Cognitive functioning after subthalamic nucleotomy for refractory Parkinson’s disease

Renee J McCarter, Nigel H Walton, Alison F Rowan, Stephen S Gill, Marina Palomo

Abstract
Objective—To evaluate whether subthalamic nucleotomy produces adverse cognitive effects in patients with Parkinson’s disease.

Method—Twelve patients with Parkinson’s disease underwent stereotactic surgery to the subthalamic nucleus. Presurgical and postsurgical neuropsychological assessment of attention, memory, executive function, language, and verbal intellect were undertaken with a battery of tests designed to minimise potential contamination of cognitive effects by motor symptoms.

Results—There was no statistically significant difference in the cognitive tests results after operation for the group as a whole. Reliable change indexes were generated for the cognitive tests. Reliable change postoperatively was found on specific tests of verbal memory, attention, and planning. Left sided operations were associated with greater incidence of deterioration post-surgery.

Conclusions—Preliminary data on the first reported cognitive changes after subthalamic nucleotomy suggested few adverse cognitive effects of the surgery although discrete neuropsychological changes were seen in some patients. These effects were consistent with current theories on the cognitive functions of the basal ganglia.

Keywords: subthalamic nucleotomy; neuropsychology; Parkinson’s disease

Motor symptoms in refractory Parkinson’s disease, either due to the disease itself or to the effects of levodopa treatment, include tremor, rigidity, and dyskinesia. Recent advances in techniques of stereotactic surgery and microelectrode stimulation of the basal ganglia have led to a resurgence in the surgical treatment of refractory Parkinson’s disease. Ventrolateral thalamotomy and posteroventral pallidotomy procedures have been reported to be of benefit, with neuropsychological dysfunction but a significant proportion do show specific or even global cognitive and memory impairment in addition to motor symptoms. The aim of surgical ablation or electrical stimulation of the basal ganglia is the relief of motor symptoms, but such procedures have the potential to produce additional neuropsychological impairments as a consequence of disruption to the corticostriatal-thalamocortical loops.

In early studies, thalamotomy and pallidotomy in Parkinson’s disease were associated with neuropsychological morbidity. More recent studies of ablation and stimulation of the globus pallidus or thalami have demonstrated variable outcomes with some reports of improved cognitive function and some reports of modest decline on specific neuropsychological measures, often related to the side of the lesion. In many reports there was either no significant effect of surgery or the decline in individual test scores was reported as being functionally insignificant. It is tempting to conclude that pallidotomy and pallidal stimulation for the relief of motor symptoms in refractory Parkinson’s disease...
Cognition after subthalamic nucleotomy for refractory Parkinson’s disease

have no or very few significant adverse effects on cognitive functioning for most patients. There are, however, individual patients who show significant cognitive morbidity after such surgery, which points to a degree of heterogeneity of outcome, possibly related to the accuracy of localisation of lesions in individual cases. Recent studies using volumetric MRI have found that lesion location, along the anteromedial to posterolateral axis within the globus pallidus interna, influences the outcome of surgery. Gross et al reported that placement of the lesion had a differential effect on relief of dyskinesia, rigidity, and gait problems in their group of 31 patients. They suggested that this was due to the presence of separate, parallel motor circuits within the globus pallidus and that symptom relief may be crucially related to precise lesion placement. Variability in outcome after pallidotomy may be partly due to location of lesions. If specific motor functions and symptoms of the disease may be differentially affected by the precise location of the lesion within the globus pallidus, then it may also be the case that specific cognitive effects might result from discrete lesions. Certainly, hemisphere specific cognitive changes have been reported after pallidotomy, suggesting the possibility of specific neuropsychological impairment as a direct consequence of the lesion.

Most recent studies have been concerned with the outcome of ablation or stimulation of the globus pallidus interna. Typically, patient numbers have been small. To our knowledge, there has been no systematic study of the effect of subthalamic nucleotomy on neuropsychological functioning in patients with Parkinson’s disease.

The current study therefore reports the first data on the cognitive outcome of subthalamic nucleotomy for refractory Parkinson’s disease. Patients were assessed just before surgery and a minimum of 3.5 months after surgery on a battery of neuropsychological tests designed to evaluate verbal and visual memory, aspects of attention and information processing, language, intellect, and executive or “frontal” functions. The tests were chosen specifically to minimise contamination from motor symptoms and to be clinically practicable for administration in this group of patients while providing a comprehensive assessment of higher cognitive function. Tests which tapped executive components of attention and memory were important to include in view of the potential for subthalamic nucleotomy to negatively impinge on frontostriatal or mammillothalamic neural circuits.

Methods

Participants

Participants in this study were 12 patients with refractory ideopathic Parkinson’s disease who were admitted to the Frenchay Hospital, Bristol, UK for stereotactic subthalamic nucleotomy (STN). These patients represented a subgroup of a larger sample of patients referred to our centre for stereotactic surgery for Parkinson’s disease, most of whom were initially considered for pallidotomy and were being investigated according to a standard pallidotomy trial protocol. Only those patients clearly identified as candidates for STN or subthalamic nucleus stimulation before referral for neuropsychological assessment are included in this study. These represent a consecutive series of 12 STN candidates who were able to complete most of the test battery before operation and to be reviewed less than 12 months after operation. Informed consent was obtained from all participants. All patients had disabling motor functioning despite optimal medication management. Patients were only included in the study if assessment by the consultant clinical neuropsychologist and consultant neurologist excluded a dementia or depressive illness. The patients were eight men and four women with an age range of 42 to 68 years at the time of operation. The mean duration of Parkinson’s disease was 16.4 years with a range of 10 to 24 years. Two patients underwent a left sided STN, seven underwent a right sided STN, and three had bilateral procedures. Four of the patients with right sided STN also had a left subthalamic nucleus stimulator implanted concurrently.

Surgical Procedure

The subthalamic nucleus was localised with high resolution MRI T2 sequences (1.5 Tesla TR 2,500, TE 150, TSE 11, NSA 12) and peroperative macrostimulation. Under general anaesthesia a modified Leksell stereotactic frame was fitted parallel to the orbitomeatal plane. The anterior-posterior commissural (AC-PC) plane was identified on a midsagittal planing scan. Axial images 2 mm thick were obtained parallel to the AC-PC plane and coronal images orthogonal to those were obtained. Boundaries of the subthalamic nucleus were then defined by reference to the magnified MRI and Schaltenbrand atlas. The dorsolateral portion of the subthalamic nucleus was the target of the stereotactic coordinates. At the time of surgery, all patients were awake and in an “off” state, their antiparkinsonian medications having been stopped 24 hours earlier. A 1.24 mm diameter electrode with a 2 mm exposed tip was guided to the dorsolateral subthalamic nucleus. The target was stimulated at 100 Hz, 1ms pulse width, and between 0.75 and 2 V during which time change in tremor, rigidity, and bradykinesia was monitored. Probe position was adjusted to gain maximal clinical improvement without development of side effects. At the optimal position one or two radiofrequency lesions were made, typically at 80°C for 60 seconds. Postoperatively a high resolution MR scan was obtained to confirm lesion position and a course of dexamethasone for 24–48 hours was administered to reduce cerebral oedema. Antiparkinsonian medication was gradually reintroduced.

Neuropsychological Evaluation

Test Battery

A comprehensive battery of tests was administered to obtain data on various cognitive and memory functions with as little contamination from motor dysfunction as possible. The test...
battery had to be practical for administration during a 2 hour window between admission, motor assessment, and cessation of medication. Even so, occasionally a test could not be completed due to the onset of akinesia or dyskinesia which affected speech or motor responses. We administered the national adult reading test–2 (NART)\textsuperscript{40} as an estimate of premorbid functioning. The following tests were administered both before surgery and at follow up: a short form of the verbal scale of the Wechsler adult intelligence scale—revised (WAIS-R)\textsuperscript{39} to measure general verbal intelligence; The *Rey auditory verbal learning test (RAVLT)\textsuperscript{42} to measure complex verbal learning (trials 1–5) and retention (trial 6); the *Rivermead behavioural memory test story memory (immediate and delayed conditions)\textsuperscript{45} to measure contextual verbal learning and retention; the *release from proactive inhibition procedure to measure verbal encoding and executive aspects of verbal memory; the *Rivermead behavioural memory test facial recognition\textsuperscript{56} to measure visual memory; the attention/concentration index from the Wechsler memory scale-revised (WMS-R A/C)\textsuperscript{66} to measure simple auditory and visual attention; the Stroop colour-word task\textsuperscript{50} to measure complex focused and sustained attention and inhibition; the paced serial addition test (PASAT 2.0 second pacing)\textsuperscript{66} to measure complex divided attention; the *controlled oral word association test (COWAT)\textsuperscript{67} to measure phonemic verbal fluency; and the five point test\textsuperscript{68} to measure design fluency. We adapted one commonly used “executive” test of spatial working memory and planning, the Tower of London test,\textsuperscript{59} so that demands on motor dexterity were minimised and we refer to this test as the Burden balls. The Tower of London test was originally devised as a test of planning and required patients to rearrange three coloured balls on three pegs of varying heights to match a presented “goal”. Moves are governed by a set of simple rules. There are different standardisations of the task and it has been adapted for computer presentation. We wanted to preserve the manual mode of presentation but to minimise confounding effects of the movement disorders of our patients. To this end, the balls and pegs were replaced with two dimensional discs and slots so that the patients could simply slide the discs from one position to another with minimal demands on fine motor dexterity. The standard instructions, items, and scoring method developed by Anderson et al\textsuperscript{69} were employed. Total time taken to complete all items and number of correct items were recorded before and after surgery for each patient. In an effort to measure complex divided attention, we utilised the paced auditory serial addition task but chose only to administer one trial of 60 items for comparison before and after surgery. Our past experience with patients with Parkinson’s disease had shown that the 2.0 second pacing was appropriate for most patients whereas slow speech and dysarthria contaminates the results for the faster 1.6 and 1.2 second pacings. Where possible, alternate forms of tests were used at pre-

surgical and postsurgical assessments to minimise practice effects. An asterisk (*) denotes which tests were given in alternate versions. The test batteries containing alternate versions of the tests were coded A and B and were administered to the patients in counterbalanced fashion.

Preoperative assessment
All patients were assessed during their hospital admission 2 days before the STN procedure. The patients were tested during an “on” phase and while taking their usual medication.

Postoperative assessment
All patients were reassessed on the test battery when they returned to the hospital for an inpatient review by the team. This occurred a minimum of 3.5 months after surgery with the longest follow up time being 12 months. The vagaries of clinical and hospital practice were responsible for the variation in follow up periods. The mean time after surgery was 6.4 months. This period is considered sufficient to minimise potential inflation of scores postsurgery due to practice effects on those tests not available in alternate form. A period of greater than 12 months after surgery may have resulted in decrements in test performance due to progression of the disease process rather than disclosing potential deficits due to surgery. Again, all patients were tested “on” and while taking their usual medication. We were not able to control for alterations in medication after surgery as a consequence of the surgical outcome. All stimulators were functioning at the time of assessment.

Results
The estimated premorbid IQ of this group was relatively high with a mean NART-2 predicted IQ of 115 (range 106–27). Patients showed deficits, compared to published normative data, on several tests before surgery. Such a finding was expected in this patient group. The objective of our study was to identify any additional neuropsychological impairment that might reflect an effect of surgery. The data were analysed using tests of statistical significance and a reliable change index.

The mean presurgical and postsurgical test results for the group as a whole are presented in table 1. There were significant differences at p<0.05 for three tests but the Bonferroni adjustment for multiple comparisons set the significance level at α=0.004. No paired comparison was significant at this level. Two tests, facial recognition and release from proactive inhibition, were not analysed using the paired $t$ test. The scores for the facial recognition memory test were not normally distributed and were subject to non-parametric Wilcoxon signed ranks tests. There was no significant difference before and after surgery for the facial recognition test. The release from proactive inhibition test was scored as normal or abnormal. Eleven patients showed normal performance preoperatively and only six showed normal performance postoperatively.

Laurent et al. showed normal performance in 11 patients (41%) and abnormal performance in 10 patients (59%].

Preliminary results indicate that surgery is associated with improvement in objective measures of motor performance as measured by the UPDRS in 10 patients (41%).

The estimated premorbid IQ of this group was relatively high with a mean NART-2 predicted IQ of 115 (range 106–27). Patients showed deficits, compared to published normative data, on several tests before surgery. Such a finding was expected in this patient group. The objective of our study was to identify any additional neuropsychological impairment that might reflect an effect of surgery. The data were analysed using tests of statistical significance and a reliable change index.

The mean presurgical and postsurgical test results for the group as a whole are presented in table 1. There were significant differences at $p<0.05$ for three tests but the Bonferroni adjustment for multiple comparisons set the significance level at $\alpha=0.004$. No paired comparison was significant at this level. Two tests, facial recognition and release from proactive inhibition, were not analysed using the paired $t$ test. The scores for the facial recognition memory test were not normally distributed and were subject to non-parametric Wilcoxon signed ranks tests. There was no significant difference before and after surgery for the facial recognition test. The release from proactive inhibition test was scored as normal or abnormal. Eleven patients showed normal performance preoperatively and only six showed normal performance postoperatively.
In this study a priori assumptions about the effect of surgery on cognitive performance could not be confidently made, hence the use of many measures. Although the Bonferroni adjustment guards against the risk of capitalising on chance differences when making multiple comparisons in repeat testing, it may also increase the risk of a type 2 error—namely, accepting the null hypothesis when it is in fact false. Further, reliance solely on statistically non-significant results may be clinically meaningless although smaller samples may produce statistically non-significant results but show large effect sizes, or magnitudes of discriminability which are clinically meaningful. Previous reports of outcome after surgery in Parkinson’s disease have suggested heterogeneity with some patients scoring worse on specific measures. The patients with unilateral lesions, or right STN plus additional left lesions, or right STN plus additional left subthalamic nucleus ablations showed more robust and clinically meaningful change as compared to patients with bilateral lesions.

The reliable change index is calculated using the SD of the control group, the normal population or pretreatment experiment group.

In this study, we used the published normative data for each test where available, using the relevant age and education subgroups. When these were not available we used the preoperative group means and SD for our patients (Burden balls, five point design fluency). The release from proactive inhibition data is based on the difference between the initial and retest scores for the normative subject sample.

Where $x_1$ is the participant’s first test score, $x_2$ is their retest score, and $S_{\text{diff}}$ is the standard error of difference:

$$RCI = \frac{x_2 - x_1}{S_{\text{diff}}}$$

The $S_{\text{diff}}$ (standard error of measurement) is calculated using the SD of the control group, the normal population or pretreatment experiment group.

The release from proactive inhibition data is based on the difference between the initial and retest scores for the normative subject sample.

In this study, we used the published normative data for each test where available, using the relevant age and education subgroups. When these were not available we used the preoperative group means and SD for our patients (Burden balls, five point design fluency). The release from proactive inhibition data is based on the difference between the initial and retest scores for the normative subject sample.

Where $x_1$ is the participant’s first test score, $x_2$ is their retest score, and $S_{\text{diff}}$ is the standard error of difference:

$$S_{\text{diff}} = \sqrt{2(S_E)}$$

The $S_E$ (standard error of measurement) is calculated using the SD of the control group, the normal population or pretreatment experiment group.

In this study, we used the published normative data for each test where available, using the relevant age and education subgroups. When these were not available we used the preoperative group means and SD for our patients (Burden balls, five point design fluency). The release from proactive inhibition data is based on the difference between the initial and retest scores for the normative subject sample.

Where $x_1$ is the participant’s first test score, $x_2$ is their retest score, and $S_{\text{diff}}$ is the standard error of difference:

$$S_{\text{diff}} = \sqrt{2(S_E)}$$

The $S_E$ (standard error of measurement) is calculated using the SD of the control group, the normal population or pretreatment experiment group.

In this study, we used the published normative data for each test where available, using the relevant age and education subgroups. When these were not available we used the preoperative group means and SD for our patients (Burden balls, five point design fluency). The release from proactive inhibition data is based on the difference between the initial and retest scores for the normative subject sample.
stimulation. Additionally, the individual tests showed a theoretically meaningful trend in their vulnerability to the procedure (table 3).

For the STROOP, test performance declined postoperatively for 30% of the patients. The RAVLT deteriorated in 20% to 30% of cases for the learning and retrieval trials, with no patients with unilateral right STN showing any decline. The Burden balls proved more difficult postoperatively for the left STN and bilateral STN groups with 100% and 33% decline respectively. This compares with no decline postoperatively for the patients with right STN and those with right STN plus left subthalamic stimulator. The only patient to deteriorate on the facial recognition test and the attention/concentration index had undergone a left STN.

The other tests did not show reliable deterioration for most patients. The Rivermead story memory—immediate and delayed recall actually improved for one third of the patients and seemed relatively impervious to STN.

Discussion
Subthalamic nucleotomy (STN) is a procedure for the relief of motor symptoms associated with Parkinson’s disease. There are few published series of outcome after subthalamic nucleotomy or subthalamic nucleus stimulation and, as far as we are aware, none which systematically examine the neuropsychological effects of such procedures. We presented data from a series of 12 patients with refractory Parkinson’s disease whom we comprehensively assessed before and after STN. Our results, although preliminary, suggested that STN did not result in global neuropsychological impairment relative to presurgical baseline. Using a conservative $\alpha$ level of 0.004, there were no statistically significant differences between the presurgical and postsurgical scores on a neuropsychological test battery which evaluated attention, memory, executive, language, and intellectual functions. The numbers of patients in this study undergoing different operative procedures and locations were too small to undertake detailed analyses of effect sizes but use of “reliable change” scores offered a conservative and robust measure of clinically meaningful change.54–56 Most measures of cognitive function in this study did not show reliable deterioration after surgery. This is not surprising given that the subthalamic nucleus has not previously been shown to play a direct role in cognition and the lesions made in our patients were small and meticulously placed using image guidance. However, there were some tests which did show deterioration for some patients.

Tests which did not show reliable deterioration in a significant number of patients included tests of verbal intelligence, contextual verbal memory, and fluency tasks. Measures of verbal intelligence tend to tap “crystallised” skills and this is thought to account for the relative resiliency of verbal IQ even after severe diffuse brain injury (see Lezak59 for review). Contextual memory, as measured by recall of short stories, is not generally held to be differentially sensitive to frontal lobe or subcortical damage, although it may be disrupted by temporal lobe damage. Previous studies of pallidotomy have not found story recall to be impaired after the procedure.32 33 The preservation of phonemic verbal fluency has been previously reported after posteroventral pallidotomy.34 Improvements in motor function as a result of the surgery could have improved performance on the five point design fluency task resulting in simple output efficiencies masquerading as improved cognitive function, or masking subtle deterioration at a cognitive level. No patients deteriorated on total number of designs produced. However, there were no deteriorations either in number of perseverative designs—potentially a more sensitive measure of frontal disinhibition or cognitive impairment—on this test.

For those tests which did show reliable deterioration, this occurred more often in those patients with left sided procedures. The patients with left ablation (left STN and bilateral STN) and to some extent the patients with right ablation with a working left sided subthalamic nucleus stimulator, deteriorated on more measures of planning, spatial working memory, verbal memory, and attention which are thought to have an executive component. The STROOP and RAVLT have previously been reported as sensitive to left hemispheric dysfunction (see Lezak and Spreen and Strauss60). The STROOP has not previously been found to be vulnerable to pallidotomy.32 33 Our finding of deterioration in the verbal list learning task (RAVLT) for patients with left sided procedures but not right sided ablations is consistent with the results for posteroventral pallidotomy reported by Trepanier et al.31 The thalamus is known to be involved in memory processes with the dorsomedial nucleus of the thalamus most often cited as important to memory integrity.61–63 although there have been case reports of amnestic deficit after damage to other thalamic nuclei.64 It is possible that surgical ablation of the subthalamic nucleus may be detrimental to memory functions either as a result of alterations in the excitation and inhibition of a more extensive corticothalamic or mamillothalamic circuit or as an iatrogenic effect of the surgical procedure. The Burden balls, an adaptation of the Tower of London test, deteriorated post-
operatively only for those patients with left ablations. This is a multicomponent test which is thought to tap spatial working memory, impulse control, verbal regulation of behaviour, and planning.\(^6\)\(^6\) The Tower of London test has been found to be sensitive to impairments in frontal neural circuits, particularly prefrontal, cingulate, and premotor cortices, although parietal and occipital cortices are also activated during the task.\(^6\) In patients with Parkinson’s disease, differential activation of the right globus pallidus interna compared with control subjects has been found during the Tower of London test, perhaps affecting the expression of prefrontal functions by disruption of the normal transmission of information through the frontostriatal circuitry.\(^6\) The release from proactive inhibition test deteriorated for a significant number of patients regardless of side of the lesion. This test is also thought to be reliant on frontal cortical circuits which may involve the basal ganglia, the thalamus in particular. Impaired release from proactive inhibition\(^6\) has been noted in amnesic disorders in which damage to the thalamus, mamillary bodies, and frontodiencephalic connections are present but not in amnesias where the primary damage is in the mesial temporal lobes.\(^6\)

We suggest, from these results, that subthalamic nucleotomy does not have global adverse cognitive consequences but it is possible that such procedures may produce discrete neuropsychological impairment in some patients. Aspects of verbal learning, attention, and planning, all of which are thought to involve an executive component and relate to frontostriatal neural circuits or frontothalamomammillary circuits, may prove vulnerable to disruption after STN due to interruption of complex information processing reliant on local circuits. Our data raise the possibility of differential vulnerability of the left and right hemispheres to this procedure. We acknowledge that the results of this study represent preliminary data on only a few patients and require replication in a larger study.

This patient group poses considerable difficulties for neuropsychological study. The motor symptoms are very debilitating, there is considerable variation in motivation and levels of fatigue in “on” and “off” states which can influence participation in cognitive assessment, and variations in general function as a consequence of medication schedules may produce variability in test performance within and across test sessions. The time after surgery was less than 12 months for our patients but it does remain possible that further progression of the Parkinson’s disease process could have produced subtle deterioration in some functions for some patients. We did not have a no surgery control group for comparison. Use of the reliable change index, although robust for normal patients and reported as reliable for clinical groups in other published studies,\(^6\)\(^6\) is affected by neuropsychological competence at baseline.\(^6\) Significantly impaired performance was found on some measures for our patients before surgery which may limit the applicability of reliable change scores to measurement of deterioration after surgery. Notwithstanding these limitations, our data suggest the possibility that operative procedures in the subthalamic nucleus may affect performance. Consideration of patient selection and decisions on the choice of ablations or stimulation may be informed by results of neuropsychological assessment and further study of the cognitive outcome of these procedures. The deficits seen in our patients were noted on complex and detailed neuropsychological assessment and correspond with current theories on the role of the basal ganglia in cognitive processes. Such deficits are not likely to be disclosed during routine neurological examination or by less complex cognitive assessments such as the Mini-Mental State examination. Informed consent for an operation which is designed to reduce disabling and distressing motor symptoms needs to take into account potential cognitive costs.

This work was conducted as part of the clinical management of patients with Parkinson’s disease at Frenchay Hospital, Bristol. We would like to acknowledge the assistance of Mr M Bunnage.


