The site of stimulation moderates neuropsychiatric symptoms after subthalamic deep brain stimulation for Parkinson's disease

Philip E. Mosley, David Smith, Terry Coyne, Peter Silburn, Michael Breakspear, Alistair Perry

ABSTRACT

Deep brain stimulation of the subthalamic nucleus for Parkinson's disease is an established advanced therapy that addresses motor symptoms and improves quality of life. However, it has also been associated with neuropsychiatric symptoms such as impulsivity and hypomania. When significant, these symptoms can be distressing, necessitating psychiatric intervention. However, a comprehensive analysis of neurocognitive and neuropsychiatric outcomes with reference to the site of subthalamic stimulation has not been undertaken. We examined this matter in a consecutive sample of 64 persons with Parkinson's disease undertaking subthalamic deep brain stimulation. Participants were assessed with a battery of neuropsychiatric instruments at baseline and at repeated postoperative intervals. A psychiatrist identified patients with emergent, clinically-significant symptoms due to stimulation. The site of the active electrode contact and a simulated volume of activated tissue were evaluated with reference to putative limbic, associative and motor subregions of the subthalamic nucleus. We studied anatomical correlates of longitudinal neuropsychiatric change and delineated specific subthalamic regions associated with neuropsychiatric impairment. We tested the ability of these data to predict clinically-significant symptoms. Subthalamic stimulation within the right associative subregion was associated with inhibitory errors on the Excluded Letter Fluency task at 6-weeks ($p = 0.023$) and 13-weeks postoperatively ($p = 0.0017$). A cluster of subthalamic voxels associated with inhibitory errors was identified in the right associative and motor subregions. At 6-weeks, clinically-significant neuropsychiatric symptoms were associated with the distance of the active contact to the right associative subregion ($p = 0.0026$) and stimulation within the right associative subregion ($p = 0.0009$). At 13-weeks, clinically-significant symptoms were associated with the distance to the right ($p = 0.0027$) and left ($p = 0.0084$) associative subregions and stimulation within the right associative subregion ($p = 0.0026$). Discrete clusters of subthalamic voxels associated with high and low likelihood of postoperative neuropsychiatric symptoms were identified in ventromedial and dorsolateral zones, respectively. When a classifier was trained on these data, clinically-significant symptoms were predicted with an accuracy of 79\%. These data underscore the importance of accurate electrode targeting, contact selection and device programming to reduce postoperative neuropsychiatric impairment. The ability to predict neuropsychiatric symptoms based on subthalamic data may permit anticipation and prevention of these occurrences, improving safety and tolerability.

1. Introduction

Subthalamic deep brain stimulation (DBS) is an advanced therapy for Parkinson's disease that reduces motor symptoms and improves quality of life (Schuepbach et al., 2013). However, the relationship between DBS and the neuropsychiatric features of Parkinson's disease is complex (Mosley and Marsh, 2015). Some individuals become more impulsive and less empathic after DBS, acting recklessly without foresight or concern for others. The incidence of this syndrome has been estimated at up to 15\% (Appleby et al., 2007; Daniele et al., 2003).
Predicting those persons who will develop this behavioural syndrome is a challenge for the clinician, given the wide spectrum of neuropsychiatric symptoms associated with PD (Weintraub and Burt, 2011). Furthermore, the predictive value of preoperative neuropsychiatric symptoms is unclear.

The emergence of postoperative impulsivity may be a neuro-stimulatory phenomenon related to the computational role of the STN in behaviour. In addition to its role as a relay nucleus that increases the inhibitory drive of the basal ganglia, the STN is a second input station to the basal ganglia, receiving direct cortical projections from the frontal lobe in the ‘hyperdirect’ pathway (Nambu et al., 2002). Here, the STN may ‘brake’ or ‘delay’ cognitive-associative circuits in the basal ganglia, suppressing impulsive and potentially error-prone responding. Functional and structural brain imaging support the role of this pathway in motor inhibition (Aron et al., 2007; Rae et al., 2015). Following STN DBS, persons with Parkinson’s disease demonstrate failures of motor inhibition (Hershey et al., 2004), action cancellation, (Obeso et al., 2013), as well as showing a failure of prepotent verbal inhibition (Thobois et al., 2007; Witt et al., 2004). Furthermore, when faced with a difficult choice, persons with Parkinson’s disease speed rather than slow their decision-making after STN-DBS (Cavanagh et al., 2011; Frank et al., 2007), where taking more time would be an optimal response in order to make an accurate decision. However, impulsivity is not the only behavioural symptom that has been reported subsequent to STN-DBS. Previous work has identified relationship discord precipitated by indifference to the emotional wellbeing of the partner of the person with Parkinson’s disease (Lewis et al., 2015; Schupbach et al., 2006). One recent investigation has suggested a role for the STN in object valuation, which offers a potential mechanism for these more complex changes (Seymour et al., 2016).

The anatomy of the STN confers vulnerability to stimulation-dependent neuropsychiatric changes. A tripartite functional organization of the STN into limbic, associative and motor subregions is suggested by primate and human studies (Haynes and Haber, 2013; Lambert et al., 2012), although with considerable topological overlap and without lobar boundaries (Alkemade and Fornsten, 2014; Lambert et al., 2015). Yet, the small size of the STN means that current diffusion from a stimulating contact in the dorsolateral sensorimotor region could modulate subthalamic regions with greater connectivity to fronto-striatal networks implicated in mood, decision-making and reward. The more ventral and medial the stimulating contact, the more likely these circuits are to be affected by DBS. Accordingly, moving stimulation between active contacts in the dorsal motor and ventral limbic aspects of the nucleus can impair response inhibition (Hershey et al., 2010) and precipitate manic symptoms (Mallet et al., 2007). One investigation has examined the influence of electrode position on psychiatric outcomes in a sample larger than a small case series (Welter et al., 2014). Here, the depth of the active contact in both the right and left hemispheres, relative to the inter-commissure plane, was associated with postoperative hypomania. However, this investigation only included one formal postoperative assessment after 1 year, at which time only depressive symptoms were evaluated. Other psychiatric symptoms, including measures of impulsive responding, were not assessed.

The locus of subthalamic stimulation has been shown to affect motor outcomes after DBS for Parkinson’s disease. Accurate targeting of the stimulating contact to the lateral aspect of the STN significantly improves motor symptoms (Wodarg et al., 2012). Recently, Akram et al. (2017) modelled the extent of neural tissue activation in the subthalamic region and regressed these against motor outcomes, delineating discrete clusters of subthalamic and peri-subthalamic voxels associated with maximal improvement in tremor, rigidity and bradykinesia. These were distributed in the posterior and superior aspects of the nucleus.

The aim of our investigation was to identify if the locus of subthalamic stimulation moderates the evolution of postoperative neuropsychiatric symptoms. Tractographic parcellation of the STN into motor, associative and limbic regions furnished precise information on the position of each electrode contact relative to these STN subregions. We simulated a volume of activated tissue (VAT) for each hemisphere, for each participant, at each follow up. This enabled us to estimate the dispersion of charge within each STN subregion at a given time, and allowed us to evaluate the contribution of stimulation parameters to emergent symptoms during titration. Finally, we applied these data to delineate STN regions significantly associated with neuropsychiatric impairment and tested the predictive validity of our data. Our hypotheses were that the position of the active electrode contact and dispersion of charge in the associative and limbic regions of the STN would be significant determinants of postoperative neuropsychiatric symptoms in individuals with Parkinson’s disease.

2. Materials and methods

2.1. Participants

Sixty-four participants were consecutively recruited at the Asia-Pacific Centre for Neuromodulation between 2013 and 2017, during the assessment of eligibility for STN-DBS. The diagnosis of Parkinson’s disease was confirmed according to the United Kingdom Queens Square Brain Bank criteria (Hughes et al., 1992). The laterality of disease onset and the Hoehn and Yahr stage (Hoehn and Yahr, 1967) at operation was recorded. The PD subtype (tremor-dominant, akinetic-rigid, mixed-type) was established based on an analysis of the dominant symptoms elicited during the Unified Parkinson’s Disease Rating Scale (UPDRS) Part III Motor Examination, as described in Spiegel et al. (2007). Candidates with cognitive impairment were excluded, as defined by a Mini Mental State Examination Score (MMSE) of 25 or less or a clinical diagnosis of PD dementia. The latter was defined according to published Movement Disorder Society criteria (Emre et al., 2007). All participants completed a psychiatric and cognitive evaluation prior to surgery.

2.2. Image acquisition

A preoperative T1-weighted MPRAGE and a T2-weighted FLAIR sequence were acquired. For participants 1–26, this took place using a 3 T GE Signa Hdx with a 32-channel head coil at St Andrews War Memorial Hospital. The acquisition parameters were as follows: T1: 1 mm³ voxel-resolution, TR = 6.13 ms, TE = 2.01 ms, flip angle = 15°, matrix size = 256 × 256, FOV = 256 × 256 × 166; T2: 1 × 1 × 2 mm voxel-resolution, TR = 9502 ms, TE = 120.54 ms, flip angle = 90°, matrix size = 256 × 256, FOV = 256 × 256 × 75. Participants 27–64 were scanned using a 3 T Siemens Prisma, with a 64-channel head coil at the Herston Imaging Research Facility. The acquisition parameters were as follows: T1, 1 mm³ voxel-resolution, TR = 2000 ms, TE = 2.38 ms, flip angle = 9°, matrix size = 256 × 256, FOV = 256 × 256 × 192; T2, 1 × 1 × 2 mm voxel-resolution, TR = 9500 ms, TE = 122.0 ms, flip angle = 120°, matrix size = 256 × 256, FOV = 256 × 256 × 70. Postoperative CT images for all participants were acquired on a Siemens Intuvo, with a resolution of 0.5 mm³.

2.3. Surgery and clinical follow-up

Bilateral implantation of Medtronic 3389 (n = 48) or Boston Vercise (n = 16) electrodes took place in a single-stage procedure using a Leksell stereotactic apparatus, with the STN having been identified as a midbrain structure on Fluid Attenuation Inversion Recovery (FLAIR) imaging. Intraoperative microelectrode recordings (MER) were employed to establish localisation within the STN and intraoperative test stimulation was performed. A CT scan confirmed satisfactory postoperative lead placement. Contact selection for initial stimulation was based upon MER signals, with titration and evaluation of stimulation over the following week as an inpatient until motor symptoms were
satisfactorily treated without adverse effects. All implanted electrode contacts were of the ring rather than segmented configuration and thus no current-steering was applied. Post-discharge, participants returned to the movement disorders clinic for further neurological and psychiatric evaluation, with further DBS manipulation according to a set schedule of visits. The predominant criterion for DBS manipulation at each visit was manifest motor symptoms of Parkinson’s disease. However, if the patient, caregiver or clinician detected neuropsychiatric symptoms (such as mood elevation, disinhibition or irritability) then a psychiatric review was initiated. Manipulation of the DBS device occurred immediately if neuropsychiatric symptoms were determined to be stimulation-related and clinically-significant (e.g. precipitating interpersonal impairment). All participants received routine psychiatric follow up as part of multidisciplinary care.

2.4. Neuropsychiatric outcomes

2.4.1. Neuropsychiatric assessments

Assessments took place prior to DBS and subsequently at 2-weeks, 6-weeks, 13-weeks and 26-weeks postoperatively, using the same battery of participant, caregiver and clinician-rated instruments, described in Mosley et al. (2018). This investigation focussed on neuropsychiatric symptoms identified as significant drivers of caregiver burden in a multivariate analysis (Mosley et al., 2018). Briefly, these included measures of impulsivity: the Barratt Impulsiveness Scale II (BIS) (Patton et al., 1995); the Questionnaire for Impulsive-Compulsive disorders in PD Rating Scale (QUIP-RS) (Weintraub et al., 2012); the Excluded Letter Fluency task (ELF) (Shores et al., 2006); the Hayling test (Burgess et al., 1997); empathy: the Empathy Quotient (EQ) (Baron-Cohen and Wheelwright, 2004); and depression: the Beck Depression Inventory II (BDI) (Beck et al., 1961). In addition, a modified version of the BIS and EQ (caregiver-rated BIS and caregiver-rated EQ) assessed these behaviours from the perspective of the caregiver, given that participant and caregiver ratings may be discrepant (Lewis et al., 2014).

Furthermore, at each visit motor symptoms were assessed with the UPDRS Part III motor examination. Dopaminergic medication was converted to a levodopa-equivalent daily dose (LEDD) value (Evans et al., 2004). DBS parameters such as active contacts, amplitude, pulsedwidth, frequency and impedance were recorded. Participants with Boston electrodes (n = 16) had their stimulation amplitude converted from milliamps to volts using impedance measurements at the active contact. Participants were ‘ON’ medication and stimulation for all assessments, in order to provide a naturalistic evaluation of symptom evolution.

2.4.2. Neuropsychiatric ‘caseness’

In addition to the above instruments, participants were assigned to the category ‘case’ or ‘non-case’ depending on whether they developed clinically-significant neuropsychiatric symptoms attributable to DBS, necessitating device manipulation (hereafter referred to as ‘caseness’). This category was operationalised as follows: a participant was brought to the attention of the psychiatrist (who had assessed all participants at baseline) through self-referral, or via a relative or a clinician with concerns about the participant’s mood or behaviour. The psychiatrist conducted a semi-structured diagnostic interview and mental state examination with attention to euphoria, irritability, disinhibition, impulsivity, compulsivity and empathy. If the psychiatrist considered that a stimulation-related neuropsychiatric presentation was likely and was causing clinically-significant impairment or distress, the neurologist adjusted the participant’s DBS settings by moving to a new active contact or reducing the amplitude of stimulation. The psychiatrist then repeated the clinical assessment and sought collateral information from the participant’s relatives. If the neuropsychiatric symptoms responded immediately to device manipulation then the participant was defined as a case.

A timeline displaying the sequence of assessments is presented in Fig. 1A.

2.5. Image processing

Preoperative MPRAGE and FLAIR images were co-registered with the postoperative CT scan using an affine transformation within FSL (FLIRT version 6.0, (Smith et al., 2004)). Each co-registration was manually checked for accuracy (Fig. 2A).

The co-registered acquisitions were spatially normalized into ICBM_2009b nonlinear asymmetric space using a fast diffeomorphic image registration algorithm (DARTEL) as implemented in SPM12 (Ashburner, 2007) (Fig. 2B and C). Using the Lead-DBS toolbox (version 1.6.4.2 (Horn and Kuhn, 2015)), (http://www.lead-dbs.org), electrodes were manually localised and corrected for brainshift by applying a refined affine transform calculated between pre- and postoperative acquisitions (Fig. 2E). A volume of activated tissue (VAT), representing the dispersion of electrical charge in neural tissue, was estimated based upon individualised stimulation parameters (Madler and Coenen, 2012) at each assessment interval (Fig. 2F).

The spatial position of each electrode contact was evaluated with reference to a tractographic parcellation of the STN into limbic, associative and motor subregions (Accolla et al., 2014). For both hemispheres we calculated: 1) the distance of the electrode contact to the centroid of each STN subregion; 2) the distance to the nearest voxel of that volume; and 3) extent of each subregion volume occupied by each participant’s simulated VAT. These variables are detailed in Table 1. The image processing pipeline is depicted in Fig. 2. Code supporting these workflows is publically available at https://github.com/AlistairPerry/DBSVATstats. Images were projected onto the BigBrain histological atlas for visualisation purposes (Amunts et al, 2013).

2.6. Statistical analysis

The emphasis of this investigation was on emergent neuropsychiatric symptoms attributable to subthalamic stimulation. By week 26, any emergent neuropsychiatric symptoms linked to stimulation had attenuated following clinical intervention. No participants in this cohort became a ‘case’ between the 13 and 26-week assessment. The assessment at week 2 was discounted as residual lesion effects from electrode implantation could not be excluded. Therefore, the focus of this analysis was the assessments at 6 and 13-weeks, relative to baseline.

At each assessment interval, data was z-normalized to account for the heterogeneity of and variance in assessment instruments. To calculate the rates of change for each neuropsychiatric outcome, normalized scores at baseline were subtracted from each follow-up assessment. Across the complete data set, < 0.03% was missing for any variable. Missing data was inferred using the pooled results of 50 iterations of longitudinal imputation by classification and regression trees, employing Gibbs sampling using mice in the R software environment. (R Core Team, 2014; van Buuren and Groothuis-Oudshoorn, 2011). The distribution of original and imputed data was checked for plausibility and results are displayed in Supplementary Fig. 1. At baseline and during follow up, the demographic and phenotypic characteristics of those participants who would become postoperative cases were contrasted with non-cases. Data with a non-Gaussian distribution was treated with the Kruskal-Wallis test. The Chi-squared test was employed for categorical variables. Prior to variable selection, a repeated-measures analysis of variance (ANOVA) was first employed to identify significant group-level longitudinal changes in neuropsychiatric assessment data. When multiple comparisons were undertaken, reported p-values were adjusted using the Benjamini and Hochberg (1995) method to control for the false discovery rate, with α = 0.05.

2.6.1. Variable selection and modelling

To reduce the dimensionality of the dataset, comprising of many candidate anatomical and neuropsychiatric covariates, resolve
unknown dependencies and remove redundant variables, a variable selection and regularisation algorithm (the Least Absolute Shrinkage and Selection Operator: LASSO) was employed to identify the combination of anatomical variables with the best predictive value for each neuropsychiatric outcome (Friedman et al., 2010). A conservative one-standard-deviation rule was chosen for the regularisation parameter ($\lambda$) to protect against overfitting (Hastie et al., 2009). A binary model was employed for binary outcomes, a Gaussian model for parametric data and a Poisson model for nonparametric data.

Subsequently, neuropsychiatric variables and their anatomical predictors as identified from the LASSO were modelled in a general linear model, with the model family defined by the nature of the dependent variable. Demographic and disease-related factors, including change in LEDD, were also entered as covariates. These analyses also took place in the R software environment, using \texttt{glmnet} for optimisation and \texttt{glm} for general linear modelling. A schematic of these steps is presented in Fig. 1B. Scanning site was included as a covariate for all significant effects (Supplementary Table 8).

2.6.2. Subthalamic voxels associated with neuropsychiatric symptoms
In order to extend the findings from the general linear model, spatial clusters of subthalamic voxels associated with the neuropsychiatric outcomes were identified using threshold-free cluster-enhancement (TFCE) (Smith and Nichols, 2009). Here, hemispheric VATs for all participants at the intervals of interest were first concatenated into a 4-dimensional image and overlapping voxels within this image were thresholded to verify that the highest probability of stimulation was within the atlas-defined STN (Accolla et al., 2014). Each voxel within this defined STN mask was entered into a general linear model against the demeaned neuropsychiatric outcome scores – with the TFCE used as the test-statistic. Using FSL Randomise (Winkler et al., 2014), non-parametric permutation inference was then conducted upon each voxel to control for family-wise error (FWE), employing 5000 permutations to build up the null distribution. Voxels significantly associated with neuropsychiatric outcome of interest were identified by $\alpha = 0.05$ (FWE-corrected). FSL Cluster was employed for cluster-based inference and local maxima extraction.

2.6.3. Classification of clinically-significant symptoms
Finally, the ability of these anatomical data to predict ‘cases’ at each interval was tested with a classifier employing an ensemble of weak-learners. In this Gradient Boosting method (Friedman, 2002), the cohort was split into a discrete training (75%) and validation (25%) set, employing 1000 trees with 10-fold cross-validation (R package \texttt{gbm}).

3. Results
3.1. Participant characteristics
The sample was predominantly male (18 females), predominantly middle-aged (mean age 61.7) with most patients being classified as the ‘Akinetic-Rigid’ (37.5%) or ‘Mixed’ phenotype (43.8%). Most patients
had moderate motor symptoms at baseline despite being ‘ON’ medication during UPDRS assessment. Summary statistics for these and selected neuropsychiatric variables are presented in Table 2.

### 3.2. Subthalamic stimulation

Analysis of the concatenated VAT volumes for the cohort demonstrated that the highest probability of stimulation was in the dorso-lateral (motor) aspect of the STN, consistent with atlas-defined boundaries (Fig. 3). The only variable demonstrating a significant change between assessment intervals was an expected increase in stimulation amplitude (right hemisphere: 95% CI 0.17–0.58, \( p = 0.0005 \), left hemisphere: 95% CI 0.17–0.60, \( p = 0.0005 \); Supplementary Table 1). However, at each interval, some anatomical variables were significantly different between right and left hemispheres. At 6- and 13-weeks the active contact was closer to the centroid of the right associative subregion (6-weeks: 95% CI 0.53–1.40 mm, \( p = 3.09 \times 10^{-4} \), 13-weeks: 95% CI 0.34–1.17 mm, \( p = 0.006 \)), the nearest voxel of the right associative subregion (6-weeks: 95% CI 0.30–1.05 mm, \( p = 0.0016 \), 13-weeks: 95% CI 0.14–0.91 mm, \( p = 0.026 \)) and the VAT occupation of the associative subregion was greater in the right hemisphere (6-weeks: 95% CI 5.89–19.99%, \( p = 0.0016 \), 13-weeks: 95% CI 3.17–18.2%, \( p = 0.02 \)). At 6-weeks the active contact was also closer to the centroid of the right limbic subregion (95% CI 0.21–1.21 mm, \( p = 0.0016 \)).

### Table 1

Summary of the neuropsychiatric variables examined in this investigation; anatomical variables entered into the LASSO for regularisation and control variables entered into the general linear model.

| Neuropsychiatric outcomes | Anatomical variables | GLM co-variates |
|---------------------------|----------------------|----------------|
| Case versus non-case      | Distance to nearest voxel in each STN subregion: | Age |
| Patient-rated BIS         | Distance to centroid of each STN subregion: | Gender |
| Caregiver-rated BIS       | Overlap of simulated VAT within each STN subregion: | Hoehn & Yahr stage |
| Patient-rated EQ          | Ratio of simulated VAT: | Clinical subtype |
| Caregiver-rated EQ        | Right Associative:Motor STN | Years since diagnosis |
| QUIP-RS                   | Left Associative:Motor STN | LEDD |
| Including subscales (sex, gambling, eating etc.) | | |
| Beck depression inventory | | |
| Hayling test              | | |
| Including: category A & B errors, AB error score | | |
| ELF task                  | | |
| Rule violations           | | |

---

Fig. 2. Image processing pipeline for the identification of electrode contact points and volume activated tissue (VAT). A: MPRAGE, FLAIR and CT images were co-registered and manually checked for accuracy, with an accurately-sited electrode indexed in the right STN. B and C: Spatial normalisation into common template space using DARTEL. D: Coronal view of the active electrode contact (asterisk) with reference to the limbic (yellow), associative (blue) and motor (maroon) subregions of the STN. E: Axial view of both electrode trajectories with reference to the STN and its subregions. Within each STN limbic = yellow, associative = blue and motor = maroon subregions. F: Oblique sagittal view with simulated VAT in each hemisphere (red sphere). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Table 2
Demographic and clinical characteristics of the study cohort at baseline:

| Categorical variable | Total (n = 64) |
|----------------------|---------------|
| Gender               |               |
| Male                 | 46 (71.9)     |
| Female               | 18 (28.1)     |
| Clinical subtype     |               |
| Parkinson’s disease  | 24 (37.5)     |
| Mixed                | 28 (43.8)     |
| Tremor               | 12 (18.8)     |

Continuous variable | Mean (SD, median (range))
|-------------------|------------------|
| Age (years)       | 61.7 (± 9.3), 64 (35–76) |
| Hoehn & Yahr stage| 2.7 (± 0.5), 2.5 (1.5–4) |
| Years since diagnosis | 9.0 (± 5.2), 7 (1–23) |
| Levodopa equiv. daily dose | 1077.3 (± 543.1), 1008 (0–3450) |
| Patient-rated BIS  | 60.6 (± 8.3), 60 (43–87) |
| Caregiver-rated BIS| 59.4 (± 11.6), 59 (40–90) |
| Patient-rated EQ   | 42.6 (± 12.2), 42 (16–68) |
| Caregiver-rated EQ | 38.9 (± 14.0), 42 (11–64) |
| QUIP-RS total      | 20.5 (± 15.0), 18 (0–63) |
| Beck depression inventory | 11.1 (± 5.1), 11 (1–22) |
| Hayling AB error score | 10.1 (± 11.2), 6 (0–45) |
| ELF rule violations | 8.8 (± 5.4), 8 (0–24) |
| UPDRS part III motor | 37.0 (± 16.6), 36 (10–91) |

3.3. Neuropsychiatric outcomes

3.3.1. Assessment data

A repeated-measures ANOVA was employed to identify group-mean change amongst variables of interest across the chosen assessment intervals (baseline, 6-weeks and 13-weeks). As expected, there was a significant decrease in motor symptoms (F = 10.38, p = 3.17 × 10^{-4}) and dopaminergic medication use (F = 79.59, p = 2.4 × 10^{-15}) across the cohort. Controlling for age, gender, clinical subtype and dopaminergic medication use, there was a significant difference between groups (summarised in Supplementary Table 3). At the 6-week and 13-week assessments, there were no significant differences between cases and non-cases amongst the neuropsychiatric assessment data (summarised in Supplementary Tables 4 and 5). There were also no significant differences between groups in stimulation amplitude at either interval (left STN: p = 0.54 at 6 weeks and p = 0.60 at 13 weeks; right STN: p = 0.29 at 6-weeks and p = 0.93 at 13-weeks).

Anatomical variables with explanatory power were identified for caseness at both the 6-week and the 13-week assessments. These then also proved highly significant when controlling for clinical and demographic covariates. At 6-weeks, developing clinically-significant symptoms was significantly associated with stimulation at a closer distance to the nearest voxel of the right associative STN subregion (z = –3.0, p = 0.0026), as well as with greater VAT overlap within this subregion (z = 3.31, p = 0.0009). At 13-weeks, caseness was significantly associated with stimulation at a closer distance to the centroid of the right STN had the greatest explanatory power at both 6 and 13-weeks. At 6-weeks, this anatomical variable was the dominant factor in classification. At 13-weeks, the distance of the active electrode contact to the right and left associative subregions were also important factors, similar to the data presented in Table 4. Using iterative cross-validation the accuracy of the model at correctly classifying cases at 6-weeks was 79% significantly associated with ELF inhibitory errors. At both 6- and 13-weeks, clusters of subthalamic voxels in the right motor, as well as in the associative subregion, were significantly associated with an increase in inhibitory failure (p < 0.05, FWE-corrected) (Fig. 4). Cluster statistics are presented in Supplementary Table 7. Thresholded statistical maps (extracted with FSL randomise) corresponding to the significant cluster-based inferences of change in inhibitory errors are provided for download.

3.3.2. Clinical cases

A total of 26 participants were identified as cases (i.e. developed clinically-significant, stimulation-dependent changes in mood, affect and behaviour) during postoperative follow up. The characteristics of cases and non-cases were compared at baseline, 6-weeks and 13-weeks (Supplementary Tables 3–5). At baseline, in all demographic, disease-related and neuropsychiatric variables, there was no significant difference between groups (summarised in Supplementary Table 3). At the 6-week and 13-week assessments, there were no significant differences between cases and non-cases amongst the neuropsychiatric assessment data (summarised in Supplementary Tables 4 and 5). There were also no significant differences between groups in stimulation amplitude at either interval (left STN: p = 0.54 at 6 weeks and p = 0.60 at 13 weeks; right STN: p = 0.29 at 6-weeks and p = 0.93 at 13-weeks).

Clinical details of all identified cases are presented in Supplementary Table 6, including the stimulation manipulation undertaken to remit symptoms.

For each case, overlap of the left and right associative subregion by the simulated VAT was calculated pre- and post-intervention and tested for significant differences. This was repeated for the distance of the active electrode contact to the centroid of the associative subregion. Neuropsychiatric cases with active symptoms had a significantly greater volume of the right associative subthalamic subregion occupied by the simulated VAT than after remission of symptoms (95% CI 8.98–32.48%, t = 3.54, p = 0.00087). Likewise, neuropsychiatric cases with active symptoms had an active electrode contact significantly closer to the centroid of the right associative subthalamic subregion than after stimulation manipulation and remission of symptoms (95% CI 0.045–1.24 mm, t = 2.17, p = 0.036), although this finding did not survive FDR correction. Interestingly, these findings were not replicated in the left STN (p = 0.15 for VAT overlap of the left associative subregion and p = 0.28 for distance to the left associative centroid).

3.4. Predicting caseness from subthalamic data

A classifier applied to a training subset of the anatomical data showed broad agreement with the data acquired from the variable selection algorithm. In this classifier (Supplementary Fig. 2a and b), the overlap of the simulated VAT with the associative subregion of the right STN had the greatest explanatory power at both 6 and 13-weeks. At 6-weeks, this anatomical variable was the dominant factor in classification. At 13-weeks, the distance of the active electrode contact to the right and left associative subregions were also important factors, similar to the data presented in Table 4. Using iterative cross-validation the accuracy of the model at correctly classifying cases at 6-weeks was 79%
95% confidence interval 54.4–94.0, sensitivity 89%, specificity 70%, positive predictive value 73%, negative predictive value 89%) and at 13-weeks was 79% (95% confidence interval 54.4–94.0, sensitivity 71%, specificity 83%, positive predictive value 71%, negative predictive value 83%).

The spatial distribution of STN voxels significantly associated with the likelihood of being a ‘case’ versus a ‘non-case’ were determined with a voxel-based analysis. The highest likelihood of being a ‘case’ and a ‘non-case’ was identified using the predicted probabilities of caseness in the machine-learning classifier trained on the whole dataset at 13-weeks (accuracy 88%, 95% confidence interval 76.9–94.5, sensitivity 90%, specificity 86%, positive predictive value 75%, negative predictive value 95%). Distinct clusters of voxels were identified in the dorsolateral STN corresponding to the highest likelihood of being a ‘non-case’, with clusters of voxels in the ventromedial STN corresponding with the highest likelihood of being a ‘case’ (Fig. 5). Cluster Statistics are presented in Supplementary Table 7. Thresholded statistical maps (extracted with FSL randomise) corresponding to the significant cluster-based inferences of change in neuropsychiatric outcomes are provided for download.

3.5. Auxiliary analyses

An additional analysis was carried out using scan group as an additional covariate for the significant findings detailed above. With the exception of ELF Rule Violations at 6-weeks post-DBS, all findings

---

**Table 3**

Modelling of anatomical variables relating to neuropsychiatric assessment data.

For each participant, standardised (z-scored) change values from baseline were calculated for each neuropsychiatric variable. These change scores were entered into a variable selection and optimisation algorithm to identify anatomical variables of interest. These anatomical variables were then tested in a general linear model, controlling for relevant disease and demographic factors.

| ELF rule violations | Anatomical variables after optimisation | LASSO coefficients (fitted) | General linear model† |
|---------------------|----------------------------------------|-----------------------------|-----------------------|
| Week 6 | VAT overlap: right associative STN | Intercept = −0.65, Coefficient = 0.0046 | t-stat = 2.349, p-value = 0.021* |
| Week 13 | VAT overlap: right associative STN | Intercept = −0.087, Coefficient = 0.0030 | t-stat = 3.30, p-value = 0.0017** |

Significance codes: ‘***’ p < 0.001 ‘**’ p < 0.01 ‘*’ p < 0.05.
* Controlled for age, gender, LEDD, clinical subtype, years since diagnosis.

(95% confidence interval 54.4–94.0, sensitivity 89%, specificity 70%, positive predictive value 73%, negative predictive value 89%) and at 13-weeks was 79% (95% confidence interval 54.4–94.0, sensitivity 71%, specificity 83%, positive predictive value 71%, negative predictive value 83%).
remained significant (Supplementary Table 8), indicating that these findings are replicated across scanning conditions. Furthermore, we repeated the main analyses using the recently released DISTAL subthalamic atlas, which is precisely co-registered to the ICBM nonlinear asymmetric MNI space and includes limbic, associative and motor subregions. Our findings were replicated in this atlas (Supplementary Table 9).

4. Discussion

We demonstrate the significance of the locus of subthalamic stimulation in the genesis of postoperative inhibitory dysfunction. We found that the spatial overlap between the inferred stimulation volume and the associative subregion of the right STN was significantly associated with inhibitory errors during the Excluded Letter Fluency (ELF) task. Rule violations in this task are a sensitive measure of inhibitory dysfunction in non-demented persons with Parkinson’s disease and are considered to represent goal-directed, selective inhibitory control. It may be a more sensitive measure of disinhibition than the Hayling sentence completion task in this population (O’Callaghan et al., 2013). We found a cluster of subthalamic voxels in the right STN associated with greater ELF rule-violations at 6 and 13-weeks post-DBS. The inclusion of motor and associative zones in these clusters emphasises that the partitioning of STN subregions is not strictly delineated and is more likely to be represented by a gradient of sensorimotor-cognitive cortical connectivity.

We also demonstrate that clinically-significant, stimulation-dependent neuropsychiatric symptoms (changes in mood, affect and behaviour from baseline, as assessed by a psychiatrist, responding immediately to stimulation manipulation) are also associated with the spatial overlap of the simulated VAT in the right associative STN, as well as the distance of the active contact from the nearest voxel of the associative subregion, and the centroid of the right and left associative subregions. We identify distinct clusters of voxels associated with high and low likelihood of stimulation-dependent neuropsychiatric symptoms and illustrate that postoperative ‘cases’ can be predicted with reasonable accuracy from these methods.

Our investigation adds to the literature in the following ways. By using a larger sample size than many recent investigations, we were able to infer more accurately on the contribution of subthalamic stimulation to neuropsychiatric outcome. Secondly, our neuropsychiatric evaluation pre- and post-DBS was extensive and allowed us to discriminate between clinically-significant syndromes (such as mania and hypomania) and subclinical changes in dimensional constructs (such as

Fig. 4. Clusters of subthalamic voxels in the right associative and motor subregions are associated with inhibitory deficits 6- and 13-weeks after subthalamic DBS. Within each STN limbic = yellow, associative = blue and motor = maroon subregions. At 6-weeks, a small cluster of FWE-corrected subthalamic voxels significantly associated with ELF rule violations is observed in the right motor subregion (pink). At 13-weeks this cluster is larger and predominates in the right associative subregion, extending into the right motor subregion (blue). A: Axial view at 6-weeks, B: oblique view at 13-weeks, C: axial view at 13-weeks, D: coronal view at 13-weeks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
general linear model, controlling for relevant disease and demographic factors. These anatomical variables were then tested in a variable selection and optimisation algorithm to identify anatomical variables relating to clinical caseness after subthalamic DBS for Parkinson’s disease. Within each STN limbic = yellow, associative = blue and motor = maroon subregions. Green: a cluster of FWE-corrected voxels significantly associated with being a non-case can be identified in the dorsolateral aspect of the right STN. Red: a cluster of FWE-corrected voxels significantly associated with being a case are identified in the ventromedial aspect of both the right and left STN. Significant clusters are found in both the right and left STN, corresponding with the known anatomy of this nucleus, with motor representations in the dorsolateral aspect of this nucleus and cognitive-associative circuits in the ventromedial region.

We employ a simplified method for estimating each VAT, which fails to account for tissue inhomogeneity and the biophysics of axonal response to DBS (Gunalan et al., 2017). However, the methods we employ benefit from being embedded in open-source rather than proprietary software and are not computationally demanding to implement.

The use of two different MRI scanners for image acquisition introduces a further variable, but the acquisition protocol was not altered and we present key data from our analyses controlling for scan group in Supplementary Table 7. Overall, the correlation of anatomical variables with inhibitory deficits and neuropsychiatric ‘caseness’, derived from the general linear model, remains highly significant despite this additional covariate. The sole exception is ELF rule violations at 6-weeks, which is no longer significantly correlated with VAT overlap within the right associative subthalamic subregion. This was the finding of weakest significance in the original analysis. However, that the remaining results hold despite the addition of scan group to the model, arguably increases the generalisability of our findings.

Table 4

| Clinical cases | Anatomical variables after optimisation | LASSO coefficients (fitted) | General linear model |
|----------------|----------------------------------------|-----------------------------|---------------------|
| Week 6         | Distance: right associative STN voxel  | Intercept = −0.72           | z-value = −3.00 p = 0.0026*** |
|                | VAT overlap: right associative STN      | Coefficient = −0.11          | p-value = 0.0026*** |
| Week 13        | Distance: right associative STN centroid| Intercept = 0.19            | z-value = −3.00 p = 0.0027** |
|                | VAT overlap: right associative STN      | Coefficient = −0.29          | p-value = 0.0084**   |

Significance codes: ***p < 0.001 **p < 0.01 *p < 0.05.

Controlled for age, gender, LEDD, clinical subtype, years since diagnosis.

![Fig. 5. Distinct clusters of subthalamic voxels are associated with high and low likelihood of developing clinically-significant neuropsychiatric symptoms after STN-DBS for Parkinson’s disease. Within each STN limbic = yellow, associative = blue and motor = maroon subregions. Green: a cluster of FWE-corrected voxels significantly associated with being a non-case can be identified in the dorsolateral aspect of the right STN. Red: a cluster of FWE-corrected voxels significantly associated with being a case are identified in the ventromedial aspect of both the right and left STN. Significant clusters are found in both the right and left STN, corresponding with the known anatomy of this nucleus, with motor representations in the dorsolateral aspect of this nucleus and cognitive-associative circuits in the ventromedial region. A: Axial view, B: Coronal view, C: Oblique view. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image-url)
Recently, several groups have employed diffusion imaging to characterise the connectivity profile of clinically-effective subthalamic stimulation (Accolla et al., 2016; Akram et al., 2017; Horn et al., 2017; Vanegas-Arroyave et al., 2016), although neuropsychiatric symptoms have not been formally examined in these approaches. Tractographic methods are able to incorporate white matter tracts adjacent to the STN, which may contribute to the balance of therapeutic versus adverse effects. Models of neural networks associated with favourable clinical outcomes offer insights into the mechanism of DBS and may, in the future, assist clinicians with selection of the optimal electrode contact and guide stimulation titration. Future work will evaluate if diffusion imaging surpasses other forms of data when predicting the evolution of postoperative neuropsychiatric symptoms.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2018.03.009.

Acknowledgements

The authors gratefully acknowledge the commitment of participants and caregivers who contributed their time to this study. The authors acknowledge the ongoing support of St Andrew's War Memorial Hospital and the Herston Imaging Research Facility. The authors thank the Lead-DBS developers for providing an open-source platform for the analysis of deep brain stimulation data.

Financial disclosures/conflicts of interest

All authors report no conflict of interest.

Funding agencies

Dr Mosley was supported by an early career fellowship from the Queensland government's 'Advance Queensland' initiative, a Royal Brisbane & Women’s Hospital Foundation Research Grant and a young investigator grant from the Royal Australian and New Zealand College of Psychiatrists. He received an unrestricted educational grant from Medtronic.

Ethics approval

Prior to the commencement of data collection, the full protocol was approved by the Human Research Ethics Committees of the Royal Brisbane & Women’s Hospital, the University of Queensland, the QIMR Berghofer Medical Research Institute and UnitingCare Health. All participants and caregivers gave written consent to participate in the study.

Author roles

Mosley: Conception and design of study, data collection, writing the first draft of the manuscript, image processing and analysis, statistical analysis.

Smith: Statistical supervision.

Coyne: Supervision of data collection, critical comments on manuscript.

Silburn: Supervision of data collection, critical comments on manuscript.

Breakspear: Supervision of study design, analytical methods and revision of manuscript.

Perry: Design and supervision of image processing and analysis, statistical analysis, revision of manuscript.

References

Accolla, E.A., Dukart, J., Helms, G., Weiskopf, N., Kherif, F., Lutti, A., et al., 2014. Brain tissue properties differentiate between motor and limbic basal ganglia circuits. Hum. Brain Mapp. 35, 5083–5092.

Accolla, E.A., Herrojo Ruiz, M., Horn, A., Schneider, G.H., Schmitz-Hubsch, T., Draganski, B., et al., 2016. Brain networks modulated by subthalamic nucleus deep brain stimulation. Brain 139, 2503–2515.
Acram, H., Sotiropoulos, S.N., Jbabdi, S., Georgiev, D., Mahlknecht, P., Hyam, J., et al., 2017. Subthalamic deep brain stimulation sweet spots and hyperdirect cortical connectivity in Parkinson’s disease. NeuroImage 158, 332–345.

Alkemade, A., Forstmann, B.U., 2014. Do we need to revise the tripartite subdivision hypothesis of the subthalamic nucleus (STN)? NeuroImage 95, 526–329.

Amunts, K., Lepage, C., Borgarat, L., Mohlig, H., Dickscheid, T., Rousseau, M.E., Bludau, S., Bänz, P.L., Lewis, L.B., Oro-Peusquens, A.M., Shah, N.J., LipPERT, T., Zilles, K., Evans, A.C., 2013. BigBrain: An ultra-high-resolution 3D human brain model. Science 340 (6139), 1427–1429.

Appleby, B.S., Duggan, P.S., Regenberg, A., Rabins, P.V., 2007. Psychiatric and neuropsychiatric adverse events associated with deep brain stimulation: a meta-analysis of ten years’ experience. Mov. Disord. 22, 1723–1728.

Aron, A.R., Robbins, T.W., Poldrack, R.A., 2004. Inhibition and the right inferior frontal cortex. Trends Cogn. Sci. 8, 170–177.

Aron, A.R., Behrens, T.E., Smith, S., Frank, M.J., Poldrack, R.A., 2007. Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging and single-voxel functional MRI. J. Neurosci. 27, 3743–3754.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38, 95–113.

Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. J. Autism Dev. Disord. 34, 163–175.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. Arch. Gen. Psychiatry 4, 561–571.

Benjaminji, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. Ser. B Methodol. 57, 289–300.

Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. J. Autism Dev. Disord. 34, 163–175.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. Arch. Gen. Psychiatry 4, 561–571.

Benjaminji, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. Ser. B Methodol. 57, 289–300.

Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. J. Autism Dev. Disord. 34, 163–175.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. Arch. Gen. Psychiatry 4, 561–571.

Benjaminji, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. Ser. B Methodol. 57, 289–300.

Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. J. Autism Dev. Disord. 34, 163–175.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. Arch. Gen. Psychiatry 4, 561–571.

Benjaminji, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. Ser. B Methodol. 57, 289–300.

Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. J. Autism Dev. Disord. 34, 163–175.