UvA-DARE (Digital Academic Repository)

Predictors of cognitive and psychosocial outcome after STN DBS in Parkinson's Disease

Smeding, H.M.M.; Speelman, J.D.; Huizenga, H.M.; Schuurman, P.R.; Schmand, B.

Link to publication

Predictors of cognitive and psychosocial outcome after STN DBS in Parkinson Disease

Harriet M M Smeding, Johannes D Speelman, Hilde M Huizenga, et al.

Updated information and services can be found at:

http://jnnp.bmj.com/content/early/2009/05/21/jnnp.2007.140012

These include:

P<P	Published online May 21, 2009 in advance of the print journal.
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To order reprints of this article go to:

http://jnnp.bmj.com/cgi/reprintform

To subscribe to *Journal of Neurology, Neurosurgery & Psychiatry* go to:

http://jnnp.bmj.com/subscriptions

Predictors of cognitive and psychosocial outcome after STN DBS in Parkinson Disease

H Smeding PhD, JD Speelman MD, PhD, HM Huizenga PhD, PR Schuurman MD, PhD,
B Schmand, PhD

From the departments of Neurology (Smeding, Speelman and Schmand) and Neurosurgery (Schuurman) of the Academic Medical Center, University of Amsterdam,
and Department of Psychology (Schmand and Huizenga), University of Amsterdam

Correspondence: H. Smeding, Department of Neurology, H2-222, Academic Medical Center, P.O. Box 22660, 1100 DD Amsterdam, the Netherlands; phone: ++31205668632;
fax: ++31206971438; email: h.m.smeding@amc.uva.nl

Abstract

Objective: To find predictors of cognitive decline and quality of life one year after bilateral subthalamic nucleus stimulation (STN DBS) in Parkinson's disease (PD).

Methods: A total of 105 patients were evaluated with a comprehensive neuropsychological assessment before and 12 months after surgery. A control group of 40 PD patients was included to control for effects of repeated testing and disease progression. We determined individual changes in cognition, mood and quality of life using a statistical method that controls for multiple comparisons. We performed logistic regression analyses to assess predictors of cognitive changes and quality of life.

Results: Twelve months after surgery, the improvement in motor function was 41% (UPDRS3 score in off). The STN group showed a large improvement in quality of life compared to the control group (Cohen's d=0.9). At the individual level, 32 percent (95% CI: 22 – 40) of the STN group showed a substantial improvement in quality of life. Thirty six percent (95% CI: 27 - 46) of the STN patients showed a profile of cognitive decline compared to the control group. Mood improved in 16 STN patients and declined in 16 subjects. Impaired attention, advanced age and a low levodopa response at baseline predicted cognitive decline, whereas a high levodopa response at baseline predicted improvement in quality of life. Postoperative decrease in dopaminergic medication was not related to cognitive decline.

Conclusions: STN DBS improves quality of life. However, a profile of cognitive decline can be found in a significant number of patients. Levodopa response, age and attention at baseline are predictors of cognitive and psychosocial outcome.

Introduction

Bilateral deep brain stimulation of the subthalamic nucleus (STN DBS) is an effective surgical treatment to reduce motor symptoms in patients with advanced Parkinson's disease (PD).[1 2 3;4] The effects of STN DBS on cognition and behaviour, however, are still the subject of controversy.[5;6] Meta-analysis of 28 reports on cognitive outcome revealed small declines in executive functions and memory, and moderate declines in verbal fluency.[6] However, most of the included studies lacked control groups. This is an important omission because cognitive decline can be expected as a result of disease progression itself, while conversely decline may be masked by retest effects. Another issue is that a statistically significant difference between groups on a cognitive test is of limited interest to understanding changes in the individual patient. Subtle changes in several cognitive domains may have more implications for daily life than a large significant deviation on a single test. Suppose, for example, that a patient shows declines in memory, attention, and executive functioning, while each of his score decrements on the corresponding tests is too small to reach statistical significance. This patient and his relatives may very well note his change in mental functioning in daily life. Finally, it is uncertain which factors increase the risk of cognitive decline after STN DBS, although cognitive impairment prior to surgery and advancing age have been suggested.[7]

In this paper we report the 12 months follow-up of the cognitive and behavioural effects of STN DBS. We determined groupwise and individual cognitive changes in STN patients compared to a nonsurgical control group.[8] Likewise, we assessed individual changes in mood and quality of life after STN DBS. Finally, we analyzed possible baseline predictors of cognitive outcome and quality of life including age, levodopa response, medication, preoperative mental status and prior stereotactic surgery.

Methods

Patients
Included were consecutive PD patients who underwent STN DBS. Twenty patients have been previously described.[9] The remaining patients were recruited between June 2001 and June 2006 in three Dutch hospitals experienced in STN stimulation for PD. Eligible patients had idiopathic PD with an unequivocal reduction in off phase symptoms on levodopa, and at least one of the following symptoms despite optimal pharmacological treatment: severe response fluctuations, dyskinesias, dystonia, tremor or bradykinesia. Exclusion criteria were predominantly unilateral symptoms without severe response fluctuations, severe brain atrophy on CT or MRI scans, Hoehn and Yahr stage 4 or 5 in the best on phase,[10] dementia (Mattis' DRS < 120),[11] psychosis or depression at inclusion, or surgical contraindications.

Alongside the STN group, we formed a control group of patients recruited from two university hospitals and a teaching hospital in the Amsterdam region. They were patients who had idiopathic PD for at least 5 years, and who did not (yet) want surgery or for whom the treating neurologist considered STN DBS not (yet) indicated. They were invited to participate as controls. Exclusion criteria were identical to those of the STN group (controls were not analyzed for surgical contraindications). Inclusion and prospective evaluation of the controls occurred parallel in time to the STN group. The 6-months follow-up of this cohort has been reported before.[12]
The medical ethics committees of the participating hospitals approved the study.

Surgical procedure
The STN group underwent stereotactic surgery as previously described,[2] using ventriculography, MRI, or CT scan to determine the coordinates of the target structure. Surgical techniques were identical in the participating centres. Microelectrode recording was used in 43 patients. After macroelectrode test-stimulation a four contact electrode (model DBS-3389, Medtronic, Minneapolis) was implanted. The electrodes were connected to implantable pulse generators (Itrel II, Soletra, or Kinetra, Medtronic, Minneapolis) in a staged procedure under general anaesthesia. We did not systematically perform MRI postoperatively, because of restrictions of the Dutch radiological departments. The location of the leads was checked at the end of surgery by bidirectional skull X-ray in the stereotactic frame. The efficacy of DBS was assessed by its effect on motor symptoms and the lack of adverse events. If possible, the medication was reduced, especially the dopamine agonists. Significant side effects were not tolerated.

Neuropsychological tests
The selection of tests was based on suggestions from the literature:[13] Mattis' Dementia Rating Scale (DRS);[11;14] category fluency (animals and occupations for 1 minute each);[15] Controlled Oral Word Association Test (COWAT);[16] alternating fluency (body parts/cities or pieces of clothing/countries);[17] Dutch Adult Reading Test (DART);[18] Paced Auditory Serial Addition Task (PASAT), speed 3.2 seconds per digit;[19] Rey Auditory Verbal Learning Test (AVLT);[20] Groningen Intelligence Test subtest Visuospatial reasoning;[15] Stroop Color Word Test;[21] Odd Man Out Test (OMO);[22] Trail Making Test parts A and B;[23] and Boston Naming test (BNT).[24]

To minimize practice effects we used alternate forms where available in a balanced order across patients.

Mood and behaviour rating scales
The DEX Questionnaire of the Behavioural Assessment of the Dysexecutive Syndrome [25] and the Memory Assessment Clinic ratings (MAC) [26;27] were completed by the patient and a proxy. The CES-Depression [28] and the Parkinson Disease quality of life (PDQL) [29] scale were filled out by the patient. We assessed neuropsychiatric changes with a Dutch translation of the Neuropsychiatric Inventory (NPI).[30;31] The NPI consists of 12 items: delusions, hallucinations, agitation /aggression, dysphoria, euphoria, anxiety, apathy, disinhibition, irritability/lability, aberrant motor behaviour, night time behaviour, eating behaviour. We constructed four more items based on the Frontal Behavioral Inventory (FBI).[32] These items are: disgust and negligence, sexual interest, language and speech, and cognitive changes.[12]

During the test session a combination of the abbreviated Profile of Mood States (POMS) [33] and the Positive Negative Affect Scale (PANAS) [34] was completed by the patient. The Montgomery & Åsberg Depression scale (MADRS) [35] was filled out by the examiner.

Assessments
At baseline scoring of motor function was done with the Unified Parkinson Disease Rating Scale part 3 (UPDRS3) and Hoehn and Yahr staging in standardized on and off phase.[4] Levodopa response was calculated as follows: UPDRS3 score in off minus UPDRS3 score in on divided by the UPDRS3 score in off.

Follow-up assessment was done in the on phase 12 months after surgery for the STN group and 12 months after baseline for the control group. Standardized on/off evaluations of motor functioning at 12 months were only available for the second half of the STN series.

Statistical analyses
We compared baseline group differences in demographic and disease characteristics with the Mann-Whitney U test, because of skewed distribution of many of these variables. Change over time was analysed in the conventional way by subtracting the test and questionnaire scores at baseline from the follow-up scores to form change scores. The distributions of change scores were first examined for deviation from normality. Next, the change scores were compared at the group level by analysis of covariance and at the individual level in a new way that compared each STN patient to the control group (see below).

Group differences: To compare the change in both groups, we analysed the change scores with ANOVAs covarying for baseline differences (if any) between the groups in demographic or disease characteristics. Furthermore, we computed effect sizes of change (Cohen's d: difference between the means of change scores of both groups divided by the pooled standard deviations of the change scores). An effect size of 0.2 reflects a small effect, 0.5 a medium, and 0.8 a large effect.[36] *p*-Values of less than 0.05 (one-tailed) were accepted as statistically significant. We did not correct the level of significance for multiple comparisons to reduce the probability of type I error

because we were mainly interested in detecting adverse effects of the surgical intervention. Under this circumstance, type II error (failure to detect an effect when it actually exists) is more serious than type I error (considering an effect to be real when it is actually not).[37]

Individual changes: We applied a new multivariate method of normative comparison to determine individual outcome.[8] Usually, when we consider cognitive impairments in an individual patient, we compare each test score with its normative data to decide whether or not it is abnormal. If multiple tests are administered, as we did in the present study, the risk of finding at least one significant deviation when in reality there is none (the familywise false positive rate) becomes very large. Moreover, the usual univariate approach is insensitive to specific patterns in deviations from the norm. Therefore, we used the multivariate normative comparison method (MNC), which tests whether a patient deviates from a control group on several characteristics simultaneously. The MNC is based on Hotelling's T^2 statistic. It is appropriate in small samples and allows testing of one-sided hypotheses, while the familywise false positive rate is adequately controlled at 5%. We performed MNC on cognitive decline and on mood changes. Included in the analysis of cognitive decline were the change scores of the Mattis' DRS, category fluency, letter fluency, alternating fluency, PASAT, AVLT immediate and delayed recall, the Stroop subtests, Trailmaking A and B, Boston naming test and GIT visuospatial subtest. Because we were mainly interested in individual decline, we tested one-sided (α=0.05). If the one-sided MNC test is significant, then the individual shows a cognitive profile that deviates in a negative sense form the control group.

In the analyses of mood we used scores on Positive affect, Negative affect, POMS subscales, and CES-D. Because mood changes can be in the direction of depression as well as in the direction of improvement, we tested multivariate mood changes two-sided (α=0.05). Individual outcome of quality of life was tested univariately with the change scores of the PDQL. Because similar studies in this field generally find improvement in quality of life after surgery, we tested one-sided (α=0.05), arbitrarily defining an improvement of more than 1.65 standard deviation of the change scores in controls as substantial.

Predictors of outcome: Finally, we conducted logistic regression analyses to find characteristics that predict outcome in cognition and quality of life. To increase the power of the analysis, we computed cognitive composite variables. Raw test scores were converted into z-scores corrected for age and education based on the published norms of each test. The variable 'verbal fluency' consisted of the mean z-score on category fluency and on letter fluency; 'memory' combined the immediate recall and the delayed recall of the AVLT; 'mental speed' combined Trailmaking A, Stroop Word and Stroop Color Card. The composite of Trailmaking part B and the Stroop Color Word Card was named 'attention' (short for divided and selective attention). Because of a lack of appropriate norms we could not integrate the PASAT, the OMO test, and the alternating fluency into these composites.

The dependent variable in the first logistic regression model was cognitive decline (yes/no) as determined with the multivariate normative comparison method. From the possible predictors (age, education, Mattis' DRS, verbal fluency, attention, memory, mental speed, levodopa response, disease duration, prior stereotactic surgery,

mood, dosage of dopaminergic medication at baseline, and application of microrecording during surgery) we chose the variables that had an association of p<0.20 with the dependent variable. A stepwise forward approach was used. In the logistic regression model for outcome in quality of life, the dependent variable was improvement of quality of life (yes/no) determined by univariate comparison. The independent variables were chosen from the same set of possible predictors according to the same procedure as described above. The predictor variables were not dichotomized.

Finally, we looked at the correlations of cognitive decline with postoperative decrease in dopaminergic medication and with postoperative increase of depressive symptoms, because these factors are often suggested to be associated with cognitive decline.

Results

Group characteristics
One hundred and eleven patients were included in the STN group. The control group consisted of 42 patients with PD. After 12 months, 6 patients from the STN group were lost to follow-up (two had an infection of the stimulator; one was imprisoned; three refused). Two patients from the control group were lost to follow-up (one deceased; one contracted an unrelated disease). Data from 105 patients of the STN group and 40 of the control group were analyzed. One STN patient had a pallidotomy during the follow-up interval, two patients had a dislocated electrode, and two patients suffered a postoperative haemorrhage. These patients were not excluded. Group characteristics are shown in table 1. The STN group was significantly younger than the control group, had fewer years of education, had longer disease duration, and used more dopaminergic medication. There were no differences in score on UPDRS part 3 or in Hoehn and Yahr score in 'on' phase.

Table 1. Demographic and disease characteristics of the patient sample at baseline

	STN (n=105)	Control (n=40)	P-values
Men/women, n	63/42	22/18	0.56
Age in years	58.4 (7.8)	63.5 (9.2)	0.01
Education in years	11.1 (2.9)	12.2 (2.9)	0.01
DART-IQ	102.0 (13.7)	106.7 (11.2)	0.06
Disease duration in years median (25, 75)	13 (10, 18)	10 (7, 15)	0.001
UPDRS part 3 "off"	43.7 (12.3)		
UPDRS part 3 "on"	21.5 (9.1)	23.9 (12.7)	0.44
Levodopa challenge test % improvement	49.7 (18.6)		
Hoehn & Yahr "off"	3.7 (0.9)		
Hoehn & Yahr "on"	2.7 (0.6)	2.7 (0.7)	0.76
Medication LEU pre-op median (25, 75)	800 (553, 1174)	545 (405, 886)	<0.001
Microrecording yes/no, n	43/62		
Prior stereotactic procedure yes/no, n	17/88		

Values are mean (SD) unless otherwise stated; SD = standard deviation; n = number of subjects; 25, 75 = 25th and 75th centile;
LEU = levodopa equivalent units; DART = Dutch Adult Reading Test;
P-values by Chi-square test (gender) or Mann-Whitney U test (remaining variables).

Motor functioning and medication
STN stimulation had a clear positive effect on motor functions. The UPDRS-3 score in off phase improved 34% after six months compared to baseline. [12] After 12 months this score was 26.4 (SD 11.9; n=52); improvement was 41% compared to baseline. After 12 months there was a large reduction in LEU in the STN group, whereas the medication dose in the control group hardly changed (median change in LEU in STN group: -150 [25th and 75th centile:
 -400, 40]; median change in LEU in control group: 0 [25th, 75th: 0, 200], the difference in median change was significant [p<0.001 Mann-Whitney U test]).

Cognitive tests
The cognitive test scores of the groups are shown in table 2. At baseline, the STN and control groups did not significantly differ in cognition except for a lower score at the delayed verbal recall in the control group and a worse score for the STN group on the Odd Man Out test.[12]

We conducted ANOVA's with change scores as the dependent variables and dopaminergic medication at baseline, and age, disease duration, and years of education as covariates. The change scores were not skewed (skewness varied between -1.26 and 0.60) except for the Mattis DRS (skewness -2.33), while the correlations between the change scores were typically low (median r = 0.25; 25th, 75th 0.18, 0.35). Twelve months after surgery the STN group showed a significant decline compared to the control group on all verbal fluency measures, Mattis' DRS, delayed recall of the AVLT, Stroop Color Card and Stroop Color Word Card. These results were comparable to those of the 6 months follow-up.[12] In addition, at 12 months follow-up the STN group showed a larger decline on immediate recall of the AVLT, Stroop Word Card, and visuospatial reasoning. Nonparametric testing of the group difference in Mattis' DRS change score was also significant (p = 0.003, one-tailed Mann-Whitney U test).

Table 2. *Cognitive test scores at baseline and change scores (mean, SD) at 12 months follow-up for STN and control groups.*

Test	STN	Control	P values	Cohen's d
Dementia Rating Scale total	135.6 (6.1)	137.2 (5.4)		
change score	-2.4 (7.5)	1.0 (3.6)	0.001	-0.5
Category fluency	38.2 (9.9)	40.5 (8.5)		
change score	-4.6 (6.4)	-0.7 (6.4)	0.005	-0.6
COWAT letter fluency	34.9 (12.9)	37.4 (12.7)		
change score	-4.1 (8.7)	3.1 (9.8)	<0.001	-0.8
Alternating fluency	46.0 (13.3)	48.8 (12.8)		
change score	-6.0 (9.7)	-0.5 (8.2)	0.001	-0.6
PASAT	42.0 (13.5)	45.1 (9.8)		
change score	-1.4 (11.3)	2.2 (8.4)	0.12	-0.6
AVLT total score	39.1 (9.3)	37.6 (10.5)		
change score	-1.0 (9.7)	3.2 (8.4)	0.004	-0.4
AVLT delayed recall	7.9 (2.8)	7.0 (2.8)		
change score	-0.7 (2.6)	0.9 (2.3)	<0.001	-0.6
GIT visuospatial reasoning	9.4 (3.6)	9.2 (3.2)		
change score	0.0 (2.8)	0.8 (2.0)	0.02	-0.3
Stroop word seconds	50.9 (13.2)	49.1 (10.4)		
change score	4.3 (10.9)	-0.9 (7.5)	0.002	-0.5
Stroop color seconds	65.9 (14.7)	66.8 (12.9)		
change score	8.0 (16.3)	0.5 (11.1)	0.01	-0.5
Stroop color word seconds	130.0 (54.8)	129.7 (42.1)		
change score	15.4 (60.4)	-6.8 (32.5)	0.01	-0.5
Odd Man Out test errors	11.9 (12.9)	9.5 (14.8)		
change score	0.3 (15.9)	-3.1 (10.4)	0.07	-0.2
Trailmaking A seconds	49.0 (17.7)	48.9 (17.2)		
change score	1.2 (19.6)	-1.7 (15.8)	0.15	-0.2
Trailmaking B seconds	136.3 (78.0)	118.7 (50.9)		
change score	22.1 (94.8)	7.3 (39.9)	0.07	-0.2
Boston Naming Test	54.0 (4.2)	54.8 (3.2)		
change score	0.7 (2.6)	0.8 (2.1)	0.23	-0.1

Negative change scores indicate decline in performance except for speeded test variables and error scores; effect size (Cohen's d) is negative if in the direction of decline in the STN group, or positive if it is in the direction of improvement.
P-values are from one-sided Anova corrected for age, years of education, amount of dopaminergic medication, and disease duration.
COWAT = Controlled Oral Word Association Test; PASAT = Paced Auditory Serial Addition Test; AVLT = Auditory Verbal Learning Test; GIT = Groningen Intelligence Test.

Mood and behaviour questionnaires
Results on mood and behaviour are presented in table 3. At baseline the STN group showed significantly more symptoms of tension and fatigue on the POMS, more negative affect on the PANAS, and lower quality of life compared to the control group.[12] Again, the change scores were not skewed (skewness varied from -0.62 to 0.39), while the correlations between the change scores were low (median r = 0.15; 25th and 75th centile: 0.08 and 0.30). However, the distribution of change scores was not normal for many of the NPI items.

When we covaried for group differences at baseline in dopaminergic medication, age, disease duration, and years of education, we found that 12 months after surgery the STN group showed a larger improvement in quality of life on the PDQL compared to baseline than the control group (Cohen's d =0.9). Relatives of the STN patients reported a larger increase in dysexecutive symptoms on the DEX at 12 months than the control group. Nonparametric testing of the NPI and additional items showed no group differences (the result of the hallucinations item, although significant, is negligible).

Table 3. *Scores on questionnaires at baseline and change scores (mean, SD) at 12 months follow-up for STN and control groups*

Questionnaire	STN	Control	P values	Cohen's d
Montgomery Asberg Depression Rating scale	6.8 (4.6)	5.5 (4.5)		
change score	0.3 (5.9)	-2.0 (3.7)	0.03	-0.5
POMS depression	5.7 (6.1)	4.9 (5.5)		
change score	-0.3 (5.3)	-0.5 (4.0)	0.70	-0.1
POMS anger	4.3 (4,4)	4.0 (4.6)		
change score	-0.2 (4.8)	0.6 (4.0)	0.57	0.2
POMS fatigue	8.2 (5.6)	6.7 (5.8)		
change score	-0.5 (6.1)	0.7 (4.3)	0.96	0.3
POMS vigor	11.8 (4.3)	10.9 (4.0)		
change score	-1.0 (4.1)	0.2 (3.1)	0.22	-0.3
POMS tension	9.3 (5.2)	6.5 (4.5)		
change score	-2.6 (4.7)	-0.6 (3.1)	0.09	0.5
PANAS positive affect	33.2 (9.1)	30.5 (11.7)		
change score	-1.5 (10.1)	2.1 (8.2)	0.19	-0.4
PANAS negative affect	20.6 (11.6)	16.0 (11.1)		
change score	-3.6 (10.9)	-0.1 (7.9)	0.23	0.3
DEX self	20.9 (10.6)	17.8 (11.3)		
change score	-1.1 (9.8)	0.5 (6.8)	0.76	0.2
DEX proxy	19.4 (11.9)	18.0 (12.3)		
change score	0.9 (10.6)	-1.6 (9.3)	0.02	-0.3
MAC self	76.5 (10.9)	74.2 (13.0)		
change score	-2.7 (9.1)	-1.5 (10.4)	0.29	0.1
MAC proxy	76.7 (13.5)	76.7 (11.3)		
change score	-4.0 (11.0)	-4.6 (10.3)	0.60	-0.1
Parkinson Quality of Life	101.2 (20.8)	85.3 (23.3)		
change score	-16.0 (21.6)	-0.1 (14.6)	0.02	0.9
CES-d	13.8 (7.4)	13.7 (7.7)		
change score	-0.0 (7.3)	-0.2 (5.2)	0.19	0.0

NPI delusions	0.2 (1.2)	0.0 (0.0)		
change score	0.2 (1.2)	-0.1 (0.7)	0.77	-0.3
NPI hallucinations	0.1 (0.5)	0.0 (0.2)		
change score	0.0 (0.5)	0.2 (1.0)	0.03	0.3
NPI agitation/aggression	0.4 (1.4)	0.3 (0.8)		
change score	0.0 (1.6)	0.1 (0.9)	0.63	0.1
NPI dysphoria	1.2 (1.9)	0.9 (1.3)		
change score	0.0 (2.2)	0.1 (1.2)	0.27	0.1
NPI anxiety	0.6 (1.4)	0.6 (1.5)		
change score	-0.1 (1.8)	0.2 (1.5)	0.72	0.2
NPI euphoria	0.0 (0.0)	0.00 (0.0)		
change score	-0.1 (0.5)	0.0 (0.2)	0.13	0.3
NPI apathy	1.0 (1.9)	0.9 (1.8)		
change score	0.2 (2.0)	0.0 (1.6)	0.99	-0.1
NPI disinhibition	0.5 (1.1)	0.1 (0.5)		
change score	0.3 (1.3)	0.1 (0.8)	0.46	-0.2
NPI irritability/lability	1.0 (1.7)	0.8 (1.7)		
change score	0.4 (2.1)	-0.2 (1.0)	0.30	-0.4
NPI aberrant motor behavior	0.4 (1.5)	0.2 (0.9)		
change score	0.1 (1.9)	0.0 (1.0)	0.69	-0.1
NPI sleep disorder	2.6 (2.1)	2.9 (2.9)		
change score	0.1 (2.7)	-0.6 (2.3)	0.20	-0.2
NPI appetite change	2.0 (2.9)	0.8 (2.3)		
change score	0.2 (3.9)	0.0 (2.7)	0.97	-0.1
Negligence	0.3 (1.2)	0.1 (0.4)		
change score	0.1 (1.2)	-0.1 (0.4)	0.96	-0.2
Sexual changes	1.1 (2.5)	0.9 (2.3)		
change score	0.1 (2.5)	0.5 (3.1)	0.46	0.1
Language/speech changes	5.0 (3.6)	3.4 (2.4)		
change score	1.2 (3.4)	0.5 (2.7)	0.20	-0.1
Cognitive changes	2.9 (3.3)	1.8 (2.1)		
change score	1.3 (3.3)	0.2 (1.6)	0.12	-0.4

The effect size (Cohen's d) is negative if in the direction of decline in the STN group, or positive if it is in the direction of improvement.

P-values are from two-sided ANOVAs corrected for age, years of education, amount of dopaminergic medication, and disease duration. P-values for NPI and additional items are from Mann-Whitney tests (two-sided), and are not corrected for covariates.

Individual outcome
The multivariate normative comparisons method revealed that 38 out of the 105 STN patients (36%; 95% CI: 27 - 46) showed cognitive decline, i.e. a pattern of test results that deviated significantly and in a negative direction from the control group (one-sided test). Furthermore, 16 patients (15%; 95% CI: 8 – 22) reported a pattern of mood changes that deviated in a positive sense from controls, and 16 patients reported a pattern of mood changes that deviated in a negative sense from controls (two-sided multivariate normative comparisons). Seventy-nine percent of the patients had a positive PDQL change score after surgery. Univariate comparison showed a

substantial improvement in quality of life in 34 of the 105 STN patients (32%; 95% CI: 22 – 40). Nine patients reported improved quality of life in spite of showing cognitive decline. Change in cognition was not related to change in quality of life ($chi^2 = 1.66$, df = 1, p = 0.20).

Predictors of cognitive decline

In the logistic regression model with cognitive decline determined by the multivariate normative comparisons method as dependent variable we entered (stepwise forward) the following independent variables: levodopa response, age, and composite scores of speed, attention, and verbal fluency. Cognitive decline after surgery was significantly associated with attention at baseline (OR 0.41 [90% CI: 0.28-0.62]), as well as with age (OR 1.07 [90% CI: 1.01-1.13]), and levodopa response (OR 0.97 [90% CI: 0.94-0.99]). The correlations between the predictors were low (r<|0.13|), and multi-collinearity was not a problem (variance inflation factors < 1.03). The results indicate that patients with impaired attention at baseline, advanced age, or with a low levodopa response at baseline had a higher risk of cognitive decline after surgery. The expected chances of cognitive decline for the ages of 55 and 65 years, given the z-scores on attention and given levodopa responses of 30%, 50% and 70% are depicted in figure 1.

Predictors of quality of life

We entered levodopa response, prior stereotactic surgery, and Mattis' DRS score into the regression analysis as independent variables and significant improvement in quality of life as the dependent variable. The odds ratio of levodopa response associated with improvement of quality of life was 1.04 (90% CI 1.02-1.06). The chances of improvement of quality of life as a function of levodopa response are depicted in figure 2.

Correlations with medication and depression

Cognitive decline in the STN group did not correlate with postoperative decrease in medication (r=0.01) and correlated slightly with postoperative increase of depressive symptoms on the MADRS (r=0.19). Cognitive decline correlated slightly with an increase in dysexecutive symptoms on the DEX (r=0.19).

Discussion

One year after STN DBS we found a large improvement in quality of life. At the individual level, 79% of the patients reported at least some improvement, while quality of life improved substantially in at least one third of the patients. Cognitive decline was found in a similar proportion of patients (36%). Some patients (9%) reported improved quality of life despite cognitive decline. Worsening of mood and improvement of mood were seen in comparable proportions (both 15%). Levodopa response at baseline predicted improvement of quality of life. Attention at baseline, age and also levodopa response predicted cognitive decline after surgery. Cognitive decline after STN DBS does not seem to be a temporary effect from which patients recover in the long term, because 12 months after surgery the effect sizes of most cognitive changes have become even larger (median -0.5) compared to the six months follow-up (median -0.3).[12]

The severity of cognitive decline ranged from slight decline that was not even noticed by the patients themselves to severe decline. For advanced PD patients who often have borderline cognitive impairments, however, even a slight cognitive decline may imply a transition into evident impairments.

From reviews in the literature [5] the picture may arise that patients decline after STN DBS on some neuropsychological tests and improve on others. In almost all studies, these changes are based on group comparisons. When individual outcome is evaluated, it is often defined as the percentage of patients that improved or declined more than a standard deviation on single tests. However, this approach suffers from a very high number of false positives. [8] The multivariate normative comparisons method is more reliable, since it controls the number of false positives at 5%. With this method we examined the profile of test performances of each of the STN patients compared to the controls. This way, slight declines on several tests can be just as important as a large decline on a single test, because score decrements that are not significant at univariate statistical testing contribute to the overall result of the multivariate test. Probably, because of the statistical method we used, and because of the comprehensive way in which we assessed our patients, the percentage of cognitive decline in our study is higher than previously described.[6] In essence, the application of this new statistical method, in combination with the use of a control group and the large sample size, is what distinguishes ours from previous studies. An advantage of this approach is that it facilitates the search for predictors because it detects more patients who have declined. A recent study that applied the usual univariate method in a similar cohort could not find predictors because it detected an insufficient number of patients as declined.[38]

Our findings are basically in agreement with what others have found. This can be seen, for example, in the decline of verbal fluency, which is the most consistent finding in this field. According to a recent meta-analysis [6] the mean decline of verbal fluency after STN DBS is 0.64 SD (95% CI 0.32 – 0.96). In our STN patients category fluency declined 4.6 words, which equals an effect size of 0.46 (i.e. 4.6/9.9; see table 2). The effect size of phonemic fluency in our patients is even smaller (4.1 words; ES = 0.32). Also, just like in almost all other studies, the decline in other cognitive domains than verbal fluency was small in our patients. With respect to demographic and disease characteristics our patients were very comparable to, for example, the patients who were included in a large clinical trial by Deuschl and colleagues.[4] Also, the amount of motor improvement, the percentage of major adverse events, and the neuropsychological findings in our study are comparable to those in the Deuschl study.[4;39]

The major limitation of our study is that it is not a randomized trial. Differences between STN patients and control patients in cognitive decline could be due to differences in demographic or disease characteristics. The STN group had three years longer disease duration than the control group and used more levodopa medication at baseline (table 1), implying more severe PD symptoms, which led to their recruitment in the surgical group. However, there were no differences in motor symptoms in the on phase. Moreover, the baseline differences did not explain a significant portion of variance in neuropsychological change scores. A second limitation is the fact that standardized on/off evaluation at 12 months was only performed in a subsample. Another limitation is that it is hard to decide when exactly a patient's quality of life has improved. Any cut-off value must be arbitrary, because there is no gold standard. A final but important limitation is the sample size of our study. Although it had the largest number of patients compared to other studies on neuropsychological aspects of DBS until now, this number is still quite small for establishing stable predictive models. Moreover, the models implicitly assume linear relations between predictors and outcome, but they may as well have a different form. The models as illustrated in the figures should be considered only crude approximations. Future studies should include more patients to find stable predictors of outcome and to assess their precise characteristics. However, it remains of importance that a comprehensive set of neuropsychological variables are studied, because subtle but clinically relevant changes will not be detected with coarse measures. Moreover, cognitive change was not restricted to verbal fluency or psychomotor speed, but was also seen on several other measures, i.e. tests of memory and attention.

It has been suggested that cognitive decline after STN DBS might be associated with apathy resulting from postoperative decrease in dopaminergic medication.[7 40] Apathy might result in an appearance of cognitive decline which is not actually there. However, this mechanism cannot explain our results, because the postoperative decrease in dopaminergic medication was not related to cognitive decline. Another hypothesis is that STN DBS stimulates the limbic area leading to an increase in depression, and subsequently to slight cognitive decline. We found a weak relationship between cognitive decline and postoperative increase in depressive symptoms, but even if depressed mood were a causative influence, it would only account for 4% of the variance in cognitive outcome. Finally, the idea that microrecording is related to cognitive decline because of the multiple tracks entering the brain leading to microlesioning,[5] was not supported by our results (microrecording was not a significant predictor of decline).

We conclude that cognitive decline is fairly common after STN DBS, despite the improvement in quality of life. Our study points out that best candidates for STN DBS are relatively young patients without impairments in attention, and with a high levodopa response. However, further replication studies in other samples are needed to validate our prognostic findings. Our results should be taken as an ingredient that doctors and patients may use when they are balancing the pro's and con's of the intervention.

Acknowledgment

We thank M. Hoogman, R. ten Dijke, A. Rienstra, and E. Wijnalda (psychologists) for the neuropsychological testing of several patients. We thank M. Postma, L. Daeter (Parkinson nurses AMC), R.M.A. de Bie, R.A.J. Esselink (neurologist UMC Nijmegen), B. Post (neurologists AMC) and A. Portman (neurologist UMC Groningen) for the scoring of motor functioning of the patients. We thank J. Kuster (neurologist from Kennemer Gasthuis Haarlem) and. J. Zijlmans (neurologist Ignatius ziekenhuis Breda) for recruitment of Parkinson patients for the control group. W.A. van Gool and R.J. de Haan made valuable comments on an earlier version of the paper.

Competing interests: The study was supported by a grant from the Prinses Beatrix Fonds and the Parkinson Patiëntenvereniging. J.D. Speelman acts as an independent consultant for Medtronic Ltd. (Minneapolis). The movement disorders group of the AMC has received an unrestricted research grant from Medtronic Ltd. No other financial interests or other potential conflicts are involved.

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in JNNP and any other BMJPGL products to exploit all subsidiary rights, as set out in our licence.

Figure legends

Figure 1

Predicted chances of cognitive decline one year after STN DBS as a function of pre-operative attention functioning for relatively young patients (55 years, upper panel) and relatively old patients (65 years, lower panel), and for three levels of pre-operative levodopa challenge test result (30%, 50% and 70% improvement in UPDRS3 score from off to on phase).
Attention is expressed as mean z-score of the Trailmaking test part B and the Stroop colour-word test interference condition, after correction for age and education. Cognitive decline is established by neuropsychological evaluation and multivariate normative comparison (see text; $p < 0.05$ compared to PD control patients).

Figure 2

Predicted chances of substantial improvement in quality of life (PDQL, $p < 0.05$ compared to PD control patients) as a function of pre-operative levodopa challenge test result (percent improvement in UPDRS3 score from off to on phase).

Reference List

1. Krack P, Batir A, Van Blercom N, Chabardes S, Fraix V, Ardouin C, Koudsie A, Limousin-Dowsey P, Benazzouz A, LeBas JF, Benabid A, Pollak P. Five-year follow-up of bilateral stimulation of the subthalamic nucleus in advanced Parkinson's disease. *N.Engl.J.Med.* 2003;**349**:1925-34.

2. Esselink R, De Bie R, de Haan RJ, Lenders M, Nijsen P, Staal MJ, Smeding H, Schuurman PR, Bosch DA, Speelman JD. Unilateral pallidotomy versus bilateral subthalamic nucleus stimulation in PD A randomized trial. *Neurol* 2004;**62**:201-7.

3. Rodriguez-Oroz MC, Obeso JA, Lang AE, Houeto JL, Pollak P, Rehncrona S, Kulisevsky J, Albanese A, Volkmann J, Hariz MI, Quinn NP, Speelman JD, Guridi J, Zamarbide I, Gironell A, Molet J, Pascual-Sedano B, Pidoux B, Bonnet AM, Agid Y, Xie J, Benabid AL, Lozano AM, Saint-Cyr J, Romito L, Contarino MF, Scerrati M, Fraix V, Van BN. Bilateral deep brain stimulation in Parkinson's disease: a multicentre study with 4 years follow-up. *Brain* 2005;**128**:2240-9.

4. Deuschl G, Schade-Brittinger C, Krack P, Volkmann J, Schafer H, Botzel K, Daniels C, Deutschlander A, Dillmann U, Eisner W, Gruber D, Hamel W, Herzog J, Hilker R, Klebe S, Kloss M, Koy J, Krause M, Kupsch A, Lorenz D, Lorenzl S, Mehdorn HM, Moringlane JR, Oertel W, Pinsker MO, Reichmann H, Reuss A, Schneider GH, Schnitzler A, Steude U, Sturm V, Timmermann L, Tronnier V, Trottenberg T, Wojtecki L, Wolf E, Poewe W, Voges J, the German Parkinson Study Group NS. A Randomized Trial of Deep-Brain Stimulation for Parkinson's Disease. *N Engl J Med* 2006;**355**:896-908.

5. Voon V, Kubu C, Krack P, Houeto JL, Troster AI. Deep brain stimulation: neuropsychological and neuropsychiatric issues. *Mov Disord.* 2006;**21 Suppl 14**:S305-S327.

6. Parsons TD, Rogers SA, Braaten AJ, Woods SP, Troster AI. Cognitive sequelae of subthalamic nucleus deep brain stimulation in Parkinson's disease: a meta-analysis. *Lancet Neurol.* 2006;**5**:578-88.

7. Saint-Cyr JA, Trépanier LL, Kumar R, Lozano AM, Lang AE. Neuropsychological consequences of chronic bilateral stimulation of the subthalamic nucleus in Parkinson's disease. *Brain* 2000;**123**:2091-108.

8. Huizenga HM, Smeding H, Grasman RP, Schmand B. Multivariate normative comparisons. *Neuropsychologia* 2007;**45**:2534-42.

9. Smeding HM, Esselink RA, Schmand B, Koning-Haanstra M, Nijhuis I, Wijnalda EM, Speelman JD. Unilateral pallidotomy versus bilateral subthalamic nucleus stimulation in PD--a comparison of neuropsychological effects. *J.Neurol.* 2005;**252**:176-82.

10. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurol* 1967;**17**:427-42.

11. Mattis, S. Dementia Rating Scale professional manual. 1973. Odessa, Florida, USA, Psychological Assement Resources, Inc.

12. Smeding HM, Speelman JD, Koning-Haanstra M, Schuurman PR, Nijssen P, van LT, Schmand B. Neuropsychological effects of bilateral STN stimulation in Parkinson disease: a controlled study. *Neurol* 2006;**66**:1830-6.

13. Saint-Cyr JA, Trépanier LL. Neuropsychological Assessment of Patiens for Movement Disorder Surgery. *Movement Disorders* 2000;**15**:771-83.

14. Lucas JA, Ivnik RJ, Smith GE, Bohac DL, Tangalos EG, Kokmen E, Graff-Radford NR, Petersen RC. Normative data for the Mattis Dementia Rating Scale. *J.Clin.Exp.Neuropsychol.* 1998;**20**:536-47.

15. Luteijn, F. and van der Ploeg, FAE. Groninger intelligentie test (Dutch Intelligence test). 1998. Lisse, the Netherlands, Swets & Zeitlinger.
Ref Type: Report

16. Rosen WG. Verbal fluency in aging and dementia. *Journal of Clinical neuropsychology* 1980;**2**:135-46.

17. Dowmes JJ, Sharp HM, Costall BM, Sagar HJ, Howe J. Alternating fluency in Parkinson's disease. *Brain* 1993;**116**:887-902.

18. Schmand, B., Lindeboom, J., and van Harskamp, F. Nederlandse leestest voor volwassenen (Dutch adult reading test). 1992. Lisse, the Netherlands, Swets & Zeitlinger.

19. Gronwall D. Paced Auditory Serial-Addition Task: a measure of recovery from concussion. *Perceptual and Motor Skills* 1977;**44**:367-73.

20. Rey, A. L' examen clinique en psychologie. 1964. Paris, France, Presses Universitaires de France.

21. Stroop JR. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 1935;**18**:643-62.

22. Flowers KA, Robertson C. The effect of Parkinson's disease on the ability to maintain a mental set. *J Neurol Neurosurg Psychiatry* 1985;**48**:517-29.

23. Reitan, RM. Trail Making Test Manual for administration and scoring. 1992. Tucson, AZ, Reitan Neuropsychological Laboratory.

24. Kaplan, E., Goodglass, H., and Weintraub, S. Boston Naming Test. 1983. Philadelphia, Lea & Febiger.

25. Wilson, BA., Alderman, N., Burgess, PW., Emslie, H., and Evans, JJ. Behavioural Assessment of the Dysexecutive Syndrome. 1996. Bury St. Edmunds, UK, Thames Valley Test Company.

26. Crook III TH, Larrabee GJ. A self-rating scale for evaluating memory in everyday life. *Psychol Aging* 1990;**5**:48-57.

27. Ponds R, Jolles J. Memory complaints in elderly people: the role of memory abilities, metamemory, depression, and personality. *Educational Gerontology* 1996;**22**:341-57.

28. Beekman A, Deeg D, van Limbeek J, Braam AW, De Vries MZ, Van Tilburg W. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D) : results from a community-based sample of older subjects in the Netherlands. *Psychological Medicine* 1997;**27**:231-5.

29. 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**:70-4.

30. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. *Neurol* 1994;**44**:2308-14.

31. Kat MG, de Jonghe JF, Aalten P, Kalisvaart CJ, Droes RM, Verhey FR. [Neuropsychiatric symptoms of dementia: psychometric aspects of the Dutch Neuropsychiatric Inventory (NPI)]. *Tijdschr.Gerontol.Geriatr.* 2002;**33**:150-5.

32. Kertesz A, Davidson W, Fox H. Frontal behavioral inventory: diagnostic criteria for frontal lobe dementia. *Can.J.Neurol.Sci.* 1997;**24**:29-36.

33. McNair, DM., Lorr, M., and Droppleman, LF. EDITS manual for the Profile of Mood States. 1981. San Diego,CA, Educational and Industrial Service.

34. Watson D, Clark LA. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology* 1988;**54**:1063-70.

35. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *British Journal of Psychiatry* 1979;**134**:382-9.

36. Zakzanis KK. Statistics to tell the truth, the whole truth, and nothing but the truth: formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Arch.Clin.Neuropsychol.* 2001;**16**:653-67.

37. Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology* 1990;**1**:43-6.

38. Tir M, Devos D, Blond S, Touzet G, Reyns N, Duhamel A, Cottencin O, Dujardin K, Cassim F, Destee A, Defebvre L, Krystkowiak P. Exhaustive, one-year follow-up of subthalamic nucleus deep brain stimulation in a large, single-center cohort of parkinsonian patients. *Neurosurgery* 2007;**61**:297-304.

39. Witt K, Daniels C, Reiff J, Krack P, Volkmann J, Pinsker MO, Krause M, Tronnier V, Kloss M, Schnitzler A, Wojtecki L, Botzel K, Danek A, Hilker R,

Sturm V, Kupsch A, Karner E, Deuschl G. Neuropsychological and psychiatric changes after deep brain stimulation for Parkinson's disease: a randomised, multicentre study. *Lancet Neurol.* 2008;**7**:605-14.

40. Funkiewiez A, Ardouin C, Krack P, Fraix V, Van Blercom N, Xie J, Moro E, Benabid A, Pollak P. Acute Psychotropic Effects of Bilateral Subthalamic Nucleus Stimulation and Levodopa in Parkinson's Disease. *Movement Disorders* 2003;**18**:524-30.

Figure legends

Figure 1
Predicted chances of cognitive decline one year after STN DBS as a function of pre-operative attention functioning for relatively young patients (55 years, upper panel) and relatively old patients (65 years, lower panel), and for three levels of pre-operative levodopa challenge test result (30%, 50% and 70% improvement in UPDRS3 score from off to on phase).
Attention is expressed as mean z-score of the Trailmaking test part B and the Stroop colour-word test interference condition, after correction for age and education. Cognitive decline is established by neuropsychological evaluation and multivariate normative comparison (see text; $p < 0.05$ compared to PD control patients).

Figure 2
Predicted chances of substantial improvement in quality of life (PDQL, $p < 0.05$ compared to PD control patients) as a function of pre-operative levodopa challenge test result (percent improvement in UPDRS3 score from off to on phase).
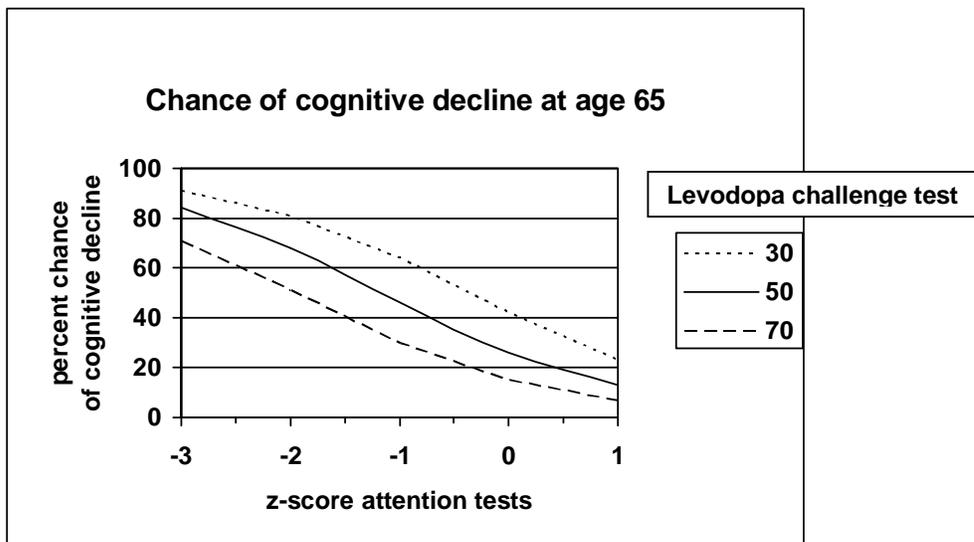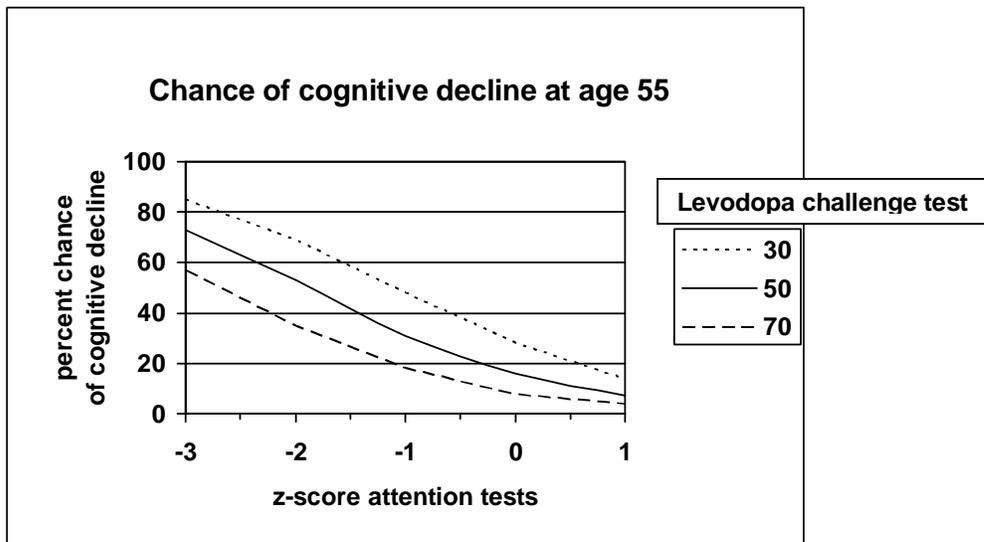
Chance of cognitive decline at age 55

Levodopa challenge test

.... 30
——— 50
– – – 70

z-score attention tests

percent chance of cognitive decline

Chance of cognitive decline at age 65

Levodopa challenge test

.... 30
——— 50
– – – 70

z-score attention tests

percent chance of cognitive decline

Figure 1.

Figure 2

**Chance of improvement in Quality of Life as a function of
levodopa challenge test result**