Subthalamic nucleus versus globus pallidus bilateral deep brain stimulation for advanced Parkinson's disease (NSTAPS study): a randomised controlled trial

Vincent J J Odekerken, Teus van Laar, Michiel J Staal, Arne Mosch, Carel F E Hoffmann, Peter C G Nijssen, Guus N Beute, Jeroen P P van Vugt, Mathieu W P M Lenders, M Fiorella Contarino, Marieke S J Mink, Lo J Bour, Pepijn van den Munckhof, Ben A Schmand, Rob J de Haan, P Richard Schuurman, Rob M A de Bie

Summary

Background Patients with advanced Parkinson's disease often have rapid swings between mobility and immobility, and many respond unsatisfactorily to adjustments in pharmacological treatment. We assessed whether globus pallidus pars interna (GPi) deep brain stimulation (DBS) gives greater functional improvement than does subthalamic nucleus (STN) DBS.

Methods We recruited patients from five centres in the Netherlands who were aged 18 years or older, had idiopathic Parkinson's disease, and had, despite optimum pharmacological treatment, at least one of the following symptoms: severe response fluctuations, dyskinesias, painful dystonias, or bradykinesia. By use of a computer-generated randomisation sequence, we randomly assigned patients to receive either GPi DBS or STN DBS (1:1), applying a minimisation procedure according to drug use (levodopa equivalent dose <1000 mg *vs* ≥1000 mg) and treatment centre. Patients and study assessors (but not those who assessed adverse events) were masked to treatment allocation. We had two primary outcomes: functional health as measured by the weighted Academic Medical Center Linear Disability Scale (ALDS; weighted by time spent in the off phase and on phase) and a composite score for cognitive, mood, and behavioural effects up to 1 year after surgery. Secondary outcomes were symptom scales, activities of daily living scales, a quality-of-life questionnaire, the occurrence of adverse events, and drug use. We used the intention-to-treat principle for all analyses. This trial is registered with www.controlled-trials.com, number ISRCTN85542074.

Findings Between Feb 1, 2007, and March 29, 2011, we enrolled 128 patients, assigning 65 to GPi DBS and 63 to STN DBS. We found no statistically significant difference in either of our primary outcomes: mean change in weighted ALDS (3·0 [SD 14·5] in the GPi group *vs* 7·7 [23·2] in the STN group; p=0·28) and the number of patients with cognitive, mood, and behavioural side-effects (36 [58%] of 62 patients in the GPi group *vs* 35 [56%] of 63 patients in the STN group; p=0·94). Secondary outcomes showed larger improvements in off-drug phase in the STN group compared with the GPi group in the mean change in unified Parkinson's disease rating scale motor examination scores (20·3 [16·3] *vs* 11·4 [16·1]; p=0·03), the mean change in ALDS scores (20·3 [27·1] *vs* 11·8 [18·9]; p=0·04), and medication (mean levodopa equivalent drug reduction: 546 [SD 561] *vs* 208 [521]; p=0·01). We recorded no difference in the occurrence of adverse events between the two groups. Other secondary endpoints showed no difference between the groups.

Interpretation Although there was no difference in our primary outcomes, our findings suggest that STN could be the preferred target for DBS in patients with advanced Parkinson's disease.

Funding Stichting Internationaal Parkinson Fonds, Prinses Beatrix Fonds, and Parkinson Vereniging.

Introduction

Patients with advanced Parkinson's disease (PD) often show rapid, seemingly unpredictable swings between mobility (the on phase), usually with dyskinesias, and immobility (the off phase). Many of these patients respond unsatisfactorily to adjustments in pharmacological treatment.[1] Bilateral deep brain stimulation (DBS) of the subthalamic nucleus (STN) for advanced PD was first used in the 1990s.[2,3] The results of subsequent studies by different groups suggested that bilateral STN DBS reduces both PD motor symptoms and dyskinesias by about 50%.[4–6] The effectiveness of bilateral DBS of the globus pallidus pars interna (GPi) was also explored.[7,8] The results of non-randomised comparative studies suggested that bilateral GPi DBS was slightly less effective than STN DBS for the treatment of PD motor symptoms and was equally effective for the treatment of dyskinesias.[4,9] However, the STN was already thought by many to be the better target for DBS in patients with PD, which might have caused a major bias in these series.[10,11] The results of two randomised controlled trials that compared bilateral STN with GPi DBS suggested that the procedures were equally effective for PD motor symptoms and dyskinesias.[12,13] DBS-associated problems in cognitive, mood, and behavioural features seemed to occur more often in the STN groups.[10,12,14]

Lancet Neurol 2013; 12: 37–44

Published Online November 16, 2012 http://dx.doi.org/10.1016/S1474-4422(12)70264-8

See Comment page 25

Department of Neurology (V J J Odekerken MD, M F Contarino MD, M S J Mink MSc, L J Bour PhD, Prof B A Schmand PhD, R M A de Bie MD), **Department of Neurosurgery** (P van den Munckhof MD, P R Schuurman MD), **and Clinical Research Unit** (Prof R J de Haan PhD), **Academic Medical Center, Amsterdam, Netherlands; Department of Neurology** (T van Laar MD) **and Department of Neurosurgery** (M J Staal MD), **University Medical Center Groningen, Groningen, Netherlands; Department of Neurology** (A Mosch MD) **and Department of Neurosurgery** (C F E Hoffmann MD), **Haga Hospital, The Hague, Netherlands; Department of Neurology** (P C G Nijssen MD) **and Department of Neurosurgery** (G N Beute MD), **St Elisabeth Hospital, Tilburg, Netherlands; Department of Neurology** (J P P van Vugt MD) **and Department of Neurosurgery** (M W P M Lenders MD), **Medisch Spectrum Twente, Enschede, Netherlands; and Department of Psychology, University of Amsterdam, Netherlands** (Prof B A Schmand)

Correspondence to: Rob M A de Bie, Department of Neurology H2-213, Academic Medical Center, PO Box 22660, 1100 DD Amsterdam, Netherlands **r.m.debie@amc.uva.nl**

The Netherlands SubThalamic and Pallidal Stimulation (NSTAPS) study was initiated in 2007 to test the hypothesis that bilateral GPi DBS would produce greater improvement in disability than would bilateral STN DBS, assuming a lower rate of cognitive, mood, and behavioural complications, with similar improvement of motor symptoms.

By contrast with previous studies that investigated the effectiveness of DBS, we chose a generic disability scale as a primary outcome measure. This was because GPi DBS and STN DBS might have different effects on the various motor symptoms and because both procedures might be accompanied by cognitive and psychiatric adverse effects.[14,15] Cognitive status and mood might have an effect on self-reported quality of life, which could lead to interpretation issues with these scales.

Methods

Study design and participants

We recruited patients from five centres in the Netherlands that were experienced in doing DBS for PD. We included patients aged 18 years or older who had idiopathic PD and, despite optimum pharmacological treatment, at least one of the following symptoms: severe response fluctuations, dyskinesias, painful dystonias, or bradykinesia. We excluded patients if they had previous functional stereotactic neurosurgery, Hoehn and Yahr stage 5 at the best moment during the day, a Mattis dementia rating scale score of 120 or lower (out of 144), active psychosis, or contraindications for the neurosurgical procedure. Each site's medical ethics committee approved the study and patients provided written informed consent.

Randomisation and masking

By use of a computer-generated randomisation sequence, patients were randomly assigned to receive either GPi DBS or STN DBS in a one-to-one ratio at the clinical trial office of the Department of Neurology, Academic Medical Center (AMC, Amsterdam, Netherlands). Randomisation was done by trial nurses who had no further involvement in the study. We applied a minimisation procedure according to drug use (levodopa equivalent dose <1000 mg *vs* ≥1000 mg) and treatment centre. Patients and the clinical, neuropsychological, and psychiatric assessors were masked to treatment allocation. Patients regularly visited a non-masked neurologist at the outpatient clinic to adjust DBS settings together with adjustment of medication.

Procedures

All centres were experienced in DBS surgery (all surgeons were performing DBS surgery for at least 3 years at the start of the trial). The DBS surgery was done according to each centre's standard protocol. The final position of the electrode was determined on the basis of MRI, macro-electrode stimulation effects, and, in three of five centres, semi-micro-electrode recordings. During the course of the study, changes in drug treatment and DBS settings were allowed in both groups.

Baseline and 12-month assessments were done during standardised off-drug and on-drug phases. The off phase was defined as the condition of the patient after withholding antiparkinsonian drugs for 12 h overnight. The on phase was a patient's condition 1 h after a supra-threshold levodopa dose. Identical doses were used at baseline and the follow-up assessment. To analyse changes in medication and to calculate the supra-threshold levodopa dose, the different drugs were pooled in levodopa equivalent doses according to the following conversion formula:

(regular levodopa dose × 1) + (slow-release levodopa × 0·75 + (bromocriptine × 10) + (apomorphine × 10) + (ropinirole × 20) + (pergolide × 100) + (pramipexole × 100) and, if taking entacapone, + 0·2 × (regular levodopa dose + [slow-release levodopa × 0·75]).[16]

The 12-month assessment was done with the stimulators turned on. No changes in outcome measures were made after the start of the trial.

The primary outcomes were disability and the number of patients with a negative composite score of cognitive, mood, and behavioural effects. We assessed disability using a 26-item version of the AMC linear disability scale (ALDS). The ALDS is a linear and generic health scale to quantify functional status in terms of the ability to do basic (eg, self-care, eating, transfer) and complex (eg, household tasks, travelling, walking long distances) activities of daily living. ALDS scores range from 0 points to 100 points, with lower scores indicating more disability. The psychometric properties (reliability, validity, responsiveness, absence of ceiling effects) of the ALDS item databank have been extensively assessed and shown to be good.[17–21] The items chosen for this study fit the expected range of disability of the PD population. Because the ALDS is a continuous measure of disability, it is possible to recalculate the scores obtained in standardised off and on phases into a weighted score for time in either phases, which then represents disability throughout the day, instead of exclusively during episodes with or without medication.

At baseline and 12 months after surgery, patients completed a diary in which they rated every period of 30 min from 06·00 to 00·00 for 3 days. Patients were instructed to rate periods under the following categories: asleep, parkinsonism, on without dyskinesias, or on with dyskinesias. We calculated patients' ALDS, weighted by time spent in either on phase or off phase, according to the following formula:

weighted ALDS = off phase ALDS × (h in off phase/[h in on phase + h in off phase]) + on phase ALDS × (h in on phase/[h in on phase + h in off phase])

For the composite score of cognitive, mood, and behavioural effects at 12 months, we assessed the following items: a clinically significant worsening on three or more cognitive tests based on the reliable change index (RCI, appendix);[22] the loss of professional activity, work, or job; the loss of an important relationship (eg, marriage); or psychosis, depression, or anxiety for 3 months or longer as defined by the mini-international neuropsychiatric interview (MINI) psychiatric assessment.[23] Death of a partner was not counted as a loss of an important relationship. If the patient fulfilled at least one of these items, we regarded the composite score to be negative.

In the off-drug phase, we recorded the following scales at baseline and at 12 months: unified Parkinson's disease rating scale (UPDRS) motor examination section (ME),[24] clinical dyskinesia rating scale (CDRS),[25] ALDS, UPDRS activities of daily living section (ADL),[24] and the Schwab and England score.[24]

We also recorded the same scales during the on-drug phase. Additionally, the Parkinson's disease sleep scale (PDSS)[26] and the PD quality of life questionnaire (PDQL)[27] were recorded in the on phase. Drugs used were recorded at baseline and follow-up. DBS settings and adverse events were also recorded during follow-up. We added the following post-hoc endpoints: hours spent in off-drug and on-drug phase, and gait and postural stability (using UPDRS ME items 27 arising from chair, 28 posture, 29 gait, 30 postural stability).

Neurologists at each treatment centre (who were aware of treatment allocation) examined the patients and recorded adverse events directly after surgery, 1 week after surgery, and 3 months, 6 months, and 12 months after surgery. Structured questionnaires, with space for events that were not specified, were used for the registration of adverse events. An independent data and safety monitoring board (DSMB), consisting of a neurologist, a neurosurgeon, and a clinical statistician, monitored the trial.

Statistical analysis

For the off-on phase weighted ALDS, we used the effect size *d* (difference between mean scores of the intervention groups divided by the pooled standard deviation [SD]) as a benchmark for assessing the relative magnitude of differences between both strategies. In the Co-morbidity and Aging in Rehabilitation Patients (CARPA) study, the SD of the ALDS scores for PD was 10.[19] The results of this study suggest that over a 3-year period the increase in disability was equivalent to a decrease of five points on the ALDS. Although an effect size of 0·50 is defined as moderate,[28] such a difference in disability scores might be clinically important. On the basis of these data, a sample size of 64 patients in each intervention group (128 in total) was required to have 80% power to detect a moderate effect size of d=0·50 in favour of GPi DBS, using a two group t test with a 0·05 two-sided significance level. Assuming that the rate of cognitive, mood, and behavioural effects would be 25% in the STN DBS group,[14] we estimated that at least 110 patients (55 patients in each treatment group) would be needed to detect a difference of 20% (25% STN *vs* 5% GPi) using a χ^2 test with α=0·05 and β=0·20.

We used the intention-to-treat principle for all analyses. We summarised baseline characteristics and outcome parameters using descriptive statistics. The main analyses of this trial consisted of a comparison between the mean change in off-on phase weighted ALDS scores from baseline to follow-up at 12 months with the two-group t test, and the proportion of patients with a negative composite score for cognitive, mood, and behavioural effects with the χ^2 test. We also analysed the 12-month ALDS scores using linear regression, including the baseline ALDS scores and the stratification variables (levodopa equivalent dose and treatment centre) into the model.

With regard to the 12-month composite scores, we also did logistic regression, using the stratification variables as covariates. The differences between treatment groups for the secondary outcomes were analysed with χ^2, Fisher's exact test, t test, or Mann-Whitney U tests, as appropriate.

Also, we created imputation models to assess possible differences in outcome due to incomplete diary data. One model used age, Schwab and England baseline off-drug scores, the MDRS at baseline, and the available diary data as predictors for missing values. The second model used type of intervention, available diary data, and outcome on the UPDRS ME as predictors to impute any missing diary data.

p values of less than 0·05 were considered statistically significant. Because the ALDS has been developed within the framework of item response theory (IRT), the calculated p values for this scale were based on the original units of measurements (logits). We did no interim analyses.

See Online for appendix

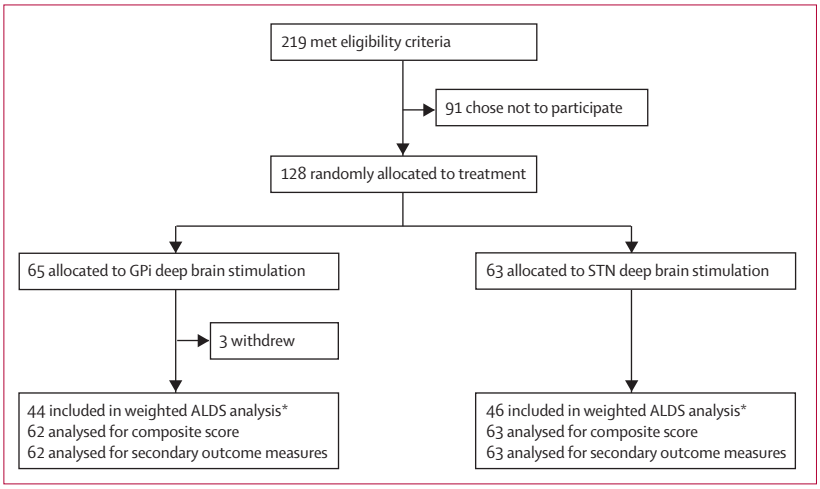

Figure: **Study profile**
GPi=globus pallidus pars interna. STN=subthalamic nucleus. *For the off-drug and on-drug phase weighted Academic Medical Center linear disability scale (ALDS), nine diaries in the GPi group and six diaries in the STN group were not available at baseline; at 12 months, 15 were not available in the GPi group and 12 were not available in the STN group.

Role of the funding source

The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, and writing of the report. Two authors (VJJO and RMAdB) guaranteed the veracity and completeness of the data analyses. VJJO and RMAdB had full access to all the data in the study and the final responsibility for the decision to submit for publication.

Results

Between Feb 1, 2007, and March 29, 2011, we enrolled 128 patients; 65 were randomly assigned to GPi DBS and 63 to STN DBS (figure). Baseline demographic and clinical characteristics are described in table 1. Three patients in the GPi group withdrew from follow-up (one patient wanted a second operation elsewhere and two considered follow-up to be too onerous). The calculations for the weighted ALDS were based on data for 90 patients that filled in the diaries at baseline and at 12-month follow-up (GPi n=44, STN n=46; see the appendix for details of missing diary data). Semi-micro-electrode recordings were used to determine the optimal location for the DBS electrodes in 88% (112 of 128) of the surgeries. We recorded no difference in mean off-on phase-weighted ALDS change score between the groups (table 2). Additional multivariable linear regression analysis showed no effect of intervention type on the 12-month ALDS scores (p=0·68).

We recorded no between-group difference in the number of patients with a negative composite score for cognitive, mood, and behavioural effects (table 2). Additional multivariate logistic regression analysis showed no between-group difference in outcome (OR 0·97, 95% CI 0·48–1·98; p=0·94). Analysis of separate components of the composite score also showed no between-group differences (table 2).

In secondary analyses, the mean change in UPDRS ME score during the off phase was lower in the GPi DBS group than in the STN DBS group (table 3). The improvement in ALDS during the off phase in the STN group was larger than in the GPi group (table 3). We found no statistically significant differences between the two groups when assessing CDRS and UPDRS ADL scores. The Schwab and England scores during the off phase improved more in the STN group than in the GPi DBS group (table 3).

In the on-drug phase, dyskinesias measured by the CDRS were reduced more in the GPi DBS group than they were in the STN DBS group (table 3). We found no differences between the groups with regard to changes in

	GPi deep brain stimulation (n=65)	STN deep brain stimulation (n=63)
Age (years)	59·1 (7·8)	60·9 (7·6)
Age at onset of Parkinson's disease (years)	48·5 (7·6)	48·6 (9·4)
Men	44 (68%)	44 (70%)
Duration of Parkinson's disease (years)	10·8 (4·2)	12·0 (5·3)
Duration of use of drugs for Parkinson's disease (years)	9·0 (3·9)	9·5 (5·6)
Hours a day spent in on-drug phase*	9·2 (3·0)	9·1 (3·3)
On-drug phase Hoehn and Yahr stage (median [range])†	2·5 (0–4)	2·5 (0–4)
Levodopa equivalent dose‡ ≥1000 mg a day	43 (69%)	43 (68%)
Treatment centre		
Academic Medical Center	37	39
University Medical Center Groningen	12	9
Haga Hospital	7	8
St Elisabeth Hospital‡	6	4
Medisch Spectrum Twente‡	3	3
Mattis dementia rating scale§	138·7 (4·0)	138·1 (5·1)

Data are mean (SD), n, or n (%), unless otherwise stated. GPi=globus pallidus pars interna. STN=subthalamic nucleus. *Calculated with a 3-day diary. †Five patients (three in GPi, two in STN) had a Hoehn and Yahr stage of 4 in on-drug phase. ‡These centres did not use perioperative micro-electrode recordings (microrecordings were done in 88% of the patients overall). §Seven patients (two in GPi, five in STN) had a score of 125–129, none had a score <125.

Table 1: **Baseline characteristics**

	Baseline		12 months		Mean change at 12 months from baseline*			
	GPi DBS	STN DBS	GPi DBS (n=62)	STN DBS (n=63)	GPi DBS	STN DBS	p value†	Effect size
Weighted ALDS (n=90)	73·8 (13·9)	68·0 (19·0)	76·8 (13·3)	75·8 (19·3)	3·0 (14·5)	7·7 (23·2)	0·28	0·24
Score for cognitive, mood, and behavioural adverse effects ≥1 (n=125)	..	..	36 (58%)	35 (56%)	..	..	0·94	..
Parts of composite score								
Decrease in neuropsychological exam‡	..	..	17 (27%)	22 (35%)	..	..	..	..
Loss of professional activity, work, or job	..	..	1 (2%)	0 (0%)	..	..	..	..
Loss of an important relationship	..	..	11 (18%)	5 (8%)	..	..	..	..
Psychosis§	..	..	4 (6%)	4 (6%)	..	..	..	..
Depression§	..	..	7 (11%)	11 (17%)	..	..	..	..
Anxiety§	..	..	9 (15%)	6 (10%)	..	..	..	..

Data are mean (SD) or n (%). ALDS=Academic Medical Center linear disability scale. DBS=deep brain stimulation. GPi=globus pallidus pars interna. STN=subthalamic nucleus. *A positive difference score indicates clinical improvement and a negative score clinical deterioration. †Two-group *t* test for weighted ALDS, χ^2 test for the composite score. ‡Decrease on neuropsychological examination is defined as a substantial worsening on three or more cognitive tests based on the reliable change index (RCI). §Psychosis, depression, or anxiety for a period of 3 months or longer.

Table 2: **Primary outcomes**

	Baseline		12 months*		Mean change at 12 months from baseline†			
	GPi DBS	STN DBS	GPi DBS	STN DBS	GPi DBS	STN DBS	p value‡	Effect size
Off phase (n=125)								
UPDRS motor examination (range 0–108)	43·8 (13·5)	44·4 (15·5)	32·4 (12·6)	24·1 (14·4)	11·4 (16·1)	20·3 (16·3)	0·03	0·55
Clinical dyskinesia rating scale (range 0–28)	0·6 (1·2)	1·0 (2·0)	0·5 (1·6)	0·8 (2·0)	0·1 (1·5)	0·2 (2·8)	0·87	0·04
ALDS (range 0–100)	53·1 (21·8)	48·8 (23·8)	64·9 (22·0)	69·1 (21·8)	11·8 (18·9)	20·3 (27·1)	0·04	0·36
UPDRS activities of daily living (range 0–52)	17·9 (6·2)	18·2 (6·5)	14·0 (6·6)	12·3 (7·9)	3·9 (6·2)	5·8 (6·2)	0·09	0·30
Schwab and England scale (range 0–100; median [range])	50 (10 to 90)	40 (10 to 90)	60 (10 to 100)	70 (10 to 90)	10 (–50 to 70)	20 (–50 to 80)	0·02	..
On phase (n=125)								
UPDRS motor examination (range 0–108)	16·0 (8·0)	17·0 (9·9)	16·0 (9·4)	14·4 (11·1)	0·0 (11·2)	3·4 (12·2)	0·17	0·13
Clinical dyskinesia rating scale (range 0–28)	5·3 (3·8)	4·8 (3·7)	2·3 (3·2)	3·8 (4·5)	3·0 (3·7)	1·1 (4·5)	0·01	–0·46
Parkinson's disease sleep scale (range 0–150)	83·6 (18·7)	81·3 (17·1)	90·8 (18·3)	94·7 (16·8)	7·2 (20·4)	13·5 (16·7)	0·07	0·35
ALDS (range 0–100)	84·2 (7·9)	81·1 (13·0)	83·4 (8·9)	80·5 (13·8)	–0·7 (9·1)	–0·7 (15·2)	0·98	..
UPDRS activities of daily living (range 0–52)	6·0 (4·9)	7·9 (5·1)	7·5 (5·4)	8·0 (6·3)	–1·4 (5·8)	0·0 (5·0)	0·16	0·11
Schwab and England scale (range 0–100; median [range])	80 (40 to 100)	80 (30 to 100)	80 (30 to 100)	80 (20 to 100)	0 (–60 to 30)	0 (–60 to 50)	0·16	..
Quality of life questionnaire (range 0–185)	86·3 (17·8)	85·4 (22·3)	96·9 (19·1)	102·0 (20·6)	10·6 (19·1)	16·5 (20·6)	0·10	0·36
Medication and DBS settings (n=125)								
Levodopa equivalent dose§	1331 (637)	1254 (473)	1122 (604)	708 (423)	–208 (521)	–546 (561)	0·01	0·30
Voltage (V)	..	..	2·9 (0·5)	2·6 (0·6)	..	..	0·004	..
Frequency (Hz)	..	..	137·5 (20·0)	135·0 (20·8)	..	..	0·52	..
Pulse width (μs)	..	..	73·0 (23·8)	63·9 (9·6)	..	..	0·008	..
Post-hoc analyses	..	..	..	..	..	..	..	..
3-day diaries (n=90)								
Time in off phase (hours a day)	6·0 (3·2)	6·2 (3·5)	4·8 (3·6)	4·5 (4·4)	–1·3 (3·4)	–1·5 (3·6)	0·71	0·06
Time in on phase without dyskinesias (hours a day)	6·5 (3·6)	6·3 (4·4)	9·5 (3·7)	9·4 (4·6)	3·0 (4·2)	3·1 (3·9)	0·92	0·02
Time in on phase with dyskinesias (hours a day)	2·5 (2·5)	2·9 (2·8)	0·5 (1·2)	0·8 (1·3)	–2·0 (2·5)	–2·1 (2·8)	0·85	0·04
Posture and gait (n=125)								
UPDRS motor examination items 27, 28, 29, 30 (range 0–16; off phase)	6·1 (2·8)	7·3 (3·7)	5·4 (2·6)	4·6 (3·3)	0·7 (3·0)	2·7 (3·3)	0·007	0·64
UPDRS motor examination items 27, 28, 29, 30 (range 0–16; on phase)	2·9 (1·5)	3·3 (2·6)	3·5 (2·1)	3·9 (3·6)	–0·5 (1·8)	–0·6 (3·3)	0·71	0·03

Data are mean (SD), unless otherwise stated. DBS=deep brain stimulation. GPi=globus pallidus pars interna. STN=subthalamic nucleus. UPDRS=unified Parkinson's disease rating scale. ALDS=Academic Medical Center linear disability scale. *Assessments with DBS on. †A positive difference score indicates clinical improvement and a negative score clinical deterioration. ‡With two group *t* test, or Mann-Whitney *U* test in case of Schwab and England scale.

Table 3: **Secondary outcomes**

UPDRS ME, PDSS, ALDS, UPDRS ADL, Schwab and England, and PDQL scores.

The mean levodopa equivalent dose reduction between baseline and 12 months was larger in the STN DBS group than in the GPi DBS group (table 3). For DBS settings, the mean amplitude and pulse width were both larger in the GPi group than in the STN group; the mean frequency settings were similar between the two groups (table 3).

In post-hoc analyses, both groups showed a similar reduction in time in off-drug phase and time in on-drug phase with dyskinesias (table 3). Gait and postural stability items improved more in the STN group than in the GPi group (table 3). The imputation models showed no changes in outcomes on the weighted ALDS (data not shown).

23 (3%) of 768 adverse-events questionnaires were not completed. There were 290 adverse events in the GPi group and 303 in the STN group; we found no statistically significant differences between the two groups in the occurrence of any adverse events (table 4). In one patient, who was allocated to STN DBS, the surgery was aborted because of a low threshold for oculomotor side-effects during macro-stimulation in the STN. The patient underwent surgery for GPi DBS 41 weeks later, but had a deep intracerebral haemorrhage during that surgery with a resultant hemiparesis. Two patients (both in the STN group) had small post-operative haemorrhages near the electrode tip that were detected by planned postoperative CT scans, without any accompanying symptoms. Semi-micro-electrode recordings were used in two of the three patients with a peri-operative haemorrhage. 64 (22%) of 290 adverse events in the GPi group and 76 (25%) of 303 in the STN group were judged to be related to active stimulation. Of all adverse events, 22 (2%) were present at one or more subsequent visits.

	GPi DBS (N=65)	STN DBS (N=63)	p value*
Cerebral infarction (perioperative)	0 (0%)	0 (0%)	..
Cerebral haemorrhage (perioperative)†	0 (0%)	3 (5%)†	0·08
Epilepsy (perioperative)	0 (0%)	1 (2%)	0·31
Epilepsy (postoperatively)	2 (3%)	1 (2%)	0·67
Implantation-site infection	2 (3%)	2 (3%)	0·89
Facial palsy	2 (3%)	4 (6%)	0·34
Dysphasia	7 (11%)	8 (13%)	0·60
Dysarthria	19 (29%)	25 (40%)	0·21
Dysphagia	13 (20%)	7 (11%)	0·34
Hiccups	10 (15%)	2 (3%)	0·06
Apraxia of eyelid opening	11 (17%)	14 (22%)	0·37
Oculomotor or visual field disturbance	6 (9%)	4 (6%)	0·70
Sensory disturbance	5 (8%)	9 (14%)	0·20
Balance disorder	23 (35%)	30 (47%)	0·19
Hypersalivation	14 (22%)	14 (22%)	0·72
Emotional lability	47 (72%)	53 (84%)	0·29
Paresis	1 (2%)	1 (2%)	0·92
Dyskinesias	17 (26%)	24 (38%)	0·16
Delirium	14 (22%)	15 (24%)	0·59
Other‡	97	87	..

Data are n (%). DBS=deep brain stimulation. GPi=globus pallidus pars interna. STN=subthalamic nucleus. *χ^2 or Fisher's exact test when appropriate. †Including one patient with a deep intracerebral haemorrhage during a second surgery for GPi DBS. ‡Including one report of hypomania (STN), ten reports of dysphoria or depression (three with GPi and seven with STN), and three haematomas at the battery implantation site (one with GPi and two with STN).

Table 4: **Adverse events**

Discussion

Our results showed no difference between GPi DBS and STN DBS in terms of our primary outcomes. In secondary analyses, however, STN DBS was associated with a better improvement in off-phase motor symptoms and disability than was GPi DBS and, by contrast with our original hypothesis, it did not cause greater cognitive, mood, and behavioural side-effects.

We detected no difference between GPi and STN DBS in drug-phase-weighted ALDS, which was one of our primary outcomes. In the off-drug phase, however, the difference in improvement on the ALDS scores between the GPi and STN DBS group was very large. We suggest three main reasons for this discrepancy. First, the fact that only 70% of patients completed the diaries at both baseline and follow-up resulted in a loss of power in our analysis of the weighted ALDS. Scheduled surgeries were not cancelled if patients had not completed the diary for logistical reasons and because such cancellation would prevent patients from receiving an effective treatment. Second, the standard deviation for the ALDS score was larger than we had anticipated: up to 23·2 points instead of the anticipated 10. The third reason concerns the weighting of the scores. The effects of treatment on the ALDS scores in on-drug phase were much the same for both procedures, and at 12-month follow-up, patients spent 70% of the time in the on phase. Hence, the large difference between the two groups in off-drug phase scores contributes only 30% to the weighted ALDS.

Nevertheless, for the off-drug phase, the difference in disability between the two groups is clinically relevant. For example, a typical patient who scored 50 on the ALDS in off phase preoperatively would only just be able to have a shower independently. With an improvement of about 12 points (GPi group) the patient would be able to walk down a flight of stairs. However, with an improvement of 20 points (STN group), this patient would be able to visit a restaurant independently.

This study is an active control: both groups receive a treatment (instead of one group receiving placebo)—an average yearly decline of the ALDS in patients with Parkinson's disease who did not receive DBS is 1·3 points.[19]

The second primary outcome, the composite score for cognitive, mood, and behavioural effects, also did not support the hypothesis of superiority of GPi DBS over STN DBS. The high number of patients with a negative composite score in both groups warrants clarification. First, previous DBS studies describe the neuropsychological tests as mean scores and SDs; our outcome measure is a dichotomous composite measure and thus cannot directly be compared with these studies. Second, if a patient had a negative score because of the loss of an important relationship, this does not necessarily imply worsening (or improvement for that matter) of disability or perceived quality of life. We chose to include this parameter because such issues might be a result of subtle cognitive or behavioural problems that have an effect on daily life but are not detected by standard psychiatric questionnaires. However, the results of the composite scores do not suggest that GPi DBS led to fewer issues regarding cognition, mood, and behaviour in direct comparison with STN DBS.

The findings of our study with respect to the effect of STN DBS on motor symptoms in the off phase are in agreement with the results of a meta-analysis of cohort studies that showed a reduction of 52% on the UPDRS ME with STN DBS.[29] Of the three previous randomised trials[13,30,31] that compared GPi DBS with STN DBS (panel) the largest was by the Veterans Administration Cooperative DBS Study group (2010, n=299), which followed patients up to 36 months postoperatively.[13,30] The primary outcome was the change in motor function (UPDRS ME). The study did not show a difference in effect on motor symptoms between GPi and STN DBS. By contrast with our findings, this trial showed an improvement in off phase motor symptoms of only 26% with STN DBS, 6 months and 24 months after surgery. A possible explanation for this discrepancy could be the fact that the physicians responsible for managing post-operative DBS settings and concurrent drug schedule adjustments in Follett and colleagues' study[13] did not

know if patients had received GPi or STN DBS. Because each target needs a different approach during postoperative management, this approach might have led to a suboptimum clinical improvement in the STN group. For example, optimum benefit from STN DBS often requires much medication reduction because of the synergistic effect of medication and STN DBS in the short term. Also, characteristics of patients and differences in targeting might account for the difference in outcome. In the Veterans Administration Cooperative DBS study,[13,30] medication use was reduced more in the STN group than in the GPi group. The level of depression worsened after STN DBS and improved after GPI DBS (p=0·02).[13,30] After 36 months, motor function was still better than it was at baseline in the off phase. Mattis dementia rating scale scores decreased faster in the STN group than in the GPi group (p=0·01); other neurocognitive measures showed a gradual decrease overall.

The COMPARE trial investigated unilateral GPi or STN DBS in 45 patients.[15,31] It showed a similar improvement of motor function and mood in both groups. However, quality of life improved more in the GPi group than it did in the STN group (38% vs 14%, respectively; p=0·03).[15,31]

The third study, by Anderson and colleagues[12] (n=23), showed an improvement of motor scores in the off phase after 12 months of both GPi and STN stimulation (39% *vs* 48%). Dyskinesia was reduced by stimulation with both GPi and STN (89% *vs* 62%). Cognitive and behavioural complications were seen only in combination with STN stimulation.

Compared with STN DBS, GPi DBS reduced dyskinesias more effectively during the standardised assessments in the on phase at 12 months. Patients received the same amount of levodopa to induce an on phase during the standardised assessment at 12 months as at baseline. However, because patients with STN DBS use less medication in daily life, they might have less severe dyskinesias than indicated by our measurements during the standardised assessments. In this respect, diaries showed similar reductions in off time as well as time that dyskinesias were present for GPi and STN DBS. Post-hoc analysis of gait and postural stability showed superiority of STN DBS in the off phase.

DBS amplitude and pulse widths were on average lower in the STN group, which is consistent with previous findings and leads to longer intervals between replacement of pulse generators.[3] The inclusion criteria for the trial were similar to the criteria used in regular clinical practice when counselling patients with Parkinson's disease for DBS treatment. In this respect, the only extra exclusion criterion for the trial was previous stereotactic functional neurosurgery. Also, the inclusion and exclusion criteria did not contain an upper age limit and the minimum MDRS score required for inclusion was low. Five of the six hospitals in the Netherlands that did DBS participated in the trial. All these factors contributed to the external validity of the study.

Panel: Research in context

Systematic review

We searched PubMed with the terms "DBS" and "deep brain stimulation" for randomised controlled trials. We also searched PubMed with the terms "DBS" and "subthalamic nucleus" and "globus pallidus internus" for reports published before Sept 11, 2012, with no language restriction. We identified three randomised controlled trials[12,13,15,30,31] that compared GPi DBS with STN DBS. All three trials found improvement of motor symptoms with both GPi DBS and STN DBS in the off-drug phase. The magnitude of the improvements differed between these trials.

Interpretation

To our knowledge, our study is the second largest randomised controlled trial comparing bilateral globus pallidus pars interna (GPi) deep brain stimulation (DBS) and subthalamic nucleus (STN) DBS. Also, our use of both disability and cognition, mood, and behaviour as primary outcomes answers important questions about the effectiveness and side-effects of GPi and STN DBS that were raised by previous trials. Although the primary outcomes did not show a difference between GPi and STN DBS on the weighted ALDS and composite score for cognition, mood and behaviour, important differences were seen on the secondary outcomes. By contrast with the Veterans Administration Cooperative DBS Study group findings, our study shows an improvement of motor symptoms of 45% in off-drug phase in the STN group, which lends support to findings from previous studies investigating STN DBS.[3,4,12,32,33]

When taking all these factors into consideration, our data suggest that the STN may be the preferred target for DBS in PD, because of more substantial improvement of symptoms and disability in the off phase, in combination with the need for fewer drugs and lower battery consumption.

Contributors

RMAdB designed the study. RMAdB, VJJO, MSJM, PRS, PvdM, TvL, MJSt, AM, CFEH, PCGN, GNB, JPPvV, and MWPML collected data. VJJO, RMAdeB, RJdH analysed and interpreted data. VJJO wrote the paper and designed the figure and tables. All other authors gave comments on drafts of the paper.

Conflicts of interest

The AMC movement disorders group received an unrestricted fellowship grant from Medtronic. RMAdB has received lecture payments for the Movement Disorder Society DBS course and the EMCT DBS course. PRS has received lecture payments from Medtronic. TvL has received consultancy fees from Novartis, Abbott, and Britannia and lecturing fees from Novartis. JPPvV has received consultancy fees from Boehringer Ingelheim NL and Teva Pharma NL and lecturing fees from Boehringer Ingelheim NL.

Acknowledgments

This study was funded by Stichting Internationaal Parkinson Fonds, Prinses Beatrix Fonds, and Parkinson Vereniging. We thank Klaus L Leenders, D Andries Bosch, and A H Koos Zwindermann for participating in the data monitoring and safety committee. We thank Nan van Geloven for advice on the statistical analysis.

References

1 Vidailhet M. Movement disorders in 2010: Parkinson disease-symptoms and treatments. *Nat Rev Neurol* 2011; **7:** 70–72.

2 Limousin P, Pollak P, Benazzouz A, et al. Bilateral subthalamic nucleus stimulation for severe Parkinson's disease. *Mov Disord* 1995; **10:** 672–74.

3 Limousin P, Krack P, Pollak P, et al. Electrical stimulation of the subthalamic nucleus in advanced Parkinson's disease. *N Engl J Med* 1998; **339:** 1105–11.

4 Deep-Brain Stimulation for Parkinson's Disease Study Group. Deep-brain stimulation of the subthalamic nucleus or the pars interna of the globus pallidus in Parkinson's disease. *N Engl J Med* 2001; **345:** 956–63.

5 Krack P, Batir A, Van Blercom N, et al. Five-year follow-up of bilateral stimulation of the subthalamic nucleus in advanced Parkinson's disease. *N Engl J Med* 2003; **349:** 1925–34.

6 Walter BL, Vitek JL. Surgical treatment for Parkinson's disease. *Lancet Neurol* 2004; **3:** 719–28.

7 Loher TJ, Burgunder JM, Pohle T, Weber S, Sommerhalder R, Krauss JK. Long-term pallidal deep brain stimulation in patients with advanced Parkinson disease: 1-year follow-up study. *J Neurosurg* 2002; **96:** 844–53.

8 Volkmann J, Allert N, Voges J, Sturm V, Schnitzler A, Freund HJ. Long-term results of bilateral pallidal stimulation in Parkinson's disease. *Ann Neurol* 2004; **55:** 871–75.

9 Krause M, Fogel W, Heck A, et al. Deep brain stimulation for the treatment of Parkinson's disease: subthalamic nucleus versus globus pallidus internus. *J Neurol Neurosurg Psychiatry* 2001; **70:** 464–70.

10 Okun MS, Foote KD. Subthalamic nucleus vs globus pallidus interna deep brain stimulation, the rematch: will pallidal deep brain stimulation make a triumphant return? *Arch Neurol* 2005; **62:** 533–36.

11 Quinn N. Progress in functional neurosurgery for Parkinson's disease. *Lancet* 1999; **13:** 1658–59.

12 Anderson VC, Burchiel KJ, Hogarth P, Favre J, Hammerstad JP. Pallidal vs subthalamic nucleus deep brain stimulation in Parkinson disease. *Arch Neurol* 2005; **62:** 554–60.

13 Follett KA, Weaver FM, Stern M, et al, and the CSP 468 Study Group. Pallidal versus subthalamic deep-brain stimulation for Parkinson's disease. *N Engl J Med* 2010; **362:** 2077–91.

14 Smeding HM, Speelman JD, Koning-Haanstra M, et al. Neuropsychological effects of bilateral STN stimulation in Parkinson disease: a controlled study. *Neurology* 2006; **66:** 1830–36.

15 Okun MS, Fernandez HH, Wu SS, et al. Cognition and mood in Parkinson's disease in subthalamic nucleus versus globus pallidus interna deep brain stimulation: the COMPARE trial. *Ann Neurol* 2009; **65:** 586–95.

16 Esselink RA, de Bie RM, de Haan RJ, et al. Unilateral pallidotomy versus bilateral subthalamic nucleus stimulation in PD: a randomized trial. *Neurology* 2004; **62:** 201–07.

17 Weisscher N, Post B, de Haan RJ, Glas CA, Speelman JD, Vermeulen M. The AMC Linear Disability Score in patients with newly diagnosed Parkinson disease. *Neurology* 2007; **69:** 2155–61.

18 Weisscher N, Glas CA, Vermeulen M, De Haan RJ. The use of an item response theory-based disability item bank across diseases: accounting for differential item functioning. *J Clin Epidemiol* 2010; **63:** 543–49.

19 Post B, Muslimovic D, van Geloven N, Speelman JD, Schmand B, de Haan RJ, and the CARPA-study group. Progression and prognostic factors of motor impairment, disability and quality of life in newly diagnosed Parkinson's disease. *Mov Disord* 2011; **26:** 449–56.

20 Holman R, Lindeboom R, Vermeulen M, de Haan RJ. The AMC Linear Disability Score project in a population requiring residential care: psychometric properties. *Health Qual Life Outcomes* 2004; **2:** 42.

21 Holman R, Weisscher N, Glas CA, et al. The Academic Medical Center Linear Disability Score (ALDS) item bank: item response theory analysis in a mixed patient population. *Health Qual Life Outcomes* 2005; **3:** 83.

22 Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991; **59:** 12–19.

23 Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998; **59** (suppl 20)**:** 22–33, quiz 34–57.

24 Fahn S, Elton RL, and the Members of the UPDRS Development Committee. Unified Parkinson's disease rating scale. In: Fahn S, Marsden CD, Calne DB, eds Recent developments in Parkinson's disease Florham Park, New Jersey, USA: Macmillan Healthcare Information; 1987 p 153–63.

25 Hagell P, Widner H. Clinical rating of dyskinesias in Parkinson's disease: use and reliability of a new rating scale. *Mov Disord* 1999; **14:** 448–55.

26 Chaudhuri KR, Pal S, DiMarco A, et al. The Parkinson's disease sleep scale: a new instrument for assessing sleep and nocturnal disability in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2002; **73:** 629–35.

27 de Boer AG, Wijker W, Speelman JD, de Haes JC. Quality of life in patients with Parkinson's disease: development of a questionnaire. *J Neurol Neurosurg Psychiatry* 1996; **61:** 70–74.

28 Cohen J, ed. Statistical power analysis for the behavioral sciences. 2nd edn. Lawrence Erlbaum Associates, Inc; 1988.

29 Kleiner-Fisman G, Herzog J, Fisman DN, et al. Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes. *Mov Disord* 2006; **21** (suppl 14)**:** S290–304.

30 Weaver FM, Follett KA, Stern M, et al, and the CSP 468 Study Group. Randomized trial of deep brain stimulation for Parkinson disease: thirty-six-month outcomes. *Neurology* 2012; **79:** 55–65.

31 Zahodne LB, Okun MS, Foote KD, et al. Greater improvement in quality of life following unilateral deep brain stimulation surgery in the globus pallidus as compared to the subthalamic nucleus. *J Neurol* 2009; **256:** 1321–29.

32 Deuschl G, Schade-Brittinger C, Krack P, et al, and the German Parkinson Study Group, Neurostimulation Section. A randomized trial of deep-brain stimulation for Parkinson's disease. *N Engl J Med* 2006; **355:** 896–908.

33 Rodriguez-Oroz MC, Obeso JA, Lang AE, et al. Bilateral deep brain stimulation in Parkinson's disease: a multicentre study with 4 years follow-up. *Brain* 2005; **128:** 2240–49.