Network modulation following sham surgery in Parkinson’s disease

Ji Hyun Ko,1 Andrew Feigin,1 Paul J. Mattis,1 Chris C. Tang,1 Yilong Ma,1 Vijay Dhawan,1 Matthew J. During,2 Michael G. Kaplitt,3 and David Eidelberg4

1Center for Neurosciences, The Feinstein Institute for Medical Research, Manhasset, New York, USA. 2Department of Molecular Virology, Immunology and Molecular Genetics, The Ohio State University, Columbus, Ohio, USA. 3Department of Neurological Surgery, Weill Cornell Medical College, New York, New York, USA.

Patient responses to placebo and sham effects are a major obstacle to the development of therapies for brain disorders, including Parkinson’s disease (PD). Here, we used functional brain imaging and network analysis to study the circuitry underlying placebo effects in PD patients randomized to sham surgery as part of a double-blind gene therapy trial. Metabolic imaging was performed prior to randomization, then again at 6 and 12 months after sham surgery. In this cohort, the sham response was associated with the expression of a distinct cerebello-limbic circuit. The expression of this network increased consistently in patients blinded to treatment and correlated with independent clinical ratings. Once patients were unblinded, network expression declined toward baseline levels. Analogous network alterations were not seen with open-label levodopa treatment or during disease progression. Furthermore, sham outcomes in blinded patients correlated with baseline network expression, suggesting the potential use of this quantitative measure to identify “sham-susceptible” subjects before randomization. Indeed, Monte Carlo simulations revealed that a priori exclusion of such individuals substantially lowers the number of randomized participants needed to demonstrate treatment efficacy. Individualized subject selection based on a predetermined network criterion may therefore limit the need for sham interventions in future clinical trials.

Introduction

Placebo effects pose a major challenge in the development of new treatments for neurodegenerative disorders. Parkinson’s disease (PD) is a case in point. Prominent placebo (or sham-surgical) responses are frequently encountered in trials of new treatments for this disorder (1, 2). The appreciable size of these effects has played no small part in the recent failure of several blinded, early-phase studies of novel interventions for refractory PD motor symptoms (3).

The neural mechanisms underlying the placebo effect are not fully understood. In PD subjects, short-term placebo responses have been associated with ventral striatal dopamine release (refs. 4, 5, and see ref. 6 for review). Indeed, the likely effects of these changes on effector pathways have been demonstrated through intraoperative recordings conducted in PD patients undergoing deep brain stimulation surgery (7). In these studies, short-term placebo responses were associated with changes in neural activity recorded at the subthalamic nucleus (STN) target and also in downstream basal ganglia and thalamic projection zones. Longer-term placebo responses have been associated, by contrast, with localized imaging changes involving the limbic and paralimbic cortex and the amygdala (8–10). It is unknown, however, whether these areas — and perhaps others as well — function in concert to mediate the placebo response as a discrete brain network.

In this study, we used network analysis in conjunction with brain imaging to identify and validate a specific metabolic topography associated with the response to sham surgery in PD subjects. Network activity measured under blinded conditions increased consistently in proportion to the individual sham response; the changes were reversed by unblinding. In individual subjects, network activity measured at baseline predicted the subsequent sham response seen under the blind. This suggested a novel image-based strategy to reduce sham effects in randomized clinical trials for brain disorders.

Results

Network characterization

Pattern identification. We studied 23 advanced PD subjects (17 men and 6 women; aged 60.3 ± 1.6 years; disease duration 11.7 ± 1.0 years; baseline off-state motor Unified Parkinson’s Disease Rating Scale [UPDRS] ratings of 39.4 ± 1.8) who were randomized to sham surgery (SHAM) as part of a 6-month blinded surgical trial of gene therapy for refractory motor symptoms (Supplemental
Network analysis identified a significant sham-related metabolic covariance pattern (SSRP), which accounted for 4.8% of the overall variance in the scan data. This network (Figure 1A) was characterized by increased activity involving the anterior cingulate cortex (Brodmann area [BA] 32/24) and subcallosal gyrus (BA 25), the hippocampus and parahippocampal gyrus, the amygdala, and the posterior cerebellar vermis (lobule VII/crus II). The voxel weights on these regions (Table 1) were found to be reliable by bootstrap estimation ($Z > 1.64$, $P < 0.05$ for the inverse coefficient of variation [ICV]; see Methods). The SSRP also included less reliable contributions from the head of the caudate and anterior putamen (top) and from the VA thalamic nucleus. While voxel weights for these clusters significantly contributed to this network (Table 1), these loadings did not meet the prespecified bootstrap reliability criteria ($z = 1.575, 1.472,$ and $1.348$ for the 3 regions, respectively; 1,000 iterations). A significant ordinal trend in SSRP expression (left) was seen in the 8 responders to SHAM who were used to identify the pattern. Each subject exhibited an increase in network expression under the blind at 6 months ($P < 0.01$, binomial test). A similar ordinal trend in SSRP expression under the blind ($P < 0.01$, binomial test) was evident in the 8 remaining sham responders (middle) who were not used for pattern identification. An ordinal trend was not observed under the blind ($P = 1.0$) for the 7 sham nonresponders (right). Three violations were evident in this group, whereas no violations were present in either SHAM$_R$ group.

Figure 1; supplemental material available online with this article; doi:10.1172/JCI75073DS1). Of the sham-operated subjects, 16 demonstrated a degree of clinical improvement, as represented by reductions in UPDRS motor ratings of 2 points or more under the blind at 6 months. They were termed sham “responders” (SHAM$_R$) (7). The remaining 7 sham-operated subjects exhibited clinical outcomes typical of PD progression, with either no change or a deterioration, as represented by an increase in blinded UPDRS motor ratings at 6 months. They were termed sham “nonresponders” (SHAM$_{NR}$). We used baseline and 6-month scans from 8 randomly selected SHAM$_R$ subjects for network identification. Scans from the remaining responders and from the nonresponders were used prospectively for testing (see Methods and Supplemental Figure 1).

Pattern validation. SSRP expression values were computed prospectively in the SHAM$_R$ testing cohort (i.e., the 8 SHAM$_R$ subjects not used for network identification) and in the 7 SHAM$_{NR}$ subjects. Network expression in the SHAM$_R$ testing set (Figure 1C, middle) increased consistently following SHAM (0 of 8 violations, $P = 0.008$, binomial test); we did not observe an analogous ordinal trend (Figure 1C, right) in the nonresponders (3 of 7 violations, $P = 1.0$).

As part of pattern validation, we tested alternative explanations for the increases in SSRP expression that were seen under the blind. First, we considered the possibility that the observed network changes resulted not from the sham effect, but from the motor improvement that characterized each of the sham responders. If so, consistent increases in SSRP expression should be evident in subjects demonstrating comparable clinical improvement under unblinded (“open-label”) conditions. We therefore measured SSRP expression in a separate group of 9 PD subjects (Figure 2A, right) scanned in the off-medication baseline state and again during an open-label levo-
dopa infusion. Treatment outcomes in these subjects were titrated to be similar to those observed in the SHAM<sub>r</sub> testing group (see Methods). The network data from these individuals (Supplemental Figure 2A) indicated that while clinically effective, unblinded levodopa treatment was not associated with consistent changes in SSRP expression (5 of 9 violations, \( P = 1.0 \), binomial test). Indeed, we observed a significant difference in network modulation (Figure 2A, light gray bars) across the 3 testing sets \( F(2,21) = 4.156, \ P = 0.030 \), with increased expression under the blind in sham responders relative to that seen in sham nonresponders \( P = 0.014, \) post-hoc least significant difference (LSD) test) as well as in responders to open-label levodopa treatment \( P = 0.036 \).

We also considered the possibility that the observed SSRP increases reflected underlying disease progression (“natural history”) effects. To this end, we measured SSRP expression in an independent group of PD subjects \( n = 15 \) who underwent serial metabolic imaging over a 2-year period (see Methods). We found that SSRP expression in these individuals (Supplemental Figure 2B) did not change with advancing disease \( (8 \ of 15 \ violations, \ P = 1.0, \ binomial \ test) \). In this vein, we compared blinded SSRP scores in the SHAM<sub>r</sub> and SHAM<sub>n</sub> testing sets with parallel changes in the expression of the PD-related metabolic pattern (PDRP) \( (11-13) \), a topographically independent metabolic brain network associated with disease progression. Indeed, the subjects exhibited a significant difference in the trajectories of the 2 networks over time under the blind \( F(1,16) = 6.966, \ P = 0.018, \) group × network interaction, \( 2 \times 2 \) repeated-measures ANOVA (RMANOVA). While SSRP expression increased in the sham responders relative to that in the nonresponders \( P = 0.016, \) post-hoc LSD test), we did not observe a corresponding group difference in PDRP expression \( P = 0.216 \) in the same subjects. In aggregate, these data show that the observed SSRP changes are unlikely to have resulted from intercurrent treatment or disease progression effects.

Changes in SSRP expression under the blind also correlated \( (r = -0.428, \ P = 0.041) \) with concurrent depression ratings according to the Beck Depression Inventory (BDI). However, mean depression ratings were not influenced by SHAM in the group as a whole \( (P = 0.471, \) paired Student’s t test), or in either set of SHAM<sub>n</sub> subjects \( (P > 0.1) \). Nonetheless, the correlation between changes in BDI ratings and network expression was not significant after adjusting for subject differences in motor outcome \( (r = 0.330, \ P = 0.133, \) partial correlation). We found that the correlation between the changes in SSRP expression and motor ratings persisted, however, after adjusting for BDI differences \( (r = -0.722, \ P < 0.001, \) partial correlation). Cognitive performance under the blind (Supplemental Table 2) was not altered by SHAM \( (P > 0.1, \) paired Student’s t tests); these measures did not correlate with the network changes that we observed.

**Nodal correlates.** We also examined the changes in local metabolic activity that occurred in SSRP nodal regions under the blind (see Methods). Following SHAM, we observed significant regional increases in the anterior cingulate cortex \( (BA 32/24; \ P = 0.015, \) paired Student’s t test) and in the posterior cerebellar vermis \( (lobe VII/crus II; \ P = 0.035) \). Metabolic changes at the cerebellar node correlated with motor outcomes under the blind \( (r = -0.460, \ P = 0.031, \) partial correlation adjusting for whole-brain activity). Changes at the other network nodes following SHAM, as well as baseline measurements, did not correlate with motor outcomes under the blind \( (P > 0.15) \). Nodal correlations with changes in BDI ratings and cognitive test measures were also not significant \( (P > 0.10, \) Pearson’s correlations).

**Network changes following gene therapy**

In the blinded surgical trial (Supplemental Figure I), 21 participants were randomized to STN gene therapy; adenoassociated viral vector-glutamic acid decarboxylase (AAV-GAD) was successfully delivered at the target site in 16 of these subjects.
Members of this group were classified as responders or nonresponders based on the same clinical rating criterion that was used to categorize the sham-operated participants. Accordingly, 14 of the gene therapy subjects were classified as “responders” (GADR); the 2 gene therapy subjects were classified as “nonresponders” (GADNR). We found that blinded motor outcomes at 6 months (Figure 4, dark gray bars) were similar for the 2 responder groups (change from baseline: SHAMR –7.88 ± 1.20 points; GADR –10.00 ± 1.28 points; $P = 0.235$, Student’s $t$ test).

Nonetheless, the degree of concurrent network modulation observed under the blind (Figure 4, light gray bars) was larger in the sham group ($P = 0.002$). Whereas changes in SSRP expression under the blind correlated with individual motor outcomes in sham responders ($r = -0.638$, $P = 0.008$, Pearson’s correlation), an analogous correlation was not present in GADR subjects ($r = -0.125$, $P = 0.670$). Moreover, in contrast to SHAM, no correlation was present ($r = 0.053$, $P = 0.845$) between baseline SSRP values and blinded 6-month changes in UPDRS motor ratings after STN AAV-GAD gene therapy. These findings point to the specificity of the SSRP network for the sham response.

Reversal of network changes by unblinding

Last, the functional relationship between SSRP expression and the sham response was demonstrated by unblinding. In this study, treatment status was revealed only after the last subject completed 6 months of blinded follow-up. Because enrollment took place continuously over nearly 1 year, a number of the early subjects completed 12 months of postoperative follow-up, Members of this group were classified as responders or nonresponders based on the same clinical ratings criterion that was used to categorize the sham-operated participants. Accordingly, 14 of the gene therapy subjects were classified as “responders” (GADR); the 2 gene therapy subjects were classified as “nonresponders” (GADNR). We found that blinded motor outcomes at 6 months (Figure 4, dark gray bars) were similar for the 2 responder groups (change from baseline: SHAMR –7.88 ± 1.20 points; GADR –10.00 ± 1.28 points; $P = 0.235$, Student’s $t$ test).

Nonetheless, the degree of concurrent network modulation observed under the blind (Figure 4, light gray bars) was larger in the sham group ($P = 0.002$). Whereas changes in SSRP expression under the blind correlated with individual motor outcomes in sham responders ($r = -0.638$, $P = 0.008$, Pearson’s correlation), an analogous correlation was not present in GADR subjects ($r = -0.125$, $P = 0.670$). Moreover, in contrast to SHAM, no correlation was present ($r = 0.053$, $P = 0.845$) between baseline SSRP values and blinded 6-month changes in UPDRS motor ratings after STN AAV-GAD gene therapy. These findings point to the specificity of the SSRP network for the sham response.

Reversal of network changes by unblinding

Last, the functional relationship between SSRP expression and the sham response was demonstrated by unblinding. In this study, treatment status was revealed only after the last subject completed 6 months of blinded follow-up. Because enrollment took place continuously over nearly 1 year, a number of the early subjects completed 12 months of postoperative follow-up,
The Journal of Clinical Investigation

Research Article

The Journal of Clinical Investigation    Research Article

Figure 4. Effects of treatment on network modulation under the blind. Changes in SSRP expression (light gray bars) were computed under the blind at 6 months in the 16 sham responders and 14 STN AAV-GAD gene therapy responders (see text). Concurrent changes in UPDRS motor scores (dark gray bars) are presented for comparison. Motor outcomes under the blind did not differ for the GAD<sub>R</sub> and SHAM<sub>R</sub> subjects [t(28) = 1.213, P = 0.235]. Nonetheless, a significant difference in SSRP modulation was observed in the 2 groups [t(28) = 3.379, P = 0.002].

Figure 5. Effects of unblinding on network expression. (A) Eleven sham (SHAM<sub>R</sub>) and 6 gene therapy (GAD<sub>R</sub>) responders were rescanned at 12 months after unblinding (see text). The time course of SSRP expression differed for the 2 groups [F(1,15) = 6.90, P = 0.019, RMANOVA]. Unblinding was associated with a significant decline in network expression in the sham-operated subjects (P = 0.031, post-hoc LSD test) but not in their gene therapy counterparts (P = 0.152). Outliers (greater than the mean +1.5 × SD) are shown by white circles. (B) After unblinding, the majority of SHAM<sub>R</sub> subjects (8 of 11 = 72.7%) exhibited a decline in network expression (left), with values falling in the range (dashed line) seen over a comparable time interval in an unblinded disease progression cohort (see text). By contrast, SSRP expression remained above this level (right) in the 4 SHAM<sub>R</sub> subjects who were still under the blind at 12 months. Dashed line represents 1.5 SD above the mean change in SSRP expression observed in an independent cohort composed of 15 PD subjects scanned twice over a 2-year period (see Supplemental Figure 2B). (C) Six of the 13 gene therapy responders (GAD<sub>R</sub>) who underwent repeat metabolic imaging at 12 months were unblinded prior to the final imaging session. Unblinding had no significant effect on SSRP expression in these subjects. Indeed, in 5 of the unblinded GAD<sub>R</sub> subjects (left), network activity at 12 months was in the open-label progression range (dashed line). Similar network values were observed (right) for the GAD<sub>R</sub> subjects who remained under the blind at the final imaging time point.
Discussion

The metabolic anatomy of the sham response. In this study, we used network analysis to identify a specific cerebello-limbic metabolic network associated with the sham response. Network activity increased in sham-operated PD subjects studied under the blind, correlating to the concurrent changes in clinical ratings. The network changes observed following SHAM were reversed by unblinding. Analogous network responses were not evident, however, during open-label levodopa treatment or following experimental STN gene therapy.

The SSRP represents a distinctive spatial covariance topography. Several regions contributing to this network have previously been noted to exhibit increases in local activity in response to placebo. For example, changes in metabolic activity in the anterior cingulate cortex (BA 32/24) and subgenual cingulate gyrus (BA 25) and in the parahippocampal gyrus and amygdala have been reported in subjects receiving placebo treatments as part of blinded antidepressant trials (8). The relationship of the placebo response to metabolic activity in the subgenual cingulate gyrus is particularly relevant, given the relationship of the network changes to concurrent, blinded BDI depression ratings (14). Of note, participants in the current study were not depressed at baseline, and SHAM was not associated with significant changes in depression ratings under the blind. Indeed, the correlation we observed following SHAM between the changes in BDI ratings and SSRP expression was driven predominantly by the motor changes, which remained significant even after controlling for individual differences in the BDI response.

Sham responses were associated with increased metabolic activity in additional components of the limbic cortico-striato-pallido-thalamo-cortical (CSPTC) loop (15), including the amygdala, VA thalamus, and parahippocampal gyrus. Indeed, changes in these and related brain regions have been found to accompany placebo effects occurring in the context of emotional processing (10, 16). Although ventral striatal dopamine release is a key mediator of the acute placebo response in PD (4, 5, 17), the relationship to the SSRP is unclear. It is tempting to associate this aspect of the network topography with specific open-loop circuits connecting the basal ganglia (particularly the ventral striatum) to motor and nonmotor cortical areas (18, 19). In this context, the SSRP circuit can be viewed as conveying reward-based signals from the basal ganglia to both the limbic and motor cortex. Nevertheless, we found that local metabolic contributions to SSRP activity from the striatum and thalamus were weaker and less reliable than the other network regions (Figure 1B, and see Table 1). Caution is needed in interpreting the role of these regions in mediating the sham response.

Nodal analysis revealed significant metabolic responses to SHAM in only 2 network regions: the anterior cingulate cortex and the posterior cerebellar vermis. We found that correlations...
between blinded motor outcomes and concurrent changes in local metabolic activity were significant only at the cerebellar node. These findings underscore the relevance of the nonmotor posterior cerebellum to the sham response. This region, which is the major cerebellar target ofafferent projections from the amygdala and from the paralimbic cingulate and parahippocampal cortex, has been linked to affective processing in both health and disease (20). Indeed, the spatial proximity of the cerebellar SSRP node to the sensorimotor zones of the anterior vermis (lobules V and VI) (21) is consistent with the motoric specificity of the sham response in this study. Nonetheless, this particular cerebellar region cannot be regarded as a general effector of the sham response. Metabolic increases localized to the cerebellar vermis are a consistent feature of the abnormal network topography of PD, even in the absence of intercurrent tremor (11, 13, 22). It is conceivable that in PD, disease-related increases in resting activity in the sensorimotor zones of the cerebellum promote susceptibility to sham effects mediated by adjacent nonmotor cerebellar regions and associated projection pathways. Moreover, the specific network topography seen in PD subjects in response to motor effects may not be generalized to other sham responses in the same population. Indeed, changes in SSRP expression were not seen in PD subjects randomized to placebo as part of a blinded clinical trial targeting the cognitive symptoms of the disorder (ref. 23 and see Supplemental Figure 2C for details). Irrespective of the precise effector, it is likely that changes in functional connectivity involving limbic and paralimbic projections to subcortical brain regions will be a common feature of sham-related network organization.

Examination of the individual SSRP trajectories of the subjects (Figure 5, B and C) illustrated the consistently increasing tendency (ordinal trend) of the network responses that occurred under the blind — and the reversal of these changes by unblinding. Indeed, increases in SSRP expression were present under the blind in all 16 sham responders. In the absence of other causes for the observed network changes, we attributed the findings to increases in the activity of expectation-related neural circuits under blinded conditions. In most cases, the network changes were reversed by unblinding, with a decline to “open-label” levels when the expectation of benefit was replaced by cognitive acceptance of the true treatment status. We note, however, that in some subjects, elevations in SSRP activity persisted after unblinding. It is interesting to consider whether these individuals were either unable or unwilling to acknowledge the sham intervention that they received under the blind. Further studies may determine, for instance, whether these subjects constitute a discrete genotypic subset of sham responders and whether synaptic dopamine concentrations are sustained at higher levels in these individuals.

We additionally note that analogous changes in SSRP expression were not observed under the blind in the gene therapy group. Indeed, despite the small differences in motor benefit observed under the blind at 6 months (Figure 4, dark gray bars) in the 2 responder groups [GAD₉; ΔUPDRS = –10.00 ± 1.28; SHAM₉; ΔUPDRS = –7.88 ± 1.20; t(28) = 1.21, P = 0.235], significant SSRP modulation occurred only in the SHAM group (Figure 4, light gray bars). Thus, the substantial motor benefit we observed in the GAD₉ subjects (who represented the majority [87.5%] of the gene therapy participants in whom the viral vector was successfully delivered to the STN target; ref. 24) was not explained by sham network effects. It is also important to recognize that the STN itself likely plays an important role in mediating the short-term placebo response in PD subjects (7). It is therefore conceivable that this structure, which represents a critical “bridge” between the cerebellum and basal ganglia (19, 25, 26), was functionally altered by local delivery of the viral vector. This, in turn, may have limited the dynamic range of SSRP modulation that was possible in the participants who received STN AAV-GAD gene therapy. Given that both the SHAM and GAD subjects were treated with similar doses of open-label levodopa/carbidopa under the blind (24), it is unlikely that the observed group differences in network modulation resulted directly from dopaminergic effects. That said, the possibility of a functional interaction between AAV-GAD and levodopa at the STN cannot be excluded.

Levodopa has been found to improve PD-related metabolic changes in key network regions, including the STN (27–29). It is also conceivable that levodopa alters the expression of other networks associated with this structure, such as the SSRP. However, the absence of SSRP modulation during open-label levodopa treatment (Figure 2A) very likely stems from the lack of uncertainty during the trial. A recent study using [11C] raclopride PET to monitor synaptic dopamine during the placebo response suggested that uncertainty plays a key role in mediating placebo effects (17). When the patients were assured that they were receiving real treatment (although they were given placebo), no significant changes were observed in striatal dopamine release. The presence of robust increases in SSRP expression under the blind in sham subjects receiving stable open-label levodopa further supports the role of uncertainty in determining the activity of this network.

Implications for trial design. Despite the limited number of subjects used to identify the SSRP in our study, it was possible to confirm the relationship of this network to the sham response by measuring its expression prospectively in independent testing datasets. Indeed, increases in pattern expression under the blind were confirmed in the prospective SHAM testing set, along with robust correlations between these changes and clinical outcome. The reversal of these changes following unblinding provided further support for the posited relationship between the SSRP network and the sham response.

The relatively limited number of randomized subjects in the current study and the relatively short period of follow-up under the blind were similar in scale to other recent sham-controlled phase II surgical trials for neurodegenerative disorders (30–32). The demonstration of therapeutic efficacy under such conditions is especially challenging, given the sizable sham effects that are typically elicited in blinded surgical trials (1, 30–33). This constraint is magnified by the logistical demands imposed by high subject throughput and the corresponding need for sufficient quantities of biological materials with which to treat participants randomized to the actual intervention. Even more concerning are the ethical issues attendant to randomizing large numbers of subjects to sham procedures that carry risk but confer little benefit to the participant (2, 3, 34).

The current findings suggest a novel approach to this issue. In general terms, placebo responses can potentially be reduced by excluding subjects found empirically to be susceptible to these and...
related sham effects (e.g., ref. 34). Alternatively, sham responders may be identified before randomization by quantifying the activity of the SSRP or related sham networks in baseline functional brain images. In the current study, baseline SSRP values (Figure 3A) correlated with blinded motor outcomes 6 months after SHAM. Indeed, baseline SSRP expression was significantly lower (Figure 3B) in responders to SHAM relative to that seen in nonresponders ($P = 0.001$, Student’s $t$ test). This suggested that individuals with low baseline SSRP expression have the capacity to deploy the network in response to SHAM, with concomitant clinical improvement. This capacity is not present, however, in individuals with high baseline SSRP expression. These subjects cannot deploy the network further and are thus unable to generate a meaningful clinical response to the sham intervention.

Based on these considerations, we used Monte Carlo simulations to determine the impact of using baseline SSRP measurements for identification of potential sham responders to randomization. The median value for the SHAM$_R$ subjects (Figure 3A, dashed line) was chosen as a cutoff; individuals with baseline expression below this level were considered to be particularly susceptible to sham effects and were withdrawn from the simulated experiment.

Simulations based on the baseline SSRP values that we observed empirically in the phase II STN AAV-GAD data (Figure 3C) suggested that 192 subjects would have to be randomized to detect a significant group difference in motor outcome in 95% of 10,000 trials ($P = 0.05$, 2-tailed Student’s $t$ test). Nonetheless, the required sample size falls to 84 subjects by excluding the potentially susceptible participants, i.e., those with baseline network expression at or below the prespecified cutoff value. In summary, prospectively computed subject scores for the SSRP or for analogous networks identified under the blind in phase II data may be used subsequently at phase III to detect and potentially exclude sham-responsive participants before randomization. Of note, baseline metabolic imaging has already been implemented as a measure to identify potential trial participants with atypical parkinsonian variant conditions (24). SSRP-related computations can easily be performed on the same scans to classify subjects according to their relative susceptibility to major sham effects under trial conditions. Approximately 35% of the participants would fall into this category, as determined by baseline SSRP expression values. However, excluding them resulted in a net reduction in sample size of over 56%.

**Methods**

**Study design**

Sixty-six patients with advanced PD were screened for eligibility to participate in a randomized, double-blind, SHAM-controlled multicenter phase II trial of STN AAV-GAD gene therapy for medically refractory motor symptoms. A flow diagram for the study is presented in Supplemental Figure 1. Prior to randomization, all subjects underwent metabolic brain imaging in the resting state with $^{18}$F-fluorodeoxyglucose (FDG) PET to exclude atypical parkinsonian “look-alike” conditions that are generally resistant to PD interventions (24, 35). Following this screening procedure, 45 PD subjects were randomized 1:1 to receive either STN AAV-GAD gene therapy ($n = 22$) or SHAM (bilateral burr hole placement, $n = 23$); the subjects and investigators were blinded to the treatment status for at least 6 months following the procedure; 1 subject in the AAV-GAD arm was unblinded early because of catheter misplacement. Finally, 5 additional AAV-GAD patients were excluded from further analysis because of failed viral vector delivery at the STN target site (see ref. 24 for details).

At baseline, no significant group differences ($P > 0.07$) were present regarding age, gender, UPDRS motor ratings (36), BDI ratings (37), or tests of memory and executive functioning assessed according to the Hopkins Verbal Learning Test (HVLT) (38), Stroop Interference Test (39), and the Symbol Digit Modality Test (SDMT) (40). The subjects were rescanned under the blind 6 months after surgery (with the exception of 1 subject in each group) and again at the conclusion of the study at 12 months. The subjects were simultaneously unblinded after the final participant completed 6 months of blinded follow-up. The surgical procedures performed in the trial were staggered over a 1-year period. Thus, although the majority of the SHAM subjects; 11 of 20 [55%] of the GAD subjects) of patients underwent imaging at 12 months after unblinding [interval 3.1 ± 0.4 months in SHAM and 3.5 ± 0.4 months in GAD subjects; ($\gamma = 0.756, P = 0.457$], the remaining 6 SHAM and 9 GAD subjects were still under the blind at this 12-month time point. The details of the surgical and imaging procedures performed, as well as the outcome of the 6-month blinded phase of the trial, have been presented previously (24).

**Metabolic imaging**

All trial participants included in our analysis (Supplemental Figure 1) underwent FDG PET at baseline and again, under the blind, at the 6-month time point. All but 2 subjects (1 SHAM$_R$ and 1 GAD$_D$) underwent repeat PET imaging at 12 months. The subjects fasted overnight before each scanning session. In all subjects and at all time points, antiparkinsonian medications were withheld for at least 12 hours before the start of imaging. In the gene therapy trial, participants were scanned at 1 of 5 imaging centers as detailed elsewhere (24). The PET tomographs used at each center, the performance features of the corresponding instruments, and the number of patients scanned on each platform are provided in Supplemental Table 1. In the test-retest disease progression and levodopa treatment validation cohorts, scanning was conducted in paired imaging sessions using the GE Advance tomographic system (GE Healthcare) at The Feinstein Institute for Medical Research.

All subjects were scanned in 3D mode for 10 minutes, beginning 30 minutes after radiotracer injection. Scanning was conducted in an awake resting state with eyes open in a dimly lit room and with minimal auditory stimulation. In each subject, the scan pairs were spatially aligned across conditions, warped into the standard Montreal Neurological Institute (MNI) space, and smoothed with a 10-mm Gaussian filter using SPM5 software (http://www.fil.ion.ucl.ac.uk/spm/software/spm5/).

**Network analysis**

**Pattern identification.** To identify a specific metabolic covariance pattern associated with the motor response to SHAM, we implemented a within-subject network mapping strategy, termed ordinal trends/canonical variates analysis (OrT/CVA; ref. 41). This computational algorithm relies on supervised principal component analysis (PCA) to identify distinct spatial covariance patterns (metabolic brain networks) with consistent changes in subject expression across experimental conditions (e.g., refs. 42–44). This approach differs from rou-
tine voxel-wise univariate analysis in that it requires that the pattern exhibit an ordinal trend (i.e., a consistent change in expression across conditions at the individual subject level) in the data. Thus, in OrT/CVA, network activity is required to increase (or decrease) monotonically in all or most of the subjects. As in other forms of spatial covariance analysis, large-scale networks are described in terms of the voxel loadings (“region weights”) on each of the relevant principal component (PC) topographies (45). Likewise, the expression of a given pattern in each scan is quantified by a specific network activity measure (“subject score”), the PC scalar multiplier for the subject at each experimental time point.

In OrT/CVA, as in other forms of spatial covariance analysis, the significance of the resulting topographies is assessed using nonparametric tests (46). In the initial network identification phase of the analysis, a permutation test of the relevant subject scores is performed to confirm that the observed monotonic changes in pattern expression did not occur by chance. An ordinal trend was considered significant in the derivation set if a consistent change in pattern expression was present for all (or most) subjects across experimental conditions ($P < 0.05$, permutation test). Likewise, the reliability of the voxel loadings on the network topography is assessed using bootstrap resampling procedures (46).

In the current study, a significant SSRP topography was sought among the linearly independent (orthogonal) PC patterns that resulted from OrT/CVA analysis of the scans acquired under the blind at baseline and 6 months following surgery in 8 sham responders. The following model selection criteria were applied to the individual patterns: (a) the analysis was limited to the first 6 PCs, which typically account for at least 75% of the subject × region variance (46); (b) subject scores for these PCs were entered singly and in all possible combinations into a series of logistic regression models, with time (before and 6 months after) as the dependent variable and the subject scores for each set of PCs as the independent variables for each model. The best model was considered to be that with the smallest Akaike information criterion (AIC) value. The selected PC(s) in this model were then used in linear combination to identify the spatial covariance pattern (if any) that exhibited significant ordinal trend in the data acquired under the blind in the 8 SHAMR derivation subjects. Once a significant SSRP was identified in these individuals, the presence of an ordinal trend was confirmed by prospectively measuring expression values for the pattern in the baseline and 6-month scans of the 8 remaining SHAMR subjects.

To minimize potential confounds stemming from concurrent effects of disease progression, we restricted the search for a sham-related metabolic network to the portion of the overall derivation space that is independent of (i.e., orthogonal to) the stereotyped functional changes that relate specifically to the underlying neurodegenerative process. In PD, the latter are represented by a distinct disease-related metabolic covariance pattern known as the PDRP (11, 12, 42). The PDRP topography is displayed in Figure 6A; the absence of a correlation between voxel weights on this network and the SSRP is demonstrated in Figure 6B. Analysis of the longitudinal scan data from the sham-operated subjects (Figure 6C) revealed significant linear increases in PDRP expression over time ($t(44) = 3.55, P = 0.0009$, multiple linear regression). Because of this network-level progression effect, we orthogonalized the scan data to the PDRP before implementing the OrT/CVA algorithm. In this way, confounds associated with disease progression were minimized in the search for the SSRP topography.

Validation studies. SSRP expression values were computed prospectively on an individual scan basis as described elsewhere (11, 45). Because of the multicenter design of the current study (24, 47), the scan data were scrutinized for potential bias caused by tomographic differences across sites. Indeed, no difference in SSRP values was found for sham-operated subjects scanned under the blind at the 5 imaging sites [baseline: $F(4,18) = 0.920, P = 0.476; 6$ months: $F(4,18) = 1.216, P = 0.338$, 1-way ANOVA]. That said, global metabolic activity varied to some degree across the sites [baseline: $F(4,18) = 2.592, P = 0.071; 6$ months: $F(4,18) = 8.058, P = 0.001$]. We therefore adjusted for individual differences in global metabolism, while correlating clinical outcomes with changes in local activity in nodal regions (see below).

It was also important to establish that pattern expression was stable in single subjects. SSRP expression, in fact, had excellent test-retest reliability (intraclass correlation coefficient [ICC]: 0.937, $P < 0.001$) when computed prospectively in scans from an independent group of PD subjects ($n = 14$) who were studied twice over a 2-month period.

Further validation was provided using metabolic scan data from independent PD reference samples to: (a) evaluate the test-retest reproducibility of prospectively computed network expression values. This was done using scan data from 14 PD subjects who underwent repeat imaging over a 2-month period (48); (b) assess the impact of open-label dopaminergic pharmacotherapy on network expression. This was done using scan data from 9 PD subjects who were studied at baseline and during an intravenous levodopa infusion (29) titrated to produce an improvement in motor ratings comparable to those of the SHAMR testing set; (c) determine the impact of disease progression on network expression. This was done using longitudinal scan data from 15 PD subjects who were studied at baseline and again 24 months later (12). Demographic and clinical features of these cohorts are provided in Supplemental Table 3.

Statistics

Network modulation and clinical correlation under the blind. After validation, SSRP expression values computed in the testing data (i.e., in the 8 SHAMR subjects not used for pattern derivation, the 7 SHAMR subjects, and in the 9 PD subjects who received open-label levodopa infusion during imaging) were compared across groups using 1-way ANOVA with post-hoc LSD tests. Regression analysis was used to determine whether the observed SSRP changes correlated with clinical outcomes under the blind, particularly with regard to concurrent changes in UPDRS motor ratings, BDI ratings, HVLT, Stroop Interference Test, and SDMT. Group-wise differences and clinical correlations were also evaluated regionally within the major SSRP nodes (Table 1) using spherical volumes-of-interest (VOI, radius = 6 mm), as described elsewhere (43).

Network changes under the blind: comparison with gene therapy. We additionally compared the changes in network expression observed under the blind in the sham-operated participants with those measured in the 16 subjects who were randomized to gene therapy with successful delivery of the viral vector to the STN target (Supplemental Figure 1 and ref. 24). Based on the UPDRS criteria for sham response described above (7), 14 of these subjects were classified as GAD$_n$, and 2 were classified as GAD$_{sh}$.

Network modulation: effects of unblinding. To determine the effect of unblinding on SSRP expression, we analyzed baseline, 6-month, and 12-month scan data from the SHAMR and GAD$_n$ subjects who
were unblinded between the latter 2 imaging sessions. RMANOVA was used to assess the time course of SSRP expression in each group and to identify potential group × time interaction effects in the network activity data. All statistical analyses were performed using SPSS 13.0 for Windows (SPSS Inc.). Results were considered significant for $P < 0.05$. Data are presented as the mean ± SEM (bars and error bars, respectively), unless stated otherwise.

Monte Carlo simulations. We first created a means of identifying a priori sham–susceptible trial participants, i.e., subjects likely to develop prominent sham effects under the blind. This was done by determining a prerandomization network criterion using the distribution (mean and SD) of baseline SSRP values that was observed in the data (see text). This criterion (dashed line in Figure 3, A and B) was fixed at −0.75, the median value for baseline SSRP expression observed in sham responders. We noted that the sham responders with subthreshold baseline values, i.e., the 50% of SHAM* subjects with low network expression at baseline, had greater improvement under the blind than did their high-expression counterparts. Based on these observations, we hypothesized that excluding such individuals improves the efficiency of randomized treatment studies by reducing the magnitude and variability of sham effects in trial populations, with concomitant lowering of sample size.

To test this hypothesis, Monte Carlo simulations were performed with or without excluding sham-susceptible individuals before randomization to the active GAD-simulated (GAD_sim) or sham-simulated (SHAM_sim) treatment categories. In both models, we estimated the minimum number of randomized subjects needed to detect a difference in GAD versus SHAM clinical outcomes equivalent to that for ∆UPDRS in the 6-month blinded phase of the STN AAV-GAD trial. In each set of simulations, subjects were randomized 1:1 to the SHAM_sim and GAD_sim categories. ∆UPDRS of SHAM_sim was modeled according to the equation: $\Delta UPDRS_{\text{SHAM}} = b_0 + b_1 \times \text{SSRP}_{\text{SHAM}} + N(0,\delta^2)$, where $b_0$, $b_1$, and $\delta$ are the regression coefficients of the observed $\Delta UPDRS_{\text{SHAM}} = b_0 + b_1 \times \text{SSRP}_{\text{SHAM}}$ and $N(0,\delta^2)$ is a normal distribution, with a mean of 0 and a SD of $\delta$ determined empirically in a separate set of 10,000 simulations. A least-squares fit of the $\Delta UPDRS_{\text{SIM}}$ data was performed to induce the same variance in $\Delta UPDRS_{\text{SIM}}$ as that observed in the $\Delta UPDRS_{\text{SHAM}}$ data, and identical correlation analysis (Pearson’s coefficient) was performed with baseline SSRP values. $\Delta UPDRS$ for GAD_sim was simulated independently to have the same mean and SD as those observed empirically for the GAD group.

The probability of observing a significant difference in outcome for the GAD_sim and SHAM_sim groups was estimated for sample sizes varying from 2 to 300 in 10,000 random simulations. For each iteration, the group difference was considered significant for $P < 0.05$ by 2-tailed Student’s $t$ test. This was then repeated after excluding subjects with baseline SSRP values below the prespecified threshold value (see above). Simulation was performed using the statistics toolbox in MATLAB 7.7.0 (The MathWorks Inc.).

Study approval

Study protocols and patient consent forms were approved by the institutional review boards of the following participating institutions: Partners Human Research Committee (Boston, Massachusetts, USA); Colorado Multiple Institutional Review Board (Aurora, Colorado, USA); University of Rochester Research Subject Review Board (Rochester, New York, USA); Henry Ford Health System, Franklin Pointe Medical Center (Southfield, Michigan, USA); Wake Forest University School of Medicine (Winston-Salem, North Carolina, USA); Biomedical Institutional Review Board Office of Responsible Research Practices (Columbus, Ohio, USA); and Stanford University Environmental Health and Safety (Stanford, California, USA). Written informed consent was obtained from all participants after they were given detailed explanation of the procedures.

Acknowledgments

This work was funded by Neurologix, the NIH (R01 NS 35069, to D. Edelberg), and the National Institute of Neurological Disorders and Stroke (PS0 071675, Morris K. Udall Center of Excellence in Parkinson’s Disease Research at The Feinstein Institute for Medical Research, to D. Edelberg). The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or of the National Institute of Neurological Disorders and Stroke. The sponsor had no role in study design, data collection, analysis, and interpretation, writing of the report, or in the decision to submit the manuscript for publication. The authors wish to thank Thomas Chaly and Claude Margouleff for technical support and Yoon Young Choi for valuable assistance in editing and manuscript preparation.

Address correspondence to: David Edelberg, Center for Neuroscience, The Feinstein Institute for Medical Research, 350 Community Drive, Manhasset, New York 11030, USA. Phone: 516.562.2498; E-mail: david1@nshs.edu.


3. Katsnelson A. Experimental therapies for experimental sham-susceptible individuals before randomization to the active GAD-simulated (GAD_sim) or sham-simulated (SHAM_sim) treatment categories. In both models, we estimated the minimum number of randomized subjects needed to detect a difference in GAD versus SHAM clinical outcomes equivalent to that for $\Delta$UPDRS in the 6-month blinded phase of the STN AAV-GAD trial. In each set of simulations, subjects were randomized 1:1 to the SHAM_sim and GAD_sim categories. $\Delta$UPDRS of SHAM_sim was modeled according to the equation: $\Delta$UPDRS_sim = $b_0 + b_1 \times$SSRP_sim + $N(0,\delta^2)$, where $b_0$ and $b_1$ are the regression coefficients of the observed $\Delta$UPDRS_sim = $b_0 + b_1 \times$SSRP_sim and $N(0,\delta^2)$ is a normal distribution, with a mean of 0 and a SD of $\delta$ determined empirically in a separate set of 10,000 simulations. A least-squares fit of the $\Delta$UPDRS_sim data was performed to induce the same variance in $\Delta$UPDRS_sim as that observed in the $\Delta$UPDRS_SHAM data, and identical correlation analysis (Pearson’s coefficient) was performed with baseline SSRP values. $\Delta$UPDRS for GAD_sim was simulated independently to have the same mean and SD as those observed empirically for the GAD group.

...