ORIGINAL ARTICLE

Effects of Perceived Treatment on Quality of Life and Medical Outcomes in a Double-blind Placebo Surgery Trial

Cynthia McRae, PhD; Eva Cherin, MA; T. Gayle Yamazaki, PhD; Gretchen Diem, PhD; Alexander H. Vo, PhD; Dan Russell, PhD; J. Heiner Ellgring, PhD; Stanley Fahn, MD; Paul Greene, MD; Sandra Dillon, RN; Hal Winfield, RN; Kimberly B. Bjugstad, PhD; Curt R. Freed, MD

Context: This study was part of a large double-blind sham surgery–controlled trial designed to determine the effectiveness of transplantation of human embryonic dopamine neurons into the brains of persons with advanced Parkinson's disease. This portion of the study investigated the quality of life (QOL) of participants during the 1 year of double-blind follow-up.

Objectives: To determine whether QOL improved more in the transplant group than in the sham surgery group and to investigate outcomes at 1 year based on perceived treatment (the type of surgery patients thought they received).

Design: Participants were randomly assigned to receive either the transplant or sham surgery. Reported results are from the 1-year double-blind period.

Setting: Participants were recruited from across the United States and Canada. Assessment and surgery were conducted at 2 separate university medical centers.

Participants: A volunteer sample of 40 persons with idiopathic Parkinson's disease participated in the transplant ("parent") study, and 30 agreed to participate in the related QOL study: 12 received the transplant and 18 received sham surgery.

Interventions: Interventions in the parent study were transplantation and sham brain surgery. Assessments of QOL were made at baseline and 4, 8, and 12 months after surgery.

Main Outcome Measures: Comparison of the actual transplant and sham surgery groups and the perceived treatment groups on QOL and medical outcomes. We also investigated change over time.

Results: There were 2 differences or changes over time in the transplant and sham surgery groups. Based on perceived treatment, or treatment patients thought they received, there were numerous differences and changes over time. In all cases, those who thought they received the transplant reported better scores. Blind ratings by medical staff showed similar results.

Conclusions: The placebo effect was very strong in this study, demonstrating the value of placebo-controlled surgical trials.

Arch Gen Psychiatry. 2004;61:412-420

A RELATIVELY NEW EXPERIMENTAL approach to treating Parkinson's disease (PD) is the stereotactic implantation of human embryonic dopamine neurons into the brains of persons with severe PD. Unblinded clinical trials[1-21] have shown that this surgical procedure ameliorates some of the symptoms of the disease and that the transplanted dopamine neurons survive. Very little research has focused on subjective changes in quality of life (QOL) that may result from the surgery.[21]

To assess the efficacy of fetal transplant surgery, the procedure was subjected to a double-blind sham surgery–controlled trial, with half the patients receiving the transplant and half receiving sham surgery.[22] The study design not only included the unusual condition of sham surgery, but it also maintained the double-blind phase for 1 year, a very long time for placebo-controlled trials.[23] These conditions combined to provide unique opportunities to investigate the effects of sham surgery as well as the benefits of transplant surgery and to determine the effects of a 12-month double-blind trial. One of the primary goals of the "parent" study was to determine whether the signs and symptoms of PD improved more in transplant recipients than in those who received the sham surgery. Similarly, one of

Author affiliations are listed at the end of this article.

the goals of this (QOL) study was to determine whether QOL improved more in the transplant group than in the sham surgery group. A related goal was to investigate outcomes at 1 year based on perceived treatment, or the type of surgery patients thought they received. The final goal of this study was to examine medical outcomes based on actual as well as perceived treatment to determine whether perceived treatment was related to ratings by medical staff at 12 months.

METHODS

A complete description of the parent study can be found in the article by Freed et al.[22] The following subsections describe the methods of the QOL study.

PATIENTS

Forty persons from across the United States and Canada were recruited to participate in this double-blind surgical trial. Twenty patients were randomly assigned to receive the fetal tissue transplant and 20 patients were randomly assigned to the sham surgery condition. Patients were told that those in the sham surgery group would have the option of receiving the transplant after the blind was lifted. Inclusion criteria for the study were (1) a diagnosis of idiopathic PD of at least 7 years' duration; (2) a continuing response to the administration of levodopa, the principal drug used to treat PD; and (3) the presence of an intractable problem, such as "off" periods, dyskinesias, or "freezing" not controlled by dopamine agonist therapy. Exclusion criteria included (1) obvious depression or cognitive impairment (as assessed by a neuropsychologist), (2) previous brain surgery, or (3) the presence of diabetes mellitus, severe cardiopulmonary disease, other severe medical disease, or magnetic resonance imaging evidence of cerebrovascular disease. After rigorous screening, patients who were accepted into the parent study were invited to participate in the QOL portion of the project. Thirty patients agreed to participate. All neurologic examinations were performed at Columbia-Presbyterian Medical Center (CPMC).

QOL Evaluation of Patients

Quality of life is a multidimensional construct that was measured using a variety of widely recognized, commonly used instruments focusing on 3 fundamental aspects of QOL: Physical, Emotional, and Social functioning. Data were collected by sending questionnaires to patients 1 week after scheduled evaluation visits to CPMC. Participants were asked to respond to the questionnaire and return it in the postage-paid envelope provided. Patients were paid $25 for each completed questionnaire. The response rate was 98%. The scales contributing to each aspect of QOL are presented in the following subsections.

Physical Functioning. *Unified Parkinson's Disease Rating Scale* (*patient version*). This scale[24] is an adaptation of the physician version of the Unified Parkinson's Disease Rating Scale (UPDRS) that is uniformly used to rate the physical functioning abilities and problems of patients with PD.[25] The patient version was developed for a health promotion program that was disseminated through the mail.[24] Four primary subscales were used in this version of the UPDRS: Activities of Daily Living (ADLs) at "best" and "worst" and Severity of Symptoms at "best" and "worst." Correlations of the patient version with a variety of the UPDRS subscale scores rated by medical staff range from $r=0.58$ to $r=0.71$.[26] Because previous results[22] have shown that scores "off" medications provide a more valid assessment of patient status, only the "off" scores (worst) from the patient UPDRS were included in the composite variable. Lower scores indicate better outcomes. The estimate of reliability (Cronbach's α) for the "worst" scores was .84 at baseline.

"Free or Restricted". This single, global item measures how free or restricted the person feels "in doing what you want to do." A Likert scale ranging from 1 (I still do everything I want to do) to 7 (I can no longer do the things I want to do) is used. Lower scores indicate better outcomes.

Emotional Functioning. *Parkinson's Disease Stress Scale*. This scale was developed for use with German patients with PD and is available to interested readers on request (J. H. Ellgring, PhD, M. Macht, PhD, R. Schwarz, MA, unpublished data, 1993). The scale has 19 items, with lower scores indicating less stress. Cronbach's α for the scale was .77.

Center for Epidemiologic Studies–Depression Scale. This scale[27] is a 20-item self-report measure of depressive symptoms. Each item is rated on a 4-point scale related to frequency; for example, 0 indicates "less than 1 day" and 3 indicates "5 to 7 days." Lower scores indicate fewer depressive symptoms. The estimate of reliability (Cronbach's α) was .89.

State-Trait Anxiety Inventory. The "State" version[28] of this well-validated instrument was designed to assess the current, situational aspects of anxiety. The 20 items are rated on a 4-point scale ranging from 1 (not at all) to 4 (very much). Lower scores indicate less anxiety. Cronbach's α for the scale was .89.

Intrusiveness of Illness Scale. This scale[29] is an adaptation of one designed to measure the degree to which a chronic illness interferes with usual life activities. A scale ranging from 1 (very little) to 7 (a great deal) was used for each item. Lower scores on this 17-item scale indicate less intrusiveness of illness. The estimated reliability of the scale (Cronbach's α) was .88.

Social Functioning. *Social Provisions Scale*. This 24-item scale[30] assesses perceived social support. Each item is rated on a 4-point scale ranging from 1 (strongly disagree) to 4 (strongly agree). Lower scores indicate less perceived support. The reliability estimate (Cronbach's α) at baseline was .91.

Social Contact. A measure of social contact was developed for this study to assess the amount of social interaction or activity experienced by each participant. Three items assessing frequency of socializing with friends, telephone communication, and participation in public activities were scored on a 6-point scale ranging from 1 (not at all) to 6 (every day). Lower scores indicate less social contact. Estimated reliability (Cronbach's α) was .69 at baseline.

Medical Evaluation of Patients

Patients were evaluated at the Irving Center for Clinical Research at CPMC twice before surgery and 4, 8, and 12 months after surgery. Each assessment lasted 3 to 4 days, allowing for evaluations on and off medications. Each patient was followed by a neurologist and by 1 of 2 research nurses (S.D. and H.W.) throughout the study. Medical staff were unaware of treatment group assignment for the duration of the study.

Global Rating Scale. The primary outcome variable of the parent study was a single item representing a subjective global rating of change in severity of disease.[22] Possible scores ranged from –3 (much worse since surgery) to +3 (much improved since surgery). One week before the follow-up evaluations at 4, 8, and 12 months, patients filled out global ratings of their health status relative to their condition just before surgery. Medical

staff completed the same evaluation at each patient visit. No reliability or validity information is available for this scale.

Perceived Surgery. Seven days after surgery and 1 week before the follow-up evaluations at 4, 8, and 12 months, patients were asked to indicate whether they thought they received the transplant or sham surgery. These evaluations were mailed to patients by the staff at CPMC. Evaluations were returned to the biostatistician at CPMC and were not revealed to the medical staff.

Unified Parkinson's Disease Rating Scale. The UPDRS[25] is a standard instrument used to assess various aspects of PD, including motor performance, muscle rigidity, tremor, speech, and gait. Lower scores indicate better performance. The UPDRS has high interrater[31] and test-retest[32] reliability.

Schwab and England Activities of Daily Living Scale. The Schwab and England[33] assesses ability to perform activities of daily living on a scale ranging from 0% to 100%, with higher scores indicating more normal performance. Descriptive anchors for the scale are presented in increments of 10, for example, 100, 90, 80, etc. This is a standard assessment instrument in PD and has well-established reliability and validity.[31]

Hoehn and Yahr. The Hoehn and Yahr[34] is the standard disease staging scale for PD. Scores range from 0 (no signs of disease) to 5 (wheelchair bound or bedridden unless aided). Lower scores indicate fewer signs of disease. High interrater reliability between neurologists, patients, and caregivers has been found.[35]

For the purpose of this study, "off" medications was defined as before administration of the first morning dose of levodopa and at least 12 hours after the last administration of levodopa the previous day.[36] As noted earlier, only "off" scores were included in data analyses. Patients and medical personnel completed the Global Rating Scale, the Schwab and England, and the Hoehn and Yahr at each assessment period. The standard version of the UPDRS was completed only by medical personnel.

TRANSPLANTATION METHOD

All patients had a stereotactic frame attached to the head for magnetic resonance imaging to establish the coordinates for the needle tracks for implantation of the tissue. Four burr holes were drilled bilaterally in the forehead to accommodate the needles. Patients were awake during the procedure. Scalp incisions and twist drill holes were done under local anesthesia. Both groups received identical preoperative evaluation and intraoperative sedation and pain control. For patients in the transplant group, implants were made into the putamen with embryonic mesencephalic tissue containing dopamine neurons. Each needle pass into the putamen contained tissue from a single embryo. Four embryos were used for the 4 sites implanted. Needles for participants in the placebo condition remained empty and did not penetrate the brain. No patient received immunosuppressive therapy. The dialogue during the surgery was limited and unscripted. Additional information related to the surgery can be found in the article by Freed et al.[22]

STATISTICAL ANALYSIS

Preliminary analyses were performed to determine relationships among the variables predicted to form the 3 aspects of QOL. Correlations for the Physical functioning variables ranged from $r=0.30$ to $r=0.75$. Relationships among the Emotional functioning variables ranged from $r=0.63$ to $r=0.77$. Reliabilities of the linear composites for these constructs at baseline and at 4, 8, and 12 months ranged from 0.72 to 0.89. The 2 variables predicted to constitute Social functioning did not correlate highly enough to be combined into one composite measure. Neither was the reliability of the linear composite high enough to justify combining the variables. Therefore, measures of perceived Social Support (Social Provisions Scale) and Social Contact were analyzed separately. For the composite variables (Physical and Emotional functioning), scores for each measure were standardized at each period based on the mean and standard deviation of the measure at baseline. Standard scores for each measure were then combined to create a total z score representing the composite variable.

Phi coefficients were calculated to determine the relationship between actual and perceived treatment at 7 days and at 4, 8, and 12 months. Independent samples t tests were used to examine differences between the transplant and sham surgery groups and between the perceived treatment groups on the 3 QOL constructs at each time. t Tests were also used to compare the same groups on the medical outcome variables. The repeated-measures generalized estimating equation was used to examine change over time. Based on the assumption that characteristics of each person are correlated over time, all observations (baseline and 4, 8, and 12 months) were entered simultaneously into these analyses. Analysis of variance techniques were used to examine differences in change scores. All P values are 2-tailed. No adjustments were made for multiple comparisons because of the exploratory nature of the analyses and because there are no comparable data from other studies.

RESULTS

The demographic information for the 30 participants in the QOL portion of the study is presented in **Table 1**. There were no statistically significant differences between the transplant and sham groups on demographic characteristics or QOL variables at baseline.

Membership in the perceived treatment groups changed at each period as patients' perceptions of which treatment they received changed over time. **Table 2** indicates the type of surgery patients thought they received at each period as well as the surgery they actually received. Results show that 22 (76%) of 29 participants thought they had received the transplant 7 days after the procedure. Based on presurgical interviews, this finding seems to reflect the belief of most patients that they would be among those who received the transplant initially. This number was reduced to 10 (33%) of 30 participants 12 months later. Phi coefficients, which measure the relationship between dichotomous variables, ranged from 0.00 to 0.15 and were not statistically significant, indicating no relationship between the type of treatment patients actually received and what they thought they received at any follow-up period.

DIFFERENCES AND CHANGES IN QOL VARIABLES

The results in this section are presented in 2 parts. First, results related to actual transplant and sham surgery groups are presented, followed by results related to perceived treatment groups, or the type of surgery patients thought they received.

Table 1. Baseline Characteristics of Participants

Characteristic	Transplant Group (n = 12)	Sham Surgery Group (n = 18)	Total (n = 30)
Sex, No. (%)			
F	7 (58)	8 (44)	**15** (50)
M	5 (42)	10 (56)	**15** (50)
Age, mean ± SD, y	59.9 ± 7.9	56.3 ± 10.4	57.8 ± 9.5
Education, mean ± SD, y	16.6 ± 2.8	16.3 ± 2.2	16.4 ± 2.4
Duration of disease, mean ± SD, y	15.5 ± 6.6	16.0 ± 3.6	15.7 ± 5.0
Married or living with partner, No. (%)	10 (83)	10 (56)	**20** (67)
Annual income >$40 000, No. (%)	9 (75)	10 (56)	**19** (63)
Currently employed, No. (%)	3 (25)	4 (22)	**7** (23)
Family history of PD, No. (%)	1 (8)	4 (22)	**5** (17)
Other chronic health problems, No. (%)	5 (42)	2 (11)	**7** (23)
Currently smoke, No. (%)	0	1 (6)	**1** (3)
Ever smoke, No. (%)	2 (17)	7 (39)	**9** (30)
Ethnicity, No. (%)			
Asian	1 (8)	0	**1** (3)
Black	0	1 (6)	**1** (3)
Hispanic	0	1 (6)	**1** (3)
White	11 (92)	16 (89)	**27** (90)

Abbreviation: PD, Parkinson's disease.

Table 2. Relationships Between Actual and Perceived Treatment

Actual Treatment	Perceived Treatment: Transplant	Perceived Treatment: Sham Surgery
7 d (Phi = 0.15)		
Transplant	10	2
Sham surgery	12	5
4 mo (Phi = 0.00)		
Transplant	8	4
Sham surgery	12	6
8 mo (Phi = –0.06)		
Transplant	4	8
Sham surgery	7	11
12 mo (Phi = 0.00)		
Transplant	4	8
Sham surgery	6	12

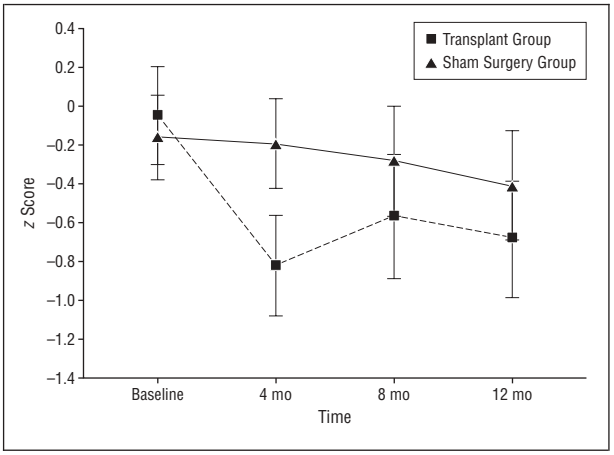

Figure 1. Mean changes in Physical functioning quality of life from baseline to 12 months: actual surgery. Decreased scores indicate improvement. Error bars represent SEM.

Actual Treatment Groups

Results from analyses at baseline and 4, 8, and 12 months indicated that the only significant difference between the transplant and sham surgery groups was in Social Contact at 4 months, with the sham surgery group reporting more social contact ($P=.03$).

Using repeated-measures analyses to examine change over time, there was a significant improvement in Physical functioning in both groups over 12 months ($P=.01$) (**Figure 1**). There was also a significant time × treatment interaction for Physical functioning between the 2 groups over the 4 assessment periods ($P=.04$). As can be seen in Figure 1, marked improvement from baseline to 4 months was seen in the transplant group. There was a downward linear trend indicating improvement over time in the sham surgery group.

Perceived Treatment Groups

When analyses were conducted based on perceived treatment groups, or the treatment patients thought they received at each assessment period, there were several differences between the groups, all favoring the perceived transplant group (**Table 3**). At 8 and 12 months, those who believed they received the transplant reported better Physical functioning than those who thought they received the sham surgery ($P=.02$ and $P=.03$, respectively). Social Support at 8 months ($P=.01$) and 12 months ($P=.03$) was better for those who thought they received the transplant vs the sham surgery.

Using repeated-measures analyses and based on perceived treatment at 12 months, there was improvement in Physical functioning over time in the perceived groups ($P=.01$) (**Figure 2**). A significant interaction between time and perceived group was found for Social Support ($P=.05$), with those who thought they received the sham

Table 3. Differences Between Perceived Groups in Quality-of-Life (QOL) Dimensions

QOL Dimension	4 mo		8 mo		12 mo	
	Transplant (n = 19)	Sham Surgery (n = 10)	Transplant (n = 11)	Sham Surgery (n = 19)	Transplant (n = 10)	Sham Surgery (n = 19)
Physical functioning*	−0.41 (0.85)	−0.24 (1.00)	−0.84 (0.69)‡	0.06 (1.03)	−0.84 (0.54)‡	−0.17 (0.98)
Emotional functioning*	0.15 (0.75)	−0.47 (0.85)	−0.13 (0.65)	0.40 (0.96)	−0.24 (0.68)	0.14 (0.74)
Social Support†	80 (8.0)	77 (9.8)	83 (7.9)§	74 (9.1)	85 (9.9)‡	77 (9.5)
Social Contact†	13 (3.2)	12 (3.6)	14 (2.9)	12 (3.2)	14 (2.2)	12 (2.9)

*Scores for composite variables are standardized (*z* scores), given as mean (SD). Lower scores indicate better functioning.
†Scores for Social Support and Social Contract are raw scores, given as mean (SD), because they are not composite variables. Higher scores indicate more social support or contact.
‡$P<.05$.
§$P<.01$.

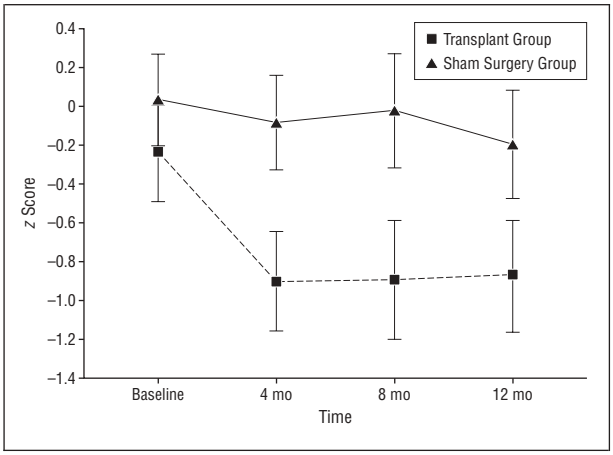

Figure 2. Mean changes in Physical functioning quality of life from baseline to 12 months: perceived surgery. Decreased scores indicate improvement. Error bars represent SEM.

surgery reporting less support at 8 and 12 months than those who thought they received the transplant.

DIFFERENCES AND CHANGES ON MEDICAL EVALUATION

To reduce the number of comparisons, and thus the possibility of type I error, only statistical analyses using the baseline and 12-month assessments were conducted.

Actual Treatment Groups

On ratings by medical staff, there were no differences between the transplant and sham surgery groups at 12 months on the Global Rating Scale, UPDRS and its subscales, Schwab and England, or Hoehn and Yahr ratings. Likewise, there were no differences between groups on medical variables rated by patients (ie, Global Rating Scale, Schwab and England, and Hoehn and Yahr).

Among medical staff, statistically significant changes over time (baseline to 12 months) were found for 2 variables. First, ratings for transplant and sham surgery groups improved for UPDRS tremor scores in the "off" state ($P=.01$). Second, an interaction effect was found on the UPDRS rigidity "off" scores ($P=.04$), with the transplant group improving and the sham surgery group getting worse. There were no changes over time in patients' self-report ratings of the medical variables.

Perceived Treatment Groups

Although blind as to the type of treatment patients actually received, medical personnel reported several differences based on the type of treatment patients thought they received at 12 months. Results presented in **Table 4** show statistically significant differences on several clinical ratings based on patients' perceived treatment. Patient ratings of medical variables also showed statistically significant differences (Table 4). In all cases, results were better for those who thought they received the transplant vs the sham surgery.

In medical staff ratings, there were several changes over time (baseline to 12 months) based on perceived treatment or the treatment patients thought they received at 12 months (**Table 5**). Similar to the actual treatment groups, medical personnel rated both perceived groups as improving on the UPDRS tremor "off" ($P=.02$). On all other measures listed in Table 5, interaction effects indicated that the perceived transplant group was getting better and that the perceived sham surgery group was getting worse. Regarding patient ratings given in Table 5, patients who believed they received the transplant at 12 months showed improvement in global rating scores over time, whereas those in the perceived sham group showed a decline in scores ($P<.001$).

These results are somewhat different from those reported for the parent study[22] because of the difference in sample size (n=39 in the parent study, with 19 persons receiving transplants, vs n=30 in the QOL study, with only 12 receiving transplants). In the parent study,[22] there was a significant difference in improvement (change scores) in UPDRS motor "off" scores between the actual transplant and sham surgery groups ($P=.04$). As shown in **Figure 3**, differences in improvement for the perceived treatment groups were also significant ($P<.05$). In the parent study, scores for Schwab and England "off" improved significantly more for the actual transplant group than for the actual sham group over 12 months ($P=.04$), as did scores for the perceived transplant group relative to the perceived sham surgery group ($P<.01$). Because no statistically

Table 4. Differences Between Perceived Treatment Groups at 12 Months*

Variable	Transplant Group (n = 10)	Sham Surgery Group (n = 20)	*P* Value
	Medical Staff Ratings		
Schwab and England "off" score	63.0 (12.2)	45.8 (21.6)	.03
Global Rating Scale score	1.3 (1.2)	−0.1 (1.1)	<.01
UPDRS score			
Walking "off"	2.7 (1.3)	5.0 (2.8)	.01
Gait "off"	4.5 (2.6)	8.4 (4.1)	<.01
Pull test "off"	1.0 (0.9)	2.3 (1.4)	.01
	Patient Ratings		
Hoehn and Yahr score	2.1 (1.1)	3.3 (0.7)	<.01
Global Rating Scale score	1.8 (1.2)	−1.3 (1.2)	<.001

Abbreviation: UPDRS, Unified Parkinson's Disease Rating Scale.
*Data are given as mean (SD).

Table 5. Changes From Baseline to 12 Months: Perceived Groups

Variable	Transplant Group (n = 10)	Sham Surgery Group (n = 20)	*P* Value
	Medical Staff Ratings*		
UPDRS score			
Tremor "off"	−1.0	−1.0	.02†
Total "off"	−10.5	2.1	.006‡
Motor "off"	−6.2	1.8	.008‡
Bradykinesia "off"	−3.7	1.6	.001‡
Gait "off"	−1.3	1.2	.004‡
Pull test "off"	-0.5	0.3	.02‡
Hoehn and Yahr "off" score	−0.2	0.3	.04‡
Schwab and England "off" score	11.4	−4.8	.004‡
	Patient Ratings*		
Global Rating Scale score	1.9	−0.7	<.001‡

Abbreviation: UPDRS, Unified Parkinson's Disease Rating Scale.
*Scores represent change from baseline to 12 months.
†Time contrast is significant.
‡Interaction contrast is significant.

significant interactions were found in these analyses, results indicate that actual and perceived treatment had independent effects on the outcomes.

COMMENT

One of the primary goals of this study was to determine whether QOL improved more in the transplant group than in the sham surgery group during the 1-year period of the double-blind study. There was only one statistically significant difference between the transplant and sham surgery groups in terms of QOL. Results of comparisons at 4, 8, and 12 months revealed a difference in Social Contact at 4 months, with the transplant group reporting less social interaction than the sham surgery group. Both groups showed statistically significant improvement in Physical functioning during 12-month follow-up (Figure 1). These results, therefore, do not indicate statistically significant improvement in QOL in the transplant group over the sham surgery group by the end of the initial phase of the study.

The reasons for the transplant group reporting less social interaction are unclear. This result may reflect a type I error, or it may be related to the subtle but prolonged recovery period reported by some patients who received the transplant.[37] Regarding both groups reporting improvement in Physical functioning (Figure 1), it is possible that hope or optimism contributed to those results. Individuals desperate enough for relief from symptoms to risk the conditions of this study may have been inclined to imagine themselves getting better over time, particularly when they did not really know what "getting better" might look or feel like.

Because the study involved the unusual element of sham surgery and the double-blind phase lasted approximately 13 months, one of the goals of the study was to investigate the effect of perceived treatment on QOL outcomes. Results indicated that the sustained improvement in Physical functioning from 4 to 12 months among the perceived transplant group (Figure 2) was accompanied by relatively stable scores in Emotional functioning and Social Support over time (both of these remained in the same range as normative scores throughout the 12 months) (Table 3). Conversely, the perceived sham surgery group improved only slightly in Physical functioning over time (Figure 2) and subsequently declined

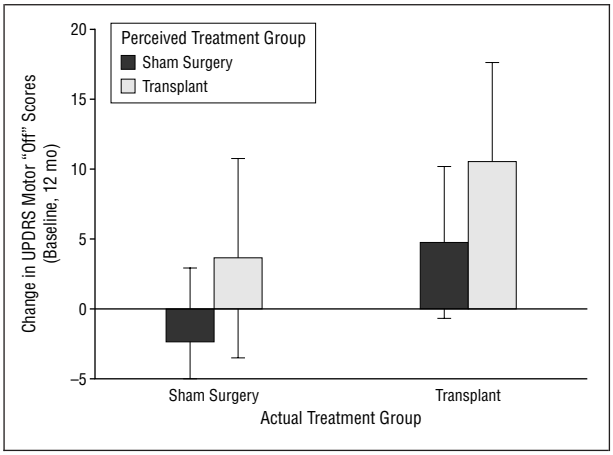

Figure 3. Mean changes in Unified Parkinson's Disease Rating Scale (UPDRS) motor "off" scores (baseline to 12 months) for the total group in the parent study (n=39). Increased scores indicate improvement. Error bars represent SEM.

in Emotional functioning and Social Support at 8 months, with a slight improvement by 12 months (Table 3). Overall, results indicated that perceived treatment was more strongly related to outcome for the duration of the 12-month double-blind study than was the actual treatment they received, with perceived transplant patients consistently reporting better QOL than the perceived sham surgery group. Thus, it seems that the placebo effect was very strong.

The final goal of this study was to examine differences and changes over time in medical outcomes based on actual and perceived treatment. There were no differences on the medical variables as rated by medical staff or patients at 12 months based on actual treatment in the subset of 30 patients. By contrast, the parent study showed statistically significant improvement in UPDRS motor "off" scores and Schwab and England "off" scores in transplant patients, with no change in sham surgery patients.[22]

When analyses were performed based on perceived treatment, several differences on medical variables at 12 months were found (Table 4). All differences, as rated by medical staff and patients, indicated that those who thought they received the transplant were doing better than those who thought they received the sham surgery. Analyses using data from the total group of 39 patients (1 patient died in a car accident before the 12-month evaluation) produced similar results, wherein medical staff rated the perceived transplant group as performing statistically significantly better than the perceived sham group.

Regarding changes over time, most were the result of interaction effects wherein those who thought they received the transplant improved and those who thought they underwent sham surgery got worse (Table 5). Similar results were found in analyses including the total group of participants in the parent study, that is, interaction effects were found.

These results suggest that expectancy regarding which type of treatment patients received had a statistically significant effect not only on subjective parameters (Emotional functioning and Social Support) but also on motor symptoms (Physical functioning). The effects on motor symptoms were also recognized by the medical staff, as shown in their clinical ratings (Tables 4 and 5 and Figure 3). Whether staff ratings of motor symptoms were affected by mood and other subjective expressions of the patients (similar to a "halo" effect) cannot be determined in this study. An alternative explanation is that actual changes in physical functioning led patients to believe that they got the transplant (or not) and resulted in changes in ratings by the medical staff.

Similar results related to the placebo effect have been found in other studies with patients with PD.[38-40] These results underscore the need for placebo controls in studies evaluating treatment for PD as the placebo effect seems to be very strong in this disease.

In terms of the larger issue of the expectancy or placebo effects, results suggest that the possible effectiveness of double-blind conditions in some studies may be longer than previously believed. The average length of time of a double-blind trial is 8 weeks.[23] The double-blind in this study was effectively maintained for 12 months. It seems plausible that because effects of the surgery were not predictable or definitive over the period of the double-blind, patients did not have clear cues as to which surgery they had received. Research[41-43] indicates that the more extreme the placebo treatment is in a clinical trial, the more susceptible participants are to the placebo effect, or believing that they are being helped by the sham medication or condition. Because this study involved brain surgery, arguably an extreme placebo treatment, and results were not clearly discerned by patients or staff for the double-blind period, it seems that conditions were favorable for evoking a strong placebo response in this study. Although this study did not include a standard care group that did not intend to receive the surgery, which would have provided the most rigorous test of the effects of transplant surgery, these results are consistent with a strong placebo effect.

The lack of relationship between actual and perceived treatment throughout the study underscores several important points. First, it indicates the success of the conditions of the sham surgery. Surgical staff performing the surgeries seem to have been convincing in their presentation of the placebo treatment (Table 2). Second, the shift from 22 persons (76%) believing that they got the transplant at 7 days to only 10 (33%) at 12 months indicates that optimism or hope may have influenced perceptions early in the study. The percentage of persons thinking that they received the transplant was similar in the parent study: 28 (72%) at 7 days and 14 (36%) at 12 months. What contributed to the change in ratings during the 12-month period is beyond the scope of this article. We speculate that it may have been perceived changes in physical condition, lack of perceived changes, or diminished optimism as results of the study failed to meet expectations.

In summary, there were more differences and changes over time in QOL among perceived treatment groups than among actual treatment groups. Medical staff, who did not know which treatment each patient received, also reported more differences and changes at 12 months based on patients' perceived treatment than on

actual treatment. These results suggest that patients' perceptions influenced their behavior, which in turn resulted in ratings among medical staff that paralleled patients' self-ratings. It is also possible that subtle changes in physical functioning influenced patients' perceptions, which then influenced behavior. These data cannot be used to determine causal relationships among these indicators of improvement. However, the results clearly show that improvement, or lack of improvement, was not solely determined by actual treatment at this point in the study. Analyses based on the type of surgery patients thought they received produced results that have implications for research related to double-blind placebo-controlled trials and the length of time a blind can be maintained.

Although the sham surgery research design is currently regarded as somewhat controversial and has raised some important ethical concerns,[44] the investigators of the parent study determined that the scientific benefits of this design outweighed potential ethical considerations. The results of this study demonstrate the importance of a double-blind design to distinguish the actual and perceived values of a treatment intervention.

Submitted for publication September 30, 2002; final revision received October 21, 2003; accepted October 29, 2003.

From the College of Education, University of Denver, Denver, Colo (Drs McRae and Yamazaki and Ms Cherin); Behavioral Medicine Department, Kaiser Permanente, Oakland, Calif (Dr Diem); Division of Neurosciences, Walter Reed Army Institute of Research, Silver Spring, Md (Dr Vo); Human Development and Family Studies, Iowa State University, Ames (Dr Russell); Institute of Psychology, University of Wuerzburg, Wuerzburg, Germany (Dr Ellgring); Neurological Institute, Columbia-Presbyterian Medical Center, New York, NY (Drs Fahn and Greene, Ms Dillon, and Mr Winfield); and the Division of Clinical Pharmacology, Department of Medicine, University of Colorado Health Sciences Center, Denver (Drs Bjugstad and Freed).

This study was supported in part by grants R29 NS32009-05 (Dr McRae) and R01 NS32368 (Dr Freed) from the National Institute of Neurological Disorders and Stroke, Bethesda, Md.

Corresponding author: Cynthia McRae, PhD, College of Education, University of Denver, Denver, CO 80208 (e-mail: cmcrae@du.edu).

REFERENCES

1. Freed CR, Breeze RE, Rosenberg NL, Schneck SA, Wells TH, Barrett JN, Grafton ST, Huang SC, Eidelberg D, Rottenberg DA. Transplantation of human fetal dopamine cells for Parkinson's disease: results at 1 year. *Arch Neurol.* 1990;47:505-512.
2. Lindvall O, Brundin P, Widner H, Rehncrona S, Gustavii B, Frackowiak R, Leenders KL, Sawle G, Rothwell JC, Marsden CD, Bjorklund A. Grafts of fetal dopamine neurons survive and improve motor function in Parkinson's disease. *Science.* 1990;247:574-577.
3. Freed CR, Breeze RE, Rosenberg NL, Schneck SA, Schroter G, Lafferty K, Talmage DW, Barrett JN, Wells T, Mazziotta JC, Huang SC, Eidelberg D, Rottenberg DA. Fetal neural implants for Parkinson's disease: results at 15 months. In: Lindvall O, Bjorklund A, Widner H, eds. *Intracerebral Transplantation in Movement Disorders.* Amsterdam, the Netherlands: Elsevier Science Publishers; 1991:69-77. *Restorative Neurology* ; vol 4.
4. Freed CR, Breeze RE, Rosenberg NL, Schneck SA, Kriek E, Qi J-X, Lone T, Zhang Y-B, Snyder JA, Wells TH, Ramiq LO, Thompson L, Mazziotta JC, Huang SC, Grafton ST, Brooks D, Sawle G, Schroter G, Ansari AA. Survival of implanted fetal dopamine cells and neurologic improvement 12 to 46 months after transplantation for Parkinson's disease. *N Engl J Med.* 1992;327:1549-1555.
5. Breeze RE, Wells TH Jr, Freed CR. Implantation of fetal tissue for the management of Parkinson's disease: a technical note. *Neurosurgery.* 1995;36:1044-1047.
6. Freed CR, Breeze RE, Rosenberg NL, Schneck SA. Embryonic dopamine cell implants as a treatment for the second phase of Parkinson's disease: replacing failed nerve terminals. In: Narabayashi H, Nagatsu T, Yanagisawa N, Mizuno Y, eds. *Parkinson's Disease: From Basic Research to Treatment.* New York, NY: Raven Press; 1993:721-728. *Advances in Neurology* ; vol 60.
7. Spencer DD, Robbins RJ, Naftolin F, Marek KL, Vollmer T, Leranth C, Roth RH, Price LH, Gjedde A, Bunney BS, Sass KJ, Elsworth JD, Kier EL, Makuch R, Hoffer PB, Redmond DE. Unilateral transplantation of human fetal mesencephalic tissue into the caudate nucleus of patients with Parkinson's disease. *N Engl J Med.* 1992;327:1541-1548.
8. Widner H, Tetrud J, Rehncrona S, Snow B, Brundin P, Gustavii B, Bjorklund A, Lindvall O, Langston JW. Bilateral fetal mesencephalic grafting in two patients with parkinsonism induced by 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP). *N Engl J Med.* 1992;327:1556-1563.
9. Lindvall O, Sawle G, Widner H, Rothwell JC, Bjorklund A, Brooks D, Brundin P, Frackowiak R, Marsden CD, Odin P, Rehncrona S. Evidence for long-term survival and function of dopaminergic grafts in progressive Parkinson's disease. *Ann Neurol.* 1994;35:172-180.
10. Freed CR, Breeze RE, Schneck SA, Bakay RAE, Ansari AA. Fetal neural transplantation for Parkinson's disease. In: Rich RR, ed. *Clinical Immunology: Principles and Practice.* St Louis, Mo: Mosby–Year Book Inc; 1995:1677-1687.
11. Ansari AA, Mayne A, Freed CR, Breeze RE, Schneck SA, O'Brien CF, Kriek EH, Zhang Y-B, Mazziotta JC, Hutchinson M, Schroter G, Bakay RA, Boyer K, Sundstrom JB. Lack of detectable systemic humoral/cellular allogeneic response in human and nonhuman primate recipients of embryonic mesencephalic allografts for the therapy of Parkinson's disease. *Transplant Proc.* 1995;27:1401-1405.
12. Freeman TB, Olanow CW, Hauser RA, Nauert GM, Smith DA, Borlongan CV, Sanberg PR, Holt DA, Kordower JH, Vingerhoets FJG, Snow BJ, Calne D, Bauger LL. Bilateral fetal nigral transplantation into the postcommissural putamen in Parkinson's disease. *Ann Neurol.* 1995;38:379-388.
13. Kordower JH, Freeman TB, Snow BJ, Vingerhoets FJG, Mufson EJ, Sanberg PR, Hauser RA, Smith DA, Nauert GM, Perl DP, Olanow CW. Neuropathological evidence of graft survival and striatal reinnervation after the transplantation of fetal mesencephalic tissue in a patient with Parkinson's disease. *N Engl J Med.* 1995;332:1118-1124.
14. Peschanski M, Defer G, N'Guyen JP, Ricolfi F, Monfort JC, Remy P, Geny C, Samson Y, Hantraye P, Jeny R, Gaston A, Kéravel Y, Deqos JD, Cesaro P. Bilateral motor improvement and alteration of L-dopa effect in two patients with Parkinson's disease following intrastriatal transplantation of fetal ventral mesencephalon. *Brain.* 1994;117:487-499.
15. Defer GL, Geny C, Ricolfi F, Fenelon G, Monfort JC, Remy P, Villafane G, Jeny R, Samson Y, Kéravel Y, Gaston A, Deqos JD, Peschanski M, Cesaro P, Nguyen JP. Long-term outcome of unilaterally transplanted parkinsonian patients, I: clinical approach. *Brain.* 1996;119:41-50.
16. Kopyov OV, Jacques D, Lieberman A, Duma CM, Rogers RL. Clinical study of fetal mesencephalic intracerebral transplants for the treatment of Parkinson's disease. *Cell Transplant.* 1996;5:327-337.
17. Freed CR, Breeze RE, Leehey MA, Schneck SA, O'Brien CF, Thompson LL, Ramig LO, McRae CA, Mazziotta JC, Miletich RS, Eidelberg D. Ten years' experience with fetal neurotransplantation in patients with advanced Parkinson's disease [abstract]. *Abstr Soc Neurosci.* 1998;24:559.
18. Wenning GK, Odin P, Morrish P, Rehncrona S, Widner H, Brundin P, Rothwell JC, Brown R, Gustavii B, Hagell P, Jahanshahi M, Sawle G, Bjorklund A, Brooks DJ, Marsden CD, Quinn NP, Lindvall O. Short- and long-term survival and function of unilateral intrastriatal dopaminergic grafts in Parkinson's disease. *Ann Neurol.* 1997;42:95-107.
19. Hauser RA, Freeman TB, Snow BJ, Nauert M, Ganger L, Kordower JH, Olanow CW. Long-term evaluation of bilateral fetal nigral transplantation in Parkinson disease. *Arch Neurol.* 1999;56:179-187.
20. Piccini P, Brooks DJ, Bjorklund A, Gunn RN, Grasby PM, Rimoldi O, Brundin P, Hagell P, Rehncrona S, Widner H, Lindvall O. Dopamine release from nigral transplants visualized in vivo in a Parkinson's patient. *Nat Neurosci.* 1999;2:1137-1140.
21. Hagell P, Crabb L, Pogarell O, Schrag A, Widner H, Brooks DJ, Oertel WH, Quinn NP, Lindvall O. Health-related quality of life following bilateral intrastriatal transplantation in Parkinson's disease. *Mov Disord.* 2000;15:224-229.
22. Freed CR, Greene PE, Breeze RE, Tsai WY, DuMochel W, Kao R, Dillon S, Winfield H, Culver S, Trojanowski JQ, Eidelberg D, Fahn S. Transplantation of em-

bryonic dopamine neurons for severe Parkinson's disease. *N Engl J Med.* 2001; 344:710-719.
23. Shapiro AK, Shapiro E. How blind is blind? In: Shapiro AK, Shapiro E, eds. *The Powerful Placebo: From Ancient Priest to Modern Physician.* Baltimore, Md: Johns Hopkins University Press; 1997:190-216.
24. Montgomery E, Lieberman A, Sing G, Fries JF. Patient education and health promotion can be effective in Parkinson's disease: a randomized controlled trial. *Am J Med.* 1994;97:429-435.
25. Fahn S, Elton R, Members of the 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.* Vol 2. Florham Park, NJ: Macmillan Health Care Information; 1987:153-163, 293-304.
26. McRae C, Cherin E, Diem G, Yamazaki TG, Russell D, Ellgring H, Fahn S, Greene P, Dillon S, Winfield H, Freed C. Changes in quality of life among persons receiving fetal transplant surgery for the treatment of Parkinson's disease. *Mov Disord.* 2000;15:1039-1040.
27. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1:385-401.
28. Spielberger CD, Gorsuch RL, Lushene RE. *State-Trait Anxiety Inventory Manual.* Palo Alto, Calif: Consulting Psychologists Press; 1970.
29. Devins GM, Binik RM, Hutchinson TA, Hollomby DJ, Barré PE, Guttmann RD. The emotional impact of end-stage renal disease: importance of patients' perceptions of intrusiveness and control. *Int J Psychiatry Med.* 1984;13:327-343.
30. Cutrona CE, Russell DW. The provisions of social relationships and adaptation to stress. *Adv Pers Relatsh.* 1987;1:37-67.
31. Ramaker C, Marinus J, Stiggelbout AM, van Hilten BJ. Systematic evaluation of rating scales for impairment and disability in Parkinson's disease. *Mov Disord.* 2002;17:867-876.
32. Siderowf A, McDermott M, Kieburtz K, Blindauer K, Plumb S, Shoulson I, Parkinson Study Group. Test-retest reliability of the Unified Parkinson's Disease Rating Scale in patients with early Parkinson's disease: results from a multicenter clinical trial. *Mov Disord.* 2002;17:758-763.
33. Schwab RS, England AC. Projection technique for evaluating surgery in Parkinson's disease. In: Gillingham FJ, Donaldson MC, eds. *Third Symposium on Parkinson's Disease.* Edinburgh, Scotland: E & S Livingstone; 1969:152-157.
34. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology.* 1967;17:427-442.
35. McRae C, Diem G, Vo A, O'Brien C, Seeberger L. Reliability of measurements of patient health status: a comparison of physician, patient, and caregiver ratings. *Parkinsonism Relat Disord.* 2002;8:187-192.
36. Langston JW, Widner H, Goetz CG, Brooks D, Fahn S, Freeman T, Watts R. Core Assessment Program for Intracerebral Transplantations (CAPIT). *Mov Disord.* 1992;7:2-13.
37. McRae C, Bowles S, Freed C. Quality of life among persons receiving neural implant surgery for Parkinson's disease. Paper presented at: Third International Congress of Behavioral Medicine; July 1994; Amsterdam, the Netherlands.
38. de la Fuente-Fernandez R, Ruth TJ, Sossi V, Schulzer M, Calne DB, Stoessl AJ. Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. *Science.* 2001;293:1164-1166.
39. Olanow CW, Fahn S, Muenter M, Klawans H, Hurtig H, Stern M, Shoulson I, Kurlan R, Grimes JD, Jankovic J, Hoehn M, Markham CH, Duvoisin R, Reinmuth O, Leonard HA, Ahlskog E, Feldman R, Hershey L, Yahr MD. A multicenter double-blind placebo-controlled trial of pergolide as an adjunct to Sinemet in Parkinson's disease. *Mov Disord.* 1994;9:40-47.
40. Watts R, Freeman TB, Hauser BA, Bakay RA, Ellias SA, Stoessel AJ, Eidelberg D, Fink JS. A double-blind randomized, controlled, multicenter clinical trial of the safety and efficacy of stereotaxic intrastriatal implantation of fetal porcine ventral mesencephalic tissue (Neurocell-PD) vs imitation surgery in patients with Parkinson's disease [abstract]. *Parkinsonism Relat Disord.* 2001;7:S87.
41. Brody H. *Placebos and the Philosophy of Medicine: Clinical, Conceptual and Ethical Issues.* Chicago, Ill: University of Chicago Press; 1977.
42. Spiro HM. *Doctors, Patients, and Placebos.* New Haven, Conn: Yale University Press; 1986.
43. Turner JA, Deyo RA, Loeser JD, Von Korff M, Fordyce WE. The importance of placebo effects in pain treatment and research. *JAMA.* 1994;271:1609-1614.
44. Boer GJ, Widner H. Clinical neurotransplantation: core assessment protocol rather than sham surgery as control. *Brain Res Bull.* 2002;58:547-553.

62. Stene J, Stene E, Stengel-Rutkowski S, Murken JD. Paternal age and Down's syndrome: data from prenatal diagnosis. *Hum Genet*. 1981;59:119-124.
63. Girard M, Couvert P, Carrie A, Tardieu M, Chelly J, Beldjord C. Parental origin of de novo MECP2 mutations in Rett syndrome. *Eur J Hum Genet*. 2001;9:231-236.
64. Gillberg C, Gillberg IC. Infantile autism: a total population study of reduced optimality in the pre- peri- and neonatal period. *J Autism Dev Disord*. 1983;13:153-166.
65. Funderburk SJ, Carter J, Tanguay P, Freeman BJ, Westlake JR. Prenatal reproductive problems and gestational hormonal exposure in autistic and schizophrenic children. *J Autism Dev Disord*. 1983;13:325-332.
66. Warren RP, Cole P, Odell D, Pingree C, Warren WL, White E, Yonk J, Singh VK. Detection of maternal antibodies in infantile autism. *J Am Acad Child Adolesc Psychiatry*. 1990;29:873-877.
67. Gee V, O'Neill MT. *Perinatal Statistics in Western Australia, 1999. Seventeenth Annual Report of the Western Australian Midwives' Notification System*. Perth, Australia: Health Department of Western Australia; 2001.
68. Hueston WJ, McClaflin RR. Variations in cesarean delivery for fetal distress. *J Fam Pract*. 1996;43:461-467.
69. Showalter E, Griffin A. All women should have a choice. *BMJ*. 1999;319:1401.
70. Humphrey MD. *The Obstetrics Manual*. Revised ed. Sydney, Australia: McGraw-Hill; 1999.
71. McCarthy A, Hunter B. *Obstetrics and Gynaecology*. London, England: Churchill-Livingstone, 1998.
72. Eftekhar K, Steer P. Women choose caesarean section. *BMJ*. 2000;320:1073.
73. Quinlivan JA, Petersen RW, Nichols CN. Patient preference the leading indication for elective caesarean section in public patients—results of a 2-year prospective audit in a teaching hospital. *Aust N Z J Obstet Gynaecol*. 1999;39:207-214.
74. Crawford JS. *Principles and Practice of Obstetric Anaesthesia*. 5th ed. London, England: Blackwell Scientific Publications; 1984.
75. Hood DD. Maternal and fetal morbidity and mortality. In: James FM, Wheeler AS, Dewan DM, eds. *Obstetric Anesthesia: The Complicated Patient*. 2nd ed. Philadelphia, Pa: FA Davis Co; 1988:57-75.
76. Newton ER. Epidural analgesia, intrapartum fever, and neonatal outcomes. *Birth*. 2000;27:206-208.
77. Impey L, MacQuillan K, Robson M. Epidural analgesia need not increase operative delivery rates. *Am J Obstet Gynecol*. 2000;182:358-363.
78. Boksa P, Zhang Y, Bestawros A. Dopamine D1 receptor changes due to caesarean section birth: effects of anesthesia, developmental time course, and functional consequences. *Exp Neurol*. 2002;175:388-397.
79. Buitelaar J, Willemsen-Swinkels S. Medication treatment in subjects with autistic spectrum disorders. *Eur Child Adolesc Psychiatry*. 2000;9:I85-I97.
80. Altshuler G. Placental insights into neurodevelopmental and other childhood diseases. *Semin Pediatr Neurol*. 1995;2:90-99.
81. Nakai A, Shibazaki Y, Taniuchi Y, Oya A, Asakura H, Kuroda S, Kishino T, Araki T. Influence of mild hypothermia on delayed mitochondrial dysfunction after transient intrauterine ischemia in the immature rat brain. *Brain Res Dev Brain Res*. 2001;128:1-7.
82. Suzuki S, Muratat T, Jiang L, Power GG. Hyperthermia prevents metabolic and cerebral flow responses to hypoxia in the fetal sheep. *J Soc Gynecol Investig*. 2000;7:45-50.
83. Wood SC, Gonzales R. Hypothermia in hypoxic animals: mechanisms, mediators, and functional significance. *Comp Biochem Physiol B*. 1996;113:37-43.
84. Ghaziuddin M, Shakal J, Tsai L. Obstetric factors in Asperger syndrome: comparison with high-functioning autism. *J Intellect Disabil Res*. 1995;39:538-543.
85. Rickerby G, Carruthers A, Mitchell M. Biological factors associated with Asperger syndrome. *J Autism Dev Disord*. 1991;21:341-348.
86. Ehlers S, Gillberg C. The epidemiology of Asperger syndrome: a total population study. *J Child Psychol Psychiatry*. 1993;34:1327-1350.
87. Mayes SD, Calhoun SL, Crites DL. Does *DSM-IV* Asperger's disorder exist? *J Abnorm Child Psychol*. 2001;29:263-271.
88. Volkmar FR, Klin A. Asperger's disorder and higher functioning autism: same or different? *Int Rev Res Ment Retard*. 2000;23:83-110.
89. Lord C, Mulloy C, Wendelboe M, Schopler E. Pre- and perinatal factors in high-functioning females and males with autism. *J Autism Dev Disord*. 1991;21:197-209.
90. Spiker D, Lotspeich LJ, Dimiceli S, Myers RM, Risch N. Behavioral phenotypic variation in autism multiplex families: evidence for a continuous severity gradient. *Am J Med Genet*. 2002;114:129-136.

Correction

Error in Figure. In the Original Article titled "Effects of Perceived Treatment on Quality of Life and Medical Outcomes in a Double-blind Placebo Surgery Trial" published in the April issue of the ARCHIVES (2004;61:412-420), an incorrect version of **Figure 3** was published. Figure 3 is published correctly here. Online versions of this article on the *Archives of General Psychiatry* Web site were corrected on April 13, 2004. The ARCHIVES regrets the error.

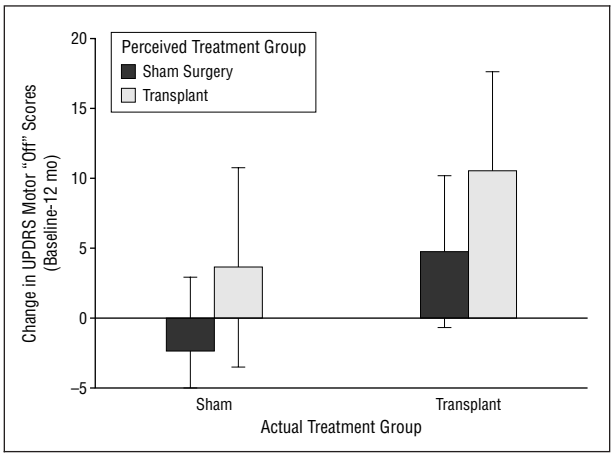

Figure 3. Mean changes in Unified Parkinson's Disease Rating Scale (UPDRS) motor "off" scores (baseline to 12 months) for the total group in the parent study (n=39). Increased scores indicate improvement. Error bars represent SEM.