Investigating Brain Dopamine Lateralization in Parkinson's Disease

Alda Kola
dept. of Information Engineering
University of Padova
alda.kola@studenti.unipd.it

Beliz Günay
dept. of Information Engineering
University of Padova
beliz.guenay@studenti.unipd.it

Arda Ertanhan
dept. of Information Engineering
University of Padova
arda.ertanhan@studenti.unipd.it

***Abstract*—Parkinson's disease is a neurologic movement disorder type that affects the brain and occurs with motor symptoms, among them problems getting the movement underway. It is first characterized by the most common motor symptoms, that are bradykinesia or slowness of movement, tremors, and muscle stiffness or rigidity.This study leverages data from the Parkinson's Progression Markers Initiative (PPMI), focusing on dopamine transporter SPECT imaging to explore and quantify dopamine function asymmetry in PD patients. Analyzing both motor and non-motor symptoms, the study examines the correlation between lateralization of dopamine function and clinical severity, aiming to understand underlying factors driving asymmetry and assess lateralization's predictive value in clinical contexts.Despite robust modeling approaches, including statistical analysis methods and machine learning techniques to predict symptom progression and treatment response based on lateralization, the study acknowledges limitations such as reliance on cross-sectional data and potential biases due to imputed missing values. These constraints highlight the need for cautious interpretation of findings. Nevertheless, results suggest significant associations between dopamine lateralization in the putamen and symptom severity, underscoring potential avenues for personalized therapeutic strategies. Further research is recommended to deepen understanding of lateralization dynamics, incorporating longitudinal designs and broader demographic sampling to enhance generalizability and clinical relevance of the findings. We conducted our study in two parts, for the first part, where we investigated brain lateralization in Healthy Controls, we obtained a significant result for ANOVA for caudate lateralization, between handedness and education years. For the second part, our principal component analysis showed relations between dopamine lateralization and mood of patients (anxious and depressed), fatigue and rest tremor amplitudes for Parkinson's disease patients. The machine learning model that showed the best predictive performance for Putamen lateralization on test data was the Gradient Boosting model. This research also emphasizes the critical role of dopamine in motor and cognitive functions in PD, providing insights that could guide more targeted interventions and advance biomarker development in Parkinson's disease.**

***Index Terms*—Parkinson's disease, Dopamine lateralization, Multivariate analysis, Machine Learning, Covariate analysis, DAT SPECT imaging**

I. Background

Parkinson's disease is a progressive neurodegenerative disease that most prominently involves motor abilities. The cardinal features of motor involvement in Parkinson's disease are tremor, rigidity, bradykinesia, and postural instability. These features represent the neurodegeneration of dopaminergic neurons in the substantia nigra pars compacta, subsequently leading to severe depletion of dopamine levels within the striatum. As an extremely important neurotransmitter, dopamine has been linked to modulating motor function and coordination. The characteristic symptoms of PD—motor impairments—have been correlated with a lack of this neurotransmitter. [1] [2] [3]

Additionally, Parkinson's disease is characterized by the degeneration of the nigrostriatal dopamine neurons that can be identified by dopamine transporter single photon emission computed tomography. Clinically, DAT-SPECT serves as a biomarker of this degeneration and is usually analyzed by visual analysis backed by semi-quantitative parameters. There is, however, a growing requirement for quantifiable analysis in order to minimize subjectivity and the variations caused by the various evaluators. In particular, asymmetric hemispheric loss of dopaminergic neurons—a characteristic of PD—can be quantitatively assessed to help in diagnosis and provide insights about disease progression. [4]

DAT-SPECT (Single-Photon Emission Computed Tomography) is a nuclear imaging technique that measures the density and activity of dopamine transporters (DAT) in the brain. In Parkinson's disease, there is a loss of dopaminergic neurons, leading to decreased DAT levels. Through the DAT SPECT examination, it is possible to assess the amount of DAT present in an individual's brain, providing information on the function and integrity of dopaminergic neurons. This information can be used to diagnose Parkinson's disease and determine its severity. [5]

A large body of literature has explored the lateralization of dopamine function in PD. Imaging studies of PD patients with PET and SPECT have repeatedly reported asymmetrical dopamine depletion in the striatum. These lateralized differences in the degree of neurodegeneration have further been associated with the side of early motor symptom onset, which may imply a close interaction between the asymmetry of dopamine loss and the clinical manifestation of PD. [6] [7]

In addition, the severity and progression of symptoms in PD have been related to which hemisphere is more impaired concerning dopamine loss, making asymmetry in dopamine function critical to clinical prognosis and treatment planning. [8]

The literature bases the following study: hypothesis in existing

work is that asymmetry in dopamine function in the striatum of patients with PD is significant, with greater depletion on the same side as the initial motor symptom onset. The degree of asymmetry of dopamine is correlated with the severity of both motor and non-motor symptoms, where the greater the asymmetry, the more severe the symptom presentation. Dopamine asymmetry increases as the disease slowly progresses, and this could become a biomarker for the progression of the disease. Clinically, both assessment and follow-up of the progression of PD severity warrant a trial of DAT-SPECT in PD. [9] [10] The objective of the current study is to explore the asymmetry of dopamine function in different brain regions of PD patients using dopamine transporter SPECT data. analyze the correlation between dopamine laterality and clinical severity in both motor and non-motor symptoms. explore what underlying factors drive the asymmetry in both anatomical regions concerning symptom severity to establish why this asymmetry develops over time; assess the predictive value of left-right lateralization of dopamine function in clinical applications regarding the progression and severity of the symptoms in PD. In summary, we want to answer two questions:

1) Is dopamine function lateralized in healthy subjects? Is there any relevant covariate associated with dopamine function lateralization?
2) Is dopamine function lateralized in PD patients? How is brain lateralization associated with PD symptoms?

II. Materials and Methods

A. Dataset

In our study we used the dataset from the Parkinson's Progression Markers Initiative (PPMI), a project created to effectively collect data throughout all stages of Parkinson's disease (PD). The aim of this large data collection project is to accelerate the development of treatment strategies and facilitate the recognition of markers for the advancement of Parkinson's disease. Individuals with baseline Dopamine Transporter (DAT) imaging data are the specific focus of the analysis. The Signal Binding Ratios (SBR) for the Caudate and Putamen regions, expressed as DATSCAN_PUTAMEN_L, DATSCAN_PUTAMEN_R, DATSCAN_CAUDATE_L, and DATSCAN_CAUDATE_R, are the main variables of focus for our study.These measures are essential in determining the density of the DAT and also in identifying, classifying, and monitoring the development of Parkinson's disease. [1]

The PPMI dataset is segmented into multiple data files, with key files relevant to our study including: Demographic and Clinical Data (Patient_Master.csv): Patient Master Data (Patient_Master.csv) file provides detailed patient data for 256 healthy controls (HC) and 1223 (Parkinson's disease) PD participants across 158 columns, including demographics, clinical assessments, and DAT scan results. Contains demographic information, clinical data, and neurological assessments. Key variables include age, gender, diagnosis details, and scores from neurological assessments. Neurological Assessments: Details on cognitive functions, motor symptoms, pain, sleep disturbances, and mood variations.

B. Research Methods

1) ***Part 1: Dopamine Brain Lateralization in Healthy Controls****:*

a) ***Preprocessing****:* Dataset was filtered to remove healthy patients according to the cohort designation (COHORT = 'HC'). In Figure 1 and Figure 2 we can see the distribution of the caudate and putamen lataralization in healthy controls. Among the covaries chosen for analysis were:

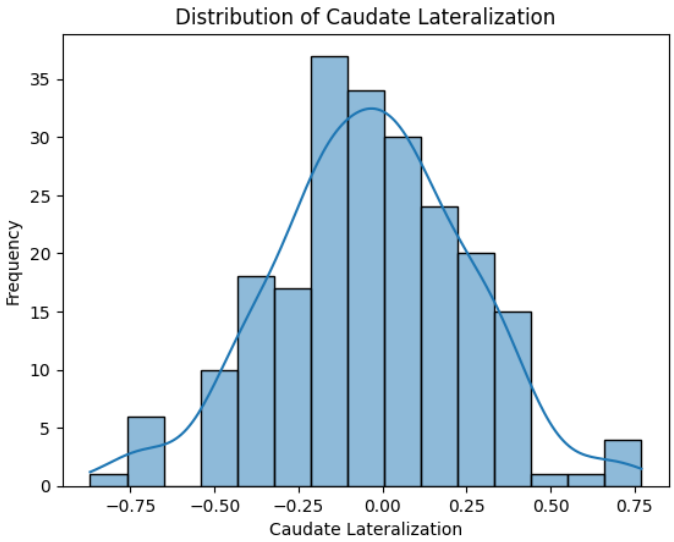

Fig. 1. Caudate Lateralization Distributions for HC

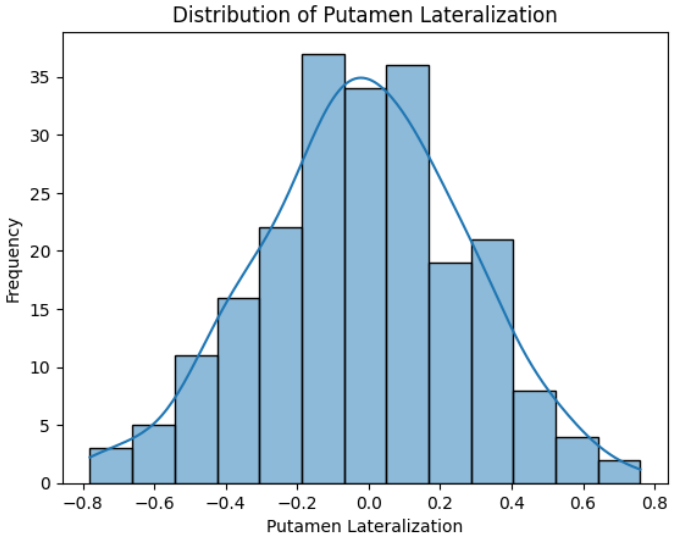

Fig. 2. Putamen Lateralization Distributions for HC

ENROLL_AGE (enrollment age), SEX, HANDED (handedness), EDUCYRS (education years), WGTKG (kilogram weight), ETHNICITY, HTCM (cm height), and MCATOT (total score on the Montreal Cognitive Assessment) Covariate rows with missing values were not included in the analysis. Mean was used to do imputation for the remaining numerical

Analysis	T-statistic	p-value
ETHNICITY (5 vs 1) and PUTAMEN LATERALIZATION	-0.22974391273651584	0.8185198411508741
ETHNICITY (5 vs 1) and CAUDATE LATERALIZATION	2.5589037094929323	0.011221322002249196
CAUDATE LATERALIZATION INDEX being different from zero	0.17319987580251195	0.8625230048477819

TABLE I
STATISTICAL T-TESTS FOR LATERALIZATION ANALYSES

columns. Because of the extensive preparation procedures, effects of missing data on the outcomes were deemed to be negligible. Categorical variables were encoded and continuous variables were binarized for comparative analysis in order to solve any potential data imbalance. Matching of the covariate distributions was aimed at by binarizing continuous variables and applying appropriate statistical tests to compare groups.

*b) **Statistical Methods**:*

- **ANOVA:** ANOVA was used to detect significant differences in lateralization between groups, examining the influence of factors like sex and handedness.
- Correlation coefficients of Pearson were calculated so as to find out the magnitude and direction of the relationship between lateralization and other continuous variable, the matrix of correlation has been prepared for knowing that whether the relationship of measurements related to dopamine exists or not with other variables.
- **OLS Regression Analysis:** Ordinary Least Squares regression was used to test dopamine lateralization in the caudate and putamen in relation to all other demographic and clinical covariates. In one set of models, sex, ethnicity, and handedness served as categorical independent variables with dopamine lateralization serving as the dependent variable and, as such, tested the main effects of each of these variables on measures of lateralization. [11].
- **Chi-square Test:** The chi-square statistic is utilized when a variable is measured at the nominal, or category level. All categories are mutually exclusive with nominal data. You could think of these data as buckets, and every piece of data can only fit in one bucket. [12] Relationship between categorical variables and dopamine lateralization in the caudate and putamen areas was evaluated in this project using the Chi-square test. The first step wasbinarizing continuous variables. Next, contingency tables between the binarized covariate and the lateralization measures (CAUDATE_LATERALIZATION and PUTAMEN_LATERALIZATION) were generated for each covariate. Finally, the Chi-square statistic and p-value were calculated to evaluate whether determinate any remarkable connections are present using the chi2_contingency function from the scipy.stats module.
- **T-Test:** It is also known as the independent t test. The unpaired t test is a test that determines whether the means of two unpaired, that is, independent groups are statistically different. It confirms whether there is statistical support for a significant difference in the population means. [13] This project employed the T-test to examine the means of categorical and continuous variables, as well as dopamine lateralization measurements between two groups. For reducing group comparisons, continuous factors were first binned into two groups (e.g., 'Low' and 'High'). Groups for each covariate were created depending on different levels of categorical covariates or depending on their binarized values. Independent two-sample T-tests were finally selected using the function ttest_ind from the scipy.stats module, and for each comparison between groups, it assures the T-statistic and p- value.
- **Wilcoxon-Mann-Whitney Test:** The Mann-Whitney U test or the Wilcoxon-Mann-Whitney test is a non-parametric test, used to check whether two independent samples are from the same distribution. This test is helpful when data does not meet the assumptions of the t-test, like normality. This test looks for the alternative hypothesis, H1 - that the two medians distinguish between the populations' - against the null hypothesis; H0- two populations' median is equal. This project ranks all observations from both groups and then compares the sum of ranks between groups. In case of tie in an observation, average of the tied ranks [14] is assigned. It is non- parametric substitute of T-test. In this project, it has been used to compare the distributions of dopamine lateralization measures between two groups. Continuous covariates were binned for the purpose of creating comparison groups. Groups were distinguished according to categorical or binarized continuous factors. This test furnished each comparison according to the p-value and U-statistic.

*2) **Part 2: Dopamine Brain Lateralization in PD**:*

*a) **Preprocessing**:* PD patients are filtered by the cohort column. SXDT and PDDXDT date columns are converted to numeric format for analytical compatibility. Also locating and choosing pertinent columns for variables and dopamine measurements is done for analytical assesment.For handling missing values, mode and median used to filled the gaps in numerical and categorical variables. One-hot encoding is used to encode categorical variables to prepare them for for statistical analysis.

*b) **Statistical Methods**:*

- **Principal Component Analysis (PCA):** is a multivariate exploratory analytic technique that is used to distinguish systematic variation in data from noise. This becomes possible while keeping all the important attributes in the dataset itself. As the PCA projects the high-dimensional data to smaller dimensions considering the variance and structure so it is possible to visualize

both objects, scores, and variables or loadings. [15] In this project, PCA has been used to find what principal component captured the variance in data and also helped in reducing the dimensionality of the dataset. First of all, standardized features of the chosen dataset have been taken. This is a crucial step since it makes every feature contribute to the PCA in an equal manner, and it does not allow the analysis to be dominated by features that have big ranges. After that, PCA was performed on this standardized data. The signal binding ratios were used to choose four primary components. The principle components are calculated after the PCA model is fitted to the scaled data. It was then necessary to calculate the variances explained by each principal component in order to find out how much variance is explained by each component. Original variables' contributions to each component are shown by the these four components: PC1 shows contributions from variables (e.g., 'NP2PTOT', 'NP3TOT') associated with motor and non-motor symptoms. PC2 shows contributions from variables relevant to diagnosis and treatment (e.g., 'PDDXDT', 'SXDT'). PC3 shows inputs from the outcomes of dopamine transporter scans (e.g., 'DATSCAN_VISINTRP_negative', 'DATSCAN_PUTAMEN_L'). PC4 states inputs from certain motor evaluations (e.g., 'NP3RTALU', 'NP3RIGLU').

- **Multivariate Analysis of Variance (MANOVA):** MANOVA is a multivariate statistical technique that is applied to analyze data from multiple variables assessed by an interval or ratio scale. Independent variable consists of two or more groups, and all of these variables are analyzed concurrently or jointly. [16] In this project, the effect of several independent variables on the total dependent variables that indicate dopamine lateralization was evaluated using MANOVA. The categorical variables were one-hot encoded to reach a numerical form. Results Can be seen on Table III. Finally, a relation had to be formulated between the independent variables that made up the covariates and dependent variables represented by the measures related to dopamine. The independent variables were placed on the right hand side of this formula and the dependent variables on its left. Different test statistics, such as Wilks' lambda, Pillai's trace, Hotelling-Lawley trace, and Roy's largest root were considered for the results of the MANOVA model. Results of independent factors affecting depending variables can be seen in the end.

*c) **Machine Learning Models**:* To enhance the analysis of dopamine function lateralization in Parkinson's disease (PD) and healthy controls, this study incorporates a range of machine learning models. Each model was selected for its capacity to manage complex medical data and offer insights into how brain lateralization indices correlate with clinical symptoms of Parkinson's disease.

- **Linear Regression:** It offers a clear connection, which is essential for our preliminary analysis, between the independent variables and the target variables. A baseline model used because it was the simplest and most interpretable-one.
- **Ridge Regression:** It is used for solving the problem of multicollinearity in the dataset. That would create a model which is not too complex yet has been prevented from overfitting. [17]
- **Lasso Regression:** This is the least absolute shrinkage and selection operator or LASSO. It is a dimensionality reduction technique to ensure variable screening and complexity reduction when fitting a generalized linear model. It makes the model more straight forward by reducing the number of features and also finds the most important features that affect lateralization. [18]
- **Random Forest:** It is a pretty robust ensemble method of many different decision trees for improving the prediction accuracy and gain control on overfitting. Its structure is with a reduced number of estimators. It gives trade-off between computing performance and complexity. [18]
- **Support Vector Regressor (SVR):** It is actually a subclass of the Support Vector Machine, as developed by Cortes and Vapnik. We used this in our study because its robust capability in handling linear as well as nonlinear regression tasks. Particularly suitable for predicting continuous variables such as the lateralization indices from DAT SPECT imaging data. To prevent overfitting we used a linear kernel with a reduced penalty factor (C=0.1), the SVR model focuses on maintaining simplicity, essential for managing the complex medical data in our dataset. The results can also be very clearly interpreted with the linear kernel, and thus informative about the relations of brain lateralization and clinical symptoms of Parkinson's disease. [19]
- **K-Nearest Neighbors (KNN):** As it is a non-parametric method, it does not make any assumption on the data distribution. It is a supervised learning procedure that relies on binary classification based on the assumption of the similar data points lying closer to each other. We used this non-parametric approach with a small number of neighbors, neighbors=3, in order to increase sensitivity towards the local structure of the data, which turned out to be really relevant to capture subtle variations in brain lateralization patterns. [20]
- **Gradient Boosting:** This model is an ensemble model of decision trees, so one would expect the output to be even more robust. We reduced the number of estimators to prevent overfitting while enhancing the model's ability to learn complex patterns. [21]
- **Hist Gradient Boosting:** This model is a more sophisticated variation of gradient boosting model, that maximizes the model's speed and performance by using histograms. As a result, it works incredibly well for large data sets with complex correlations. [22]

	Caudate Lateralization				Putamen Lateralization			
	sum_sq	df	F	PR(>F)	sum_sq	df	F	PR(>F)
C(SEX)	0.007030	1.0	0.085665	0.770052	0.037701	1.0	0.447154	0.504427
C(HANDED)	0.167962	2.0	1.023365	0.361177	0.123192	2.0	0.730571	0.482860
C(ETHNICITY)	0.722432	5.0	1.760662	0.122332	0.358324	5.0	0.849993	0.515767
Residual	17.151305	209.0	NaN	NaN	17.621251	209.0	NaN	NaN

TABLE II
ANOVA RESULTS FOR CAUDATE AND PUTAMEN LATERALIZATION

*d) **Model Evaluation**:* The effectiveness of each model was assessed using the following metrics:

- Mean Squared Error (MSE): It is the average of the squares of the errors and represents the average squared difference between the estimated values and the actual value. A smaller MSE implies a better fit.
- R-squared (R^2): Represents the proportion of the variance in the dependent variable that is predictable from the independent variables. Higher values indicate a better fit and more explanatory power.

*e) **Model Validation**:* We applied cross-validation techniques alongside with a train-test split methodology to validate the models' generalizability and reliability. This guarantees models are reliable and accurate across various data subsets.

III. RESULTS

ANOVA Results

The ANOVA results indicate that the effects of sex, handedness, and ethnicity on both caudate and putamen lateralization are generally not statistically significant, as shown by the high p-values. For instance, in the caudate lateralization, the p-values for sex, handedness, and ethnicity are 0.770, 0.361, and 0.122 respectively, indicating no strong evidence against the null hypothesis of no effect. This suggests that these factors may not play a significant role in influencing the lateralization of these brain regions within the studied cohort.

However, the interaction terms in the extended ANOVA model for caudate lateralization, specifically the interaction between handedness and education years, show a significant effect with a p-value of 0.0215, suggesting that the impact of handedness on caudate lateralization may vary by the educational level of individuals. Table II presents the results for the caudate and putamen lateralization.

Pearson Correlation Results

The Pearson correlation matrix reveals mostly weak correlations between the lateralization indices and other continuous variables like age, education years, and body measurements. Notable is the absence of strong correlations, indicating that these covariates alone do not explain much of the variance in lateralization.

Linear Regression Results

The linear regression results, showing negative R^2 values in the initial models, indicate that the models don't fit well, suggesting that these models are not suitable for predicting lateralization based on the provided covariates. However, when interaction terms are added, the R^2 values improve slightly

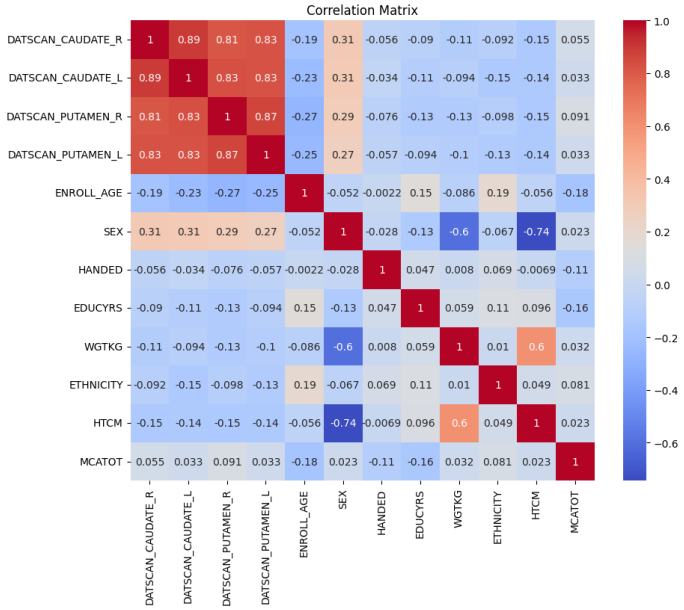

Fig. 3. Correlation matrix to identify relationships between dopamine function and covariates

(0.01799 for caudate and 0.01959 for putamen with interaction terms), suggesting that complex relationships involving interactions might be more informative.

Evaluation of Different Models

For the analysis of lateralization in the caudate and putamen regions we utilized various machine learning models to predict lateralization indices from clinical and imaging data. The models tested included Linear Regression, Random Forest, Support Vector Regression (SVR), K-Nearest Neighbors (KNN), and Gradient Boosting, each evaluated for their performance in terms of Mean Squared Error (MSE) and R-squared (R^2) values, alongside cross-validation metrics. Models show varying degrees of effectiveness, with the Gradient Boosting model performing notably well for Putamen lateralization with an R^2 of 0.8732 on test data, indicating strong predictive power.

Caudate Lateralization Evaluation

- Linear Regression: Achieved a mean squared error (MSE) of 0.008221 and an R^2 of 0.351, showing moderate predictive ability.
- Random Forest: Displayed a slightly lower performance with an MSE of 0.008964 and an R^2 of 0.292.
- Support Vector Regressor: Comparable to Linear Regression with an MSE of 0.008290 and an R^2 of 0.345.

Covariate	Wilks' lambda	F-value	p-value
DATSCAN	0.9876	3.4993	0.0075
DATSCAN_QUALITY_RATING	0.9830	4.8043	0.0008
DXRIGID	0.9907	2.6176	0.0338
ENROLL_AGE	0.9909	2.5367	0.0386
NP2PTOT	0.9914	2.4229	0.0466
NP2SALV	0.9914	2.4082	0.0477
NP2SPCH	0.9887	3.1617	0.0135
NP2FREZ	0.9900	2.8131	0.0243
NP3RTALU	0.9876	3.4850	0.0077
NP3RIGLU	0.9914	2.4169	0.0471
NP3FTAPL	0.9913	2.4423	0.0451

TABLE III
MANOVA RESULTS SHOWING SIGNIFICANT COVARIATES WITH WILKS' LAMBDA, F-VALUE, AND P-VALUE

- K-Nearest Neighbors: Demonstrated an MSE of 0.004961 and an R^2 of 0.500, reflecting a better fit compared to other models for caudate lateralization.

Putamen Lateralization Evaluation

- Linear Regression: Shows significant explanatory power with a MSE of 0.009312 and a R2 of 0.784.
- Random Forest: This model demonstrated strong performance, with a MSE of 0.010670 and a R2 of 0.752.
- Support Vector Regressor: Shows moderate efficacy with a MSE of 0.015254 and a R^2 of 0.646.
- K-Nearest Neighbors: This model was less appropriate for the task, as seen by its lower R^2 of 0.213 and higher MSE of 0.033909.
- Gradient Boosting: Gradient Boosting was the model that performed the best with a MSE of 0.005462 and R^2 of 0.873.

Descriptive Statistics

From the results of the descriptive statistics, in the evaluation for caudate lateralization, the mean index across all participants was approximately 0.000554 with a standard deviation of 0.111885. While, for the putamen lateralization, the mean index was slightly higher at 0.013102 with a standard deviation of 0.204885. These statistics suggest variability in lateralization across the study population in our study.

Statistical Tests

The results associated to the statistical tests in our study are: Caudate Lateralization: T-tests for caudate lateralization index being different from zero yielded a T-statistic of 0.1732 with a P-value of 0.8625, suggests no significant lateralization from zero.

Putamen Lateralization: For putamen lateralization, the T-test provided a T-statistic of 2.2364 with a P-value of 0.0255, which indicates significant lateralization from zero.

Linear Regression Model Summaries

The results we obtained in our study associatted to the linear regression model are as listed below: Caudate Lateralization: The model accounted for 38.2% of the variance (Adj. R-squared: 0.376) with significant predictors including rigidity and finger-tapping scores.

Putamen Lateralization: This model explained 39.2% of the variance (Adj. R-squared: 0.386), similarly highlighting significant predictors related to motor symptoms.

The correlation matrix on Figure 3, suggests a strong bilateral symmetry in dopamine function within both the caudate and putamen regions, with moderate influences from age and sex, and weaker influences from other covariates.

Chi-Square

There were no significant relationships between the categorical covariates (gender, handedness, and ethnicity) and the dopamine measurements.

T-test

Dopamine lateralization assessments based on continuous or categorical variables doesn't differ significantly, with the exception of a few particular ethnicity pairs that displayed substantial variations in caudate lateralization (Table I).

Wilcoxon-Mann-Whitney

There were no noticeable variations in dopamine levels among the different continuous covariate categories.

Principal Component Analysis (PCA)

Twenty-two percent of the variance in the dataset is cumulatively explained by the four main components, as shown in Table IV, while the main variables are shown in Table V. This proportion shows that many additional components continue to contribute to the overall variance even with a moderate reduction in dimensionality, visualized in Figure 4.

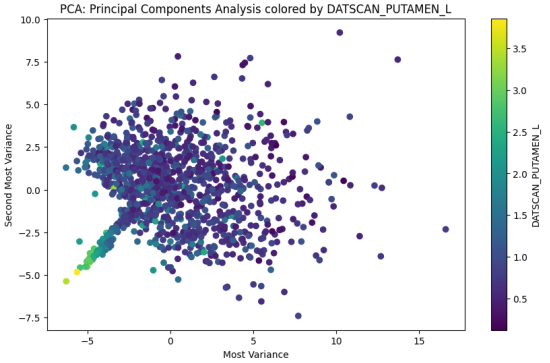

Fig. 4. Principal Component 1 (highest percentage of explained variance)

IV. DISCUSSION

In our study we investigated the lateralization of dopamine function in both healthy controls (HC) and individuals with

Principal Component 1	10.14
Principal Component 2	5.09
Principal Component 3	4.57
Principal Component 4	4.40
Total Explained Variance	24.20

TABLE IV
EXPLAINED VARIANCE FOR PCA

PC	Variables
1	NP2PTOT, NP3TOT, NP2RISE, NP2DRES, NP2HOBB, NP1PTOT, NP2WALK, NP2TURN, NP2EAT, NP3SPCH
2	PDDXDT, SXDT, PDTRTMNT, NP1DPRS, GENETICS_LRRK2, NP1RTOT, DATSCAN_VISINTRP_positive, NP1ANXS, NP1COG, NP1APAT
3	DATSCAN_PUTAMEN_L, DATSCAN_VISINTRP_negative, NP1FATG, NP1PTOT, DATSCAN_CAUDATE_L, DATSCAN_PUTAMEN_R, NP1PAIN, NP1SLPN, NP1RTOT, NP1SLPD
4	NP3RTALU, NP3RIGLU, NP3PRSPL, NP3FTAPL, NP3HMOVL, NP3RIGLL, NP3TTAPL, NP3PTRML, NP3LGAGL, NP3KTRML

TABLE V
TOP CONTRIBUTING VARIABLES TO PRINCIPAL COMPONENTS 1 TO 4

Parkinson's Disease (PD). We focused on the relationship between lateralization of the brain and a number of clinical symptoms in PD patients. For determining if dopamine function is lateralized in this population, we focused on healthy subjects in the first task. Aim of this research was to find any covaries connected to dopamine lateralization. In the second part, we expanded the study to include individuals with Parkinson's disease (PD), analyzing metrics related to dopamine function and investigating the connection between lateralization and PD symptoms. Comprehensive dataset including clinical, demographic, and neuroimaging data from Parkinson's disease (PD) patients was used for this thorough project. The results showed that different clinical symptoms in Parkinson's disease patients and dopamine function lateralization interact significantly. Numerous covariates revealed this link, stating various variables affect dopamine activity in the brain.

Part 1: Dopamine Brain Lateralization in Healthy Controls: Statistical results indicate that there is no significant lateralization of dopamine function in healthy cohorts. Dopamine function of the caudate and putamen was not significantly influenced by the left or right hemisphere, as seen by the distributions of these measurements' centers around zero. There was no notable significant correlation between any factors and the lateralization of dopamine function. Chi-square, T-test and Wilcoxon-Mann-Whitney tests showed that in healthy cohorts, variables including age, sex, ethnicity, handedness, education, height, and weight had no significant effect on the lateralization measurements of dopamine function. Our analysis did not find statistically significant effects of sex, handedness, and ethnicity on dopamine lateralization in the caudate and putamen regions. These demographic factors alone may not play a decisive role in the lateralization processes within this cohort. However, the high p-values may indicate that other, unmeasured variables could interact with these factors to influence brain structures. Importantly, the interaction between handedness and education years revealed a significant effect , suggesting that educational progress might modify how handedness affects caudate lateralization. This finding suggests that in this case, higher cognitive capacity, associated with education years can influence neuroplastic responses, and potentially impact the asymmetry of dopamine neurotransmission.

Part 2: Dopamine Brain Lateralization in PD: In our study, we found that the lateralization of dopamine in PD patients is significantly related to the severity of motor symptoms, as strongly demonstrated also by results obtained in predicting the Gradient Boosting model for putamen lateralization, and also PCA and MANOVA analysis. Gradient Boosting model performance underscores the clinical relevance of understanding lateralization patterns, which may help predict the progression of PD symptoms or the response to specific treatments. From the comparison of the healthy controls (HC) and Parkinson's Diease (PD) caudate lateralization, we can see no significant differences, as shown by the T-test. Caudate lateralization doesn't differ for the two groups. However, if we check high standard deviations and ranges in PD, we can notice that putmane lateralization shows more variability in PD patients compared to healthy controls. This is also shown in Figure 5. Moreover, the significant lateralization observed in the putamen, contrary to the non-significant results from the caudate, suggests that the putamen may be more sensitive to changes associated with PD or more directly involved in the motor pathways affected by the disease. These results are consistent with current knowledge of PD pathology insofar as the putamen is more severely affected by PD pathology given its central role in regulating motor functions and active connectivity to dopaminergic pathways. However, the results from MANOVA, shown in Table III, and PCA among patients with PD, brain lateralization is significantly linked to a string of motor and non-motor symptoms. Of which stiffnes, trouble speaking, freezing of gait, and overall UPDRS motor scores.

Study Limitations

During our work we observed some limitations. These limitations of our study are from its reliance on cross-sectional data from the PPMI dataset, which limits the ability to determine causal relationships or observe changes in dopamine lateralization over time. The study's scope is limited to the early stages of the disease and may not fully reflect the course of Parkinson's disease due to its emphasis on baseline DAT imaging. Other biases that result from missing values in the dataset can also have an impact on the final result. Furthermore, negative R2 values in the original models suggest possible problems with model fit. This implies that the intricate dynamics of dopamine lateralization might not be sufficiently captured by the chosen features and model configurations. The

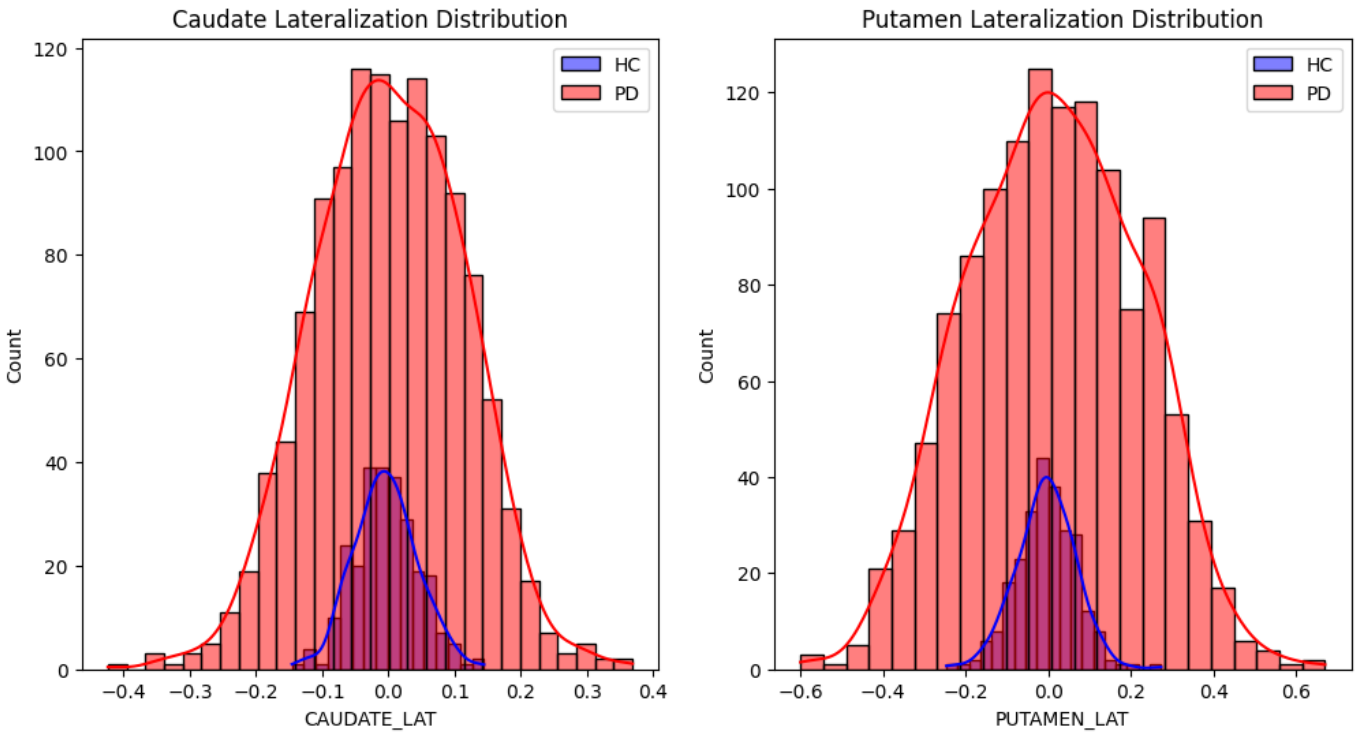

Fig. 5. Comparison of Lateralization betweeen HC and PD

high p-values in the ANOVA analysis also suggest that demographic factors such as sex, handedness, and ethnicity might not have been sufficiently varied within the cohort to detect significant effects, or that other unmeasured variables could be influencing the results. Moreover, the generalizability of the findings may be limited by the demographic composition of the study population, which might not represent wider, more diverse populations.

Future Work

The findings of the present study would be beneficial for other future research. Lifestyle and cognitive factors are possibly affecting brain health in ways that are emphasized by the relationship of handedness, together with education, in healthy subjects. This makes it possible to propose some preventive measures against neurogenerative processes. Association of the putamen lateralization with the severity of early stage Parkinson's disease underlines ways of giving treatment a more individualized profile so that the therapeutic outcome would benefit from being tailored to individual patterns of lateralization. The complex relationships between genetic, environmental, and lifestyle factors in brain lateralization requires more investigation. It is possible to further improve the machine learning models in terms of interpretation and prediction power. We can say that, the importance of brain asymmetry can be better understood through longitudinal research that examines neuroinflammatory markers and use a variety of data.

REFERENCES

[1] Dauer, W., & Przedborski, S. (2003b). Parkinson's Disease. Neuron, 39(6), 889–909. https://doi.org/10.1016/s0896-6273(03)00568-3

[2] Kish, S. J., Shannak, K., & Hornykiewicz, O. (1988). Uneven Pattern of Dopamine Loss in the Striatum of Patients with Idiopathic Parkinson's Disease. New England Journal of Medicine/the New England Journal of Medicine, 318(14), 876–880. https://doi.org/10.1056/nejm198804073181402

[3] Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. Journal of Neurology, Neurosurgery and Psychiatry, 79(4), 368–376. https://doi.org/10.1136/jnnp.2007.131045

[4] Kagi, G., Bhatia, K. P., & Tolosa, E. (2009). The role of DAT-SPECT in movement disorders. Journal of Neurology, Neurosurgery and Psychiatry, 81(1), 5–12. https://doi.org/10.1136/jnnp.2008.157370

[5] Shiiba, T., Takano, K., Takaki, A., & Suwazono, S. (2022). Dopamine transporter single-photon emission computed tomography-derived radiomics signature for detecting Parkinson's disease. EJNMMI Research, 12(1). https://doi.org/10.1186/s13550-022-00910-1

[6] Djaldetti, R., Ziv, I., & Melamed, E. (2006). The mystery of motor asymmetry in Parkinson's disease. Lancet Neurology, 5(9), 796–802. https://doi.org/10.1016/s1474-4422(06)70549-x

[7] Booij, J., Tissingh, G., Boer, G. J., Speelman, J. D., Stoof, J. C., Janssen, A. G., Wolters, E. C., & Van Royen, E. A. (1997). [123I]FP-CIT SPECT shows a pronounced decline of striatal dopamine transporter labelling in early and advanced Parkinson's disease. Journal of Neurology, Neurosurgery and Psychiatry, 62(2), 133–140. https://doi.org/10.1136/jnnp.62.2.133

[8] Haaxma, C. A., Bloem, B. R., Borm, G. F., Oyen, W. J. G., Leenders, K. L., Eshuis, S., Booij, J., Dluzen, D. E., & Horstink, M. W. I. M. (2007). Gender differences in Parkinson's disease. Journal of Neurology, Neurosurgery and Psychiatry, 78(8), 819–824. https://doi.org/10.1136/jnnp.2006.103788

[9] Morrish, P. K., Sawle, G. V., & Brooks, D. J. (1996). An [18F]dopa–PET and clinical study of the rate of progression in Parkinson's disease. Brain, 119(2), 585–591. https://doi.org/10.1093/brain/119.2.585

[10] Fiorenzato, E., Antonini, A., Bisiacchi, P., Weis, L., & Biundo, R. (2021). Asymmetric dopamine transporter loss affects cognitive and

motor progression in Parkinson's disease. Movement Disorders, 36(10), 2303–2313. https://doi.org/10.1002/mds.28682

[11] Fauth, E. B., Schaefer, S. Y., Zarit, S. H., Ernsth-Bravell, M., & Johansson, B. (2016). Associations between fine motor performance in activities of daily living and cognitive ability in a non-demented sample of older adults: Implications for geriatric Physical rehabilitation. Journal of Aging and Health, 29(7), 1144–1159. https://doi.org/10.1177/0898264316654674

[12] Connelly, L. M. (n.d.). Chi-Square Test. Nursing Research, 50(3), 199-201.

[13] Mishra, P., Singh, U., Pandey, C., Mishra, P., & Pandey, G. (2019). Application of student's t-test, analysis of variance, and covariance. Annals of Cardiac Anaesthesia/Annals of Cardiac Anaesthesia, 22(4), 407. https://doi.org/10.4103/aca.aca9419

[14] Iuliano, A., & Franzese, M. (2019). Introduction to Biostatistics. https://www.semanticscholar.org/paper/Introduction-to-Biostatistics-Iuliano-Franzese/8974f815f0efb19d98cd3767f3b4e2cf950338d5

[15] Geladi, P., & Linderholm, J. (2020). Principal component analysis. In Elsevier eBooks (pp. 17–37). https://doi.org/10.1016/b978-0-12-409547-2.14892-9

[16] Saleh, H., Hasanah, S. I., & Subaidi, A. (2019). Implementation of Multivariate Analysis of Variance (MANOVA) in experiments factorial two factors (Study: Growth and development of soybean germination). Journal of Physics. Conference Series, 1375(1), 012013. https://doi.org/10.1088/1742-6596/1375/1/012013

[17] Rauschenberger, A., & Glaab, E. (2021). Predicting correlated outcomes from molecular data. Bioinformatics, 37(21), 3889–3895. https://doi.org/10.1093/bioinformatics/btab576

[18] Zhao, S., Zhang, L., Ji, W., Shi, Y., Lai, G., Chi, H., Huang, W., & Cheng, C. (2022). Machine learning-based characterization of cuprotosis-related biomarkers and immune infiltration in Parkinson's disease. Frontiers in Genetics, 13. https://doi.org/10.3389/fgene.2022.1010361

[19] Nilashi, M., Abumalloh, R. A., Minaei-Bidgoli, B., Samad, S., Ismail, M. Y., Alhargan, A., & Zogaan, W. A. (2022). Predicting Parkinson's disease progression: Evaluation of ensemble methods in machine learning. Journal of Healthcare Engineering, 2022, 1–17. https://doi.org/10.1155/2022/2793361

[20] Li, A., & Li, C. (2022). Detecting Parkinson's Disease through Gait Measures Using Machine Learning. Diagnostics, 12(10), 2404. https://doi.org/10.3390/diagnostics12102404

[21] Lee, S., Kim, Y., Hwang, S., Son, H., Lee, S. K., Park, K., & Kim, Y. (2022). Predicting Parkinson's disease using gradient boosting decision tree models with electroencephalography signals. Parkinsonism & Related Disorders (Online)/Parkinsonism & Related Disorders, 95, 77–85. https://doi.org/10.1016/j.parkreldis.2022.01.011

[22] Liu, W., Jia, L., Xu, L., Yang, F., Guo, Z., Li, J., Zhang, D., Liu, Y., Xiang, H., Cheng, H., Hou, J., Li, S., & Li, H. (2024). Prediction of early neurologic deterioration in patients with perforating artery territory infarction using machine learning: a retrospective study. Frontiers in Neurology, 15. https://doi.org/10.3389/fneur.2024.1368902