HHS Public Access

Author manuscript

Neurosurgery. Author manuscript; available in PMC 2026 March 10.

Published in final edited form as:

Neurosurgery. 2026 May 01; 98(5): 1139–1148. doi:10.1227/neu.0000000000003721.

Technical Feasibility of Quantitative Susceptibility Mapping Radiomics for Predicting Deep Brain Stimulation Outcomes in Parkinson's Disease

Alexandra G. Roberts, MS[1,2], **Jinwei Zhang, PhD**[3], **Ceren Tozlu, PhD**[2], **Dominick Romano, MS**[2], **Sema Akkus, MD**[4], **Heejong Kim, PhD**[2], **Mert R. Sabuncu, PhD**[1,2,5], **Pascal Spincemaille, PhD**[2], **Jianqi Li, PhD**[6], **Yi Wang, PhD**[1,2], **Xi Wu, MD**[7], **Brian H. Kopell, MD**[4]

[1]Electrical and Computer Engineering, Cornell University, Ithaca, NY, United States

[2]Department of Radiology, Weill Cornell Medicine, New York, NY, United States

[3]Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, United States

[4]Department of Neurosurgery, Mount Sinai Hospital, New York, NY, United States

[5]Electrical and Computer Engineering, Cornell Tech, New York, NY, United States

[6]School of Physics and Electronic Science, East China Normal University, Shanghai, China

[7]Department of Neurosurgery, Changhai Hospital, Shanghai, China

Abstract

Background and Objectives: Parkinson's disease (PD) patients with motor complications are often considered for deep brain stimulation (DBS) surgery. Predicting symptom improvement to separate DBS responders and non-responders remains an unmet need. Currently, DBS candidacy is evaluated using the levodopa challenge test (LCT) to confirm dopamine responsiveness and diagnosis. However, prediction of DBS success by measuring presurgical symptom improvement associated with levodopa dosage changes is highly problematic. Quantitative susceptibility mapping (QSM) is a recently developed MRI method that depicts brain iron distribution. As the substantia nigra and subthalamic nuclei are well-visualized, QSM has been used in presurgical planning of DBS. Spatial features resulting from iron distribution in these nuclei have been previously linked with disease progression and motor symptom severity. Given its clear target depiction and prior findings regarding susceptibility and PD, this study demonstrates the technical feasibility of predicting DBS outcomes from presurgical QSM.

Methods: A novel presurgical QSM radiomics approach using a regression model is presented to predict DBS outcome according to spatial features in QSM deep gray nuclei. To overcome limited and noisy training data, data augmentation using label noise injection or "compensation" was used to improve outcome prediction of the regression model. The QSM radiomics model was evaluated on 67 PD patients who underwent DBS at two medical centers.

Corresponding authors: Yi Wang, PhD, Department of Radiology, Weill Cornell Medicine, 525 East 68th Street New York, NY 10065, yiwang@med.cornell.edu, Fax: +1 646 962 0122, Brian H. Kopell, MD, Department of Neurosurgery, Mount Sinai Health System, 1000 10th Avenue, 10G-46 New York, New York 10019, brian.kopell@mountsinai.org, Fax: +1 212 523 8342.

Results: The QSM radiomics model predicted DBS improvement in the unified Parkinson's disease rating scale (UPDRS-III) at Center 1 and Center 2 with Pearson's correlation $r = .75$, ($p = 1.1 \times 10^{-7}$) and $r = .70$, ($p = 1.8 \times 10^{-5}$), respectively. LCT failed to predict DBS improvement at Center 1 and Center 2 with Pearson's correlation $r = .12$, ($p = .48$) and $r = -.14$, ($p = .45$), respectively.

Conclusion: QSM radiomics has potential to accurately predict DBS outcome in treating PD patients, offering a valuable alternative to the time-consuming and low-accuracy LCT.

Keywords

Deep brain stimulation; machine learning; Parkinson's disease; quantitative susceptibility mapping; radiomics; regression

INTRODUCTION

Deep brain stimulation (DBS) is a surgery to recalibrate neural circuitry in movement and psychiatric disorders including essential tremor, dystonia, and obsessive-compulsive disorder[1]. DBS in the subthalamic nucleus (STN) is a widespread treatment for advanced Parkinson's disease (PD) motor symptoms[2]. Selection of candidates for DBS involves assessing responsiveness to levodopa through the levodopa challenge test (LCT). The LCT has shown poor prediction of DBS outcomes[3,4] and is burdensome, requiring time-consuming assessments and medication withdrawal. A consistent and convenient method for predicting DBS outcomes in PD is an important unmet clinical need.

Quantitative susceptibility mapping (QSM), reconstructed from common multi-echo gradient echo (mGRE) MRI data[5,6], provides superior contrast in deep gray nuclei during presurgical planning[7–9], and aids in monitoring disease progression[10]. Radiomic features characterizing spatial distribution include first-order histogram-based features describing voxel values, second-order features representing joint probability distributions of voxel values, and region of interest (ROI) shape[11]. Deep gray nuclei QSM radiomic features potentially improve PD diagnosis [12,13], and correlate with both PD motor symptoms[14,15,16,17] and DBS binary outcomes[18].

The purpose of this work is to demonstrate the potential of presurgical QSM radiomic features for predicting numerical outcomes of DBS needed in patient selection.

METHODS

Overview

For the radiomics prediction model based on presurgical QSM, postsurgical Parkinson's disease rating scale III (UPDRS-III) improvement is used as "ground truth" or "label" data. UPDRS-III metrics $u_{presurgical, on\text{-}medication}$ and $u_{postsurgical,\ on\text{-}stimulation,\ off\text{-}medication}$ were used to compute label data in Equation 1. As UPDRS-III is measured on an interval scale, a linear regression with least absolute shrinkage operator (Lasso) penalization is proposed using QSM radiomic features to predict DBS outcomes. DBS outcome is defined as the postoperative to preoperative UPDRS-III ratio, or UPDRS-III improvement. UPDRS-III

ratings contain variability despite good agreement, described in supervised learning as "noise" ubiquitous in human measurements[19]. "Label noise" degrades regression models and supervised learning in general; compensating for this via label noise injection has been shown to improve model performance[20]. This study proposes a QSM radiomics Lasso model with label noise injection for DBS outcome prediction.

Patient Cohorts

This retrospective study was approved by institutional review boards and ethics committees of relevant institutions. Informed consent was obtained for 40 anonymized patients undergoing bilateral STN-DBS at Center 1, and 31 anonymized bilateral STN-DBS patients at Center 2. Participants met the Movement Disorder Society (MDS) diagnostic criteria for idiopathic PD[21]. Patients were included based on the following criteria: 1) diagnosis of idiopathic PD, 2) demonstrated good response to levodopa, 3) decreased medication efficacy or serious motor complications impacting quality-of-life despite optimal medical management, 4) absence of significant cognitive impairment as determined by neuropsychological assessment (patients with mild cognitive impairment were evaluated individually considering their ability to participate in postoperative programming and management), and 5) presence of motor fluctuations or medication-refractory tremors. Exclusion criteria were: 1) atypical parkinsonian disorders, 2) active severe psychiatric illness, 3) dementia, 4) severe diffuse cerebral ischemic changes on brain MRI, and 5) systemic comorbidities precluding surgical candidacy.

Data Acquisition

QSMs from mGRE data were reconstructed with MEDI-L_1[22], with an overview of QSM implementation[23] available in Supplemental Digital Content (SDC) 1. Presurgical UPDRS-III scores u on and off-medication from LCT and postsurgical UPDRS-III scores on-stimulation and off-medication were collected[24] and the UPDRS-III improvement was computed as:

$$y = \frac{u_{presurgical,\,on\text{-}medication} - u_{postsurgical,\,on\text{-}stimulation,\,off\text{-}medication}}{u_{presurgical,\,on\text{-}medication}} \quad (1)$$

ROIs were acquired by manual or automatic atlas segmentation. The substantia nigra (SN) and STN were segmented with radiologist supervision from Center 1. Collected sample size was 40, with 3 cases excluded because of motion artifacts. From Center 2, a patient-specific atlas[25] was created using Advanced Normalization Tools (ANTs)[26–28] to acquire SN and STN masks. Atlas construction is detailed in SDC 2. Collected sample size was 31, with 1 case excluded via registration failure. The pipeline (Figure 1) is discussed in Data Preprocessing.

Data Preprocessing

Figure 1 depicts the proposed pipeline with QSM and STN/SN ROI inputs. Features were extracted from ROIs (joint entropy feature map shown) and passed to a Lasso model after

normalization, label augmentation, and recursive feature selection. A total of 93 features were extracted over 17 "image types" (various filtering operations on the original image) in PyRadiomics[29]. Additionally, 14 shape features were extracted on the original image, totaling 1595 features per ROI. Shape features describe ROI characteristics rather than underlying susceptibility distributions. The marching cubes algorithm approximates a mask mesh to compute volume, surface area, sphericity, etc. First-order radiomic features refer to voxel intensity (here, susceptibility) statistics throughout STN and SN masks. These features (percentiles, range, mean, etc.) are functions of the ROI susceptibility histogram. Second-order features are derived from gray level co-occurrence, size zone, run length, and dependence matrices (GLCM, GLSZM, GLRLM, and GLDM). These represent joint occurrences of susceptibility by voxel, zone, run, and dependency descriptors. The GLCM computes the joint average, difference entropy, and contrast between pairs of susceptibility values. Similarly, susceptibilities can be described via zone scale (GLSZM), run scale (GLRLM) and include a range of values "dependent" on a central susceptibility value (GLDM). Such matrices generate texture features. In addition to the original susceptibility, gradient, square, and wavelet images enable feature extraction. A feature list is linked in SDC 3.

Regression Model

Linear regression with Lasso penalization implemented with SciKit-Learn was performed to predict UPDRS-III improvement using 1595 features per ROI[30].

$$w^* = \operatorname{argmin} \frac{1}{2N} \| y - X_\Phi w \|_2^2 + \lambda \| w \|_1 \tag{2}$$

UPDRS-III improvements y are predicted using optimal weights w^* and test features $\widehat{X}_\Phi$, regularization parameter λ, and $N = \{37,30\}$ denotes center sample size. Given the dataset size, a leave-one-out approach for subject-specific models[31] was used with each patient withheld and the remaining $N - 1$ patients used for training.

Label Noise Injection Data Augmentation

The Lasso training dataset can be augmented using computed label noise statistics. A Gaussian distribution was fit to $N - 1$ original UPDRS-III improvements as $y_t^0 \sim \mathcal{N}(\mu_{N\text{-}1}, \sigma_{N\text{-}1})$, with mean $\mu_{N\text{-}1}$ and standard deviation $\sigma_{N\text{-}1}$. For each model with original features X_t^0, noise instances ϵ_q were drawn from the distribution $\mathcal{N}(0, z\sigma_{N\text{-}1})$, with $z = 1.96$, and added to labels in the dataset y_t^0:

$$\left\{X_t^Q, y_t^Q\right\} = \left\{X_t^0, \ldots, X_t^0,\ y_t^0, y_t^0 + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_1 \end{bmatrix}, y_t^0 + \begin{bmatrix} \epsilon_2 \\ \vdots \\ \epsilon_2 \end{bmatrix}, \ldots, y_t^0 + \begin{bmatrix} \epsilon_Q \\ \vdots \\ \epsilon_Q \end{bmatrix}\right\} \tag{3}$$

This newly augmented dataset of size $(Q + 1)(N - 1)$ has labels $y_r^Q = \boldsymbol{Y} + \boldsymbol{\epsilon} = \begin{bmatrix} y_0 & y_0 & \cdots & y_0 \\ \vdots & \vdots & \ddots & \vdots \\ y_{N-1} & y_{N-1} & \cdots & y_{N-1} \end{bmatrix} + \begin{bmatrix} 0 & \epsilon_1 & \cdots & \epsilon_Q \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \epsilon_1 & \cdots & \epsilon_Q \end{bmatrix}$. Original labels are preserved in the first column. This use of noise injection augmentation is referred to as “label noise compensation.”

For comparison, Lasso was trained with augmentation strategies including bootstrap, wild bootstrap[32], and synthetic minority oversampling with Gaussian noise (SMOGN)[33].

Training

Nested cross-validation was used to train and test models for each subject. In the outer loop, the model was trained with $N - 1$ subjects and tested with one subject. For the inner loop, k-fold cross-validation was performed.

Recursive feature selection determined the optimal number of normalized features m^* from total features $\not{p} = 6380$, (step size $\Delta m = 1000$ features) before Lasso input.

Splits $k = \{2, 3, 4, 5\}$, selected features m, and regularization λ were chosen by minimizing training error. The minority class $(y_{train} < .3)$ was resampled for stratified cross-validation.

The dataset was augmented $(Q = 10)$ with label noise. Least angle regression solver weights (machine epsilon regularization $\epsilon_c = .1$) were initialized at 0. The model was retrained over the entire $(N - 1)$ dataset using m^* features, split k^*, and regularization λ^* from the inner loop.

Harmonization commonly corrects bias and variation in multi-center datasets[34]. Training was applied to Center 1 and harmonized[35] Center 2 data. The Center 2 noise compensation factor was increased to $2z$ to accommodate its smaller sample size[36]. Each patient prediction $\hat{y}$ was a linear combination of features and weights, $\hat{y} = \widehat{X}_{\Phi} w^*$. Training was repeated with standard and noise compensated classifiers.

Statistical Analysis

Performance of LCT and Lasso regressors was evaluated using linear regression (Pearson correlation r, slope $\mathscr{M}$, intercept b, Wald test significance p) between observed UPDRS-III improvement and predicted UPDRS-III improvement. Model accuracy was measured by correlation between observed and predicted outcome. Significance of correlation differences between ground truth and estimators were assessed using Steiger’s test[37]. To evaluate augmentations over different initial noise states, models were evaluated under 10 additional seeds.

The frequency of features with nonzero weights was compared for models, with predictive power measured by the number of nonzero weights w^* across N patient-specific models. Feature maps were compared between non-responders $(y < .3)$ and strong responders $(y > .85)$, a threshold selected to balance with non-responders. Feature values between

groups were evaluated using the Wilcoxon signed-rank test; the minimum number of predictive features to significantly separate the responders from non-responders was evaluated with a Bonferroni correction.

Subtype Analysis

Model and LCT accuracy was further evaluated in a variety of population subtypes. These include degree of levodopa responsiveness according to improvement threshold (. 33), symptom indication (motor fluctuations versus medication-refractory tremor), and improvement considering non-tremor cardinal manifestations.

Data Availability Statement

The minimal information clinical artificial intelligence modeling (MI-CLAIM) checklist was implemented[38] with code and radiomic data available upon publication (https://github.com/agr78/RadDBS-QSM). Deidentified data is available upon request, pending specific data-sharing agreements.

RESULTS

Demographic Information

Patient demographics at Center 1 were: age 63.3 ± 7.45 years, sex ratio of 22 males to 18 females and disease duration 8.45 ± 2.87 years. Center 1 LCT procedure entailed 12 hours of abstaining from medication for the “off-medication” condition and 90 minutes following 1.5 LEDD for the “on-medication” condition. The off-medication preoperative UPDRS-III was 59.63 ± 20.01. With LEDD 728.27 ± 296.01 mg/day, the on-medication preoperative UPDRS-III was 23.65 ± 17.54 . Postoperative UPDRS-III (on-stimulation, off-medication) was 25.00 ± 13.30. Absolute off-medication UPDRS-III improvement was 34.63 ± 17.11 and the postoperative LEDD was 323.78 ± 253.98 mg/day.

Patient demographics at Center 2 were: age 63.1 ± 7.84 years, sex ratio of 23 males to 7 females and disease duration 8.6 ± 3.88 years. The “on-medication” condition at Center 2 occurs 30 minutes following the standard dose of LEDD. Off-medication preoperative UPDRS-III was 49.80 ± 12.17. Preoperative UPDRS-III and LEDD was 20.97 ± 9.89 at 857.40 ± 486.92 mg/day at and postoperative UPDRS-III (on-stimulation, off-medication) was 17.03 ± 9.23. Absolute off-medication UPDRS-III improvement 32.77 ± 14.85 and postoperative LEDD was 172.37 ± 303.04 mg/day.

At both centers, bilateral STN-DBS was performed using frame-based stereotactic approaches with microelectrode recording for target refinement. Initial programming occurred 4 weeks after lead insertion (2 weeks post-pulse generator implantation), with 4–6 programming sessions over 6 months and parameters optimized based on clinical response. Postoperative outcomes were assessed 6 months post-surgery. Programming parameters and profiles of responders and non-responders and their influence on model predictions at both centers are in SDC 4.

Statistical Assessment of Model Performance

Figure 2 illustrates UPDRS-III improvement predictions. LCT failed to correlate ($r = .12,\ p = .48,$ Figure 2a) with true outcome (Equation 1). Radiomic predictions correlated with improvements at Center 1, with $r = .75,\ p = 1.1 \times 10^{-7}$ for noise compensated Lasso (Figure 2b). Compared augmentations with radiomic models also showed significant but smaller correlations: $r = .65,\ p = 1.4 \times 10^{-5}$ for Lasso (Figure 2c), $r = .68,\ p = 4.5 \times 10^{-6}$ for bootstrap Lasso (Figure 2d), $r = .65,\ p = 1.5 \times 10^{-5}$ for wild bootstrap Lasso (Figure 2e), and $r = .67,\ p = 5.2 \times 10^{-6}$ for SMOGN Lasso (Figure 2f). The noise compensated Lasso model correctly classified all outcomes: "response" with improvement $y \geq .3$ (first quadrant) or "non-response" (third quadrant) in Figure 2b. Improved correlations from noise compensation persist over varying random seeds (SDC 5).

Similarly, Center 2 LCT was inaccurate (Figure 3a) and noise compensation separated responders from non-responders with $r = .70,\ p = 1.8 \times 10^{-5}$ (Figure 3b).

Improved correlations observed between radiomic models as compared to LCT were highly significant: at Center 1, $p = .00040$ for noise compensated Lasso, $p = .0012$ for standard Lasso, $p = .0017$ for bootstrap Lasso, $p = .0022$ for wild bootstrap Lasso, and $p = .0030$ for SMOGN Lasso and $p = .00013$ for noise compensated Lasso at Center 2.

Noise compensation also improved logistic regression classifier performance (Supplemental Figure (SF) 1a) and regression error characteristic (SF 1b) at Center 1 as well as Center 2 (SF 1c, 1d), in SDC 6.

Statistical Assessment of Predictive Features

The most predictive features for the considered ROIs were on the STN QSM Coiflet wavelet decomposition and the SN gray level dependence matrix (GLDM) at Center 1 (Figure 4). Specific features were SN large dependence high gray level emphasis (measuring homogeneity) and STN gray level size zone matrix (GLSZM) entropy (measuring heterogeneity) shown in Figure 5 for strong responders ($y > .85$) and non-responders ($y < .3$), where the motor response is defined in Equation 1. Center 2 models used SN wavelet decomposition features and GLZSM entropy on the local binary pattern STN image. The ability of individual features to separate responders from non-responders and the minimum number of features required for significant separation is discussed in SDC 7.

Subtype Analysis

At Center 1, the model mean-squared error (MSE) for the 5 levodopa non-responders was .035 and .024 for the remaining 32 levodopa responders. At Center 2, the MSE for the 2 levodopa non-responders was .012 and .048 for the remaining 28 levodopa responders. At Center 1, one patient was indicated solely by medication-refractory tremor with MSE .016 and .027 for remaining motor fluctuations-indicated patients. At Center 2, model prediction for patients indicated by motor fluctuations was $r = .74,\ p = .0006,\ n = 17$; compared to medication-refractory tremor indications $r = .63,\ p = .021,\ n = 13$. At Center

1, model prediction excluding tremor response was $r = .53$, $p = .00067$ and LCT was $r = .07$, $p = .67$. At Center 2, model prediction excluding tremor response was ($r = .54$, $p = .0002$, $n = 17$) and LCT was $r = -.02$, $p = .91$.

DISCUSSION

Key Results

Preliminary data demonstrates that a QSM radiomics model can predict numerical DBS UPDRS-III improvements with significantly more accuracy than LCT at both centers. This QSM radiomic prediction is consistent with previous findings of ROI analyses including texture features on PD symptom burden[12,16,39]. This study is significant in that a numerical prediction of UPDRS-III improvement can be obtained from QSM with high contrast-to-noise target visualization[8,9,17,40], currently acquired for presurgical MRI.

Prior Works

There are large uncertainties in LCT predictions[24,41]. LCT fails to predict levodopa-resistant symptoms such as tremor improvement[42] and certain dyskinesias[43]. LCT induces burdensome withdrawal symptoms[42], while suffering inaccuracies dependent on dosage[44] and disease duration[3], though it is useful to evaluate levodopa responsiveness and confirm diagnoses.

Magnetoencephalography[45] and transcranial direct current stimulation (tDCS)[46] have been investigated for predicting DBS improvement, yet neither is routinely acquired in presurgical planning. MRI-based brain morphology and functional connectomes have been used to predict DBS motor outcome, providing limited accuracy with correlations ranging from .37 to .65[47–51], and the inclusion of more clinical data improves the prediction[47,49].

Of the compared data augmentations, wild bootstrap is the most similar method to noise injection compensation. Though it improves non-responder estimates, modeling label variability in noise injection compensation is more effective. The Lasso model provides a significant and accurate prediction of DBS improvements using routinely acquired QSM under typical levodopa regimen.

Model Interpretation

QSM measures spatial iron distributions in deep gray nuclei, linked to diffusion-tensor-imaging-defined structural connectivity[52]. This work demonstrates that QSM deep gray nuclei iron spatial distribution features potentially predict DBS outcome. The prediction is a weighted sum of relevant radiomic features in UPDRS-III improvements. The frequency of nonzero weights across patient-specific models indicate features important in outcome prediction. Comparison of responder and non-responder feature maps demonstrate distinction in both the STN and SN. Dysregulation of SN and STN iron disrupts locomotor activation and inhibition[53], and subsequent motor symptoms measured by UPDRS-III. Specifically, the effect of iron on dopamine, serotonin, gamma-aminobutyric acid (GABA), and glutamate mechanisms is discussed below.

Clinical Feature Interpretation

QSM SN texture features may reflect degeneration of nigral neurons, reducing dopamine receptor stimulation where iron is a cofactor in dopamine synthesis[54]. QSM SN homogeneity and prevalence of high susceptibility regions may suggest advanced tissue damage or less viable circuitry. There is substantial association between SN iron accumulation and disease progression in PD[55]. This interpretation is consistent with DBS patient selection guidance that considers “very late stages” of PD to be an impediment to symptom improvement[56]. In addition to its role as a cofactor in tyrosine hydroxylase enzyme preceding dopamine, iron is a cofactor for tryptophan hydroxylase, an enzyme responsible for serotonin synthesis, tied to motor fluctuations and dyskinesia symptoms[57].

Loss of GABAergic inhibition, (also associated with motor fluctuations and dyskinesia), in the STN may perturb iron deposition texture features[58]. Loss of QSM STN gradient, captured in predictive wavelet entropy features[29], correlates with loss of motor function[39], which is consistent with the wavelet decomposition features identified in previous work[18].

In both deep gray nuclei, levodopa action is restricted to dopaminergic pathways[59], while iron is involved in dopaminergic, serotonergic, GABAergic, and glutamatergic pathways[60]. As such, the predictive power of LCT is limited. Conversely, the cofactor role of iron in precursors of serotonin and dopamine[61] in the SN and metabolism of GABA[62] and glutamate in the STN[53] encode features predictive of symptom improvement in the iron distribution depicted on QSM.

These findings affirm both the observation that heterogeneity of the STN iron distribution and elevated susceptibility in the SN iron distribution may be predictive of motor symptom severity and patient responsiveness to DBS. Yet, these features alone are insignificant between responders and non-responders, underscoring the need for multivariable models such as Lasso for patient selection. QSM radiomics provides accurate predictions in this preliminary study $(r = .70, .75)$, better than MRI-based functional and structural connectivity features. These measures along with clinical data may further improve the prediction power[47–51].

Limitations

This work has several limitations. As the prediction is preoperative, lead placement and programming factors are not considered in the model. Also, as a feasibility study, further validation is required. Eligibility requires LCT improvement, biasing against DBS patients without LCT improvement and precluding healthy subjects. Additionally, sample sizes are small compared to the number of features, adversely affecting performance at Center 2 versus Center 1. The Gaussian distribution is chosen for label noise modeling but may oversimplify UPDRS-III interval scale. Finally, QSM quality in deep gray nuclei may be improved by including the voxel sensitivity function[63,64]. Iron distribution depiction may be improved with paramagnetic and diamagnetic susceptibility source separation[65–69].

Generalizability

The feasibility of predicting DBS outcome response from presurgical QSM was demonstrated at two sites with differing programming parameters, populations and segmentation methods. The predictive potential of QSM was shown across regression and classification models.

Future Work

Alongside motor symptoms, non-motor predictions may be further explored (apathy - linked to microstructure[70] - and depression[71] inventories). Additionally, postoperative factors[72], symptom indication and degree of DBS response effects on model accuracy may be evaluated. The patient population featured STN-DBS, but the ventral intermediate nucleus and globus pallidus pars interna are also well-visualized on QSM for presurgical targeting and should be considered in future studies[73]. Beyond deep gray nuclei, whole brain ROIs may be included[74]. Label noise distributions should be investigated to reflect multiple raters present in clinical settings.

CONCLUSION

QSM radiomics predicted DBS outcome better than LCT in two centers. This reveals an additional role for QSM, currently used for presurgical DBS planning.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

1. Lozano AM, Lipsman N. Probing and regulating dysfunctional circuits using deep brain stimulation. Neuron. Feb 6 2013;77(3):406–24. doi:10.1016/j.neuron.2013.01.020 [PubMed: 23395370]
2. Bronstein JM, Tagliati M, Alterman RL, et al. Deep Brain Stimulation for Parkinson Disease: An Expert Consensus and Review of Key Issues. Archives of Neurology. 2011;68(2):165–165. doi:10.1001/archneurol.2010.260 [PubMed: 20937936]
3. Lachenmayer L, Mürset M, Antih N, et al. Subthalamic and pallidal deep brain stimulation for Parkinson's disease-meta-analysis of outcomes. npj Parkinson s Disease. 09/06 2021;7doi:10.1038/s41531-021-00223-5
4. Wolke R, Becktepe JS, Paschen S, et al. The Role of Levodopa Challenge in Predicting the Outcome of Subthalamic Deep Brain Stimulation. Movement Disorders Clinical Practice. 2023;10(8):1181–1191. 10.1002/mdc3.13825 [PubMed: 37635781]
5. Wang Y, Liu T. Quantitative susceptibility mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. Magnetic Resonance in Medicine. 2015;73(1):82–101. doi:10.1002/mrm.25358 [PubMed: 25044035]
6. Bilgic B, Costagli M, Chan KS, et al. Recommended implementation of quantitative susceptibility mapping for clinical research in the brain: A consensus of the ISMRM electro-magnetic tissue properties study group. Magn Reson Med. May 2024;91(5):1834–1862. doi:10.1002/mrm.30006 [PubMed: 38247051]
7. Dimov AV, Gupta A, Kopell BH, Wang Y. High-resolution QSM for functional and structural depiction of subthalamic nuclei in DBS presurgical mapping. Journal of Neurosurgery. 2019;131(2):360–367. doi:10.3171/2018.3.jns172145 [PubMed: 30095333]

8. Liu T, Eskreis-Winkler S, Schweitzer AD, et al. Improved subthalamic nucleus depiction with quantitative susceptibility mapping. Radiology. Oct 2013;269(1):216–23. doi:10.1148/radiol.13121991 [PubMed: 23674786]
9. Rasouli J, Ramdhani R, Panov FE, et al. Utilization of Quantitative Susceptibility Mapping for Direct Targeting of the Subthalamic Nucleus During Deep Brain Stimulation Surgery. Operative neurosurgery (Hagerstown, Md). Apr 1 2018;14(4):412–419. doi:10.1093/ons/opx131 [PubMed: 28531270]
10. Guan X, Lancione M, Ayton S, Dusek P, Langkammer C, Zhang M. Neuroimaging of Parkinson's disease by quantitative susceptibility mapping. NeuroImage. 2024/04/01/ 2024;289:120547. 10.1016/j.neuroimage.2024.120547 [PubMed: 38373677]
11. Materka A Texture analysis methodologies for magnetic resonance imaging. Dialogues in Clinical Neuroscience. 2004;6(2):243–250. doi:10.31887/dcns.2004.6.2/amaterka [PubMed: 22033841]
12. Li G, Zhai G, Zhao X, et al. 3D texture analyses within the substantia nigra of Parkinson's disease patients on quantitative susceptibility maps and R2(*) maps. Neuroimage. Mar 2019;188:465–472. doi:10.1016/j.neuroimage.2018.12.041 [PubMed: 30578927]
13. Kang JJ, Chen Y, Xu GD, et al. Combining quantitative susceptibility mapping to radiomics in diagnosing Parkinson's disease and assessing cognitive impairment. Eur Radiol. Oct 2022;32(10):6992–7003. doi:10.1007/s00330-022-08790-8 [PubMed: 35461376]
14. Roberts AG, Zhang J, Tozlu C, et al. Quantitative Susceptibility Mapping Radiomics with Label Noise Compensation for Predicting Deep Brain Stimulation Outcomes in Parkinson's Disease. medRxiv. 2024:2024.12.26.24319663. doi:10.1101/2024.12.26.24319663
15. Roberts A, Zhang J, Tozlu C, et al. Radiomic prediction of Parkinson's disease deep brain stimulation surgery outcomes using quantitative susceptibility mapping and label noise compensation. neuromodecJ. 2024/10/1 2024;doi:10.31641/nmj-ogms6387
16. Zhao W, Yang C, Tong R, et al. Relationship Between Iron Distribution in Deep Gray Matter Nuclei Measured by Quantitative Susceptibility Mapping and Motor Outcome After Deep Brain Stimulation in Patients With Parkinson's Disease. J Magn Reson Imaging. Aug 2023;58(2):581–590. doi:10.1002/jmri.28574 [PubMed: 36594513]
17. Roberts AG, Zhang J, Kim H, et al. Radiomics for Deep Brain Stimulation outcome prediction using Quantitative Susceptibility Mapping (RadDBS-QSM). International Society of Magnetic Resonance in Medicine. 2024; 10.58530/2024/4718
18. Liu Y, Xiao B, Zhang C, et al. Predicting Motor Outcome of Subthalamic Nucleus Deep Brain Stimulation for Parkinson's Disease Using Quantitative Susceptibility Mapping and Radiomics: A Pilot Study. Front Neurosci. 2021;15:731109. doi:10.3389/fnins.2021.731109 [PubMed: 34557069]
19. Frenay B, Verleysen M. Classification in the Presence of Label Noise: A Survey. IEEE Transactions on Neural Networks and Learning Systems. 2014;25(5):845–869. doi:10.1109/tnnls.2013.2292894 [PubMed: 24808033]
20. Natarajan N, Dhillon IS, Ravikumar PK, Tewari A. Learning with Noisy Labels. 2013; https://proceedings.neurips.cc/paper_files/paper/2013/file/3871bd64012152bfb53fdf04b401193f-Paper.pdf
21. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. Movement Disorders. 2015;30(12):1591–1601. doi:10.1002/mds.26424 [PubMed: 26474316]
22. Liu J, Liu T, De Rochefort L, et al. Morphology enabled dipole inversion for quantitative susceptibility mapping using structural consistency between the magnitude image and the susceptibility map. NeuroImage. 2012;59(3):2560–2568. doi:10.1016/j.neuroimage.2011.08.082 [PubMed: 21925276]
23. Bilgic B, Costagli M, Chan KS, et al. Recommended implementation of quantitative susceptibility mapping for clinical research in the brain: A consensus of the ISMRM electro-magnetic tissue properties study group. Magnetic Resonance in Medicine. 2024;91(5):1834–1862. doi:10.1002/mrm.30006 [PubMed: 38247051]
24. Lin Z, Zhang X, Wang L, et al. Revisiting the L-Dopa Response as a Predictor of Motor Outcomes After Deep Brain Stimulation in Parkinson's Disease. Frontiers in Human Neuroscience. 2021;15 doi:10.3389/fnhum.2021.604433

25. Bennett LS, Warfield. Grand Challenge and Workshop on Multi-Atlas Labeling. 2012:
26. Avants BB, Yushkevich P, Pluta J, et al. The optimal template effect in hippocampus studies of diseased populations. NeuroImage. 2010;49(3):2457–2466. doi:10.1016/j.neuroimage.2009.09.062 [PubMed: 19818860]
27. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: Improved N3 Bias Correction. IEEE Transactions on Medical Imaging. 2010;29(6):1310–1320. doi:10.1109/tmi.2010.2046908 [PubMed: 20378467]
28. Avants B, Tustison NJ, Song G. Advanced Normalization Tools: V1.0. Insight J. 2009/7/29 2009;doi:10.54294/uvnhin
29. Van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research. 2017;77(21):e104–e107. doi:10.1158/0008-5472.can-17-0339 [PubMed: 29092951]
30. Roberts AG, Romano D, Zhang J, et al. QRadAR: An Open-source Toolbox for Quantitative Magnetic Resonance Radiomics Analysis and Reproducibility. International Society of Magnetic Resonance in Medicine. 2024; 10.58530/2024/0249
31. Li X, Xiong Y, Liu S, et al. Predicting the Post-therapy Severity Level (UPDRS-III) of Patients With Parkinson's Disease After Drug Therapy by Using the Dynamic Connectivity Efficiency of fMRI. Frontiers in Neurology. 2019;10 doi:10.3389/fneur.2019.00668
32. Wu CFJ. Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. The Annals of Statistics. 1986;14(4):1261–1295. doi:10.1214/aos/1176350142
33. Branco P, Torgo L, Ribeiro RP. SMOGN: a pre-processing approach for imbalanced regression. PMLR; 2017:36–50.
34. Stamoulou E, Spanakis C, Manikis GC, et al. Harmonization Strategies in Multicenter MRI-Based Radiomics. Journal of Imaging. 2022;8(11):303. doi:10.3390/jimaging8110303 [PubMed: 36354876]
35. Orlhac F, Eertink JJ, Cottereau A-S, et al. A Guide to ComBat Harmonization of Imaging Biomarkers in Multicenter Studies. Journal of Nuclear Medicine. 2022;63(2):172–179. doi:10.2967/jnumed.121.262464 [PubMed: 34531263]
36. Roberts AG, Romano D, Zhang J, et al. Radiomic Prediction of Parkinson's Disease Deep Brain Stimulation Surgery Motor and Non-motor Outcomes using Quantitative Susceptibility Mapping. Joint Workshop on MR Phase, Magnetic Susceptibility and Electrical Properties Mapping; 2024
37. Steiger JH. Tests for comparing elements of a correlation matrix. Psychol Bull. 1980/3 1980;87(2):245–251. doi:10.1037/0033-2909.87.2.245
38. Norgeot B, Quer G, Beaulieu-Jones BK, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. Nature Medicine. 2020;26(9):1320–1324. doi:10.1038/s41591-020-1041-y
39. Huang W, Ogbuji R, Zhou L, Guo L, Wang Y, Kopell BH. Motoric impairment versus iron deposition gradient in the subthalamic nucleus in Parkinson's disease. J Neurosurg. Jul 1 2021;135(1):284–290. doi:10.3171/2020.5.JNS201163 [PubMed: 32764171]
40. Roberts AG, Kovanlikaya I, Kopell B et al. Improved Visualization of the Medial Medullary Lamina with Phase Prior Reconstruction in Quantitative Susceptibility Mapping. International Society of Magnetic Resonance in Medicine. 2023; 10.58530/2023/1391
41. Wolke R, Becktepe JS, Paschen S, et al. The Role of Levodopa Challenge in Predicting the Outcome of Subthalamic Deep Brain Stimulation. Movement Disorders Clinical Practice. 2023;10(8):1181–1191. doi:10.1002/mdc3.13825 [PubMed: 37635781]
42. Saranza G, Lang AE. Levodopa challenge test: indications, protocol, and guide. Journal of Neurology. 2020;doi:10.1007/s00415-020-09810-7
43. Albanese A, Bonuccelli U, Brefel C, et al. Consensus statement on the role of acute dopaminergic challenge in Parkinson's disease. Movement Disorders. 2001;16(2):197–201. doi:10.1002/mds.1069 [PubMed: 11295770]
44. Kleiner-Fisman G, Herzog J, Fisman DN, et al. Subthalamic nucleus deep brain stimulation: Summary and meta-analysis of outcomes. Movement Disorders. 2006;21(S14):S290–S304. doi:10.1002/mds.20962 [PubMed: 16892449]

45. Hirschmann J, Steina A, Vesper J, Florin E, Schnitzler A. Neuronal oscillations predict deep brain stimulation outcome in Parkinson's disease. Brain Stimulation. 2022;15(3):792–802. doi:10.1016/j.brs.2022.05.008 [PubMed: 35568311]
46. Goede LL, Oxenford S, Kroneberg D, et al. Linking Invasive and Noninvasive Brain Stimulation in Parkinson's Disease: A Randomized Trial. Movement Disorders. 2024;doi:10.1002/mds.29940
47. Chen Y, Zhu G, Liu Y, et al. Predict initial subthalamic nucleus stimulation outcome in Parkinson's disease with brain morphology. CNS neuroscience & therapeutics. May 2022;28(5):667–676. doi:10.1111/cns.13797 [PubMed: 35049150]
48. Shang R, He L, Ma X, Ma Y, Li X. Connectome-Based Model Predicts Deep Brain Stimulation Outcome in Parkinson's Disease. Front Comput Neurosci. 2020;14:571527. doi:10.3389/fncom.2020.571527 [PubMed: 33192428]
49. Horn A, Reich M, Vorwerk J, et al. Connectivity Predicts deep brain stimulation outcome in Parkinson disease. Ann Neurol. Jul 2017;82(1):67–78. doi:10.1002/ana.24974 [PubMed: 28586141]
50. Wang J, Shang R, He L, et al. Prediction of Deep Brain Stimulation Outcome in Parkinson's Disease With Connectome Based on Hemispheric Asymmetry. Front Neurosci. 2021;15:620750. doi:10.3389/fnins.2021.620750 [PubMed: 34764846]
51. Lai Y, He N, Wei H, et al. Value of functional connectivity in outcome prediction for pallidal stimulation in Parkinson disease. J Neurosurg. Jan 1 2023;138(1):27–37. doi:10.3171/2022.3.JNS212732 [PubMed: 35523258]
52. Dimov A, Patel W, Yao Y, Wang Y, O'Halloran R, Kopell BH. Iron concentration linked to structural connectivity in the subthalamic nucleus: implications for deep brain stimulation. J Neurosurg. Jan 1 2020;132(1):197–204. doi:10.3171/2018.8.jns18531 [PubMed: 30660115]
53. Kosta P, Argyropoulou MI, Markoula S, Konitsiotis S. MRI evaluation of the basal ganglia size and iron content in patients with Parkinson's disease. Journal of Neurology. 2006;253(1):26–32. doi:10.1007/s00415-005-0914-9 [PubMed: 15981079]
54. Ma L, Gholam Azad M, Dharmasivam M, et al. Parkinson's disease: Alterations in iron and redox biology as a key to unlock therapeutic strategies. Redox Biology. 2021/05/01/ 2021;41:101896. 10.1016/j.redox.2021.101896 [PubMed: 33799121]
55. Fu X, Deng W, Cui X, et al. Time-Specific Pattern of Iron Deposition in Different Regions in Parkinson's Disease Measured by Quantitative Susceptibility Mapping. Frontiers in Neurology. 2021;12 doi:10.3389/fneur.2021.631210
56. Hartmann CJ, Fliegen S, Groiss SJ, Wojtecki L, Schnitzler A. An update on best practice of deep brain stimulation in Parkinson's disease. Therapeutic Advances in Neurological Disorders. 2019;12:175628641983809. doi:10.1177/1756286419838096
57. Magrinelli F, Picelli A, Tocco P, et al. Pathophysiology of Motor Dysfunction in Parkinson's Disease as the Rationale for Drug Treatment and Rehabilitation. Parkinson's Disease. 2016;2016:1–18. doi:10.1155/2016/9832839
58. Huang W, Ogbuji R, Zhou L, Guo L, Wang Y, Kopell BH. Motoric impairment versus iron deposition gradient in the subthalamic nucleus in Parkinson's disease. Journal of Neurosurgery. 2020;135(1):284–290. doi:10.3171/2020.5.jns201163 [PubMed: 32764171]
59. Freitas ME, Fox SH. Nondopaminergic Treatments for Parkinson's Disease: Current and Future Prospects. Neurodegenerative Disease Management. 2016;6(3):249–268. doi:10.2217/nmt-2016-0005 [PubMed: 27230697]
60. Agarwal KN. Iron and the brain: neurotransmitter receptors and magnetic resonance spectroscopy. British Journal of Nutrition. 2001;85(S2):S147. doi:10.1049/bjn2000307 [PubMed: 11509103]
61. Zucca FA, Segura-Aguilar J, Ferrari E, et al. Interactions of iron, dopamine and neuromelanin pathways in brain aging and Parkinson's disease. Progress in Neurobiology. 2017;155:96–119. doi:10.1016/j.pneurobio.2015.09.012 [PubMed: 26455458]
62. Hill JM, Switzer RC. The regional distribution and cellular localization of iron in the rat brain. Neuroscience. 1984/03/01/ 1984;11(3):595–603. 10.1016/0306-4522(84)90046-0 [PubMed: 6717804]
63. Zhou D, Cho J, Zhang J, Spincemaille P, Wang Y. Susceptibility underestimation in a high-susceptibility phantom: Dependence on imaging resolution, magnitude contrast, and other

parameters. Magn Reson Med. Sep 2017;78(3):1080–1086. doi:10.1002/mrm.26475 [PubMed: 27699883]

64. Roberts AG, Dimov A, Nguyen T, et al. Inhomogeneity-informed Field-fitting for Quantitative Susceptibility Mapping (if-QSM). International Society of Magnetic Resonance in Medicine. 2024; 10.58530/2024/4897
65. Chen J, Gong NJ, Chaim KT, Otaduy MCG, Liu C. Decompose quantitative susceptibility mapping (QSM) to sub-voxel diamagnetic and paramagnetic components based on gradient-echo MRI data. Neuroimage. Nov 15 2021;242:118477. doi:10.1016/j.neuroimage.2021.118477 [PubMed: 34403742]
66. Dimov AV, Nguyen TD, Gillen KM, et al. Susceptibility source separation from gradient echo data using magnitude decay modeling. Journal of neuroimaging : official journal of the American Society of Neuroimaging. Sep 2022;32(5):852–859. doi:10.1111/jon.13014 [PubMed: 35668022]
67. Shin HG, Lee J, Yun YH, et al. χ-separation: Magnetic susceptibility source separation toward iron and myelin mapping in the brain. Neuroimage. Oct 15 2021;240:118371. doi:10.1016/j.neuroimage.2021.118371 [PubMed: 34242783]
68. Li Z, Feng R, Liu Q, et al. APART-QSM: An improved sub-voxel quantitative susceptibility mapping for susceptibility source separation using an iterative data fitting method. Neuroimage. Jul 1 2023;274:120148. doi:10.1016/j.neuroimage.2023.120148 [PubMed: 37127191]
69. Roberts AG, Sisman M, Dimov A, et al. Whole Brain Source Separation for Neurodegeneration. International Society of Magnetic Resonance in Medicine. 2024;10.58530/2024/4370
70. Loehrer PA, Bopp MHA, Dafsari HS, et al. Microstructure predicts non-motor outcomes following deep brain stimulation in Parkinson's disease. npj Parkinson's Disease. 2024;10(1)doi:10.1038/s41531-024-00717-y
71. Roberts AG, Akkus S, Spadaccia M, et al. Joint Prediction of Motor and Non-motor Deep Brain Stimulation Outcomes using Quantitative Susceptibility Mapping. International Congress of Parkinson's Disease and Movement Disorders. https://www.mdsabstracts.org/abstract/joint-prediction-of-motor-and-non-motor-deep-brain-stimulation-outcomes-using-quantitative-susceptibility-mapping/
72. Boutet A, Madhavan R, Elias GJB, et al. Predicting optimal deep brain stimulation parameters for Parkinson's disease using functional MRI and machine learning. Nature Communications. 2021;12(1)doi:10.1038/s41467-021-23311-9
73. Roberts AG, Avecillas-Chasin J, Spadaccia M, et al. χ-DBS: An Open-Source Susceptibility Atlas Tool for Deep Brain Stimulation Target Visualization and Segmentation. International Congress of Parkinson's Disease and Movement Disorders. https://www.mdsabstracts.org/abstract/χ-dbs-an-open-source-susceptibility-atlas-tool-for-deep-brain-stimulation-target-visualization-and-segmentation/
74. Roberts AG, Romano DJ, Şişman M, et al. Maximum spherical mean value filtering for whole-brain QSM. Magnetic Resonance in Medicine. 2024;91(4):1586–1597. doi:10.1002/mrm.29963 [PubMed: 38169132]

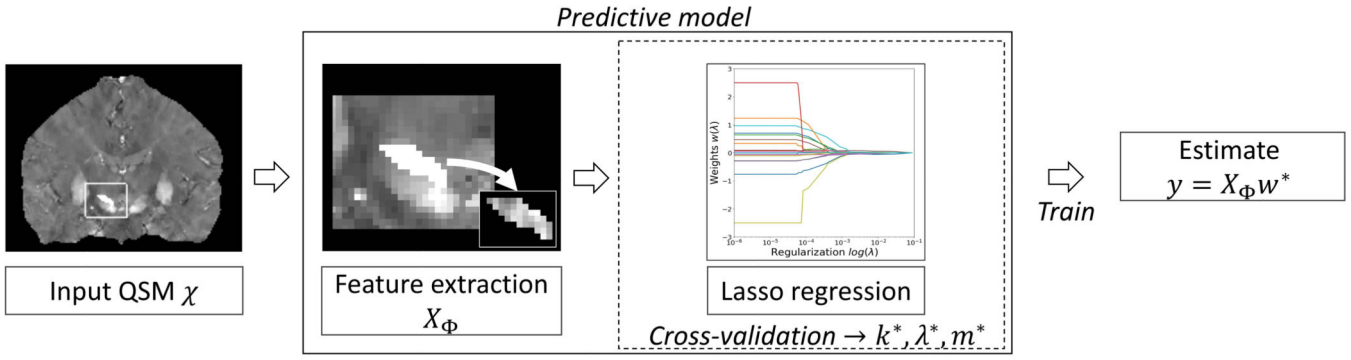

Figure 1.

Proposed pipeline with inputs (quantitative susceptibility maps) QSM and ROIs. Features were extracted from the input QSM χ over each ROI (joint entropy feature map shown) and passed to a Lasso regression model after normalization, label noise compensation injection, and recursive feature selection. A total of 93 features were extracted over 17 image types available in PyRadiomics. Additionally, 14 shape features were extracted on the original image for a total of 1595 features per ROI.

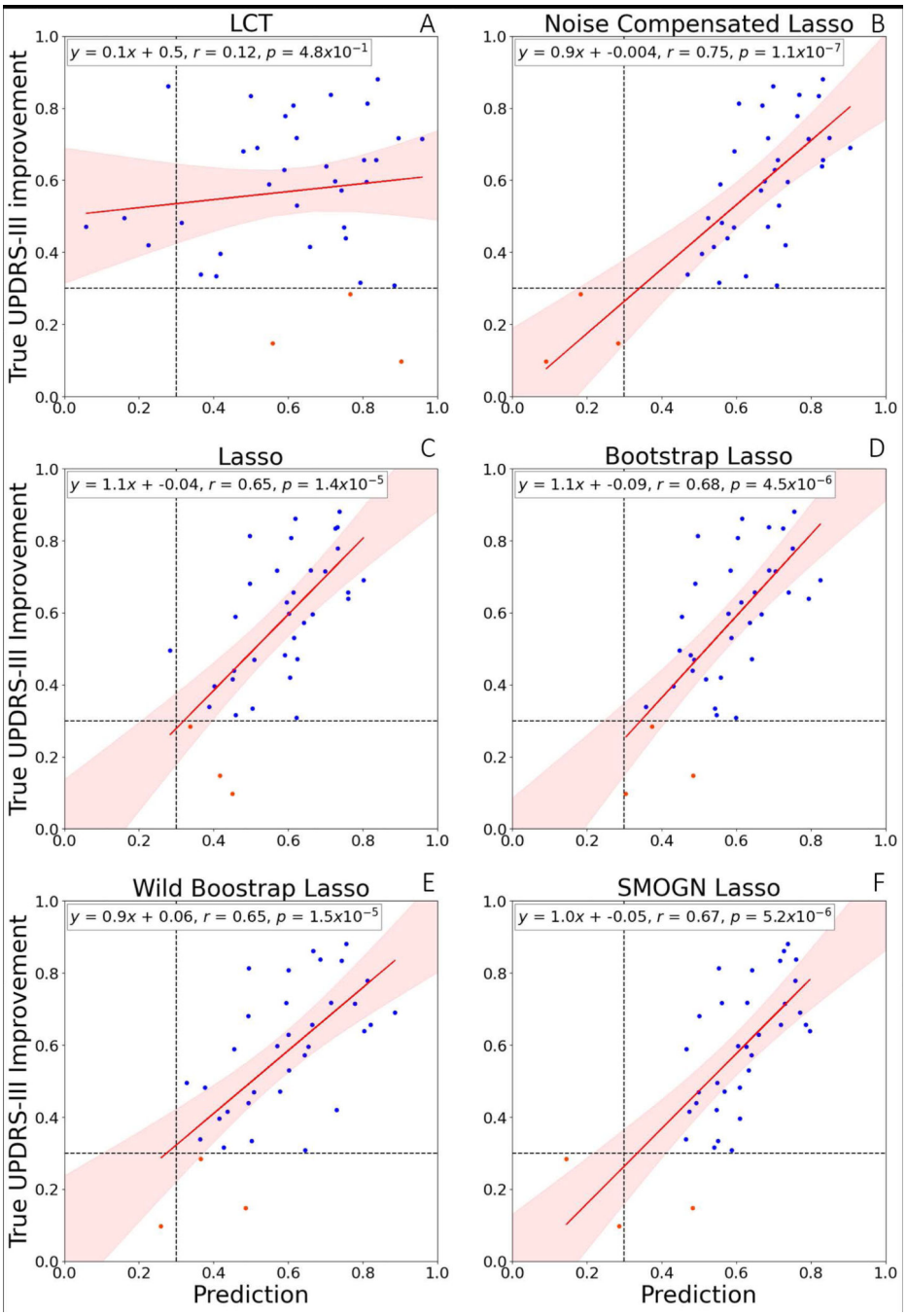

Figure 2.

Predicted (horizontal axis) Center 1 UPDRS-III improvement and true outcomes (vertical axis) using predictors: a) levodopa challenge test or LCT, b) Noise Compensated Lasso (proposed), c) Standard Lasso, d) Bootstrap Lasso, E) Wild Bootstrap Lasso, f) SMOGN Lasso. Blue points indicate a responsive surgical outcome (responder) while red points indicate a non-responsive surgical outcome (non-responder). Note the noise compensated Lasso model correctly classifies all cases.

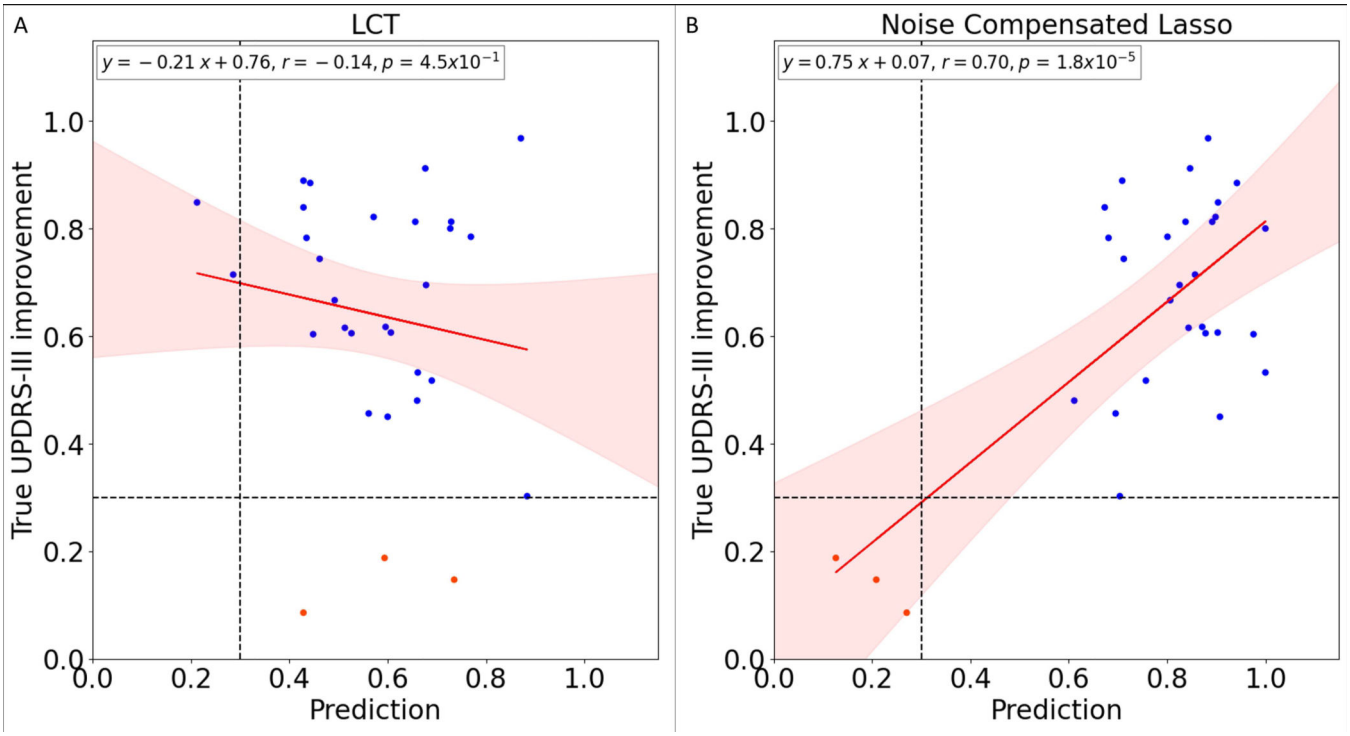

Figure 3.

Center 2 noise compensated results are comparable to Center 1, with inaccurate LCT prediction (a) and noise compensated prediction (b). Note the noise compensated model also correctly separates responders from non-responders at Center 2.

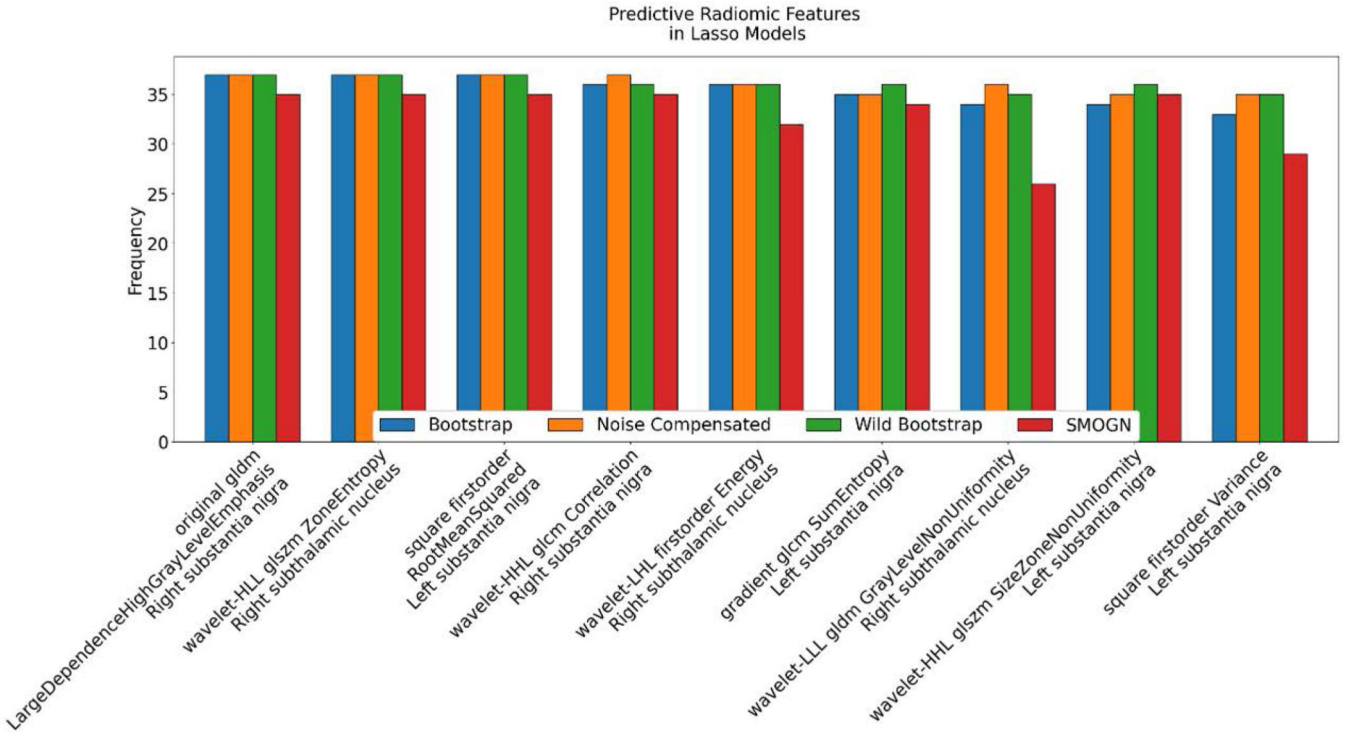

Figure 4.

Predictive radiomic features sorted by frequency in *N* models from Center 1. All models use the first feature, GLDM large dependence high gray level emphasis, which measures the homogeneity of the substantia nigra. In the subthalamic nucleus, the wavelet GLZSM entropy is most predictive, which measures the heterogeneity of susceptibility distribution. Center 2 models used GLZSM entropy on the local binary pattern image as well as wavelet decomposition features.

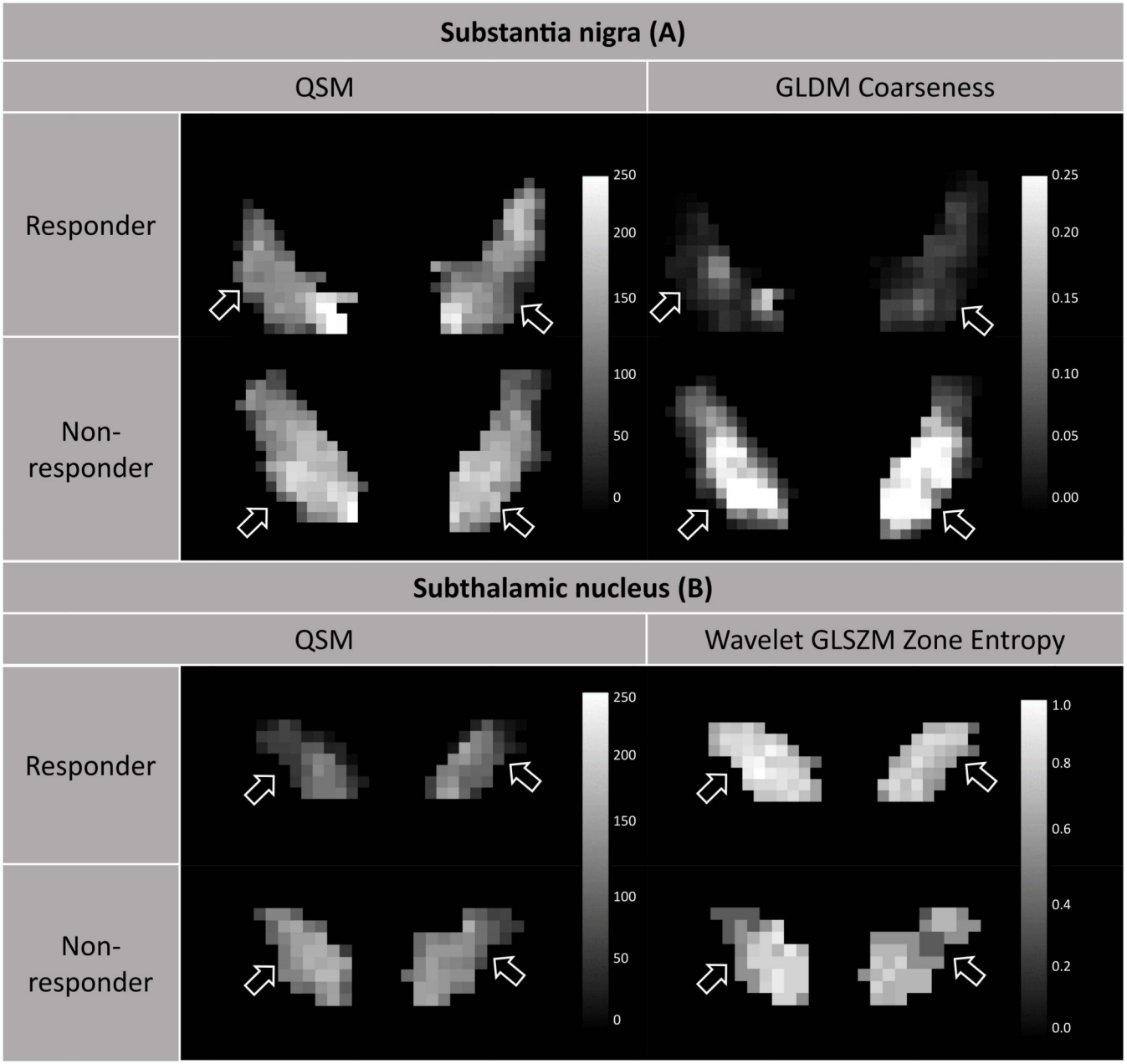

Figure 5.

QSM and normalized predictive feature maps for responders (first row) and non-responders (second row) in the substantia nigra (left, A, GLDM large dependence high gray level emphasis, or homogeneity) and subthalamic nucleus (right, B, wavelet GLZSM zone entropy). The increased gray level emphasis (A) and the decreased zone entropy (B) in non-responders versus responders were consistent with the ROI feature trends in SDC 7.