Bilateral Globus Pallidus Internus Deep Brain Stimulation for Parkinson’s Disease: Therapeutic Effects and Motor Outcomes Prediction

Dingding Shen  
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

Linhao Cao  
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

Yun Ling  
GYENNO SCIENCE CO., LTD.

Dianyou Li  
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

Kang Ren  
GYENNO SCIENCE CO., LTD.

Xiaoliu Yu  
GYENNO SCIENCE CO., LTD.

Zhonglue Chen  
GYENNO SCIENCE CO., LTD.

Haiyan Zhou  
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

Jun Liu  
Department of Neurology & Institute of Neurology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine  
https://orcid.org/0000-0001-8300-8646

Research

Keywords: deep brain stimulation, globus pallidus interna, levodopa challenge test, machine learning

DOI: https://doi.org/10.21203/rs.3.rs-135557/v1

License: ☺️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

**Background:** Deep brain stimulation (DBS) has emerged as a highly effective surgical treatment for advanced Parkinson's disease (PD). Good response in levodopa challenge test has suggested as criterion to identify optimal candidates for surgery. However, the response to levodopa and DBS is not always congruent, and predictive value of the levodopa test remains controversial. This study was set out to identify predictors of response to DBS and develop a novel prediction model evaluating DBS candidacy.

**Methods:** Herein, we retrospectively analyzed 62 consecutive PD patients who underwent bilateral globus pallidus interna (GPi) DBS from 2016 to 2019. The changes in UPDRS-III (Unified Parkinson's Disease Rating Scale part III) total and subscores after surgery at one-year follow-up were evaluated and potential predictor variables were also collected. In the training cohort of 29 patients, we developed a novel machine learning method with 5-fold cross validations implementing on these variables to predict GPi DBS treatment outcomes in a multivariate linear analysis. Furthermore, the machine learning model was externally validated with another cohort of 33 GPi DBS PD patients.

**Results:** GPi DBS significantly improved postoperative motor function of PD patients. The overall UPDRS-III scores improved by 30.4%, with highest improvement in tremor (75.0%), followed by limb bradykinesia (27.5%), rigidity (27.3%) and axial bradykinesia (22.4%). Most intriguingly, improvement in tremor can be predicted with high accuracy using this prediction model (adjusted $R^2 = 0.82$ for absolute improvement, and adjusted $R^2 = 0.76$ for relative improvement), in which off medication tremor subscore was identified as the most powerful preoperative predictor. In the external validation cohort, the machine learning method showed good predictive performance.

**Conclusions:** We confirmed the effects of bilateral GPi-DBS with a one-year follow-up. The good performance of the present prediction model demonstrated the utility of machine-learning based motor response prediction after GPi DBS, based on clinical preoperative variables.

**Background**

Parkinson's disease (PD) is the second most common neurodegenerative disorder, with a prevalence of approximately to be 1.7% among Chinese people older than 65 years(1), compared to the world average of 1%(2). It is pathologically characterized by progressive loss of nigrostriatal dopaminergic innervation; the typical motor symptoms include resting tremor, rigidity, bradykinesia, gait and balance dysfunction(3). Although dopaminergic therapies are initially effective for management of motor symptoms, patients with advanced PD often experience severe motor fluctuations between dyskinesias (the on-medication phase) and immobility (the off-medication phase)(4, 5).

In recent years, deep brain stimulation (DBS) has evolved as an effective surgical therapy for management of medication-refractory motor fluctuations of PD(6). Currently, the subthalamic nucleus (STN) and the globus pallidus interna (GPi) are two main targets for DBS in PD. Increasing evidence across randomized clinical trials demonstrated similar, marked motor benefits in both two targets(7–9).
However, subtle target difference exists and determination of the single best surgical target for DBS remains controversial despite extensive research in the field(10).

Careful selection of applicable patients is the initial step in optimizing the efficacy of DBS and avoiding the risk of unsatisfactory outcome after surgery. Candidacy for DBS in PD is typically assessed by the preoperative motor response to levodopa using the levodopa challenge test (LCT) along with an interdisciplinary evaluation(9). According to the core assessment program for surgical interventional therapies in PD (CAPSIT-PD protocol), a levodopa-induced reduction of motor symptoms by > 30% of the UPDRS III (Unified Parkinson Disease Rating Scale III) has been suggested as criterion to identify optimal candidates for surgery(11). However, the response to levodopa is not always congruent with the effect of DBS, and predictive value of the levodopa test remains controversial(12, 13). While preoperative levodopa responsiveness has proved to be predictive of DBS efficacy on motor function and activities of daily living(14–16), it failed to show predictive value for improvement of disease-specific quality of life in a recent study(17). Therefore, it is critical to develop novel prediction models to reliably predicts postoperative motor response and evaluate DBS candidacy for individual patients. Establishing such a prediction tool facilitates the clinician to improve patient counselling, expectation management, and postoperative patient satisfaction. Notably, while non-conclusive results from previous prediction studies demonstrating preoperative factors that may potentially affect surgical outcomes of STN GBS, data on predictors of surgical outcomes of GPi DBS in PD patients is lacking so far.

Machine learning approaches based on routinely collected clinical data exclusively, without the need for any manual processing, are increasingly used in medical practice to predict clinical outcomes(18). In contrast to traditional statistics, predictive machine learning models generates outcome predictions for new, individual patients, instead of correlations between pre- and postoperative variables on a group level. Instead of only reproducing valid clinical decisions, machine learning on PD has been suggested to help making challenging clinical decisions(19). For example, one study implemented on age and neuroimaging data to predict individual patient response to dopaminergic therapy for PD(20).

In the present retrospective study, we aimed to report the development and proof-of-concept of a machine learning prediction model that generates motor outcomes one year after GPi DBS for individual PD patient.

**Methods**

**Patients**

We retrospectively studied 62 consecutive PD patients who underwent bilateral GPi-DBS in Ruijin Hospital (Shanghai, China) from November 2016 to July 2019. A total of 29 patients assessed by movement disorder specialists in department of neurology were studied as the training group. Another 33 PD patients assessed by different movement disorder specialists in department of neurosurgery were
enrolled in the validation study. All patients enrolled in the study have written informed consent, and the hospital ethics committee approved the study.

The inclusion criteria for surgery were as follows: (1) diagnosis of idiopathic PD based on the 2015 MDS-PD Criteria\(^\text{(21)}\); (2) response to L-dopa following a preoperative L-dopa challenge test (see below); (3) disabling motor fluctuations or dyskinesia despite all drug strategies; (4) informed consent for the surgery; and (5) good general health and accommodation of regular postoperative programming and follow-ups. The exclusion criteria were as follows: (1) contraindication for neurosurgery or high-field magnetic resonance imaging (MRI); 2) severe dementia or neuropsychiatric disorders, and (3) organic cerebral abnormalities.

**Surgical procedure and programming**

All patients underwent 3.0 T MRI before surgery. Surgical procedures were performed under general anesthesia. We applied the Leksell stereotactic frame to the patient’s head followed by a head CT scan. The specific target coordinates and trajectory were defined using the SurgiPlan system after the coregistration of MRI-CT images, targeting the posterior GPi. The implantable pulse generator (IPG) was placed subclavicularly and was connected with electrodes via subcutaneous wires. Postoperative imaging scan was performed to confirm satisfactory electrode placement of DBS leads and absence of complications.

The initial IPG programming was performed on the following day after surgery. Parameters including voltage, pulse-width, and frequency were optimized within the first 3 months of surgery and adjusted by two experienced movement specialists who referred to the Chinese standardized protocol\(^\text{(22)}\). The contacts were individually tested to inspect patients’ motor response and assess side effects. Preferably, the initial parameters were set to monopolar mode, with a pulse-width of ~ 90 µs and a frequency of ~ 135 Hz, and a stepwise increase in amplitude according to the patient’s response.

**Clinical Assessment**

Patients were evaluated preoperatively and 12 months postoperatively. A detailed medical history of the participants was completed. This involved confirming the patient’s age, gender, date of PD diagnosis, date of first intervention with antiparkinsonian medication, and current medications. The L-dopa equivalent dose (LED) in the preoperative medication regimen was calculated according to Tomlinson et al.\(^\text{(23)}\). Motor function was evaluated preoperatively using the Movement Disorder Society Unified Parkinson Disease Rating Scale-Motor Part (MDS UPDRS-III)\(^\text{(24)}\) and was scored in both off (MedOff) and on antiparkinsonian medication (MedOn) conditions. The motor examination portion (part III) of the UPDRS was performed to provide a clinically defined- “medication OFF” motor score (PreOFF). Then they were valued in a “medication ON” state when the best clinical response was obtained following a dose of levodopa (PreON).
Patients underwent the levodopa challenge test according to the CAPSIT-PD protocol. This test involved participants visiting the research center after at least 12 h without intake of PD medication (practically defined “medication OFF” state), allowing for an appropriate washout of levodopa. A single supra-threshold dose of L-dopa (the usual effective dose taken in the morning × 1.5) (25) was administered subsequently for the MedOn condition, in which the patient and the investigator agreed that the best functional benefits were achieved. Preoperative response to levodopa was calculated as the difference in UPDRS-III off and on drug: Absolute improvement from levodopa = \( \text{PreOFF} - \text{PreON} \).

Postoperatively, all patients were tested under two conditions: medication and stimulation both ‘OFF’, and medication ‘OFF’ and stimulation ‘ON’. The evaluation for the MedOff / StimOff condition was performed following overnight dopaminergic medication cessation and turning stimulation off for 1 h. For the MedOff / StimOn condition, evaluations were performed 1 h after restarting stimulation. The relative improvement of motor symptoms from GPi DBS was defined as:

\[
\text{Response to DBS (\%)} = \frac{\text{off drug off stimulation (PostOFF)} - \text{off drug on stimulation (PostON)}}{\text{off drug off stimulation (PostOFF)}} \times 100\%
\]

To compare specific symptom improvement, the UPDRS-III (motor) section was broken up into 4 composite symptom scores as follows: 1) tremor (Sect. 3.15 to 3.18); 2) rigidity (Sect. 3.3); 3) axial bradykinesia (Sect. 3.1, and 3.9 to 3.13); and 4) limb bradykinesia (Sect. 3.2, 3.4 to 3.8, and 3.14).

**Statistical analysis**

Statistical analysis was performed using SAS JMP 13 and Python3.6. Continuous variables were presented as mean ± SD, while classifying variables were presented as percentage. Wilcoxon test (paired samples) with Bonferroni correction was used to determine whether there was a significant improvement of motor symptoms.

The prediction models were implemented on the preoperative predictors and trained in 29 patients. We used 5-fold cross-validation for stepwise predictor selection. Briefly, the original samples were randomly partitioned into 5 equal sized subsamples. Of the 5 subsamples, a single subsample was retained as the validation data for testing the model, and the remaining 4 subsamples were used as training data. In the process of feature selection, we took the maximum goodness of fit of regression model as the criterion. When a new variable is put in, a new regression model with a new R-Square is generated. After we have traversed all variables, we selected the model with the largest R-square as the final model. The candidate predictors tested in our study include: 1) patient’s age at time of PD; 2) disease duration; 3) absolute values of \( \text{PreOFF} \) and \( \text{PreON} \) UPDRS-III total and subscores; 4) The combined of the above predictors such as predictor A multiply predictor B as a new predictor for enhancement of fitting. F statistic test was performed for the total model by comparing with the null model. The coefficient of the predictors was tested by Students T test. The adjusted R-Square is used for measurement of the fitness. RMSE is used for measurement of error. The predictor subset leading max R-Square was selected to build the prediction model.
Models are retrievable by assembling the coefficients to linear equations. The models are validated by the extra 33 samples and the results are described in detail by different error percentiles.

**Results**

**Demographic of recruited patients**

A total of 62 PD patients with complete data were enrolled in the study. The main clinical and demographic characteristics of patients at baseline were presented in Table 1. 40 (64.5%) patients were men and 22 (36.5%) were women. Across all patients, the mean age at disease onset was 51.5 ± 7.6 years. The mean age at the time of surgery received bilateral GPi stimulator implantation was 65.4 ± 7.8 years. The mean disease duration at the time of GPi DBS was 154.5 ± 60.0 months. Patients were followed up around 12 months after operation.
Table 1
Characteristics of patients at baseline

|                          | Total | Training | Validation |
|--------------------------|-------|----------|------------|
| Sex (Male, %)            | (40, 64.5%) | (20, 69.0%) | (20, 60.6%) |
| Age at the time of DBS (y)| 65.4 ± 7.8  | 65.5 ± 7.3  | 63.5 ± 8.2  |
| Age at disease onset (y) | 51.5 ± 7.6  | 50.8 ± 7.9  | 52.1 ± 7.3  |
| Disease duration (m)     | 154.5 ± 60.0 | 174.5 ± 66.9 | 137.0 ± 47.6 |
| UPDRS III Medication off |       |          |            |
| Total                    | 48.0 ± 11.4 | 47.2 ± 12.4 | 48.7 ± 10.6 |
| Tremor                   | 6.4 ± 6.1   | 6.4 ± 6.0   | 6.4 ± 6.3   |
| Rigidity                 | 9.8 ± 3.9   | 8.6 ± 3.0   | 10.1 ± 4.4  |
| Axial bradykinesia      | 13.8 ± 5.1  | 13.2 ± 5.3  | 14.0 ± 5.0  |
| Limb bradykinesia       | 18.0 ± 5.1  | 19 ± 4.9    | 17.1 ± 5.2  |
| UPDRS III Medication on  |       |          |            |
| Total                    | 25.5 ± 8.5  | 26.6 ± 9.0  | 24.6 ± 8.1  |
| Tremor                   | 1.4 ± 2.3   | 1.8 ± 2.6   | 1.0 ± 1.9   |
| Rigidity                 | 5.6 ± 3.7   | 5.0 ± 3.2   | 6.1 ± 4.4   |
| Axial bradykinesia      | 7.1 ± 3.8   | 6.7 ± 3.9   | 7.5 ± 3.6   |
| Limb bradykinesia       | 11.4 ± 4.2  | 13.1 ± 3.8  | 10.0 ± 4.0  |

Data are expressed as numbers, with percentages in parentheses, or as means ± SE.

The demographic and clinical characteristics of the training cohort (29 patients) and validation cohort (33 patients) were similar at baseline (Table 1). No significant differences were observed in the baseline measures.

**GPi-DBS significantly improved the postoperative motor function of PD patients in the entire cohort**

As illustrated in Fig. 1, all motor functions including total score of UPDRS III and subscores for every section all demonstrated a significant improvement in postoperative “off-medication/on-stimulation” (*PostON*) state compared with postoperative “off-medication/off-stimulation” (*PostOFF*) state.
GPi stimulation induced a significant improvement of 30.4% (mean change, 15.0 points; 95% CI, 12.0–19.3; \( p < 0.001 \)) in the overall UPDRS-III score. Tremor was significantly improved by 75.0% (mean change, 4.2 points; 95% CI, 2.5–6.2; \( p < 0.001 \)), rigidity by 27.3% (mean change, 2.8 points; 95% CI, 1.9–4.1; \( p < 0.001 \)), axial bradykinesia by 22.4% (mean change, 2.8 points; 95% CI, 2.1–3.6; \( p < 0.001 \)), and limb bradykinesia by 27.5% (mean change, 5.2 points; 95% CI, 3.9–7.1; \( p < 0.001 \)). In particular, improvement of tremor was the most prominent among the components of UPDRS III.

**Prediction of the effect of GPi-DBS operation on motor function**

The candidate predictors tested in this study are: 1) patient’s age at the onset of PD; 2) disease duration; 3) absolute values of *PreOFF* and *PreON* UPDRS-III total and subscores; 4) The combined predictors derived from the above predictors.

The correlations between preoperative predictors and postoperative outcomes for the UPDRS-III total- and sub-scores were measured by regression analysis (Table 2 and Fig. 2). After discarding factors with low predictive power, we chose the model with the highest R-Square and obtained the following formulas.

Table 2. Multivariate linear model in predicting GPi DBS motor outcome

1) Prediction of absolute improvement in motor outcome
| Term                                      | coefficient | p-value | adjusted R2 | RMSE  |
|-------------------------------------------|-------------|---------|-------------|-------|
| **Total score**                           |             |         |             |       |
| intercept                                 | 11.136      | 0.043   | 0.74        | 4.95  |
| PreON limb bradykinesia                  | -0.672      | 0.122   |             |       |
| PreOFF tremor                             | 1.056       | 0.001   |             |       |
| (PreON limb bradykinesia -13.07) *       | -0.499      | <.001   |             |       |
| (PreOFF tremor-6.38)                      |             |         |             |       |
| PreON tremor                              | -0.406      | 0.615   |             |       |
| (PreON limb bradykinesia -13.07) *       | 0.638       | 0.000   |             |       |
| (PreON tremor-1.79)                      |             |         |             |       |
| PreON rigidity                            | -0.219      | 0.489   |             |       |
| PreOFF limb bradykinesia                 | 0.369       | 0.168   |             |       |
| (PreON rigidity-5) * (PreOFF tremor-6.38)| -0.158      | 0.030   |             |       |
| **Tremor**                                |             |         |             |       |
| intercept                                 | 1.369       | 0.285   | 0.82        | 2.05  |
| PreOFF rigidity                           | -0.179      | 0.215   |             |       |
| PreON tremor                              | 0.286       | 0.158   |             |       |
| (PreOFF rigidity-8.62) *(PreON tremor-1.79)| -0.175      | 0.006   |             |       |
| PreOFF tremor                             | 0.647       | <.001   |             |       |
| **Rigidity**                              |             |         |             |       |
| intercept                                 | -5.723      | 0.131   | 0.23        | 2.64  |
| PreOFF rigidity                           | 0.548       | 0.014   |             |       |
| age at disease onset                      | 0.110       | 0.118   |             |       |
| PreON axial bradykinesia                 | 0.005       | 0.974   |             |       |
| (age at disease onset-50.76)*(PreON axial bradykinesia-6.72) | 0.035 | 0.042 |
| PreON rigidity                            | -0.339      | 0.103   |             |       |
| **Axial bradykinesia**                    |             |         |             |       |
| intercept                                 | -0.365      | 0.736   | 0.22        | 1.78  |
| disease duration                          | 0.014       | 0.010   |             |       |
| PreOFF tremor                             | 0.112       | 0.060   |             |       |
| **Limb bradykinesia**                     |             |         |             |       |
| intercept                                 | 0.122       | 0.968   | 0.08        | 4.01  |
| PreOFF limb bradykinesia                 | 0.282       | 0.077   |             |       |
2) Prediction of relative improvement in motor outcome

| Term | coefficient | p-value | adjusted R2 | RMSE |
|------|-------------|---------|-------------|------|
| Total score | intercept | 0.582 | 0.005 | 0.35 | 0.15 |
| | age at disease onset | -0.007 | 0.079 |
| | PreOFF tremor | 0.015 | 0.005 |
| | PreON axial bradykinesia | -0.003 | 0.665 |
| | (age at disease onset-50.76) * PreON (axial bradykinesia-6.72) | 0.002 | 0.043 |
| Tremor | intercept | 1.582 | <.001 | 0.76 | 0.09 |
| | PreON limb bradykinesia | -0.080 | <.001 |
| | PreOFF limb bradykinesia | 0.009 | 0.135 |
| | PreON tremor | 0.037 | 0.004 |
| | (PreOFF limb bradykinesia -18.9) * (PreON tremor-2.4) | 0.008 | 0.005 |
| Rigidity | intercept | 0.550 | <.001 | 0.20 | 0.31 |
| | PreON rigidity | -0.038 | 0.063 |
| | PreON tremor | -0.018 | 0.477 |
| | (PreON Rigidity-5.17857) * (PreON tremor-1.85714) | 0.018 | 0.023 |
| Axial bradykinesia | intercept | 0.800 | 0.001 | 0.28 | 0.15 |
| | age at disease onset | -0.011 | 0.004 |
| | PreON limb bradykinesia | 0.002 | 0.814 |
| | (age at disease onset-50.7586) * (PreON limb bradykinesia-13.069) | 0.003 | 0.031 |
| Limb bradykinesia | intercept | 0.658 | 0.087 | 0.12 | 0.21 |
| | PreOFF rigidity | 0.011 | 0.415 |
| | age | -0.007 | 0.248 |
| | (age at disease onset-65.469) * (PreON rigidity-5) | 0.005 | 0.033 |
Prediction of improvement in UPDRS-III total scores

Total or subtotal scores would be calculated as the sum of the products of the corresponding entries of the team column and the coefficient column in Table 2. For example, UPDRS-III total score estimated as follows: UPDRS-III total = 11.14 – 0.22 × PreON rigidity + 0.37 × PreOFF axial bradykinesia – 0.67 × PreON axial bradykinesia + 1.06 × PreOFF tremor - 0.41 × PreON tremor - 0.16 × (PreON rigidity - 5) × (PreOFF tremor - 6.38) − 0.50 × [(PreON limb bradykinesia − 13.07) × (PreOFF tremor - 6.38)] + 0.63 × [(PreON limb bradykinesia − 13.07) × (PreON tremor - 6.38)].

The absolute improvement in UPDRS-III total score was significantly correlated with preoperative predictors (adjusted R² = 0.736, p < 0.001). In particular, PreON axial bradykinesia × PreOFF tremor subscores demonstrated the most significant prediction value (p = 0.0003). Combined predictors derived from PreON axial bradykinesia and PreOFF tremor, PreOFF axial bradykinesia and PreON tremor, as well as PreON axial bradykinesia and PreOFF tremor were incorporated to fit the nonlinear relationship.

Additionally, the relative improvement in UPDRS-III total score was significantly correlated with preoperative predictors (adjusted R² = 0.345, p = 0.0061). In particular, PreOFF tremor subscores demonstrated the most significant prediction value (p = 0.0047). Combined predictor derived from age at disease onset and PreON axial bradykinesia was incorporated to fit the nonlinear relationship.

Prediction of improvement in tremor subscores

The absolute improvement in UPDRS-III tremor score could be reliably predicted with preoperative predictors (adjusted R² = 0.82, p < 0.00001). In particular, PreOFF tremor subscores demonstrated the most significant prediction value (p < 0.001). Combined predictor derived from PreOFF rigidity and PreON tremor (p = 0.0057) was incorporated to fit the nonlinear relationship.

Similarly, the relative improvement in UPDRS-III total score could be reliably predicted with preoperative predictors (adjusted R² = 0.76, p < 0.00001). In particular, PreON limb bradykinesia and tremor subscores both demonstrated significant prediction value (p < 0.001). Combined predictor derived from PreOFF limb bradykinesia and PreON tremor (p = 0.005) was incorporated to fit the nonlinear relationship.

Prediction of improvement in rigidity, axial bradykinesia and limb bradykinesia subscores

In contrary to UPDRS total scores and tremors subscores, improvement in rigidity, axial bradykinesia and limb bradykinesia subscores were only weakly or poorly associated with potential predictors (adjusted R² < 0.5).

External validation

Given the relatively small size of the training cohort, we did external retrospective verification using to confirm the robustness of machine learning model. In the validation analysis, another cohort of 33 PD
patients who received GPi DBS were included.

The clinical characteristics of the training and validation groups of patients were similar at baseline level and postoperative state (Table 1). As shown in Table 3, prediction of the absolute improvement of tremor resulted in high accuracy. Most of the sample errors were symmetrically distributed around 0, and 50% of the sample prediction errors were between −1.4 and 1.2, and 80% of the sample prediction errors were between −2.4 and 2.5.

Table 3

| External validation of the model in predicting tremor improvement |
|---------------------------------------------------------------|
| 10 percentile | 25 percentile | 50 percentile | 75 percentile | 90 percentile |
| Prediction error of absolute improvement | -2.4 | -1.42 | -0.24 | 1.18 | 2.53 |
| Prediction error of relative improvement | -0.49 | -0.21 | 0 | 0.1 | 0.22 |

**Discussion**

**The beneficial effect of GPi deep brain stimulation on cardinal motor symptoms of PD**

In this large, single-center, retrospective study, bilateral GPi DBS provided substantial improvement in motor function of advanced PD patients at one year after surgery. We observed improvement in the off-medication on stimulation UPDRS III total scores by 30.4% at one year after GPi implantation compared to baseline of off-stimulation state. This finding was comparable with those from previous studies(7–9). The significant effect on the tremor, rigidity, limb, and axial symptoms was also consistent with literature reporting beneficial effect of GPi DBS in controlling cardinal motor symptoms(26). In particular, the most dramatic benefit was observed in tremor symptoms.

In general, bilateral GPi and STN DBS may be equally effective for treating PD motor symptoms(7). At the present time, The subthalamic nucleus is used more commonly as the target, despite the evidence from available randomized trials has not confirmed an advantage or disadvantage for either target(10). Ideally, longer comparative follow-up is needed to clarify the long-term effect of GPi DBS and potential benefits in nonmotor domains.

**Machine learning method using absolute values as predictors**

Preoperative levodopa responsiveness (LR) in the levodopa challenge test, a well-established predictor of motor improvements after STN DBS therapy(14), has been used to screen PD patients for DBS since its incorporation into the CAPSIT-PD protocol in 1999(11). Despite the widespread use of the LR, little
attention has been paid as to how it should be calculated. Whether the absolute change (aLR) or percentage change (%LR) in the UPDRS-III is a better measure remains unsupported (27). A traditional approach to quantifying motor improvement due to DBS treatment was simply implementing on the percent OFF to ON change in the UPDRS-III motor section scores (relative levodopa responsiveness). We term this the “relative” approach, because it indicates that only the relative magnitude of improvement is relevant, whereas the baseline score (PreOFF) is in itself irrelevant. For example, improvement from a baseline score of 100 down to 50 is equivalent to improvement from 20 down to 10, as both present an improvement by 50%. This strategy may potentially miss preoperative information with powerful predictive values. Use of absolute values (both PreOFF and PreON values) as predictors was a novel aspect of the present study, which increased prediction accuracy.

**The prediction model performance and its relevant predictive variables**

In our study, improvement in UPDRS-III total and tremor scores could be reliably predicted with preoperative predictors (adjusted $R^2 = 0.82$ for absolute improvement, and adjusted $R^2 = 0.76$ for relative improvement). *PreOFF* tremor subscores demonstrated the most significant prediction value. These results of the presented model confirm the proof-of-concept of machine learning prediction of postoperative motor outcome based on preoperative clinical variables. To the best of our knowledge, our work is the first to systematically describe the predictive value of preoperative L-dopa responsiveness for GPI-DBS responsiveness.

We failed to establish reliable models to predict improvement in rigidity and bradykinesia based on potential predictors obtained from levodopa challenge test. We speculate that biochemical mechanisms of these motor symptoms were not only related to degeneration of the dopaminergic system, but also greater involvement of other neurotransmitter systems, such as loss of noradrenaline in the locus coeruleus, glutamatergic hyperactivity and loss of cholinergic pedunculopontine nucleus neurons (28).

We stress that the clinical preoperative factors incorporated in our prediction model are limited. In the future, the accuracy of the prediction model could be improved if other potential variables such as gender, disease subtype before DBS, duration of motor fluctuations (29) are included.

**Limitations**

Several limitations should temper the strength of these results. First, the main limitation of our study and its interpretation was its retrospective character, although we have systemically evaluated pre- and postoperatively the clinical features of all study subjects. Second, another potential shortcoming of this study is that the predictive value of tremor may have been overestimated because of a selection bias toward DBS patients who have medication-responsive tremor. The current evidence remains insufficient to clarify predictive value of medication-resistant tremor in PD patients who are considered candidates for GPI DBS. Third, the sample size of our training cohort and validation cohort is relatively small. Larger, prospective and multicenter studies in the future may help confirm our findings.
Conclusions

In the present retrospective study, bilateral GPi-DBS provided substantial improvement in motor function of advanced PD patients with a one-year follow-up. Furthermore, the good performance of the present prediction model demonstrated the utility of machine-learning based motor response prediction after GPi DBS, based on clinical preoperative variables.

Abbreviations

PD: Parkinson's disease

DBS: Deep brain stimulation

STN: Subthalamic nucleus

GPi: Globus pallidus interna

UPDRS-III: Unified Parkinson's Disease Rating Scale part III

Declarations

Ethics approval and consent to participate

All patients enrolled in the study have written informed consent. This study was approved by the Research Ethics Committee of Ruijin Hospital affiliated to Shanghai Jiaotong University school of medicine.

Consent to publication

Not applicable.

Availability of data and material

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding
This work was supported by grants from the National Key Research and Development Program (2016YFC1306505) and the National Natural Science Foundation of China (81471287, 81501097, 61901256, 91949120).

**Authors' contributions**

Jun Liu, Haiyan Zhou and Zhonglue Chen: study concept and design, interpretation of data, manuscript revision. Dingding Shen, Linhao Cao and Yun Ling: study design, statistical analyses, interpretation of data, drafting/ manuscript revision. Dianyou Li, Kang Ren and Xiaoliu Yu: statistical analyses, interpretation of data, review of manuscript. All authors read and approved the final version of the manuscript.

**Acknowledgement**

Not applicable.

**References**

1. Zhang ZX, Roman GC, Hong Z, Wu CB, Qu QM, Huang JB, et al. Parkinson's disease in China: prevalence in Beijing, Xian, and Shanghai. Lancet. 2005;365(9459):595-7.
2. Kalia LV, Lang AE. Parkinson's disease. Lancet. 2015;386(9996):896-912.
3. Armstrong MJ, Okun MS. Diagnosis and Treatment of Parkinson Disease: A Review. JAMA. 2020;323(6):548-60.
4. Ahlskog JE, Muenter MD. Frequency of levodopa-related dyskinesias and motor fluctuations as estimated from the cumulative literature. Mov Disord. 2001;16(3):448-58.
5. Vidailhet M. Movement disorders in 2010: Parkinson disease-symptoms and treatments. Nat Rev Neurol. 2011;7(2):70-2.
6. Weaver FM, Follett K, Stern M, Hur K, Harris C, Marks WJ, Jr., et al. Bilateral deep brain stimulation vs best medical therapy for patients with advanced Parkinson disease: a randomized controlled trial. JAMA. 2009;301(1):63-73.
7. Anderson VC, Burchiel KJ, Hogarth P, Favre J, Hammerstad JP. Pallidal vs subthalamic nucleus deep brain stimulation in Parkinson disease. Arch Neurol. 2005;62(4):554-60.
8. Follett KA, Weaver FM, Stern M, Hur K, Harris CL, Luo P, et al. Pallidal versus subthalamic deep-brain stimulation for Parkinson's disease. N Engl J Med. 2010;362(22):2077-91.
9. Odekerken VJ, van Laar T, Staal MJ, Mosch A, Hoffmann CF, Nijssen PC, et al. Subthalamic nucleus versus globus pallidus bilateral deep brain stimulation for advanced Parkinson's disease (NSTAPS study): a randomised controlled trial. Lancet Neurol. 2013;12(1):37-44.
10. Ramirez-Zamora A, Ostrem JL. Globus Pallidus Interna or Subthalamic Nucleus Deep Brain Stimulation for Parkinson Disease: A Review. JAMA Neurol. 2018;75(3):367-72.

11. Defer GL, Widner H, Marie RM, Remy P, Levivier M. Core assessment program for surgical interventional therapies in Parkinson's disease (CAPSIT-PD). Mov Disord. 1999;14(4):572-84.

12. Fasano A, Romito LM, Daniele A, Piano C, Zinno M, Bentivoglio AR, et al. Motor and cognitive outcome in patients with Parkinson's disease 8 years after subthalamic implants. Brain. 2010;133(9):2664-76.

13. Floden D, Cooper SE, Griffith SD, Machado AG. Predicting quality of life outcomes after subthalamic nucleus deep brain stimulation. Neurology. 2014;83(18):1627-33.

14. Charles PD, Van Blercom N, Krack P, Lee SL, Xie J, Besson G, et al. Predictors of effective bilateral subthalamic nucleus deep brain stimulation for PD. Neurology. 2002;59(6):932-4.

15. Welter ML, Houeto JL, Tezenas du Montcel S, Mesnage V, Bonnet AM, Pillon B, et al. Clinical predictive factors of subthalamic stimulation in Parkinson's disease. Brain. 2002;125(Pt 3):575-83.

16. Abboud H, Genc G, Thompson NR, Oravivattanakul S, Alsallom F, Reyes D, et al. Predictors of Functional and Quality of Life Outcomes following Deep Brain Stimulation Surgery in Parkinson's Disease Patients: Disease, Patient, and Surgical Factors. Parkinsons Dis. 2017;2017:5609163.

17. Schuepbach WMM, Tonder L, Schnitzler A, Krack P, Rau J, Hartmann A, et al. Quality of life predicts outcome of deep brain stimulation in early Parkinson disease. Neurology. 2019;92(10):e1109-e20.

18. Meyer A, Zverinski D, Pfahringer B, Kempfert J, Kuehne T, Sundermann SH, et al. Machine learning for real-time prediction of complications in critical care: a retrospective study. Lancet Respir Med. 2018;6(12):905-14.

19. Cerasa A. Machine learning on Parkinson's disease? Let's translate into clinical practice. J Neurosci Methods. 2016;266:161-2.

20. Ballarini T, Mueller K, Albrecht F, Ruzicka F, Bezdicek O, Ruzicka E, et al. Regional gray matter changes and age predict individual treatment response in Parkinson's disease. Neuroimage Clin. 2019;21:101636.

21. Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord. 2015;30(12):1591-601.

22. Chen S, Gao G, Feng T, Zhang J, Chinese Medical Association neurosurgery branch of functional neurosurgery g, Chinese Medical Association neurology branch of Parkinson's d, et al. Chinese expert consensus on programming deep brain stimulation for patients with Parkinson's disease. Transl Neurodegener. 2018;7:11.

23. Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. Mov Disord. 2010;25(15):2649-53.

24. Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord. 2008;23(15):2129-70.
25. Pollak P. Deep brain stimulation for Parkinson's disease - patient selection. Handb Clin Neurol. 2013;116:97-105.

26. Dankers F, Traverso A, Wee L, van Kuijk SMJ. Prediction Modeling Methodology. In: Kubben P, Dumontier M, Dekker A, editors. Fundamentals of Clinical Data Science. Cham (CH)2019. p. 101-20.

27. Pieterman M, Adams S, Jog M. Method of Levodopa Response Calculation Determines Strength of Association With Clinical Factors in Parkinson Disease. Front Neurol. 2018;9:260.

28. Devos D, Defebvre L, Bordet R. Dopaminergic and non-dopaminergic pharmacological hypotheses for gait disorders in Parkinson's disease. Fundam Clin Pharmacol. 2010;24(4):407-21.

29. Su XL, Luo XG, Lv H, Wang J, Ren Y, He ZY. Factors predicting the instant effect of motor function after subthalamic nucleus deep brain stimulation in Parkinson's disease. Transl Neurodegener. 2017;6:14.

Figures

![Figure 1](image)

**Figure 1**

Effects of GPi DBS on motor outcome Patients were assessed at 12-months follow-up. Means are plotted with the error bar representing the standard deviation. ****p < 0.0001 for significant difference between conditions.
Figure 2

Scatter plot and linear regression analysis showing the relationship between the preoperative predictors and the postoperative DBS response. Preoperative predictors were significantly correlated with improvement in UPDRS-III total and tremor subscores. Individual data points are represented as black dots. Black line indicates line of best fit.