Cognition following bilateral implants of embryonic dopamine neurons in PD

A double blind study

C.T. Trott, PhD; S. Fahn, MD; P. Greene, MD; S. Dillon, RN; H. Winfield, RN; L. Winfield, RN; R. Kao; D. Eidelberg, MD; C.R. Freed, MD; R.E. Breeze, MD; and Y. Stern, PhD

Abstract—*Objectives:* To determine if bilateral transplantation of embryonic mesencephalic dopamine cells into the putamen of patients with PD significantly affected their cognitive functioning when compared with patients receiving sham surgery and to examine the effect of age on cognitive performance after implantation. *Methods:* Forty patients (19 women, 21 men; age 34 to 75 years) with idiopathic PD of at least 7 years' duration (mean 14 years) who had disabling motor signs despite optimal drug management were randomly assigned to tissue implants or sham craniotomies in a double-blind design. Neuropsychological tests assessing orientation, attention, language, verbal and visual memory, abstract reasoning, executive function, and visuospatial and construction abilities were administered before and 1 year after surgery. Treatment groups did not differ at baseline in demographic, neuropsychological, motor, depression, or levodopa equivalent measures. *Results:* Postsurgical change in cognitive performance was not significantly different for real or sham surgery groups. Performance in both groups remained unchanged at follow-up for most measures. *Conclusions:* Embryonic dopamine producing neurons can be implanted safely into the putamen bilaterally without impairing cognition in patients with PD, but within the first year, improved cognition should not be expected.

NEUROLOGY 2003;60:1938–1943

Following transplantation of fetal mesencephalic dopaminergic neurons into the striatum, a number of researchers have reported a reduction in the severity of motor symptoms in MPTP-induced parkinsonism[1] and in PD.[2] Although several reports of fetal implantation have included information on cognition after the surgery,[3] the data are limited and the results inconclusive. All of the previous studies of fetal implants were unblinded open trials, most had only a few patients in each, and none had a sham surgery control group.

Results of the first double-blind, sham surgery–controlled trial of bilateral implantation of embryonic dopamine neurons into the putamen of patients with PD were recently published.[4] Surgery carries the risk of further damage to the subcortical/cortical pathways that support cognition. Two basal ganglia thalamic cortical loops that are thought to be critical for cognition involve the caudate nucleus. Even though dopaminergic cells were transplanted into the putamen in this study, the growth of fetal tissue and the increased production of dopamine carries the possibility of improved cognition. This article reports on the performance of the same patients on a battery of neuropsychological tests pre and post surgery and looks at changes in cognitive performance after 1 year as a function of real vs sham surgery and as a function of age. The purpose of this study was to determine if the transplantation of dopaminergic cells into the putamen of patients with PD significantly affected their cognitive functioning when compared with patients receiving sham surgery. Because patients were stratified by age,[4] a second goal was to examine the effect of age on cognitive performance after implantation.

Methods. *Subjects.* A total of 40 patients (19 women, 21 men) participated in the study. Ages ranged from 34 to 75 years at the time of enrollment; the majority (34) were white. All patients had idiopathic PD with at least two of the following symptoms: bradykinesia, rigidity, resting tremor. The average duration of illness was 14 years (range 7 to 21 years). Potential patients were excluded if they had a Mini-Mental State Examination[5] score less than 24 out of 30, hallucinations or delusions while on levodopa, major depression, dementia, other severe medical disease (e.g., diabetes mellitus, epilepsy, cerebrovascular disease), previous brain surgery, or evidence on MRI of any other neurologic disorder. One 66-year-old woman, who was in the real surgery group, was killed in a traffic accident 7 months after surgery.[6] Her data were not included in any analysis. The real surgery group had 10 younger and 9 older patients; the sham surgery group, 11 younger and 9 older patients. On the basis of their performance on the neuropsychological test battery described below, three patients (one younger, real surgery; two older, one each in real and sham surgery) were judged independently by two of the authors (C.T.T. and Y.S.) to have developed dementia[7] by the 1-year follow-up and were therefore excluded from some analyses as indicated. Dementia was assessed on the basis of performance on neuropsychological tests using criteria described in a previous publication.[7]

From the Cognitive Neuroscience Division (Drs. Trott and Stern), G.H. Sergievsky Center, and Department of Neurology (Drs. Fahn, Greene, and Stern, and S. Dillon, H. Winfield, L. Winfield, and R. Kao), Columbia University, New York, NY; North Shore University Hospital (Dr. Eidelberg), Manhasset, NY; and Department of Neurosurgery (Drs. Freed and Breeze), University of Colorado Medical Center, Denver.

Supported by RO1 NS32368, T32 AG00261, and the Parkinson's Disease Foundation.

Received December 7, 2001. Accepted in final form March 9, 2003.

Address correspondence and reprint requests to Dr. Yaakov Stern, Sergievsky Center, 630 W 168 St., New York, NY 10032; e-mail: ys11@columbia.edu

Procedures. Patients underwent a number of procedures before and after surgery. Only those relevant to the current analyses are described below. Further procedural details, including those pertaining to the surgeries, can be found in a previous publication.[6] The sham surgical procedure consisted of similar twist-drill craniotomies as in the real surgery, but the dura mater was not penetrated.

Neuropsychological evaluation. Neuropsychological testing was done on two successive days before and 1 year after surgery. Patients were on antiparkinson medication at times of testing. All but one patient had at least a high school education and were fluent in English. One younger, Spanish-speaking patient in the real surgery group had only 3 years of education. Tests were administered in Spanish for this patient and in English for all others. Estimated level of intellectual functioning was obtained from the North American Adult Reading Test.[8]

The following test battery was administered on both visits:

Ten items from the Mini-Mental State Examination[5] were used to assess orientation to time and place.

Time to completion and omissions on the diamond shape from the Cancellation Test[9] were used to assess attention.

Visuospatial matching was assessed using the 10 designs of Form C from the Benton Visual Retention Test (BVRT).[10] Patients matched a large design to one of four smaller designs.[11]

Construction ability was assessed by having patients copy five drawings of varying difficulty from the Rosen Drawing Test.[12] Number correct was the measure.

Immediate verbal memory was assessed with the forward digit span from the Wechsler Adult Intelligence Scale–Revised (WAIS-R).[13] Working memory was assessed with the backward digit span from the WAIS-R, and reordering span and subtraction span tests.[14] The reordering span was the number of digits a patient could mentally reorder and repeat in numerical order. The subtraction span was the number of digits a patient could give after subtracting two from each.

Learning and verbal memory were assessed with the six-trial, 12-word Bushke Selective Reminding Test (SRT),[15] and with the California Verbal Learning Test (CVLT).[16] The SRT total recall was used to assess short-term memory. Long-term recall was used as a measure of learning. Long-term storage was used to assess encoding ability, and consistent long-term retrieval to assess retrieval efficiency. Long-term verbal recall was assessed 15 minutes after completing the SRT. Recognition of words that were not recalled was then tested using four-choice arrays. All patients learned the CVLT, 16-item List A and its alternate form. One list was presented with words blocked into their categories, and the other with categories intermixed. The order of presentation of the two lists and conditions (blocked vs unblocked) was counterbalanced across patients. For each list, there were five immediate recall trials, and one 20-minute delayed recall. The T score for trials one through five, and the number of words retrieved on the delay trial, were used as measures of learning and memory for blocked and unblocked presentations.

A multiple-choice version of the BVRT[11] was used to assess immediate nonverbal memory. The patient studied a design for 10 seconds. Then it was removed and the patient was asked to recognize the design in a four-choice array. The 10 designs from Form D of the original test[15] were used.

Language repetition was assessed using the high-frequency items from the Repeating Phrases subtest of the Boston Diagnostic Aphasia Evaluation (BDAE).[17] Comprehension of verbal material was assessed using the first six items of the Complex Ideational Material subtest of the BDAE.[17] Naming was assessed using 15 items of varying difficulty from the Boston Naming Test (BNT).[18]

Verbal fluency was assessed using both the Controlled Oral Word Association test (COWA)[19] and animal naming (BDAE).[17] For the COWA, the letters in English were C, F, and L, and in Spanish they were P, S, and V. Percentile scores were based on age- and education-adjusted norms for the COWA, and age norms for animal naming.

Two tests were used to assess abstract reasoning: the Similarities subtest of the WAIS-R[13] and the Identities and Oddities subtest of the Mattis Dementia Rating Scale (DMS).[20] The Odd Man Out (OMO)[21] and a modified version of the Wisconsin Card Sorting Test (WCST)[22] assessed executive function. For the modified WCST, patients were asked to choose a sorting criterion. After correctly sorting 10 cards according to that criterion, they were asked to choose a different sorting principle. The third time they were asked to chose a sorting criterion that was different from either of the previous two. This procedure was repeated for three additional trials of 10 correct sorts or until all the cards were exhausted. The number of categories achieved was used as the outcome measure on this test.

Clinical evaluation. Severity of parkinsonism was evaluated with the Unified PD Rating Scale (UPDRS).[23] Measurements were obtained when patients were at their best ("on") and worst ("off"), both on and off optimal medication. Scores used in the following analyses were the average of assessments on several days, obtained during practically defined "off" states[24] for each patient at baseline and follow-up. The Hoehn & Yahr[25] and the Schwab & England[26] scales were administered during the same periods. Depression was assessed at baseline and follow-up using the Hamilton Rating Scale for Depression (HRS-D).[27]

PET scans. Fluorodopa PET scans were obtained at baseline and follow-up. A measure of cell growth (ratio of putamen or caudate to occipital uptake of [^{18}F]fluoro-L-dopa) was used in some of the analyses.

Data analysis. Demographic, depression, motor, medication, and cognitive differences between treatment groups at baseline were assessed using one-way analysis of variance (ANOVA) for all patients, followed by ANOVA for each age group. The significance of between-group changes in depression was assessed using a repeated measure ANOVA (test time $\times$ treatment group $\times$ age group). To describe changes in test scores from baseline to follow-up, means and SD at baseline were computed for each cognitive measure within each age group. Using the baseline means and SD, Z scores were computed for all cognitive measures for each patient at baseline and 1-year follow-up. Differences in Z scores from baseline to follow-up testing were then obtained. When there were multiple measures per test, Z score differences on the most global measure were reported, resulting in 21 scores for each patient. The significance of change in performance on the neuropsychological tests from baseline to 1-year follow-up was assessed using test time $\times$ treatment group $\times$ age group repeated measure ANOVA for each of 29 measures with education as a covariate. Significant interactions were analyzed with one-way ANOVA and paired sample *t*-tests. To investigate further the differences between real and sham surgery groups at follow-up, change scores were created by subtracting baseline scores from scores at follow-up for neuropsychological measures. Change scores were also created for the levodopa equivalent and the UPDRS, the Hoehn & Yahr, the Schwab & England, and the depression scales. Dopamine uptake change scores were created by subtracting the ratio of putamen to occipital and caudate to occipital uptake of [^{18}F]fluoro-L-dopa at baseline from the ratios at 1-year follow-up. Differences between treatment groups in the neuropsychological change scores were then assessed with multivariate ANOVA (change score $\times$ treatment group $\times$ age group) with education as a covariate. One analysis was computed using the 32 patients who had scores on all 29 measures. This did not include any demented patients. The second analysis included all 39 patients and the 13 measures on which all had scores. The third included all 36 nondemented patients and the 24 measures on which all had scores. Significant interactions were analyzed with one-way ANOVA. When possible, analyses were first calculated with all patients included, followed by analyses excluding the three patients who became demented. Level of significance was set at 0.01. In order to evaluate the relationship of cognitive change to other aspects of patients' functioning, 29 neuropsychological change scores were correlated (two-tailed) with change scores for two motor measures (total UPDRS and Hoehn & Yahr scores), the Schwab & England assessment of activities of daily living (ADL), the HRS-D, four dopamine uptake change scores (ratio of uptake in the right and left putamen and the right and left caudate), and the levodopa equivalent change score. Because of the large number of correlations, some were likely to be significant by chance. Because this research was exploratory, *p* values were reported when correlations were greater than 0.001. Correlations were calculated for each treatment group, and for age and treatment stratified groups with and without the demented patients.

Results. Treatment groups did not differ at baseline on age ($p = 0.950$), education ($p = 0.297$), estimated IQ ($p = 0.059$), duration

Table 1 *Number of tests on which performance changed one SD or more*

Subject, real surgery	Declined	Improved	Subject, sham surgery	Declined	Improved
Younger subjects					
1	2	4	1	1	1
2	3	3	2	2	1
3	5	3	3	2	1
4	3	4	4	5	0
5	1	1	5	1	1
6	0	3	6	1	5
7	1	2	7	10	1
8	6	2	8	2	2
9	4	0	9	1	4
10	1	3	10	2	3
			11	10	1
Older subjects					
1	2	1	1	0	2
2	5	3	2	2	2
3	4	0	3	1	2
4	7	1	4	3	1
5	2	1	5	4	3
6	3	1	6	7	0
7	7	5	7	1	4
8	0	1	8	2	1
9	13	0	9	2	1

There were a total of 21 tests.

of illness ($p = 0.831$), depression ($p = 0.249$), or levodopa equivalent ($p = 0.956$). When analyzed by age group, older patients in the sham surgery group were somewhat more educated than those in the real surgery group ($p = 0.015$). No other differences approached significance. There also were no significant differences between treatment groups on any of the baseline neuropsychological or motor measures whether analyzed by overall groups or stratified by age.

There was no change in depression scores from baseline to 1-year follow-up ($p = 0.762$). One younger patient was mildly depressed at baseline (Hamilton = 12) and one older patient was moderately depressed at follow-up (Hamilton = 17). Both were in the real surgery group. Performance of patients on the majority of neuropsychological measures did not change from baseline to 1-year follow-up. When a patient's performance on a test changed one SD or more from baseline to follow-up, scores declined more often than they improved (table 1). There was no difference in the number of tests that improved or declined as a function of treatment group. Performance of patients who became demented did not decline on more tests than did the performance of other patients. None of the patients who showed the most decline were demented.

The three patients who became demented did, however, have lower baseline scores and the most missing data at follow-up (two, three, and nine tests) owing to their inability to perform the tasks. Two older patients in the real surgery group also had missing data on one or two tests owing to inability to perform. Data were missing for one test each on two older patients in the sham surgery group for other reasons. One score was missing on each of the following tests: orientation, BNT, BDAE repetition, BDAE comprehension, reordering span, and subtraction span. Two scores were missing on WCST, CVLT, shape time & omit, and Identities and Oddities.

Test time × treatment group × age group repeated measure analyses revealed no main effects of treatment group. There was a main effect of age for Identities and Oddities ($p = 0.008$). A follow-up one-way ANOVA indicated an age group difference at 1 year postsurgery ($p = 0.009$) but not at baseline ($p = 0.280$). Paired sample t-tests examining the change from baseline to follow-up for each age group indicated no change for either age group separately (younger: $t = -0.389$, $p = 0.701$; older: $t = 1.370$, $p = 0.191$). After removing the demented patient who had scores on this measure from the analyses, the main effect of age remained ($p = 0.014$), but not the age difference at follow-up ($p = 0.02$). As would be expected, younger nondemented patients (mean = 15.5, SD = 1.1) performed better than older nondemented patients (mean = 14.7, SD = 1.6). There was a within-patient test time × age group interaction ($p = 0.000$) for category fluency. Follow-up ANOVA indicated an age difference at baseline ($p = 0.005$), but not at follow-up ($p = 0.795$). Paired sample t-tests revealed a decrease for the younger ($t = 4.217$, $p = 0.000$) but no change for the older group ($t = -1.2$, $p = 0.247$). Treatment groups did not differ for either age group at baseline or follow-up. Removing the patients who became demented from the analyses did not change the results. There were within-patient test time × treatment group × age group interactions for Subtraction Span ($p = 0.002$) and Rosen drawings ($p = 0.008$). Follow-up analyses revealed only a decline in the Rosen score from baseline to follow-up for older patients in the real surgery group ($t = 3.5$, $p = 0.008$). When the three patients who became demented were removed from the analyses, neither of the three-way interactions was significant. Test performance declined from baseline (mean = 9.5, SD = 0.7) to follow-up (mean = 9.3, SD = 1.1) on orientation ($p = 0.011$) and improved from baseline (mean = 7.8, SD = 0.5) to follow-up (mean = 7.9, SD = 0.3) on repetition ($p = 0.004$). When the demented patients were removed from these analyses, neither change remained significant: orientation baseline (mean = 9.5, SD = 0.7) to follow-up (mean = 9.4, SD = 0.99) [$F(1,31) = 0.424$]; repetition baseline (mean = 7.8, SD = 0.5) to follow-up (mean = 7.9, SD = 0.3) [$F(1,31) = 0.070$] (table 2).

Multivariate analyses of neuropsychological change scores revealed no main effects of treatment group in any of the three analyses. There was a main effect of age for category fluency in all three analyses, confirming the findings of the repeated measure analysis. The performance of the younger group declined (mean = −23.2, SD = 25.2) while that of the older group improved (mean = 5.9, SD = 21). In the analysis of all 39 patients with 13 measures, there was a treatment group × age group interaction for Rosen drawings ($p = 0.008$). However, as in the repeated measure analysis, when the demented patients were removed, the interaction was no longer significant ($p = 0.023$). When the analysis was computed with all 36 nondemented patients on 24 measures, there was a treatment group × age group interaction for the SRT delayed recognition test ($p = 0.002$). One-way ANOVA indicated no difference between treatment groups for either age group (younger, $p = 0.134$; older, $p = 0.070$) or treatment group (real surgery group, $p = 0.024$; sham surgery group, $p = 0.170$). Absolute performance declined from baseline to follow-up in the young sham (mean = −0.6, SD = 1.1) and older real surgery groups (mean = −0.9, SD = 0.99) while it improved in the young real (mean = 0.1, SD = 0.6) and older sham groups (mean = 0.3, SD = 1.3). No significant differences in change scores were found for the UPDRS, Hoehn & Yahr, or Schwab & England scales.

The 29 neuropsychological change scores were then correlated with two motor, one ADL, one depression, four dopamine uptake, and one levodopa equivalent change scores. In general, negative correlations indicate a positive relationship between measures—i.e., better performance on the cognitive test was found with better performance on the other measure. However, positive correlations with the Schwab & England and with the ratio of cell growth indicate a positive relationship with the neuropsychological task. Seven correlations reached levels greater than 0.001 in the real surgery group and two in the sham surgery group. In the real surgery group, the significant correlations were as follows: unblocked CVLT with UPDRS ($r = -0.653$) and with Schwab & England ($r = 0.620$), blocked CVLT T Score with Schwab & England ($r = 0.593$), Identities and Oddities with UPDRS ($r = -0.639$) and with Hoehn & Yahr ($r = -0.636$), Rosen drawings with Schwab & England ($r = 0.591$), and CFL with Schwab & England ($r = 0.580$). In the sham surgery group, the significant

Table 2 *Neuropsychological test means and SD for all patients in each treatment group at baseline and 1-year follow-up*

		Baseline		1-Year follow-up	
Test	Measure	Real	Sham	Real	Sham
Mini-Mental State Examination (orientation)	Total correct	9.4 (0.7)	9.6 (0.7)	9.2 (1.4)	9.4 (0.9)
Cancellation test	Diamond time, s	65.1 (20.9)	64.6 (33.1)	70.5 (23.8)	65.4 (28.4)
	Diamond omissions	3.8 (3.6)	2.8 (2.1)	3.6 (4.0)	4.4 (3.9)
BDAE—repetition	Total correct	7.7 (0.6)	7.9 (0.5)	7.8 (0.4)	8.0 (0.0)
BDAE—comprehension	Total correct	5.6 (0.6)	5.9 (0.4)	5.2 (1.4)	5.7 (0.6)
Boston Naming Test	Total correct	14.2 (1.4)	14.9 (0.3)	14.1 (1.7)	14.9 (0.4)
COWA (CFL)	Percentile	69 (29)	68 (26)	60 (38)	59 (31)
Animal naming	Percentile	43 (33)	54 (26)	35 (31)	43 (30)
WAIS-R similarities	Age scaled score	10.7 (3.2)	11.8 (2.1)	10.6 (3.0)	12.0 (1.7)
WCST	Number of categories	4.7 (1.9)	5.4 (1.5)	4.9 (1.9)	4.9 (2.2)
Identities & Oddities	Total correct	15.4 (1.1)	15.0 (1.3)	14.6 (1.9)	15.3 (1.3)
Odd Man Out	Total correct	32.4 (7.7)	34.5 (5.3)	31.8 (8.8)	33.0 (6.4)
WAIS-R DSF	Span	6.8 (1.3)	7.4 (1.5)	6.5 (1.5)	6.8 (1.6)
WAIS-R DSB	Span	4.95 (1.4)	5.15 (1.2)	5.37 (1.6)	5.00 (1.2)
Reordering span	Span	3.3 (0.5)	3.5 (0.6)	3.3 (0.7)	3.7 (0.7)
Subtraction span	Span	4.4 (1.5)	4.7 (1.2)	4.0 (1.9)	4.7 (1.1)
Bushke SRT	Total recall	46.7 (12.0)	46.7 (10.0)	42.0 (13.7)	44.8 (13.6)
	LTR percentile	35.0 (16.3)	34.6 (14.3)	29.5 (15.4)	32.4 (18.2)
	LTS percentile	38.0 (31.7)	39.7 (31.3)	28.7 (28.3)	34.1 (33.2)
	CLTR percentile	35.5 (28.5)	36.4 (29.8)	23.8 (27.5)	33.0 (32.9)
	Delayed recall	6.7 (3.1)	7.1 (2.9)	6.2 (2.7)	5.4 (3.8)
	Delayed recognition	11.4 (1.4)	11.5 (0.8)	11.4 (1.1)	11.2 (1.2)
CVLT (blocked)	T score trials 1–5	58.2 (18.1)	62.0 (11.3)	56.9 (20.2)	58.7 (17.0)
	No. correct—delay	12.3 (3.3)	13.1 (2.2)	12.0 (3.2)	11.74 (3.6)
CVLT (unblocked)	T score trials 1–5	45.8 (14.1)	45.5 (13.6)	41.8 (18.8)	42.6 (16.7)
	No. correct—delay	10.4 (4.3)	10.4 (3.5)	9.2 (4.6)	10.0 (3.7)
BVRT (delay)	Total correct	8.6 (1.8)	8.7 (1.3)	7.8 (2.0)	8.2 (2.0)
BVRT (matching)	Total correct	9.5 (1.2)	9.8 (0.6)	9.1 (1.9)	9.5 (1.4)
Rosen Drawing Test (five items)	Total correct	3.2 (0.9)	3.5 (0.9)	2.6 (1.3)	3.2 (1.2)

Values are mean (SD).
BDAE = Boston Diagnostic Aphasia Examination; COWA = Controlled Oral Word Association Test (CFL); WAIS–R = Wechsler Adult Intelligence Scale–Revised; WCST = Wisconsin Card Sort Test; DSF = Digit Span Forward; DSB = Digit Span Backward; SRT = Selective Reminding Test; LTR = Long Term Retrieval; LTS = Long Term Storage; CLTR = Consistent Long Term Retrieval; CVLT = California Verbal Learning Test; BVRT = Benton Visual Retention Test.

correlations were as follows: OMO with the ratio of dopamine uptake left in caudate ($r = 0.609$) and subtraction span with Hoehn & Yahr ($r = -0.611$). When the correlations were repeated in the age by treatment stratified groups, the only correlations that reached a significance level greater than 0.001 occurred with the older real surgery group. Without the demented patient these included: unblocked CVLT delayed recall with the UPDRS ($r = -0.883$), unblocked CVLT T-scores with Schwab & England ($r = 0.858$), CFL with Schwab & England ($r = 0.879$), and Rosen with Hamilton Depression ($r = -0.842$). When the demented patient was included, all of these except CFL remained significant. In addition, unblocked CVLT T-scores correlated with UPDRS ($r = -0.827$). The above correlations suggest that performance on these neuropsychological tests of learning and memory, and verbal fluency may be positively correlated with motor and ADL scores for older patients in the real surgery group. In addition, performance on a drawing test may be positively correlated with change scores on the depression inventory.

Discussion. In this study, the postsurgical change in cognitive performance was not significantly different for patients who received bilateral implantation of embryonic dopamine producing neurons into the putamen and those who had sham surgery. Neither treatment group showed a significant change in performance on most measures. This was true for the group as a whole, as well as for each age group separately. When performance did change, it was more likely to decline in all groups. This is what

could be expected as a result of the disease process itself. The three patients who became demented did not have a precipitous decline in cognitive performance, and their dementia was considered to be a result of the disease process. There was the suggestion that improved motor (UPDRS) and daily (Schwab & England) functioning may be positively correlated with improvement in some aspects of cognitive functioning (CVLT, CFL) in the older patients who received the neural transplants. A decrease in scores on the depression inventory may also correlate with better performance on a drawing test (Rosen) in this group. However, because the majority of patients were not clinically depressed, the significance of this correlation is not clear. In addition, because of the large number of correlations, these results can only be considered suggestive.

Previous reports of the cognitive performance of patients following fetal implantation have found no change[28] and improved[3] abilities. Because none of these studies had a control group of patients, the changes in performance could be due to other factors such as variability between patients in cognitive symptoms and progression of the disease, daily fluctuations in functioning, and test practice effects. By comparing patients who had implantation with a control group who had sham surgery, the current study was able to show that changes in cognitive performance were not significantly different for the two groups. This suggests that whatever cognitive changes occurred in individual patients were due to factors other than the surgery and cell growth. Inclusion of the control group did not rule out the possibility of differences between treatment groups due to local neurobiologic changes at the site of implantation because there was no disruption to the subcortical or cortical tissue in the sham surgery group. However, because there were no cognitive differences between the treatment groups, this issue is moot.

In the current research, no significant differences between treatment groups were seen 1 year after surgery on the global rating scale, diary scores, drug doses, total UPDRS, or Schwab & England scores when patients were on medication. (Presurgery levels of medication were maintained postsurgery to the extent possible.) The only significant findings on UPDRS and Schwab & England scores were when patients were off medication. In that case, younger patients with real surgery improved more than those with sham surgery.[4] Given these findings, it may not be surprising that there were no significant differences between treatment groups on neuropsychological test measures after surgery because the cognitive testing was carried out while patients were on medication. The combination of medication and dopamine replacement can increase the variability of the scores on outcome measures, thus reducing the probability of finding significant group differences.[4] However, when the variability of change scores on a particular cognitive test was large for one group, it tended to be equally large for the other group. Thus, the implanted cell dopamine did not significantly increase variability compared with the sham surgery group. Nevertheless, the large variability in both groups on several test change scores could have reduced the ability to detect significant between-group differences.

There is some evidence that motor and cognitive deficits in PD may be dissociable,[29] and that the basal ganglia thalamic cortical loops involved in cognition may be separate from those involved in motor activity.[30] Two loops that are thought to be critical for cognition involve the caudate nucleus, with projections to the dorsolateral prefrontal and lateral orbitofrontal cortex. Postmortem studies of patients with PD show that neuron depletion is much greater in the putamen than in the caudate.[31] In the patients in the current study, the uptake of ^{18}F-fluorodopa was more diminished in the putamen than in the caudate before surgery. Because the dopamine producing neurons were implanted into the putamen in this trial, it is possible that the increased level of dopamine was not sufficient to have a significant impact on cognition via the caudate loops. It is possible that a longer time frame would result in levels of dopamine great enough to influence cognition significantly via the caudate loops.

The effect of dopamine on cognition in PD is not entirely clear. Various studies have reported treatment with levodopa to improve, impair, and not affect various aspects of cognition in patients with PD. The effect of levodopa on cognition in PD "may be a function of the level of dopamine depletion in different parts of the basal ganglia and the prefrontal cortex."[32(p365)] The ability of levodopa to improve cognition in PD appears to occur in the early stage of the disease, but seldom returns the patient to premorbid levels of cognitive functioning. When patients are stabilized medically, additional levodopa appears not to improve cognitive performance. When the "wearing off" motor response to levodopa occurs, additional levodopa can impair cognition.[32] The patients in the current study had had the disease for at least 7 years before surgery, and were taking clinically optimal levels of medication. Thus, according to the above findings, an increase in available dopamine would not necessarily be expected to result in a change in cognition.

Although the primary pathology in PD involves the dopaminergic system, other neurotransmitter systems (i.e., cholinergic,[33] noradrenergic[34]) are also involved. Several studies have shown the influence of anticholinergics on executive functions,[35] memory,[36] and attention[37] in PD. It is likely that cognitive deficits in PD are the result of some interaction of the dopaminergic, cholinergic, and noradrenergic systems and their pathology.[38]

Possible explanations for the lack of treatment group differences include 1) the large variability in some test change scores, 2) a dissociation between motor and cognitive basal ganglia thalamic frontal loops, 3) the lack of responsiveness of cognition to

increased dopamine in stabilized patients, and 4) the influence and pathology of cholinergic and noradrenergic neurotransmitter systems on cognition.

References

1. Widner H, Tetrud J, Rehncrona S, et al. Fifteen months follow-up on bilateral embryonic mesencephalic grafts in two cases of severe MPTP-induced parkinsonism. In: Narabayashi H, Nagatsu T, Yanagisawa N, Mizuno Y, eds. Advances in Neurology, vol. 60. New York: Raven Press, 1993:729–733.
2. Hauser RA, Freeman TB, Snow BJ, et al. Long-term evaluation of bilateral fetal nigral transplantation in Parkinson disease. Arch Neurol 1999;56:179–187.
3. Sass KJ, Buchanan CP, Westerveld M, et al. General cognitive ability following unilateral and bilateral fetal ventral mesencephalic tissue transplantation for treatment of Parkinson's disease. Arch Neurol 1995; 52:680–686.
4. Freed CR, Greene PE, Breeze RE, et al. Transplantation of embryonic dopamine neurons for severe Parkinson's disease. N Engl J Med 2001; 344:710–719.
5. Folstein MF, Folstein SE, McHugh PR. "Mini-mental State": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198.
6. Freed CR, Breeze RE, Greene PE, et al. Double-blind controlled trial of human embryonic dopamine cell transplants in advanced Parkinson's disease: study design, surgical strategy, patient demographics and pathological outcome. Neurology 1999;52(suppl 2):A272–A273. Abstract.
7. Stern Y, Andrews H, Pittman J, et al. Diagnosis of dementia in a heterogeneous population: development of a neuropsychological paradigm-based diagnosis of dementia and quantified correction for the effects of education. Arch Neurol 1992;49:453–460.
8. Blair JR, Spreen O. Predicting premorbid IQ: a revision of the National Adult Reading Test. Clin Neuropsychol 1989;3:129–136.
9. Sano M, Rosen W, Mayeux R. Attention deficits in Alzheimer's disease. Washington, DC: American Psychiatric Association, 1984.
10. Benton AL. The Visual Retention Test. New York: The Psychological Corporation, 1955.
11. Sivan AB, Spreen O. Der Benton-Test, 7th ed. Berne, Switzerland: Verlag Hans Huber, 1906.
12. Rosen W. The Rosen Drawing Test. Bronx, NY: Veterans Administration Medical Center, 1981.
13. Wechsler D. Wechsler Adult Intelligence Scale–Revised. New York: The Psychological Corporation, 1981.
14. Salthouse TA. The role of processing resources in cognitive aging. In: Howe ML, Brainerd CJ, eds. Cognitive development in adulthood. New York: Springer-Verlag, 1988:185–239.
15. Bushke H, Fuld PA. Evaluating storage, retention, and retrieval in disordered memory and learning. Neurology 1974;24:1019–1025.
16. Delis DC, Massman PJ, Kaplan E, McKee R, Kramer JH, Gettman D. Alternate form of the California Verbal Learning Test: development and reliability. Clin Neuropsychol 1991;5:154–162.
17. Goodglass H, Kaplan E. The assessment of aphasia and related disorders, 2nd ed. Philadelphia: Lea & Febiger, 1983.
18. Kaplan E, Goodglass H, Weintraub S. The Boston Naming Test. Philadelphia: Lea & Febiger, 1983.
19. Benton AL, Hamsher K. Multilingual Aphasia Examination. Iowa City, IA: AJA Associates, 1989.
20. Mattis S. Mental status examination for organic mental syndrome in the elderly patient. In: Bellak L, Karasu TB, eds. Geriatric psychiatry. New York: Grune & Stratton, 1976:77–121.
21. Flowers K, Robertson C. The effect of Parkinson's disease on the ability to maintain a mental set. J Neurol Neurosurg Psychiatry 1985;48:517–529.
22. Heaton RK. Wisconsin Card Sorting Test manual. Odessa, FL: Psychological Assessment Resources, 1981.
23. Fahn S, Elton RL, UPDRS Development Committee. The Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Calne DB, Goldstein M, eds. Recent developments in Parkinson's disease. Florham Park, NJ: Macmillan Healthcare Information, 1987;2:153–163, 293–304.
24. Langston JW, Widner H, Goetz CG, et al. Core assessment program for intracerebral transplantations (CAPIT). Mov Disord 1992;7:2–13.
25. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology 1967;17:427–442.
26. Schwab RS, England AC. Projection technique for evaluating surgery in Parkinson's disease. In: Gillingham FJ, Donaldson IML, eds. Third symposium on Parkinson's disease. Edinburgh, Scotland: Livingstone, 1969:152–157.
27. Hamilton M. Rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–61.
28. Lindvall O, Rehncrona S, Brundin P, et al. Human fetal dopamine neurons grafted into the striatum in two patients with severe Parkinson's disease. Arch Neurol 1989;46:615–631.
29. Cooper JA, Sagar HJ, Jordan N, Harvey NS, Sullivan EV. Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. Brain 1991;114:2095–2122.
30. Middleton FA, Strick PL. Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. Brain Cogn 2000;42: 183–200.
31. Kish SJ, Shannak K, Hornykiewicz O. Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. N Engl J Med 1988;318:876–880.
32. Kulisevsky J. Role of dopamine in learning and memory. Implications for the treatment of cognitive dysfunction in patients with Parkinson's disease. Drugs Aging 2000;16:365–379.
33. Dubois B, Ruberg M, Javoy-Agid F, Ploska A, Agid Y. A subcortico-cortical cholinergic system is affected in Parkinson's disease. Brain Res 1983;288:213–218.
34. Stern Y, Mayeux R, Cote L. Reaction time and vigilance in Parkinson's disease. Possible role of altered norepinephrine metabolism. Arch Neurol 1984;41:1086–1089.
35. Bedard MA, Pillon B, Dubois B, Duchesne N, Masson H, Agid Y. Acute and long-term administration of anticholinergics in Parkinson's disease: specific effects on the subcortico-frontal syndrome. Brain Cogn 1999;40:289–313.
36. Cooper JA, Sagar HJ, Doherty SM, Jordan N, Tidswell P, Sullivan EV. Different effects of dopaminergic and anticholinergic therapies on cognitive and motor function in Parkinson's disease. Brain 1992;115:1701–1725.
37. Dubois B, Pillon B, Lhermitte F, Agid Y. Cholinergic deficiency and frontal dysfunction in Parkinson's disease. Ann Neurol 1990;28:117–121.
38. Levin ED, McGurk SR, Rose JE, Butcher LL. Cholinergic-dopaminergic interactions in cognitive performance. Behav Neural Biol 1990;54:271–299.

Cognition following bilateral implants of embryonic dopamine neurons in PD: A double blind study

C. T. Trott, S. Fahn, P. Greene, et al.

Neurology 2003;60;1938-1943
DOI 10.1212/01.WNL.0000070181.28651.3B