disease Rating Scale (UPDRS) (Fahn, 1986), the Progressive Supranuclear Palsy Rating Scale (PSPRS) (Golbe and Ohman-Strickland, 2007) and the Brixton test of executive function. The Brixton test is an untimed test of executive function with minimal motor demands (Burgess and Shallice, 1997). Visual face perception was assessed with the famous faces test (Calder et al., 1996). In this test, pictures of the faces of 30 famous people and 10 unfamiliar people are intermixed. Participants were asked to state which faces were familiar and to state their occupation and name. Scores were given for correct answers for each of the four aspects of the test (i.e. correctly stating someone was famous, giving their correct occupation and name and correctly rejecting those who were not famous).

Auditory thresholds were assessed in each ear with an automated program with five frequencies. Pure tones (250, 500, 1000, 2000 and 4000 Hz) were presented through a calibrated sound card on a Dell Latitude D520 laptop with Sony MDR-7506 headphones.

Social cognition testing

The ability to recognize emotion in voices was tested using the voice emotion recognition task, developed from the Montreal Affective Voices (Belin et al., 2008). Seventy vocal ‘affect bursts’ (Juslin and Scherer, 2005) recorded by 10 different actors (five male and five female) were played to the participants. These sounds were short, ranging from 240 ms to 4310 ms and portrayed the emotion in some way—for example, a retch for disgust. The stimuli have been previously rated as representing one of the six basic emotions, or a neutral sound devoid of emotion (Belin et al., 2008). The names of the six basic emotions (happy, sad, fear, anger, disgust, surprise and neutral) were displayed on the screen and participants were asked to choose the best descriptor for the sound heard. The sounds were played through a Dell Latitude D520 laptop and Sony MDR-7506 headphones. Stimuli were presented using E-Prime (Psychology Software Tools, Inc.).

Theory of mind was tested with TASIT (McDonald et al., 2002, 2003). This presents participants with short videos of exchanges between actors. The actors exhibit sincerity, sarcasm or paradoxical sarcasm in these exchanges. Paradoxical sarcasm refers to an exchange where the words do not make sense unless the observer is aware that the actor is being sarcastic (Supplementary material). Participants were asked to relay the meaning behind the exchanges in response to fixed questions asked by the tester. Videos typically lasted a minute and were played more than once if requested. The videos were played using Windows Media Player on a Dell Latitude D520 laptop. Performance of each social cognition subtest is given in Supplementary Tables 1 and 2.

Statistics

Statistical analysis of behavioural data used SPSS v15 (SPSS Inc.). Parametric data for patients and controls were compared with t-tests or repeated measures ANOVA with post hoc t-tests, and Greenhouse–Geisser correction where necessary. Non-parametric data were investigated with Mann Whitney or $χ^2$ tests if categorical. Significantly skewed data were transformed using arcsin transformations. Bonferroni correction was used for multiple comparisons where appropriate. Pearson’s correlations were used for correlations, with correlations only being carried out between patients. When voice emotion was used as a composite measure, such as in correlations or voxel-based morphometry, this was the average of all emotional stimuli without neutral. Similarly, references to TASIT scores as a composite measure refer to the average score on the sarcastic stimuli, without the sincere stimuli.

Voxel-based morphometry

The voxel-based morphometry (Ashburner, 2009) analysis used T1-weighted MP-RAGE images (repetition time 2300 ms; echo time 2.86 ms; inversion time 900 ms; flip angle 9°; matrix dimensions 192 x 192 in 144 slices with isotropic voxels 1.25 mm) acquired on a 3 T Siemens MAGNETOM TrioTim Syngo MR B17 scanner (Siemens Medical Systems). Skull stripping and correcting for non-uniformities in the images can improve the performance of voxel-based morphometry in SPM5 (Acosta-Cabronero et al., 2008). Accordingly, the Hybrid Watershed Algorithm, using atlas information (Ségonne et al., 2004) (Freesurfer v.4.05, surfer.nmr.mgh.harvard.edu), was used to remove the skull and other non-brain tissue except CSF. Field inhomogeneities were corrected using a non-parametric non-uniform intensity normalization algorithm (N3 v.1.10 with Freesurfer default...
settings) (Sled et al., 1998). Venous sinuses and CSF were extracted using the brain extraction tool (BET) v.2.1 (Smith, 2002) in FSL v.4.1 (www.fmrib.ox.ac.uk/fsl). Fractional intensity threshold, $t$, was set to 0.2 and the vertical gradient, $g$, set to 1.

Subsequent preprocessing and analysis used SPM5 (www.fil.ion.ucl.ac.uk/spm). Images were spatially normalized and segmented into different tissue classes using the unified segmentation model in SPM5 (Ashburner and Friston, 2005); default settings were used throughout. The resulting modulated images were smoothed with a 16-mm full-width at half-maximum kernel to accommodate anatomical variation among subjects and enable Gaussian random field theory for statistical inferences.

Whole brain voxel-wise analyses used SPM5 with total intracranial volume as a nuisance covariate. Total intracranial volume was calculated by summing the number of voxels in each segmented tissue class over a threshold of 0.5 and multiplying the total by the voxel volume. This method is reproducible and accurate in older individuals (Pengas et al., 2009). An explicit mask was used in the grey and white matter analysis, respectively, averaged over all controls and patients in the 'control versus patient' contrasts, or patients only in the regression analyses, using a threshold of 0.1 for voxel inclusion.

Multiple regression analyses examined the relation between specified test scores and volumes of voxel grey or white matter. We predefined regions of interest based on prior studies of emotion recognition and theory of mind. For emotion recognition, the regions of interest included the left and right amygdala, bilateral insula and the right inferior frontal gyrus (Wildgruber et al., 2005; Schirmer and Kotz, 2006).

The definition of the regions of interest for the theory of mind task was based on two converging lines of evidence. First, we recognize that there are a multitude of different theory of mind tasks with different modes of presentation (Gregory et al., 2002; Frith and Frith, 2003; Saxe and Kanwisher, 2003; Snowden et al., 2003; Amadio and Frith, 2006; Gilbert et al., 2006; Hirao et al., 2008; Carrington and Bailey, 2009; Kipps et al., 2009b; Rankin et al., 2009; Shamay-Tsoory et al., 2009). However, there is a consensus that the medial frontal cortex, the temporoparietal area and the posterior superior temporal sulcus are core areas associated with theory of mind. In a review of 40 studies of theory of mind, Carrington and Bailey (2009) found that the medial frontal cortex was implicated in 88% of studies, with the posterior superior temporal sulcus and the temporoparietal area found in 45%. Secondly, MRI-based measures of atrophy in progressive supranuclear palsy commonly identify the frontal rather than temporal areas (Cordato et al., 2000, 2002, 2005; Brenneis et al., 2004; Paviour et al., 2004). Therefore, we hypothesized that atrophy in the medial frontal cortex would be a significant contributor to theory of mind deficits and used the anterior rostral medial frontal cortex [as defined by Amadio and Frith (2006)] as our region of interest. We used the WFU picatlas toolbox (Lancaster et al., 1997, 2000; Tzourio-Mazoyer et al., 2002; Maldjian et al., 2003, 2004) to define the regions of interest with $x = 0 \pm 20 \text{mm}$, as in other studies (Supplementary Fig. 1; Gilbert et al., 2006; Carrington and Bailey, 2009; Van Overwalle, 2009).

Two statistical thresholds are reported. First, at $P < 0.05$ with family wise error (FWE) correction for multiple comparisons and second, an exploratory threshold of $P < 0.001$ (uncorrected). For region of interest analysis, only areas reaching significance with FWE correction for multiple comparisons are reported.

### Results

#### Participants

Demographic data are detailed in Table 1. There were no differences between patients and controls in gender ($\chi^2 = 0.262, df = 1, P > 0.6$), age ($t_{(43)} < 1, P > 0.9$) or education years ($U = 182.5, P > 0.1, r = -0.24$). The Brixton test of executive function was significantly different between patients and controls ($U = 133.5, P = 0.03, r = -0.3$).

There were no differences between patients and controls for any of the perceptual control tasks. In the famous faces test, a $2 \times 4$ (familiar faces, occupations, names, unfamiliar faces) repeated measures ANOVA of the TASIT (arcsin transformed) confirmed impaired scores showed a main effect for the test [$F(1.75,71.7) = 39.3, P < 0.001, n = 21$] but no main effect for group [$F(1.41, n = 21) = 2.4, P > 0.1$]. There were no differences between patients and controls in gender ($t_{(41)} = 1.6, P > 0.1$). For auditory thresholds, repeated measures ANOVA showed no main effect for group [$F(1,43) = 2.5, P > 0.1, n = 23$] or Group × Frequency interaction [$F(4.6,196.1) = 1.5, P > 0.1$].

#### Social cognition

Patients with progressive supranuclear palsy were impaired on the social cognition tasks. For the voice emotion recognition task, repeated measures ANOVA of arcsin transformed scores showed a main effect for group [$F(1,43) = 40.8, P < 0.001$], a main effect of emotion type [$F(4.1,176.4) = 86.6, P < 0.001$] and an Emotion × Group interaction [$F(4.1,176.4) = 3.9, P = 0.004$]. Post hoc analysis showed no significant difference (Bonferroni corrected) in the happy condition but significant deficits for other emotions (each $P < 0.007$; Fig. 1).

ANOVA of the TASIT (arcsin transformed) confirmed impaired performance by patients with progressive supranuclear palsy [main effect for group, $F(1,40) = 15.9, P < 0.001$, main effect for context (sincere, sarcastic and paradoxical sarcasm), $F(1.7,66.7) = 4.3$, $P < 0.05$ with family wise error (FWE) correction for multiple comparisons and second, an exploratory threshold of $P < 0.001$ (uncorrected). For region of interest analysis, only areas reaching significance with FWE correction for multiple comparisons are reported.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>M:F</th>
<th>Age (years) Mean (Range)</th>
<th>Education (years) Mean (Range)</th>
<th>Disease duration Mean (Range)</th>
<th>UPDRS Mean (Range)</th>
<th>PSPRS Mean (Range)</th>
<th>Brixton Mean (Range)</th>
<th>TIV (ml) Mean (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>23</td>
<td>14:9</td>
<td>71.1 (8.6) 13 (9–20)</td>
<td>2.5 (1–17)</td>
<td>33.8 (15.7)</td>
<td>38.2 (18.1)</td>
<td>2 (1–7)</td>
<td>1435 (272)</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>22</td>
<td>15:7</td>
<td>71.4 (7.6) 11 (9–19)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>5.5 (1–7)</td>
<td>1383 (176)</td>
<td></td>
</tr>
</tbody>
</table>

Mean values are given for age, UPDRS, PSPRS and total intracranial volume (TIV) with standard deviation in parentheses. Median values are given for education years, estimated symptomatic disease duration and Brixton with range in parentheses. F = female; M = male.
P = 0.02 and a Group \times Context interaction, F(1.7, 66.7) = 5.8, P = 0.007. Post hoc analysis showed patients were not worse at understanding sincere scenarios [t(33.4) = −0.8, n.s.], but were worse at understanding sarcastic [t(40) = 3.4, P = 0.001] and paradoxically sarcastic scenarios [t(38.9) = 3.5, P < 0.001; Fig. 2].

The voice emotion recognition test was highly correlated with the TASIT (r = 0.73, P < 0.001, n = 20). Voice emotion also correlated with the motor section of the UPDRS (r = −0.49, P = 0.02, n = 22) and the PSPRS (r = −0.44, P = 0.047, n = 21). However, TASIT did not correlate with either UPDRS (r = −0.16, P = 0.5, n = 20) or the PSPRS (r = −0.17, P = 0.5, n = 19). When the correlation between TASIT and voice emotion was reassessed, partialling out UPDRS or PSPRS as a proxy for disease severity, the correlation was still significant [UPDRS (r = 0.74, P < 0.001, df = 17) and PSPRS (r = 0.72, P = 0.001, df = 16)].

We investigated the possible role of executive function in the social cognition tests by carrying out a repeated measure ANOVA for each of the social cognition tests as the dependant variables and the Brixton test as a covariate. The social cognition tests remained significantly different between patients and controls. For voice emotion (arcSin transformed), there was a main effect for group [F(1,39) = 24.5, P < 0.001], main effect for voice emotion test [F(4,2,162.5) = 18.0, P < 0.001], no interaction between voice emotion test and Brixton [F(4,2,162.5) = 1.2, P = 0.3] but a weak interaction between voice emotion test and group [F(4,2,162.5) = 2.4, P = 0.047]. For TASIT (arcSin transformed) there was a main effect for group [F(1,38) = 7.3, P = 0.01], but no main effect for TASIT [F(1,6,589) = 3.0, P = 0.07]. There was an interaction between TASIT and Brixton [F(1,6,589) = 4.1, P = 0.03] and an interaction between TASIT and group [F(1,6,589) = 3.9, P = 0.04].

**Voxel-based morphometry**

**Simple comparison between patients and controls**

The results are summarized in Tables 2 and 3 and Fig. 3. Progressive supranuclear palsy was associated with grey matter atrophy in the right insula/frontal operculum and precentral gyrus, and in the left superior frontal gyrus, postcentral gyrus and superior parietal lobule. Bilateral or midline atrophy was seen in the vermis, middle frontal gyri and cerebellum. White matter atrophy was also present in right orbitofrontal and superior frontal regions. This pattern of atrophy replicates previous reports (Brenneis et al., 2004; Price et al., 2004; Padovani et al., 2006; Paviour et al., 2006b; Nilsson et al., 2007; Josephs et al., 2008; Nicoletti et al., 2008; Rizzo et al., 2008).

**Voxel-based morphometry correlates of neuropsychological performance**

**Voice emotion**

There were associations between averaged voice emotion performance and the right inferior frontal gyrus in both grey and white matter (significant at P < 0.001 uncorrected on whole brain analysis and P < 0.05 FWE corrected within regions of interest). In addition, grey matter atrophy correlated with voice emotion scores in the cerebellum and the left middle frontal gyrus (Table 4 and Fig. 4). There were no significant suprathreshold clusters that correlated with individual voice emotions in grey or white matter. This included a lack of significant correlations in the a priori regions of interest for fear (amygdala) and disgust (insula).

**Theory of mind task**

TASIT performance correlated negatively with grey matter atrophy in the anterior rostral medial frontal cortex (significant at P < 0.001 uncorrected on whole brain analysis and P < 0.05 FWE corrected within regions of interest; Fig. 5) as well as the right superior temporal gyrus and the left supramarginal gyrus (equivalent to the left temporoparietal junction; Table 5 and Fig. 5). An association was also seen with the right inferior frontal gyrus (as seen in the voice emotion analysis) and the middle and superior frontal gyri, postcentral gyrus and cerebellum (Table 5 and Fig. 5). There were suprathreshold white matter clusters (P < 0.001 uncorrected) correlated with TASIT scores, in the...
rostral medial frontal areas and lateral superior and inferior frontal areas on the right (Table 6 and Fig. 5). No areas in grey or white matter regressed with performance on the UPDRS or the PSPRS ($P < 0.001$).

**Discussion**

This study confirms that patients with progressive supranuclear palsy have significant impairments in multiple tests of social cognition. This study extends previous work (Ghosh et al., 2009) to show that emotion recognition is affected across modalities and that more complex theory of mind abilities are also impaired. In other words, progressive supranuclear palsy prevented the patients from properly understanding what another person is thinking or feeling. These cognitive deficits were not attributable to differences in simple visual or auditory perceptual functions, and are observed with both static and dynamic stimuli.

The social cognition deficits seen are similar to that seen in behavioural variant frontotemporal dementia, a closely related
disorder to progressive supranuclear palsy, which often has a similar underlying tau neuropathology (Esiri et al., 2004; Lough et al., 2006; Kipps et al., 2009; Rankin et al., 2009; Adenzato et al., 2010). However, in contrast to behavioural variant frontotemporal degeneration, the social cognitive deficits have been under-recognized in progressive supranuclear palsy. This is perhaps because the combination of physical disability, immobility, communication and cognitive problems (Litvan et al., 1996; Bak and Hodges, 1998; Schrag et al., 2003), which were all observed in our patients, reduces the expression of socially inappropriate behaviours resulting from poor social cognition.

The analysis of MRI confirmed the group effect of progressive supranuclear palsy on regional brain structures, including grey matter atrophy of medial frontal and insula regions (Brenneis et al., 2004; Price et al., 2004; Padovani et al., 2006; Josephs et al., 2008) and severe white matter atrophy of the brainstem.

### Figure 3

Figure 3 SPM(\(t\)) of atrophy in grey (A) and white (B) matter in patients with progressive supranuclear palsy versus controls (threshold \(P < 0.001\) uncorrected, overlaid on representative T1 brain in standard space using the MNI template).

### Figure 4

Figure 4 SPM(\(t\)) map of atrophy in grey (A) and white (B) matter seen in patients with progressive supranuclear palsy when regressed with voice emotion scores. Illustrated at \(P < 0.001\) uncorrected.

### Table 4 Areas of grey and white matter atrophy in patients when regressed with voice emotion scores

<table>
<thead>
<tr>
<th>Brain area</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Peak Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grey matter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus (R)</td>
<td>62</td>
<td>24</td>
<td>8</td>
<td>3.19*</td>
</tr>
<tr>
<td>Middle frontal gyrus (L)</td>
<td>-42</td>
<td>14</td>
<td>32</td>
<td>3.16</td>
</tr>
<tr>
<td>Lateral cerebellum (R)</td>
<td>56</td>
<td>-56</td>
<td>-44</td>
<td>3.33</td>
</tr>
<tr>
<td>White matter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base of inferior frontal gyrus (R)</td>
<td>36</td>
<td>22</td>
<td>12</td>
<td>3.13</td>
</tr>
</tbody>
</table>

The table reports peak voxel location and Z-score, exceeding \(P < 0.001\) (uncorrected). *Peaks at which \(P < 0.05\) with a FWE correction for whole brain multiple comparison. a The use of an a priori region of interest for Brodmann areas 45 and 47. Coordinates (\(x, y, z\)) are given according to standard anatomic space using the Montreal Neurological Institute template. L = left; R = right.
Neuropsychology of social dysfunction

Progressive supranuclear palsy impaired the recognition of multiple emotions in the auditory (vocal) domain. In contrast to the deficits in the recognition of most emotions, happiness recognition was preserved. This might be because happiness is a positive emotion, the others being either negative (anger, disgust, sad, fear) or intermediate (surprise). It may also have been a ceiling effect, as happiness is least often confused with the other emotions when using visual stimuli (Ekman and Friesen, 1976; Scherer, 2003) and is more readily recognized than the other basic emotions in health (Russell, 1994). Nonetheless, the preservation of happiness recognition provides additional evidence that the patients understood the tasks and were able to respond appropriately.

Progressive supranuclear palsy also impaired the TASIT theory of mind task. Patients did not have difficulties interpreting sincere statements, suggesting that they understood the task, were able to engage with it and could manage its demands on working memory. However, they found the sarcastic statements difficult, particularly the paradoxical sarcasm. This suggests that they were unable to interpret the mental state of the protagonists in the video and have a theory of mind deficit.

Is there a role for executive dysfunction in the social cognition deficits? Some have argued that they are independent (Gregory et al., 2002; Lough et al., 2006; Torralva et al., 2007), while others have proposed that executive function is necessary to understand theory of mind tests (Channon and Crawford, 2000; Aboulafia-Brakha et al., 2011). To investigate this in our group, we used the Brixton test of executive function as it is a non-motor and untimed test of executive function. Although executive function was impaired in patients with progressive supranuclear palsy, we found that social cognition was still impaired when the Brixton test was included as a covariate. This suggests that executive functions are not sufficient to account for the social cognitive deficit in progressive supranuclear palsy, at least as measured by the tests used here.

The social cognition deficits that we have described are similar in range to those found in behavioural variant frontotemporal dementia. Patients with behavioural variant frontotemporal dementia have a profound deficit in both emotion recognition and theory of mind, which has been shown in a variety of tests including TASIT used here (Lough et al., 2006; Kipps et al., 2009b; Rankin et al., 2009; Adenzato et al., 2010; Shany-Ur et al., 2011). As pathology (Esiri et al., 2004) and patterns of atrophy (Rosen et al., 2002a; Brenneis et al., 2004; Price et al., 2004; Williams et al., 2005; Padovani et al., 2006; Josephs et al., 2008; Pereira et al., 2009) overlap between progressive supranuclear palsy and behavioural variant frontotemporal degeneration, it is likely that the aetiology of the social cognition deficits is similar in these two diseases.

Neural correlates of social cognition

A major aim of this study was to understand the neuroanatomical basis of social cognition deficits in progressive supranuclear palsy. The regression of grey and white matter volumes against voice emotion scores confirmed our prediction of a behaviourally relevant structure–function relationship in the right inferior frontal...
gyrus, as well as regressing with the left middle frontal gyrus. These areas have been associated with the explicit decoding of the prosody of sounds and non-linguistic vocalizations in order to determine their emotional content (George et al., 2006). Coordinates \((x, y, z)\) are given according to standard anatomic space using the Montreal Neurological Institute template. L = left; R = right.

Table 5 Areas of grey matter atrophy in patients when regressed with TASIT scores

<table>
<thead>
<tr>
<th>Brain area</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Peak Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inferior frontal gyrus (R)</td>
<td>62</td>
<td>24</td>
<td>8</td>
<td>4.1</td>
</tr>
<tr>
<td>Middle frontal gyrus (L)</td>
<td>−46</td>
<td>16</td>
<td>32</td>
<td>4</td>
</tr>
<tr>
<td>Middle frontal gyrus (R)</td>
<td>26</td>
<td>28</td>
<td>58</td>
<td>3.97</td>
</tr>
<tr>
<td>Posterior cerebellum (R)</td>
<td>44</td>
<td>−76</td>
<td>−52</td>
<td>3.86</td>
</tr>
<tr>
<td>Superior frontal gyrus (L)</td>
<td>−24</td>
<td>8</td>
<td>64</td>
<td>3.77</td>
</tr>
<tr>
<td>Postcentral gyrus (L)</td>
<td>−64</td>
<td>−10</td>
<td>40</td>
<td>3.71</td>
</tr>
<tr>
<td>Anterior medial frontal cortex (R)</td>
<td>2</td>
<td>40</td>
<td>36</td>
<td>3.49*</td>
</tr>
<tr>
<td>Superior temporal gyrus (R)</td>
<td>74</td>
<td>−18</td>
<td>−2</td>
<td>3.46</td>
</tr>
<tr>
<td>Posterior cerebellum (R)</td>
<td>−42</td>
<td>−60</td>
<td>−62</td>
<td>3.41</td>
</tr>
<tr>
<td>Supramarginal gyrus (L)</td>
<td>−52</td>
<td>−44</td>
<td>56</td>
<td>3.38</td>
</tr>
</tbody>
</table>

The table reports peak voxel location and Z score, exceeding \(P < 0.001\) (uncorrected).

*Peaks at which \(P < 0.05\) with a FWE correction for multiple comparison.
a The use of an a priori region of interest for the anterior rostral medial frontal cortex (Amodio and Frith, 2006). Coordinates \((x, y, z)\) are given according to standard anatomical space using the Montreal Neurological Institute template. L = left; R = right.

Table 6 Areas of white matter atrophy in patients when regressed with TASIT scores

<table>
<thead>
<tr>
<th>Brain area</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Peak Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>White matter to inferior frontal gyrus (R)</td>
<td>38</td>
<td>22</td>
<td>10</td>
<td>3.65</td>
</tr>
<tr>
<td>Cingulate area (L)</td>
<td>−2</td>
<td>16</td>
<td>44</td>
<td>3.35</td>
</tr>
<tr>
<td>Anterior rostral medial frontal area (R)</td>
<td>12</td>
<td>52</td>
<td>34</td>
<td>3.19*</td>
</tr>
<tr>
<td>White matter to superior frontal gyrus (R)</td>
<td>24</td>
<td>18</td>
<td>38</td>
<td>3.14</td>
</tr>
<tr>
<td>Posterior rostral medial frontal area (R)</td>
<td>20</td>
<td>26</td>
<td>40</td>
<td>3.12</td>
</tr>
</tbody>
</table>

The table reports peak voxel location and Z score, exceeding \(P < 0.001\) (uncorrected).

*Peaks at which \(P < 0.05\) with a FWE correction for multiple comparison.
a The use of an a priori region of interest for the anterior rostral medial frontal cortex (Amodio and Frith, 2006). Coordinates \((x, y, z)\) are given according to standard anatomical space using the Montreal Neurological Institute template. L = left; R = right.

Despite our predictions, we did not find a correlation between the insula and disgust or the amygdala and fear stimuli. We estimated that we had sufficient statistical power to detect large or medium effects of association. One possible explanation is that these areas are not essential for recognizing emotion stimuli, despite activation in health. Alternatively, the atrophy in these regions may already have been too severe to show further correlations with performance: the group comparison confirmed that the insula and amygdala were severely atrophic in progressive supranuclear palsy, both in grey and surrounding white matter.

The regressions with TASIT performance revealed structural correlates of theory of mind in the anterior rostral medial frontal cortex. This area has been widely implicated in the function of theory of mind both by analysing theory of mind function in normal participants using functional MRI, as well as theory of mind deficits in patients using voxel-based morphometry (Amodio and Frith, 2006; Gilbert et al., 2006; Shamay-Tsoory et al., 2009). The finding of this association in our patients adds greater credence to our behavioural findings and adds further weight to the link between the anterior rostral medial frontal cortex and theory of mind ability.

In our study, the peak of association was on the superior border of the anterior rostral medial frontal cortex. A meta-analysis (Amodio and Frith, 2006) suggests that this border region of anterior rostral medial frontal cortex is used particularly when people judge actions or thoughts of unfamiliar others, as opposed to the inferior section of the anterior rostral medial frontal cortex, used when assessing the feelings of familiar others. In keeping with this anatomical specialization, TASIT involves assessing the intentions of unfamiliar people, and hence, our results are in keeping with this anatomical distinction.

Consistent with previous studies (Saxe and Kanwisher, 2003; Brunet-Gouet and Decety, 2006; Carrington and Bailey, 2009; Van Overwalle, 2009; Frith and Frith, 2010), the posterior superior temporal sulcus and the temporoparietal junction correlated with social cognition performance in progressive supranuclear palsy. These were not specified as a priori regions of interest for the purposes of small volume correction of multiple comparisons. Their identification at the exploratory (uncorrected) threshold is nonetheless consistent with their posited role in theory of mind, and suggests that theory of mind deficits in our patients are linked to the atrophy of these areas seen pathologically (Schofield et al., 2011).

Interestingly, atrophy in the right inferior frontal gyrus and the left middle frontal gyrus correlated with TASIT score, as well as the voice emotion score. The voice emotion and TASIT tests were highly correlated with each other behaviourally. This correlation was still significant when the UPDRS or PSPRS, as proxy measures for general disease progression, were partialled out of the correlation. This suggests that performance on both social cognition tests may represent an underlying neuropsychological mechanism supported by the middle frontal gyrus and inferior frontal gyrus. One possibility could be that these areas are needed to interpret the auditory emotions from the stimuli in both the voice emotion recognition and theory of mind tasks. Previous research has shown that the ability to interpret voice prosody is important for the interpretation of sarcasm (Rockwell, 2007; Cheang and Pell, 2010).
2009), and that impairments in these two abilities are correlated in schizophrenia (Leitman et al., 2006). This provides a parsimonious explanation for the finding that atrophy in the right inferior frontal gyrus was correlated with both the voice emotion task and the theory of mind task. Both sets of stimuli necessitate the interpretation of prosody, a function which has been linked to the right inferior frontal gyrus (George et al., 1996; Buchanan et al., 2000; Adolphs et al., 2002; Wildgruber et al., 2002, 2005; Gandour et al., 2003; Ethofer et al., 2006; Johnstone et al., 2006; Schirmer and Kotz, 2006; Alba-Ferera et al., 2011).

The right inferior frontal gyrus might also be related to social cognition deficits as part of a putative mirror neuron system dysfunction. Mirror neurons are involved in the execution of an action, but are also activated while observing the same action when carried out by another (Gallese et al., 2011 for review). In progressive supranuclear palsy, mirror neurons have already been proposed to explain a deficit in naming objects that are manipulable, rather than those that are not (Chow et al., 2010). The authors relate this to patients’ progressive difficulty with manual dexterity and found that naming scores for manipulable items regressed with atrophy in left pre-motor areas. Several groups have gone further to show a link between deficits in the production and recognition of emotions (Wicker et al., 2003; Gallese et al., 2004).

Other studies have looked at contextually specific, observed or heard motor tasks and found activation in cerebral areas that are necessary for carrying out social cognitive tasks (Gazzola et al., 2006; Kaplan and Iacoboni, 2006). Increased activation in the putative mirror neuron system has also been related to higher scores on empathy tests. For example, activity in the right inferior frontal gyrus increased when observed actions were correct contextually, suggesting the subjects used a ‘cognitive empathy’ approach for interpretation of the task (Kaplan and Iacoboni, 2006).

In addition, that study found that the right inferior frontal gyrus activity correlated with several measures of empathy, including the ability to identify with characters in a fictional situation, clearly an ability that would be relevant in our assessment of theory of mind in progressive supranuclear palsy [but see Tavares et al. (2011) for contrasting findings, and Gallese et al. (2011) for broader issues with mirror neuron accounts of social cognition]. In our study, this interpretation is partially supported by the correlation of the voice emotion score to tests with a strong motor component, the UPDRS and the PSPRS (although we note that the TASIT did not correlate to either the UPDRS or the PSPRS).

Limitations of the study

There are several limitations to this study. First, we relied on clinical not pathological diagnosis, and this might in principal affect the specificity of the results. However, the clinical diagnostic criteria convey >90% accuracy in large trials (Bensimon et al., 2009) and in our local pathology series, so the associations we identify are unlikely to be driven by a pathology other than progressive supranuclear palsy. All nine of the patients who have undergone post-mortem examination by the time of writing had confirmation of progressive supranuclear palsy as the principal diagnosis. Second, we used a cross-sectional design, with patients who were both early and late in the disease. We cannot therefore make inferences about the within-subject progression of social cognition over time, or its relation to progressive atrophy. Regarding neuropsychological assessment, one must also consider confounds of performance, including fatigue. We took several steps to limit this, including short assessment periods over several days, preferentially in the morning or with frequent rest periods. The normal performance on control tasks/conditions speaks against a generic confound such as fatigue.

Voxel-based morphometry has potential confounds, including extra-axial tissue confounds, mis-registration to templates or tissue misclassification. To reduce these confounds, we used a voxel-based morphometry protocol that has been shown to be optimal in other age-related neurodegenerative diseases, and we replicated the simple group comparison results from previous studies of grey and white matter change in progressive supranuclear palsy (Brenneis et al., 2004; Price et al., 2004; Paviour et al., 2005, 2006a, b; Padovani et al., 2006; Nilsson et al., 2007; Josephs et al., 2008; Nicoletti et al., 2008; Rizzo et al., 2008). Our statistical approach was to emphasize control of type I error, and to use pre-specified regions of interest to enhance statistical power at the expense of anatomical coverage for key contrasts. Several areas, such as the temporal and temporoparietal cortex, were identified using the exploratory threshold. Although these areas have sometimes been associated with normal social cognition, we did not include them in our a priori regions of interest, highlighting the compromises inherent in region of interest analyses. We therefore also report at a more liberal threshold (0.001 uncorrected) as this may increase reliability across studies (Thirion et al., 2007). Larger groups would increase power, but our study was similar or larger than many voxel-based morphometry studies of progressive supranuclear palsy and related disorders (Whitwell et al., 2005; Kipps et al., 2007, 2009b; Noppeney et al., 2007; Werner et al., 2007; Rohrer et al., 2010).

We suggest that the deficits in social cognition arise from focal atrophy of critical brain regions and that they are likely to contribute to the reduced quality of life and psychiatric comorbidity among patients and carers (Schrag et al., 2003; Herting et al., 2007; Bak et al., 2010). However, in this study, we did not assess quality of life among our patient cohort, nor monitor the progression of these factors over time, and further work would be required to demonstrate any causal association.

Conclusion

We have shown that progressive supranuclear palsy impairs the recognition of emotion and impairs theory of mind. In addition to replicating grey and white matter atrophy patterns in progressive supranuclear palsy, we have found structural correlations with social cognition performance. Specifically, atrophy in the right inferior frontal gyrus correlated negatively with scores on both the voice emotion test and TASIT, and atrophy of the anterior rostral medial frontal cortex correlated with theory of mind impairments. These results, together with the preserved performance on control tasks, indicate a generic social cognition deficit due to progressive supranuclear palsy pathology in critical brain networks for social cognition.
Set in the context of previous research, showing that 50% of patients complain of a negative impact on their quality of life due to social impairment (Schrag et al. 2003) and research indicating that patients present with behavioural change, which includes disinhibition and aggressiveness (Donker Kaat et al., 2007), a social cognition deficit should be considered by those managing and caring for patients with progressive supranuclear palsy. Greater acknowledgment of the patients’ impaired ability to empathize with those around them may help to alleviate patient–carer relationship difficulties at a time when the patient needs them most.

Acknowledgements
We are grateful to Dr Nimmo-Smith of the MRC Cognition and Brain Sciences Unit for providing advice in relation to the statistics and to our patients for assisting us with this research.

Funding
This work was supported by the Medical Research Council [G0700503 to B.G., MC_US_A060_0017 to A.C.]; the Guarantors of Brain (to B.G.); the Raymond and Beverley Sackler Trust (to B.G.); Parkinson’s UK (to A.L. and P.P.). J.R. is supported by the Wellcome Trust [088324] and the Cambridge National Institute of Health Research Comprehensive Biomedical Research Centre. A.D.L. is supported by the Wales Institute of Cognitive Neuroscience (WICN).

Supplementary material
Supplementary material is available at Brain online.

References


