



Predicting Motor Outcome of Subthalamic Nucleus Deep Brain Stimulation for Parkinson's Disease Using Quantitative Susceptibility Mapping and Radiomics: A Pilot Study

OPEN ACCESS

Edited by:

Hua Feng,
Army Medical University, China

Reviewed by:

Fangang Meng,
Capital Medical University, China
Xiaofei Hu,
Army Medical University, China

*Correspondence:

Naying He
i@nayinghe.com
Dianyou Li
ldy11483@rjh.com.cn
Qian Wang
wang.qian@sjtu.edu.cn

† These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Neurodegeneration,
a section of the journal
Frontiers in Neuroscience

Received: 26 June 2021

Accepted: 17 August 2021

Published: 07 September 2021

Citation:

Liu Y, Xiao B, Zhang C, Li J, Lai Y,
Shi F, Shen D, Wang L, Sun B, Li Y,
Jin Z, Wei H, Haacke EM, Zhou H,
Wang Q, Li D, He N and Yan F (2021)
Predicting Motor Outcome
of Subthalamic Nucleus Deep Brain
Stimulation for Parkinson's Disease
Using Quantitative Susceptibility
Mapping and Radiomics: A Pilot
Study. *Front. Neurosci.* 15:731109.
doi: 10.3389/fnins.2021.731109

Yu Liu^{1†}, Bin Xiao^{2†}, Chencheng Zhang^{3†}, Junchen Li⁴, Yijie Lai³, Feng Shi⁵,
Dinggong Shen^{5,6,7}, Linbin Wang³, Bomin Sun³, Yan Li¹, Zhijia Jin¹, Hongjiang Wei⁸,
Ewart Mark Haacke⁹, Haiyan Zhou¹⁰, Qian Wang^{2*}, Dianyou Li^{3*}, Naying He^{1*} and
Fuhua Yan¹

¹ Department of Radiology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, ² School of Biomedical Engineering, Institute for Medical Imaging Technology, Shanghai Jiao Tong University, Shanghai, China,

³ Department of Neurosurgery, Center for Functional Neurosurgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, ⁴ Department of Radiology, Changshu Hospital Affiliated to Nanjing University of Chinese

Medicine, Changshu, China, ⁵ Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China, ⁶ School of Biomedical Engineering, ShanghaiTech University, Shanghai, China, ⁷ Department of Artificial Intelligence, Korea University, Seoul, South

Korea, ⁸ School of Biomedical Engineering, Med-X Research Institute, Shanghai Jiao Tong University, Shanghai, China, ⁹ Department of Radiology, Wayne State University, Detroit, MI, United States, ¹⁰ Department of Neurology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Background: Emerging evidence indicates that iron distribution is heterogeneous within the substantia nigra (SN) and it may reflect patient-specific trait of Parkinson's Disease (PD). We assume it could account for variability in motor outcome of subthalamic nucleus deep brain stimulation (STN-DBS) in PD.

Objective: To investigate whether SN susceptibility features derived from radiomics with machine learning (RA-ML) can predict motor outcome of STN-DBS in PD.

Methods: Thirty-three PD patients underwent bilateral STN-DBS were recruited. The bilateral SN were segmented based on preoperative quantitative susceptibility mapping to extract susceptibility features using RA-ML. MDS-UPDRS III scores were recorded 1–3 days before and 6 months after STN-DBS surgery. Finally, we constructed three predictive models using logistic regression analyses: (1) the RA-ML model based on radiomics features, (2) the RA-ML+LCT (levodopa challenge test) response model which combined radiomics features with preoperative LCT response, (3) the LCT response model alone.

Results: For the predictive performances of global motor outcome, the RA-ML model had 82% accuracy (AUC = 0.85), while the RA-ML+LCT response model had 74% accuracy (AUC = 0.83), and the LCT response model alone had 58% accuracy (AUC = 0.55). For the predictive performance of rigidity outcome, the accuracy of

the RA-ML model was 80% (AUC = 0.85), superior to those of the RA-ML+LCT response model (76% accuracy, AUC = 0.82), and the LCT response model alone (58% accuracy, AUC = 0.42).

Conclusion: Our findings demonstrated that SN susceptibility features from radiomics could predict global motor and rigidity outcomes of STN-DBS in PD. This RA-ML predictive model might provide a novel approach to counsel candidates for STN-DBS.

Keywords: quantitative susceptibility mapping, radiomics, deep brain stimulation, Parkinson's disease, motor outcome, prediction

INTRODUCTION

Deep brain stimulation (DBS) targeting the subthalamic nucleus (STN) is a surgical therapy with class I evidence for improving motor symptoms of Parkinson's disease (PD) (Odekerken et al., 2013). Overall, STN-DBS yields improvements in motor symptoms by approximately 52% when observed in the DBS-on with medication off (DBS-ON/med-OFF) state at 6-month follow-up compared to preoperative med-OFF state (Kleiner-Fisman et al., 2006). However, the global and specific motor outcomes of DBS therapy remain highly variable (Okun et al., 2005) and the underlying mechanisms have not been fully unveiled. Several factors can contribute to the variable outcome after DBS surgery, including candidate selection, target localization, postoperative programming, and medication adjustment (Okun et al., 2005). Most of these key factors can be well-controlled at experienced sites. However, owing to the heterogeneity of patient-specific disease profile, predicting motor response remains challenging. Considering the cost, invasiveness and potential adverse effects of DBS surgery, it is vitally important to predict surgical efficacy accurately when counseling candidates for DBS (Lang and Widner, 2002).

In current clinical settings, satisfactory levodopa challenge test (LCT) responsiveness on motor symptoms has been widely used to select suitable PD candidates for STN-DBS. However, there is currently an ongoing debate regarding whether preoperative LCT response is a reliable predictor or not. Some studies demonstrated its predictive value (Charles et al., 2002) but not the others (Piboolnurak et al., 2007). In addition, the threshold values for the LCT responsiveness have not been standardized for DBS, ranging from 25 to 50% in published surgical series (Vingerhoets et al., 2002; Welter et al., 2002).

PD is a complex movement disorder with different responses to treatment, probably reflecting different pathological and biological contributions (Espay et al., 2017). One of the reasons for the heterogeneous therapeutic effects among PD patients could be the heterogeneity underlying PD pathology (Espay et al., 2017). With the advent of novel imaging techniques, it would be necessary to assess pathophysiological process and approach imaging biomarkers to facilitate the development of patient-specific interventions. The neuropathological hallmark of PD is dopaminergic neuronal loss in the iron-rich substantia nigra (SN) (Oakley et al., 2007). The neurodegenerative process is triggered long before the onset of motor disability in PD patients. Since the first report of elevated nigral iron content

in a PD patient in 1924. Consistent findings have demonstrated that, over time, excessive iron accumulates in the pigmented neurons of the substantia nigra pars compacta, especially in the form of a neuromelanin (NM)-iron complex (Jellinger et al., 1992; Zecca et al., 1996). Biologically, NM serves as an iron chelator leaving behind a NM-iron complex that activates microglia. Increased ferritin-loaded microglia in the SN initiates and prompts α -synuclein aggregation which links to PD development and progression (Li et al., 2010). The aggregation of α -synuclein and formation of Lewy bodies may lead to a continuing protective response which results in selective neuronal loss in the SN-dopaminergic (DA) system (Zucca et al., 2017). It is believed that one of the most promising neuroimaging biomarker candidates for PD is iron content as assessed with MRI. Several imaging studies observed decreased T2 relaxation times and increased relaxation rates (R2/R2*) in the SN in PD patients compared to healthy controls (Dexter et al., 1989; Martin et al., 2008; Péran et al., 2010; Yan et al., 2018; Ghassaban et al., 2019). An ultra-high field 7 Tesla (7T) MRI study based on T2*-weighted images delineated the nigrosome-1 territory of the SN with increased spatial resolution and new contrast (Lehéricy et al., 2014). Elevated nigral iron content was interpreted as the primary source of imaging changes in PD patients. However, these T2*/R2* measurements have a non-local component to them owing to the dipole effects of the iron sources (Schweser et al., 2011).

Neuroimaging has revolutionized quantitative iron detection *in vivo* with the advent of quantitative susceptibility mapping (QSM) (Haacke et al., 2015). Today QSM is able to directly represent the source of the iron and, hence, its concentration (He et al., 2015, 2021). Using phase data from multi-echo gradient recalled echo (GRE) sequences, QSM deconvolves the phase to produce a source susceptibility map. Post-mortem studies have validated that QSM provided a reliable and sensitive measurement to deep gray matter iron content (Langkammer et al., 2012). Importantly, QSM opens an avenue to investigate the role of iron deposition in PD pathology (Lotfipour et al., 2012) and clinical outcomes. Furthermore, radiomics with machine learning (RA-ML) analysis based on QSM precisely captures spatial heterogeneity in nigral iron distribution and subsequently has practical clinical applications in PD diagnosis (Li et al., 2019).

Since Langley first proposed the notion of machine learning (ML) in 1996, MRI-based ML for PD research has been widely used in disease classification (Salvatore et al., 2014; Adeli et al., 2016; Huppertz et al., 2016). In addition to the

traditional QSM-based analysis of volume and susceptibility in deep gray nuclei between PD and healthy controls (HC) (He et al., 2015), researchers have applied imaging features for further investigation. One recent QSM study observed significant difference of histogram features in several deep gray nuclei between PD patients and HC (Zhang Y. et al., 2020). They also demonstrated that combined model of the 10th percentile of SN (SN_{P10}) and the 75th percentile of Putamina (PUT_{P75}) successfully distinguished PD patients from HC far outweighing than the mean value (Zhang Y. et al., 2020). In addition to the first-order features of the deep gray nuclei such as mean value and percentile, other radiomics features, for example texture features, were also investigated for PD diagnosis. The texture features reflect the variation and distribution of local tissue properties, which are otherwise not apparently observed by a human reader (Zhang et al., 2012). These features can capture subtle changes within the structure and can be useful for detecting alteration from normal tissue in PD patients. The texture analyses of the SN showed the superiority of QSM to $R2^*$ map in discriminating PD and HC in one study (Li et al., 2019). In order to apply radiomics to patient-specific diagnosis instead of intergroup comparison, the researchers took these radiomics features as prior knowledge and constructed a novel diagnostic model with the approach of ML (Xiao et al., 2019).

Currently, there are no efficient indicators for predicting specific motor outcome, warranting the search for other imaging biomarkers to predict surgical efficacy and screen for suitable candidates. We hypothesized that individual iron distribution in the SN, which may reflect patient-specific disease trait, could potentially account for some variability in motor outcome of STN-DBS surgery. In this study, we introduce a computer-aided PD prognostic framework to develop a RA-ML model to predict motor outcome of STN-DBS when counseling PD candidates for STN-DBS.

MATERIALS AND METHODS

Patients

A total of thirty-three PD patients were prospectively recruited for bilateral STN-DBS surgery from December 2017 to September 2019. This study was approved by the Ethics Committee of Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine. All patients provided written informed consent prior to recruitment.

The PD patient selection criteria were: (i) clinically diagnosed idiopathic PD based on the United Kingdom PD Society Brain Bank Criteria (Hughes et al., 1992) at the time of surgery; (ii) absence of severe cognitive impairment [Montreal Cognitive Assessment (MoCA) score ≥ 17 , or Mini-mental State Examination (MMSE) score ≥ 24] (Hoops et al., 2009); and (iii) no significant neuropsychiatric disorders, such as depression [Beck Depression Inventory (BDI) score ≥ 15], anxiety disorders (e.g., panic disorder and obsessive-compulsive disorder), psychotic disorders (e.g., schizophrenia) or post-traumatic stress disorder; and (iv) no MRI contraindications.

Clinical Evaluation

Demographic and clinical information including gender, age, disease duration, levodopa equivalent daily dosage (LEDD), and LCT responsiveness were recorded prior to STN-DBS operation. Preoperative Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale Part III (MDS-UPDRS III) scores were recorded in the off-medication (med-OFF, 12 h discontinuation of levodopa) state. Postoperative MDS-UPDRS III scores and LEDD were obtained in the DBS-ON/med-OFF state at 6-months follow-up. MDS-UPDRS III total scores, corresponding to global motor outcome, were then subdivided to measure improvements in specific motor outcome, including rigidity (item 3.3), bradykinesia (item 3.4–3.8), and tremor (item 3.15–3.17) (Malaga et al., 2021). Motor improvement was calculated as follows,

Improvement(%)

$$= \frac{(\text{preoperative MDS-UPDRS III score}) - (\text{postoperative MDS-UPDRS III score})}{\text{preoperative MDS-UPDRS III score}} \times 100\%$$

Optimal surgical outcome was defined as having achieved no less than 30% improvement, consistent with previously published classification (Rodriguez et al., 2007) and suboptimal outcome was defined as less than 30% improvement. In this work, tremor was not included in the RA-ML analysis given that there were 100% improvement of tremor symptoms in most PD patients (16/33) in the postoperative DBS-ON/med-OFF state.

MRI Data Acquisition

Prior to STN-DBS operation, all patients were imaged at 3T MRI scanner (Ingenia, Philips Healthcare, Netherlands) with a 15-channel phased head coil. Thin foam pads both under and beside head were used to reduce potential motion artifacts. The imaging parameters were chosen as follows: (1) multi-echo GRE sequence: voxel resolution = $1 \times 1 \times 1 \text{ mm}^3$, flip angle (FA) = 12° , repetition time (TR) = 25 ms, eight echoes with the first echo times (TE1) = 3.3 ms, echo spacing (ΔTE) = 2.6 ms, sampling bandwidth = 673 Hz/pixel, number of slices = 136, and an acquisition time of 6 min and 16 s; (2) T2-weighted fast spin echo (FSE T2W): TR = 4,000 ms, TE = 106 ms, FA = 90° , voxel resolution = $0.75 \times 0.75 \times 1.5 \text{ mm}^3$, bandwidth = 440 Hz/pixel, and an acquisition time of 5 min and 36 s. Fluid attenuated inversion recovery (FLAIR) and diffusion-weighted imaging (DWI) data were also collected to ensure there were no cerebrovascular diseases or space-occupying lesions in the brain. The transverse plane was set to be parallel to the anterior commissure-posterior commissure line.

Targeting Procedure for STN-DBS

Each patient was defined as a potential candidate for STN-DBS after reaching group consensus among functional neurosurgeons, neurologists, neuropsychologists, and psychiatrists. All patients received bilateral STN-DBS following the standard procedure of screening for DBS eligibility at Ruijin Hospital. The targeting procedure for STN-DBS was performed as described previously

(Zhang C. et al., 2020). The electrodes of two manufacturers (model 3387–40, Medtronic, Minneapolis, MN, United States or model L302, PINS, Beijing, China) were used in the recruited PD patients, which had the same electrode size and physical parameters. For each patient, correct electrode placement was verified by postoperative CT imaging merged with preoperative high-resolution FSE T2W images.

Data Processing and Model Construction

Image Pre-processing and Segmentation

QSM reconstruction was performed in MATLAB (MathWorks, Inc., Natick, MA, United States). Original DICOM (Digital Imaging and Communications in Medicine) data from multi-echo GRE sequences were separated into magnitude and phase images. The magnitude images were used to extract brain tissue by using the Brain Extraction Tool (BET) (Smith, 2002) in FSL (FMRIB, University of Oxford, Oxford, United Kingdom). Then, a Laplacian-based algorithm was used to unwrap the raw phase data. Lastly, the susceptibility map was reconstructed through dipole inversion using STAR-QSM (streaking artifacts reduction for QSM) algorithm (Wei et al., 2015).

The regions-of-interest (ROIs) for the SN were manually drawn by two experienced radiologists using SPIN software (SpinTech, Inc., Bingham Farms, MI, United States). They were blinded to clinical details and drew the initial ROIs independently. When tracing the ROIs of SN, the QSM images were zoomed by a factor of four to make sure the objects could be accurately drawn. The ROIs were drawn to cover the bilateral SN on all slices where the SN signal was clearly visible on QSM images. The final ROIs were obtained from the initial ROIs using a semi-automated dynamic programming approach in SPIN software (He et al., 2021). Finally, we combined the final ROIs of all slices unilaterally to obtain the SN volume-of-interest (VOI).

Radiomics Feature Extraction

The radiomics feature extraction were performed by PyRadiomics, an open-source platform described in a previous publication (van Griethuysen et al., 2017). In total, 1,328 radiomics features were extracted from the SN VOIs of each patient, including 26 shape features, 252 first-order features, 336 Gray Level Co-occurrence Matrix (GLCM) features, 224 Gray Level Run Length Matrix (GLRLM) features, 224 Gray Level Size Zone Matrix (GLSZM) features, 196 Gray Level Dependence Matrix (GLDM) features, and 70 Neighboring Gray Tone Difference Matrix (NGTDM) features. The shape features were directly extracted from the SN VOIs. The first-order and texture features were extracted from the original QSM image or from the filtered QSM images with one of three specific filter types: (i) Wavelet filtering: Applying high (H) or low (L) pass filters to a 3D image in each of the three dimensions. We performed LLL and HHH to filter QSM images with relative high stability. (ii) LoG: Laplacian of Gaussian filter for highlighting edges. (iii) Gradient: Calculating and returning the gradient magnitude of the image.

Feature Selection by Stability

The intraclass correlation coefficient (ICC) was used to quantify the stability degree of each feature between the two raters. ICC

is an established tool to assess inter-rater reliability. In this study, ICC estimates was based on an absolute-agreement and two-way random-effects model, which follows the guidelines for choosing the appropriate form of the ICC (Shrout and Fleiss, 1979). Here, ICC (2,1) estimates were calculated for feature values using the Pingouin python package (Vallat, 2018). It is recommended that ICC estimates should be >0.80 , which indicates good reliability between the raters. In this study we selected 657 radiomics features whose ICC values were all >0.85 , which is a reasonable selection strategy (van Griethuysen et al., 2017). Since ICC is a statistical indicator, it cannot be assumed that each feature with $ICC >0.85$ will not have a large error for each case. Therefore, a relative error of 10% was used for further feature selection. Relative error is the ratio of an error in a measured quantity to the magnitude of that quantity, which indicates the credibility of the measurement. We preserved 149 features with relative error <0.1 , which aimed to select the features with a relatively high degree of credibility of the measurement.

Feature Selection by Prior Knowledge

Due to the limited sample size and high-dimensional feature vector, it is important to remove noisy features and avoid potential overfitting problems. As suggested in one study (Amoroso et al., 2018), another validation of the proposed methodology is that the selected features should correspond to the ones whose relationship with the disease has already been established. Therefore, we took prior knowledge into consideration. We assume that the features applied to classify optimal and suboptimal surgical outcomes in PD patients are related to PD pathophysiological changes. Among those 149 features, 10 had the same feature types with those from the feature clusters associated with PD and HC classification in our previous QSM studies (Cheng et al., 2019; Xiao et al., 2019), which were subsequently used for model construction.

Logistic Regression With Recursive Feature Elimination

Logistic regression (LR) was employed as a linear model for the classification task (Optimal vs. Suboptimal). After a linear transformation and a nonlinear activation function, all input radiomics features were transformed into classification probability values, ranging from 0 to 1. The value describes the confidence that a PD subject will have optimal surgical outcome. In the training stage, the cost function of LR was given by:

$$\min_{w,c} \frac{1}{2} \|w\|_p + C \underbrace{\sum_{i=1}^n \log(\exp(-y_i(X_i^T w + b)) + 1)}_{\text{loss function}}, p \in \{1, 2\}$$

where w and b represent the trainable parameters. $X_i \in \mathbb{R}^n$ and $y_i \in \{-1, +1\}$ is the radiomics feature vector and target label representing the i th sample. In the loss function, there are two hyperparameters: $C \in \{1, 3, 5, 7, 9, 13, 15\}$ and $p \in \{1, 2\}$, which controls the weight of the loss function and the complexity of the model, respectively.

Combined with LR, recursive feature elimination (RFE) was also employed for feature selection. RFE is a commonly used method of feature selection (Wottschel et al., 2019). And hyperparameters of the features are determined by the algorithm itself, therefore there is little manual intervention. Given that the LR model assigns weights to features, the goal of RFE is to select the most important ones by recursively selecting smaller and smaller sets of features. Firstly, the LR was trained on the initial set of features and the importance of each feature was assessed by its feature coefficient w . Then, the least important features were excluded. This procedure was recursively repeated on the stepwise-changed set until an optimal model was obtained. We set the number of desired features as a hyperparameter $N \in \{5, 7, 9, 11, 13\}$.

Development and Validation of the Model

Statistical tests were performed on Scikit-learning software when we developed the LR model. After feature selection by stability and prior knowledge, 10 features were finally used for model construction. All features were normalized by removing the mean and scaling to unit variance. We used Leave-One-Out cross-validator (LOOCV) to test the generalization performance of the model. For each experiment, we randomly selected one sample for testing and the rest samples for training. In the training stage, we selected the best hyperparameter based on the following criteria: selecting the model with the simplest complexity among those with the best training accuracy. Specifically, when accuracies of different models were the same, we chose the one with smaller $N \in \{5, 7, 9, 10\}$, $C \in \{1, 3, 5, 7, 9\}$, and $p \in \{1, 2\}$. A schematic flowchart of the RA-ML processing is shown in **Figure 1**.

In addition to the RA-ML model based on SN radiomics features, we also constructed the models as follows: (1) the RA-ML+LCT response model, which combined SN radiomic features with preoperative LCT response, (2) the LCT response model alone.

Statistical Analysis

Demographic information and clinical details, such as age at surgery, disease duration, preoperative LCT responsiveness, pre- and post- surgical LEDD were assessed in the two groups (Optimal vs. Suboptimal) using an independent sample t -test. Patient's gender was assessed using a Fisher's exact test. We applied Mann-Whitney U test for intergroup comparisons regarding the values of radiomics features. The level of significance was set at $p < 0.05$ for a two-tailed test. The statistical analyses mentioned above were performed in IBM SPSS Statistic Version 26.0 (International Business Machines Corporation, Armonk, NY, United States). Receiver operating characteristic (ROC) curves were constructed and area under the curve (AUC) values were used for assessment of predictive performances. Statistical comparisons of ROC curves among the three models were performed by DeLong's test (DeLong et al., 1988) in MedCalc Statistical Software version 19.7.2 (MedCalc Software Ltd, Ostend, Belgium). The Bonferroni correction was used to correct the error of multiple comparisons in ROC curves. The

adjusted p -value was set at $0.05/n$, where n = times of comparison (The adjusted p -value threshold for significance: 0.0167).

RESULTS

Participants, Baseline Characteristics, and Follow-Up Outcomes

Demographic and clinical information of all recruited patients are summarized in **Table 1**. No significant differences of any demographic or clinical characteristics were observed between the two groups (Optimal vs. Suboptimal).

Predictive Performances of the Three Models

The predictive performance was evaluated *via* accuracy, balance accuracy, AUC, sensitivity, and specificity (**Table 2**). Balance accuracy was used to overcome the problem with an imbalanced data set. It normalizes true positive and true negative predictions by the number of the positive and negative samples, which is in accordance with the division of our data into optimal and suboptimal cases, respectively. For the predictive performances of global motor outcome, the RA-ML model had an accuracy of 82% (AUC = 0.85), while the RA-ML+LCT response model had 74% accuracy (AUC = 0.83), and the LCT response model alone had 58% accuracy (AUC = 0.55). For the predictive performance of rigidity outcome, the accuracy of the RA-ML model was 80% (AUC = 0.85), superior to those of the RA-ML+LCT response model (76% accuracy, AUC = 0.82), and the LCT response model alone (58% accuracy, AUC = 0.42). For the predictive performance of bradykinesia outcome, the RA-ML model yielded an accuracy of 50% in comparison with the RA-ML+LCT model (accuracy = 50%) and the LCT response model alone (accuracy = 58%). **Figure 2** showed the ROC curves for the three models predictive of global and specific motor outcomes.

As shown in **Table 3**, among the ROC curves of the three models predictive of global motor outcome, the RA-ML model and the RA-ML+LCT response model were significantly higher than the LCT response model alone ($p = 0.001$ and $p = 0.005$, respectively), but the difference was not significant between the RA-ML model and the RA-ML+LCT response model ($p = 0.359$). Multiple ROC comparisons of the three models predictive of rigidity improvement all had significant differences: $p < 0.001$ for the RA-ML and the LCT response model comparison, $p = 0.011$ for the RA-ML and the RA-ML+LCT response model comparison, and $p = 0.003$ for the RA-ML+LCT response and the LCT response model comparison. As for the ROC curves predictive of bradykinesia improvement, there was no significant difference between any two of the models ($p > 0.0167$).

Radiomics Features With Significant Intergroup Differences

Among the 10 radiomics features selected for model construction, the values of Wavelet-LLL-GLRLM-RunEntropy ($p = 0.036$) and Wavelet-LLL-GLCM-IDN (Inverse Difference Normalized) ($p = 0.039$) were both significantly higher in

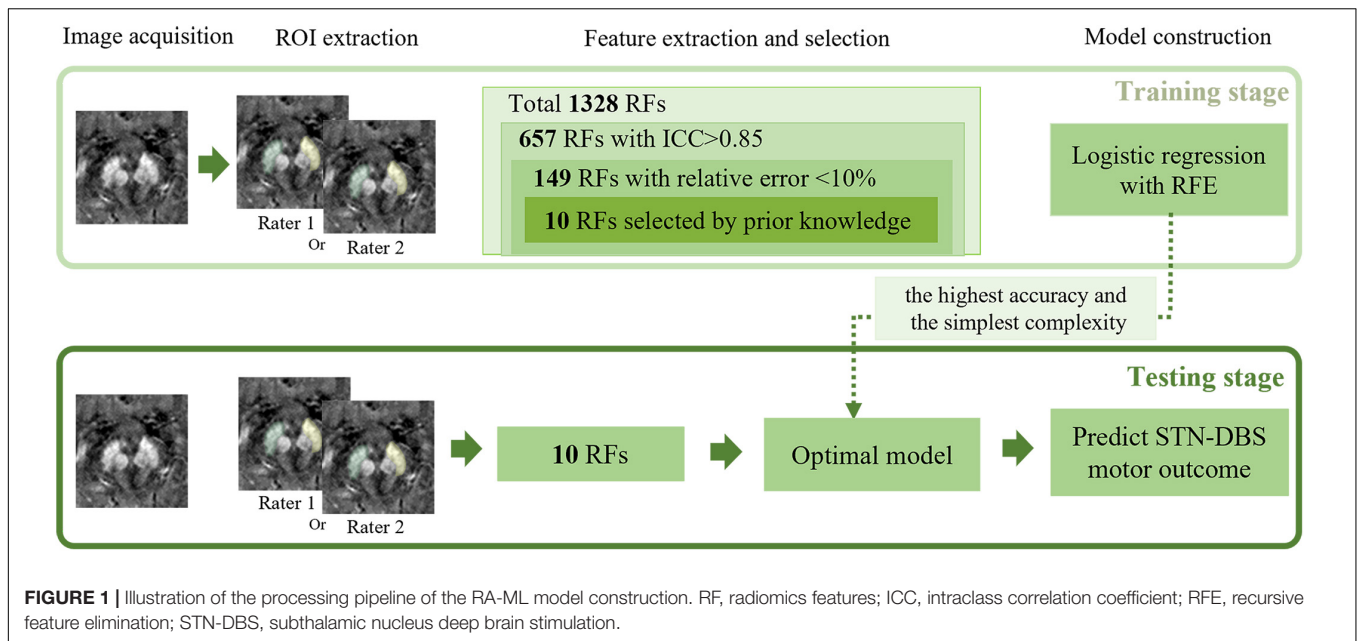


TABLE 1 | Demographic and clinical characteristics of patients with optimal and suboptimal motor outcome.

	Improvement	N	Age at surgery (Mean ± SD)	Gender (M/F)	LCT response (%)	Disease duration (Years)	Pre-LEDD (mg)	Post-LEDD (mg)
All patients	/	33	60.0 ± 10.1	21/12	48.2 ± 13.6	10.6 ± 4.3	848.1 ± 446.0	433.7 ± 250.6
Global	≥30%	20	59.0 ± 9.6	15/5	49.5 ± 15.0	11.0 ± 4.6	834.5 ± 362.4	389.0 ± 204.3
	<30%	13	61.7 ± 11.1	6/7	46.2 ± 11.5	10.1 ± 4.1	869.0 ± 567.3	498.0 ± 297.7
	<i>P</i>	/	0.46	0.14	0.51	0.58	0.83	0.220
Rigidity	≥30%	23	59.1 ± 9.7	15/8	50.7 ± 12.8	10.9 ± 4.3	911.5 ± 494.0	421.9 ± 247.7
	<30%	10	62.1 ± 11.3	6/4	42.3 ± 14.3	9.9 ± 4.4	702.2 ± 277.6	455.2 ± 256.7
	<i>P</i>	/	0.45	0.78	0.11	0.54	0.22	0.73
Bradykinesia	≥30%	10	57.4 ± 12.6	8/2	45.4 ± 13.1	12.5 ± 5.3	812.7 ± 402.4	404.0 ± 216.8
	<30%	23	61.2 ± 9.0	13/10	49.4 ± 14.0	9.8 ± 3.6	863.5 ± 471.5	444.1 ± 262.5
	<i>P</i>	/	0.33	0.26	0.45	0.10	0.77	0.68

Data are presented as mean ± standard deviation unless otherwise noted. LCT, levodopa challenge test; LEDD, levodopa equivalent daily dosage.

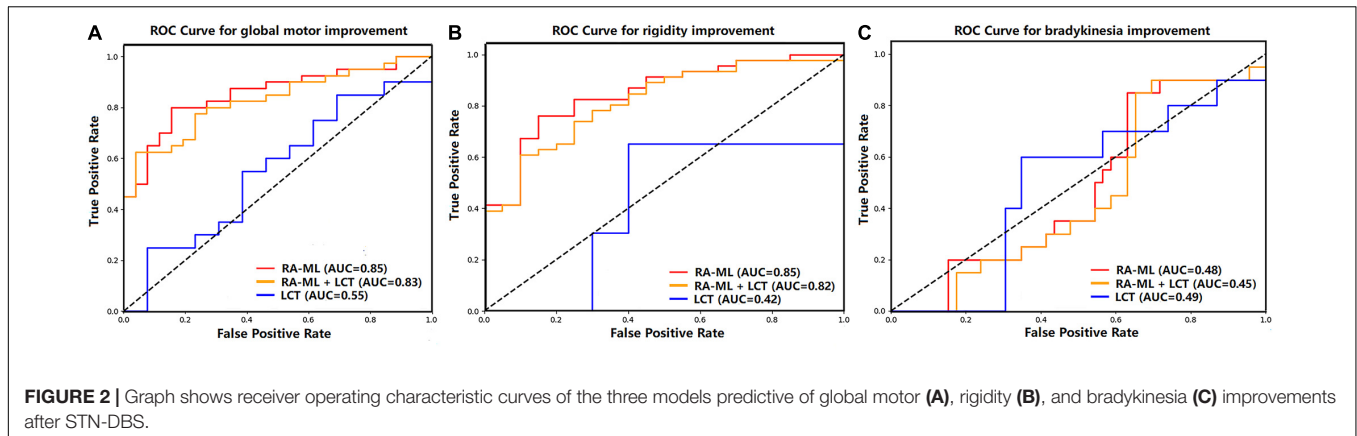
TABLE 2 | Predictive performances of the three models.

Motor outcome	Predictive model	Accuracy	Balance accuracy	AUC	Sensitivity	Specificity
Global	RA-ML	0.82	0.82	0.85	0.80	0.85
	LCT	0.58	0.57	0.55	0.60	0.54
	RA-ML+LCT	0.74	0.75	0.83	0.73	0.77
Rigidity	RA-ML	0.80	0.79	0.85	0.75	0.83
	LCT	0.58	0.58	0.42	0.60	0.57
	RA-ML+LCT	0.76	0.73	0.82	0.65	0.80
Bradykinesia	RA-ML	0.50	0.44	0.48	0.59	0.30
	LCT	0.58	0.58	0.49	0.57	0.60
	RA-ML+LCT	0.50	0.44	0.45	0.59	0.30

RA-ML model: the model based on SN susceptibility radiomics features.

LCT model: the model based on preoperative levodopa challenge test response.

RA-ML+LCT model: the model combining SN susceptibility radiomics features with preoperative levodopa challenge test response.



patients with optimal global motor outcome as compared to those with suboptimal efficacy (shown in A and B in **Figure 3**). No significant differences of any other radiomics features were observed between the two groups (Optimal vs. Suboptimal) regarding rigidity or bradykinesia improvements ($p > 0.05$).

DISCUSSION

To the best of our knowledge, this is the first study to predict motor outcome in PD patients based on individual pre-DBS quantitative images using RA-ML approach. It suggests that RA-ML analysis can capture the subtle differences in iron distribution within the SN structure, which could be identified as a potential imaging biomarker to understand the variability in global motor and rigidity improvements of STN-DBS for PD patients.

Advantages of the RA-ML Model

In this study, the LCT responsiveness models failed to predict global motor, rigidity and bradykinesia improvements, suggesting that the value of preoperative LCT response for predicting the short-term motor outcome of STN-DBS was limited. However, LCT is widely used to determine whether a patient may benefit from DBS surgery in current clinical settings

(Pollak, 2013). These conflicting results may be partly attributed to the enrolled PD patients of different stage in different studies. Importantly, a recent study observed that levodopa medication, rather than DBS itself, had a modulatory effect on basal ganglia activity during finger movement execution (Mueller et al., 2020), which implies different therapeutic mechanisms for levodopa and DBS. Therefore, predicting STN-DBS outcome *via* preoperative LCT responsiveness may be underpowered. Future studies must be carried out for clarification. Importantly, it is reported that levodopa and STN-DBS responses are not congruent (Zaidel et al., 2010). Researchers clarified the methodological shortcomings of previous studies and questioned the validity of using LCT responsiveness as a selection criterion for STN-DBS (Zaidel et al., 2010).

Different from LCT responsiveness, individual nigral iron measurements assessed with QSM data could reflect patient-specific disease trait and potentially help to understand PD prognosis (Bergsland et al., 2019). The predictive performance of the RA-ML model outweighs that of the LCT response model for both global motor and rigidity improvements, which offers a promising alternative to general-applicable LCT responsiveness in counseling PD candidates for STN-DBS. It can be partly explained by the explicit characterization of PD-related pathophysiological changes in the deep gray nuclei by radiomics features, and well-designed modeling between radiomics features and clinical presentation by the ML model. In addition, LCT has its own disadvantages: (1) discontinuation of levodopa for at least 12 h would raise the risk of motor complications; (2) levodopa may prime drug-induced dyskinesias; and (3) LCT is time-consuming.

More recently, imaging studies demonstrated the predictive value of structural and functional connectivity for motor outcome of STN-DBS in PD patients based on diffusion tensor imaging (DTI) (Horn et al., 2017) and resting-state functional MRI (rs-fMRI) (Younce et al., 2020). Smaller ventricular volumes measured on T1 weighted images were associated with more favorable motor outcome in one report (Younce et al., 2019) but not in others (Price et al., 2011). Nevertheless, predicting specific motor outcome based on individual level remains challenging. This study addressed the issues and provided a simple and convenient approach to predict global motor and

TABLE 3 | Comparison of ROC curves of the three predictive models.

	RA-ML ~ LCT	RA-ML+LCT ~ RA-ML	RA-ML+LCT ~ LCT
Global	0.001 [0.117, 0.483] ***a	0.359 [−0.027, 0.075]	0.005 [0.084, 0.468] ***c
Rigidity	<0.001 [0.114, 0.425] ***a	0.011 [0.007, 0.054] *b	0.003 [0.082, 0.396] ***c
Bradykinesia	0.901 [−0.208, 0.236]	0.303 [−0.022, 0.069]	0.740 [−0.187, 0.263]

Data are presented as p -value [95% confidence interval] unless otherwise noted.

*Indicates $p < 0.05$; ** indicates $p < 0.01$; *** indicates $p < 0.001$.

^aIndicates RA-ML performing better than LCT.

^bIndicates RA-ML performing better than RA-ML+LCT.

^cIndicates RA-ML+LCT performing better than LCT.

The adjusted p -value threshold for significance after Bonferroni correction is 0.0167.

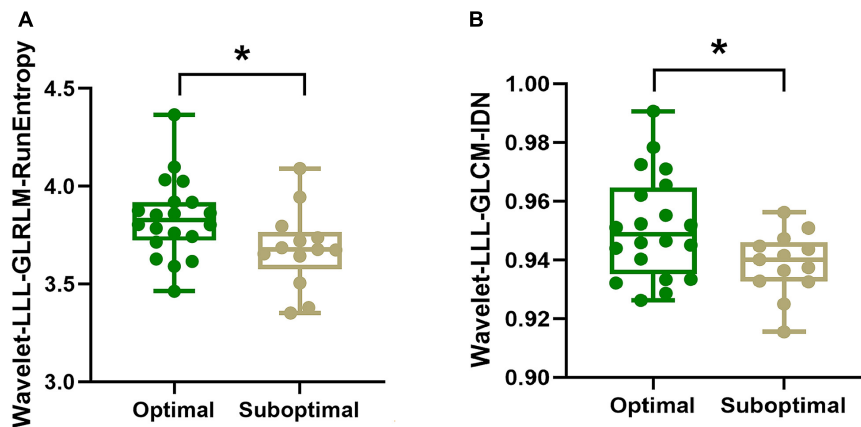


FIGURE 3 | Intergroup comparisons of radiomics features using Mann-Whitney U test. As shown in panels (A,B), respectively. The values of Wavelet-LLL-GLRLM-RunEntropy ($p = 0.036$) and Wavelet-LLL-GLCM-IDN (Inverse Difference Normalized) ($p = 0.039$) for the SN are higher in patients with optimal global motor outcome than those who had suboptimal improvement following STN-DBS surgery. In panels (A,B), the box denotes the 25th and 75th percentiles with the horizontal line denoting the median value. * indicates $p < 0.05$.

rigidity outcomes for individual PD patients. Furthermore, since QSM is easily accessible and imaging acquisition is relatively fast, the RA-ML model developed in this study has promising clinical applications to counsel STN-DBS candidates for neurosurgeons' decision-making.

A previous imaging study reported that the maximum T2-relaxation times (T2) for the SN failed to predict motor outcome of STN-DBS (Lönnfors-Weitzel et al., 2016). This conflicting finding with our work could be partly explained by the fact that change in water content can also change T2 acting as a confounder, but it does not affect the susceptibility measurements. Hence QSM outperforms T2 in measuring nigral iron content and potentially reflecting patient-specific disease trait. In addition, they calculated the absolute instead of the relative improvement in MDS-UPDRS III score. The relative improvement may be preferred to measure surgical efficacy since the absolute improvement calculation only indicates the magnitude of the improvement (Zaidel et al., 2010).

Interpretability of the RA-ML Model

The level of interpretability is about understanding how the RA-ML model makes decisions, which could help to comprehend the outcome based on the key radiomics features. Among the 10 features selected for model construction, the values of Wavelet-LLL-GLRLM-RunEntropy and Wavelet-LLL-GLCM-IDN for the SN were both significantly higher in patients with optimal global motor outcome as compared to those with suboptimal efficacy. GLRLM-RunEntropy in each VOI measures the randomness of run length and gray distribution. GLCM-IDN is another measure of the local homogeneity of an image. For both these radiomics features, a lower value indicates more homogeneity in the texture patterns, suggesting that those PD patients exhibiting more uniform iron signal with decreased iron heterogeneity may respond worse to STN-DBS treatment.

This finding is comparable to the iron heterogeneity changes underlying PD pathology. Different from the iron-rich SN pars

reticulata (SNpr), the lateral-ventral part of the SNpc is packed with melanized dopaminergic neurons with less iron content (Damier et al., 1999). Imaging studies have demonstrated that iron increases predominantly occurred in the lateral-ventral SNpc (Langley et al., 2017; He et al., 2020) in PD; this is in line with the pathological findings (Damier et al., 1999). Therefore, this region-specific high iron accumulation changes the spatial patterns of nigral iron content which could lead to decreased heterogeneity of iron distribution in PD patients. Importantly, susceptibility measurements using RA-ML could capture this subtle difference in spatial iron distribution within the SN structure.

Taken together, this study suggests that the heterogeneity of spatial iron distribution in the SN may be pertinent to patient-specific pathological patterns, which could help to understand the variability of motor outcome in PD patients after STN-DBS.

Prediction of Specific Motor Outcome

This study conducted RA-ML analysis of specific motor outcome which is not available in most previous studies (Price et al., 2011; Lönnfors-Weitzel et al., 2016; Horn et al., 2017; Younce et al., 2019, 2020). Cardinal motor symptoms in PD patients include rigidity, bradykinesia, and tremor. At least 89% of PD patients suffered from rigidity causing stiffness or inflexibility in muscles (Mutch et al., 1986). In this study, the predictive performance of the RA-ML model was superior to the LCT response model regarding rigidity improvement. The mechanisms underlying the high-precision performance of the RA-ML model predictive of rigidity improvement are not yet fully understood; however, this finding may suggest that susceptibility measurements of nigral iron accumulation could help to understand the variability in rigidity outcome of STN-DBS for PD patients. Bradykinesia is one of the early signs of PD characterized by a reduced ability to move. However, none of the three models in this study were able to predict bradykinesia improvement. This may partly be

explained by the fact that STN-DBS treatment affected each motor symptom *via* different mechanisms (Temperli et al., 2003).

Limitations

There are several limitations to this study. Firstly, the sample size is relatively small, which might be the reason for the poor performance of bradykinesia prediction. In order to prevent the potential overfitting problems resulting from small sample size, we used two sets of VOIs, took prior knowledge into consideration and applied recursive feature elimination to preserve those features with the topmost importance. Clearly, future investigation in a larger cohort is needed. Secondly, The RA-ML analyses of specific motor symptom excluded tremor given that the absence of tremor symptoms in most PD patients (16/33) postoperatively. In order to investigate tremor prediction, more PD patients with suboptimal tremor improvement should be collected in the future. In addition to improving motor symptoms of patients with PD, STN-DBS aims to treat motor complications, such as increasing on-time without dyskinesia and improving motor fluctuation during off periods. Owing to the relatively low presence of levodopa induced dyskinesia (LID) as a result of lower medication dose, the RA-ML analysis of related motor complications was not performed. Future studies with a much larger sample size including more PD patients with motor complications is needed to address this issue. Finally, it would be of interest to perform longer term follow-up analyses, incorporating continued disease progression to explore the fading of DBS efficacy after the surgical honeymoon period (Fasano and Merello, 2020).

CONCLUSION

Our study demonstrated that individual SN susceptibility features derived from radiomics can predict global motor and rigidity outcomes of STN-DBS in PD patients. This RA-ML predictive model may provide a novel and practical approach to counsel suitable candidates for neurosurgeons' decision-making. Our research suggests that individual spatial heterogeneity of nigral iron distribution may be pertinent to patient-specific pathological patterns, which could help to understand the variability of STN-DBS motor benefits.

REFERENCES

- Adeli, E., Shi, F., An, L., Wee, C. Y., Wu, G., Wang, T., et al. (2016). Joint feature-sample selection and robust diagnosis of Parkinson's disease from MRI data. *Neuroimage* 141, 206–219. doi: 10.1016/j.neuroimage.2016.05.054
- Amoroso, N., La Rocca, M., Monaco, A., Bellotti, R., and Tangaro, S. (2018). Complex networks reveal early MRI markers of Parkinson's disease. *Med. Image Anal.* 48, 12–24. doi: 10.1016/j.media.2018.05.004
- Bergsland, N., Zivadinov, R., Schweser, F., Hagemeyer, J., Lichter, D., and Guttuso, T. Jr. (2019). Ventral posterior substantia nigra iron increases over 3 years in Parkinson's disease. *Mov. Disord.* 34, 1006–1013. doi: 10.1002/mds.27730
- Charles, P. D., Van Blercom, N., Krack, P., Lee, S. L., Xie, J., Besson, G., et al. (2002). Predictors of effective bilateral subthalamic nucleus stimulation for PD. *Neurology* 59, 932–934. doi: 10.1212/wnl.59.6.932
- Cheng, Z., Zhang, J., He, N., Li, Y., Wen, Y., Xu, H., et al. (2019). Radiomic features of the nigrosome-1 region of the substantia nigra: using quantitative susceptibility mapping to assist the diagnosis of idiopathic Parkinson's disease. *Front. Aging Neurosci.* 11:167. doi: 10.3389/fnagi.2019.00167
- Damier, P., Hirsch, E. C., Agid, Y., and Graybiel, A. M. (1999). The substantia nigra of the human brain. II. patterns of loss of dopamine-containing neurons in Parkinson's disease. *Brain* 122(Pt. 8), 1437–1448. doi: 10.1093/brain/122.8.1437
- DeLong, E. R., Delong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44, 837–845. doi: 10.2307/2531595
- Dexter, D. T., Wells, F. R., Lees, A. J., Agid, F., Agid, Y., Jenner, P., et al. (1989). Increased nigral iron content and alterations in other metal ions occurring in brain in Parkinson's disease. *J. Neurochem.* 52, 1830–1836. doi: 10.1111/j.1471-4159.1989.tb07264.x

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Ruijin Hospital affiliated to Shanghai Jiao Tong University School of Medicine. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YLiu, BX, and CZ: writing—original draft, writing—review and editing, investigation, data curation, and formal analysis. JL, YLai, FS, DS, LW, BS, YLi, ZJ, HW, EH, and HZ: data curation, formal analysis, and writing—review and editing. QW, DL, and NH: conceptualization, supervision, funding acquisition, resources, and writing—review and editing. FY: supervision, funding acquisition, resources, and writing—review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported, in part, by the National Natural Science Foundation of China (Grant Number: 81801652 for NH; Grant Number: 81971576 for FY, NH, and CZ), a grant from the Science and Technology Commission of Shanghai Municipality (17411952700 for FY, NH, CZ, and YLi), a grant from the Shanghai Sailing Program (18YF1414700 for NH and 20YF1426500 for CZ), the Innovative Research Team of High-level Local Universities in Shanghai, a grant from the Science and technology development plan project of Changshu City (CS202021 for JL), and CZ was supported by the fellowship of Shanghai Research Center for Brain Science and Brain-Inspired Intelligence.

- Espay, A. J., Brundin, P., and Lang, A. E. (2017). Precision medicine for disease modification in Parkinson disease. *Nat. Rev. Neurol.* 13, 119–126. doi: 10.1038/nrneurol.2016.196
- Fasano, A., and Merello, M. (2020). Fading of deep brain stimulation efficacy versus disease progression: untangling a gordian knot. *Mov. Disord. Clin. Pract.* 7, 747–749. doi: 10.1002/mdc3.13041
- Ghassaban, K., Liu, S., Jiang, C., and Haacke, E. M. (2019). Quantifying iron content in magnetic resonance imaging. *Neuroimage* 187, 77–92. doi: 10.1016/j.neuroimage.2018.04.047
- Haacke, E. M., Liu, S., Buch, S., Zheng, W., Wu, D., and Ye, Y. (2015). Quantitative susceptibility mapping: current status and future directions. *Magn. Reson. Imaging* 33, 1–25. doi: 10.1016/j.mri.2014.09.004
- He, N., Ghassaban, K., Huang, P., Jokar, M., Wang, Y., Cheng, Z., et al. (2021). Imaging iron and neuromelanin simultaneously using a single 3D gradient echo magnetization transfer sequence: combining neuromelanin, iron and the nigrosome-1 sign as complementary imaging biomarkers in early stage Parkinson's disease. *Neuroimage* 230:117810. doi: 10.1016/j.neuroimage.2021.117810
- He, N., Langley, J., Huddleston, D. E., Chen, S., Huang, P., Ling, H., et al. (2020). Increased iron-deposition in lateral-ventral substantia nigra pars compacta: a promising neuroimaging marker for Parkinson's disease. *Neuroimage Clin.* 28:102391. doi: 10.1016/j.nicl.2020.102391
- He, N., Ling, H., Ding, B., Huang, J., Zhang, Y., Zhang, Z., et al. (2015). Region-specific disturbed iron distribution in early idiopathic Parkinson's disease measured by quantitative susceptibility mapping. *Hum. Brain Mapp.* 36, 4407–4420. doi: 10.1002/hbm.22928
- Hoops, S., Nazem, S., Siderowf, A. D., Duda, J. E., Xie, S. X., Stern, M. B., et al. (2009). Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. *Neurology* 73, 1738–1745. doi: 10.1212/WNL.0b013e3181c34b47
- Horn, A., Reich, M., Vorwerk, J., Li, N., Wenzel, G., Fang, Q., et al. (2017). Connectivity predicts deep brain stimulation outcome in Parkinson disease. *Ann. Neurol.* 82, 67–78. doi: 10.1002/ana.24974
- Hughes, A. J., Daniel, S. E., Kilford, L., and Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J. Neurol. Neurosurg. Psychiatry* 55, 181–184. doi: 10.1136/jnnp.55.3.181
- Huppertz, H. J., Möller, L., Südmeyer, M., Hilker, R., Hattingen, E., Egger, K., et al. (2016). Differentiation of neurodegenerative parkinsonian syndromes by volumetric magnetic resonance imaging analysis and support vector machine classification. *Mov. Disord.* 31, 1506–1517. doi: 10.1002/mds.26715
- Jellinger, K., Kienzl, E., Rumpelmair, G., Riederer, P., Stachelberger, H., Ben-Shachar, D., et al. (1992). Iron-melanin complex in substantia nigra of parkinsonian brains: an x-ray microanalysis. *J. Neurochem.* 59, 1168–1171. doi: 10.1111/j.1471-4159.1992.tb08362.x
- Kleiner-Fisman, G., Herzog, J., Fisman, D. N., Tamma, F., Lyons, K. E., Pahwa, R., et al. (2006). Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes. *Mov. Disord.* 21(Suppl. 14), S290–S304. doi: 10.1002/mds.20962
- Lang, A. E., and Widner, H. (2002). Deep brain stimulation for Parkinson's disease: patient selection and evaluation. *Mov. Disord.* 17(Suppl. 3), S94–S101. doi: 10.1002/mds.10149
- Langkammer, C., Schweser, F., Krebs, N., Deistung, A., Goessler, W., Scheurer, E., et al. (2012). Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. *Neuroimage* 62, 1593–1599. doi: 10.1016/j.neuroimage.2012.05.049
- Langley, J., Huddleston, D. E., Sedlacik, J., Boelmans, K., and Hu, X. P. (2017). Parkinson's disease-related increase of T2*-weighted hypointensity in substantia nigra pars compacta. *Mov. Disord.* 32, 441–449. doi: 10.1002/mds.26883
- Lehéricy, S., Bardin, E., Poupon, C., Vidailhet, M., and François, C. (2014). 7 Tesla magnetic resonance imaging: a closer look at substantia nigra anatomy in Parkinson's disease. *Mov. Disord.* 29, 1574–1581. doi: 10.1002/mds.26043
- Li, G., Zhai, G., Zhao, X., An, H., Spincemaille, P., Gillen, K. M., et al. (2019). 3D texture analyses within the substantia nigra of Parkinson's disease patients on quantitative susceptibility maps and R2(*) maps. *Neuroimage* 188, 465–472. doi: 10.1016/j.neuroimage.2018.12.041
- Li, W. J., Jiang, H., Song, N., and Xie, J. X. (2010). Dose- and time-dependent alpha-synuclein aggregation induced by ferric iron in SK-N-SH cells. *Neurosci. Bull.* 26, 205–210. doi: 10.1007/s12264-010-1117-7
- Lönnfors-Weitzel, T., Weitzel, T., Slotboom, J., Kiefer, C., Pollo, C., Schübach, M., et al. (2016). T2-relaxometry predicts outcome of DBS in idiopathic Parkinson's disease. *Neuroimage Clin.* 12, 832–837. doi: 10.1016/j.nicl.2016.09.019
- Lotfipour, A. K., Wharton, S., Schwarz, S. T., Gontu, V., Schäfer, A., Peters, A. M., et al. (2012). High resolution magnetic susceptibility mapping of the substantia nigra in Parkinson's disease. *J. Magn. Reson. Imaging* 35, 48–55. doi: 10.1002/jmri.22752
- Malaga, K. A., Costello, J. T., Chou, K. L., and Patil, P. G. (2021). Atlas-independent, N-of-1 tissue activation modeling to map optimal regions of subthalamic deep brain stimulation for Parkinson disease. *Neuroimage Clin.* 29:102518. doi: 10.1016/j.nicl.2020.102518
- Martin, W. R., Wieler, M., and Gee, M. (2008). Midbrain iron content in early Parkinson disease: a potential biomarker of disease status. *Neurology* 70, 1411–1417. doi: 10.1212/01.wnl.0000286384.31050.b5
- Mueller, K., Urgošik, D., Ballarini, T., Holiga, Š, Möller, H. E., Růžicka, F., et al. (2020). Differential effects of deep brain stimulation and levodopa on brain activity in Parkinson's disease. *Brain Commun.* 2:fcaa005. doi: 10.1093/braincomms/fcaa005
- Mutch, W. J., Strudwick, A., Roy, S. K., and Downie, A. W. (1986). Parkinson's disease: disability, review, and management. *Br. Med. J.* 293, 675–677. doi: 10.1136/bmj.293.6548.675
- Oakley, A. E., Collingwood, J. F., Dobson, J., Love, G., Perrott, H. R., Edwardson, J. A., et al. (2007). Individual dopaminergic neurons show raised iron levels in Parkinson disease. *Neurology* 68, 1820–1825. doi: 10.1212/01.wnl.0000262033.01945.9a
- Odekerken, V. J., Van Laar, T., Staal, M. J., Mosch, A., Hoffmann, C. F., Nijssen, P. C., et al. (2013). Subthalamic nucleus versus globus pallidus bilateral deep brain stimulation for advanced Parkinson's disease (NSTAPS study): a randomised controlled trial. *Lancet Neurol.* 12, 37–44. doi: 10.1016/s1474-4422(12)70264-8
- Okun, M. S., Tagliati, M., Pourfar, M., Fernandez, H. H., Rodriguez, R. L., Alterman, R. L., et al. (2005). Management of referred deep brain stimulation failures: a retrospective analysis from 2 movement disorders centers. *Arch. Neurol.* 62, 1250–1255. doi: 10.1001/archneur.62.8.noc40425
- Péran, P., Cherubini, A., Assogna, F., Piras, F., Quattrocchi, C., Peppe, A., et al. (2010). Magnetic resonance imaging markers of Parkinson's disease nigrostriatal signature. *Brain* 133, 3423–3433. doi: 10.1093/brain/awq212
- Piboolnurak, P., Lang, A. E., Lozano, A. M., Miyasaki, J. M., Saint-Cyr, J. A., Poon, Y. Y., et al. (2007). Levodopa response in long-term bilateral subthalamic stimulation for Parkinson's disease. *Mov. Disord.* 22, 990–997. doi: 10.1002/mds.21482
- Pollak, P. (2013). Deep brain stimulation for Parkinson's disease: patient selection. *Handb. Clin. Neurol.* 116, 97–105. doi: 10.1016/b978-0-444-53497-2.00009-7
- Price, C. C., Favilla, C., Tanner, J. J., Towler, S., Jacobson, C. E., Hass, C. J., et al. (2011). Lateral ventricle volume is poor predictor of post unilateral DBS motor change for Parkinson's disease. *Parkinsonism Relat. Disord.* 17, 343–347. doi: 10.1016/j.parkrel.2011.01.018
- Rodriguez, R. L., Fernandez, H. H., Haq, I., and Okun, M. S. (2007). Pearls in patient selection for deep brain stimulation. *Neurologist* 13, 253–260. doi: 10.1097/NRL.0b013e318095a4d5
- Salvatore, C., Cerasa, A., Castiglioni, I., Gallivanone, F., Augimeri, A., Lopez, M., et al. (2014). Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and progressive supranuclear palsy. *J. Neurosci. Methods* 222, 230–237. doi: 10.1016/j.jneumeth.2013.11.016
- Schweser, F., Deistung, A., Lehr, B. W., and Reichenbach, J. R. (2011). Quantitative imaging of intrinsic magnetic tissue properties using MRI signal phase: an approach to in vivo brain iron metabolism? *Neuroimage* 54, 2789–2807. doi: 10.1016/j.neuroimage.2010.10.070
- Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* 86, 420–428. doi: 10.1037/0033-2909.86.2.420
- Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155. doi: 10.1002/hbm.10062
- Temperli, P., Ghika, J., Villemure, J. G., Burkhard, P. R., Bogousslavsky, J., and Vingerhoets, F. J. (2003). How do parkinsonian signs return after

- discontinuation of subthalamic DBS? *Neurology* 60, 78–81. doi: 10.1212/wnl.60.1.78
- Vallat, R. (2018). Pingouin: statistics in Python. *J. Open Source Soft.* 3:1026. doi: 10.21105/joss.01026
- van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., et al. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 77, e104–e107. doi: 10.1158/0008-5472.Can-17-0339
- Vingerhoets, F. J., Villemure, J. G., Temperli, P., Pollo, C., Pralong, E., and Ghika, J. (2002). Subthalamic DBS replaces levodopa in Parkinson's disease: two-year follow-up. *Neurology* 58, 396–401. doi: 10.1212/wnl.58.3.396
- Wei, H., Dibb, R., Zhou, Y., Sun, Y., Xu, J., Wang, N., et al. (2015). Streaking artifact reduction for quantitative susceptibility mapping of sources with large dynamic range. *NMR Biomed.* 28, 1294–1303. doi: 10.1002/nbm.3383
- Welter, M. L., Houeto, J. L., Tezenas Du Montcel, S., Mesnage, V., Bonnet, A. M., Pillon, B., et al. (2002). Clinical predictive factors of subthalamic stimulation in Parkinson's disease. *Brain* 125, 575–583. doi: 10.1093/brain/awf050
- Wotschel, V., Chard, D. T., Enzinger, C., Filippi, M., Frederiksen, J. L., Gasperini, C., et al. (2019). SVM recursive feature elimination analyses of structural brain MRI predicts near-term relapses in patients with clinically isolated syndromes suggestive of multiple sclerosis. *Neuroimage Clin.* 24:102011. doi: 10.1016/j.nicl.2019.102011
- Xiao, B., He, N., Wang, Q., Cheng, Z., Jiao, Y., Haacke, E. M., et al. (2019). Quantitative susceptibility mapping based hybrid feature extraction for diagnosis of Parkinson's disease. *Neuroimage Clin.* 24:102070. doi: 10.1016/j.nicl.2019.102070
- Yan, F., He, N., Lin, H., and Li, R. (2018). Iron deposition quantification: applications in the brain and liver. *J. Magn. Reson. Imaging* 48, 301–317. doi: 10.1002/jmri.26161
- Younce, J. R., Campbell, M. C., Hershey, T., Tanenbaum, A. B., Milchenko, M., Ushe, M., et al. (2020). Resting-state functional connectivity predicts STN DBS clinical response. *Mov. Disord.* 36, 662–671. doi: 10.1002/mds.28376
- Younce, J. R., Campbell, M. C., Perlmutter, J. S., and Norris, S. A. (2019). Thalamic and ventricular volumes predict motor response to deep brain stimulation for Parkinson's disease. *Parkinsonism Relat. Disord.* 61, 64–69. doi: 10.1016/j.parkreidis.2018.11.026
- Zaidel, A., Bergman, H., Ritov, Y., and Israel, Z. (2010). Levodopa and subthalamic deep brain stimulation responses are not congruent. *Mov. Disord.* 25, 2379–2386. doi: 10.1002/mds.23294
- Zecca, L., Shima, T., Stroppolo, A., Goj, C., Battiston, G. A., Gerbasi, R., et al. (1996). Interaction of neuromelanin and iron in substantia nigra and other areas of human brain. *Neuroscience* 73, 407–415. doi: 10.1016/0306-4522(96)00047-4
- Zhang, C., Wang, L., Hu, W., Wang, T., Zhao, Y., Pan, Y., et al. (2020). Combined unilateral subthalamic nucleus and contralateral globus pallidus interna deep brain stimulation for treatment of parkinson disease: a pilot study of symptom-tailored stimulation. *Neurosurgery* 87, 1139–1147. doi: 10.1093/neuros/nyaa201
- Zhang, J., Yu, C., Jiang, G., Liu, W., and Tong, L. (2012). 3D texture analysis on MRI images of Alzheimer's disease. *Brain Imaging Behav.* 6, 61–69. doi: 10.1007/s11682-011-9142-3
- Zhang, Y., Yang, M., Wang, F., Chen, Y., Liu, R., Zhang, Z., et al. (2020). Histogram analysis of quantitative susceptibility mapping for the diagnosis of Parkinson's Disease. *Acad. Radiol.* doi: 10.1016/j.acra.2020.10.027 [Epub ahead of print].
- Zucca, F. A., Segura-Aguilar, J., Ferrari, E., Muñoz, P., Paris, I., Sulzer, D., et al. (2017). Interactions of iron, dopamine and neuromelanin pathways in brain aging and Parkinson's disease. *Prog. Neurobiol.* 155, 96–119. doi: 10.1016/j.pneurobio.2015.09.012

Conflict of Interest: FS and DS were employed by the Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China. The company has no role in designing, performing the surveillances, analyzing, and interpreting the data.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Liu, Xiao, Zhang, Li, Lai, Shi, Shen, Wang, Sun, Li, Jin, Wei, Haacke, Zhou, Wang, Li, He and Yan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.