OVERVIEW

This manuscript investigates signatures of brain structure which are associated with treatment response in individuals with Parkinson’s Disease (PD). Using a dataset of individuals with PD who received deep brain stimulation of the subthalamic nucleus (STN-DBS), they apply a recently developed approach for deriving individualized measures of brain structural covariance. Through a series of analyses incorporating machine learning, network theory, and case-control designs, the authors characterize the morphological features which are associated with better or worse response to stimulation. They find several regions and network descriptors which are associated with clinical improvement. Their results have potential for informing the prediction of treatment response in this population.

There are some notable strengths of this work. The authors outline a practical problem of STN-DBS monitoring, which is the costly and time-consuming use of fMRI to predict treatment response. As such, their approach addresses a clear gap in the field. There are many places where the authors have interesting and nuanced descriptions of neuroscientific and neurological background related to PD, which helps contextualize their findings and broadens the appeal of the work. They also apply advanced methodology for investigating their questions of interest.

However, I have some concerns with the work, particularly in terms of the methodological detail and the presentation and interpretation of findings. The methodological descriptions are unclear or lacking in several places, which limits the replicability of the work as well as its general appeal. My comments below further expand upon these and other concerns.

ABSTRACT

1. Between the “Background” and “Methods”, there is not a clear motivation for how the method applied (characterization of individualized structural network covariance) addresses the problem (differential efficacy of DBS). Why should one expect that individual measures of structural covariance affect DBS efficacy? Clarification of this may improve the appeal of the work to more readers.

2. In the “Results”, there is an opened parenthesis which is never closed: “six edges (left Middle Frontal …”).

INTRODUCTION

3. The sentence “However, LCT is effective in excluding patients with a non-responsive reaction to levodopa, its predictive accuracy for postoperative long-term motor outcomes is considerably wide and not precise [7-10]” seems to be missing an “and” or semicolon.

4. The authors provide clear practical motivation for using structural MRI instead of fMRI. However, the introduction seems to indicate that fMRI is what is currently used for assessing
brain connectivity in this setting. Is there evidence to suggest that structural connectivity/covariance can be used as a drop-in method for fMRI? The time/cost is a practical advantage; are there any biological advantages as well? Or disadvantages?

METHODS

5. The authors mention medication-off and medication-on states, but don’t describe the medication use of participants. Are all PD participants taking medication? Is it L-DOPA? When are they on medication, and when are they off medication?

6. It would be helpful if the authors could indicate what values the MDS-UPDRS takes on. Is there a minimum or maximum score? Do higher values indicate improvement or impairment? This would help make the results more interpretable and clarify the formula for calculating percentage improvement.

7. The description of the methodology for constructing IDSCNs is unclear. I was not able to understand how the method worked without referring to the original paper (Liu et al, 2021). I particularly found the notion of “individualized structural covariance” confusing, as covariance typically needs to be estimated from repeated measurements (this is something that is motivated better by Liu et al). I would suggest adding more details to the “Individual Differential Structural Covariance Network (IDSCN) Measures” section in order to provide readers with a better intuition for how the approach works, what are the inputs/outputs, and its advantages relative to alternative approaches. Some specific suggestions are as follows:
   a. The authors mention that each patient added to create PCCn+1 is “denoted as ‘k’”. However, this ‘k’ is never used throughout the manuscript, besides in Figure 1 where it is also superfluous. If the authors do not use this extra symbol, I would remove it because it is confusing extra information.
   b. The term “incrementally added to the PCCn” is ambiguous, as it evokes both concatenation and mathematical addition. I would rephrase this to be clearer. My understanding is that the correlation matrix (PCCn) is recomputed with the inclusion of a single PD participant (to create PCCn+1). Saying “added to the PCCn” sounds like the authors are mathematically adding a value or values to the correlation matrix, not recomputing it.
   c. The use of the word “iteration” also makes it sound like the PCCn is being iteratively worked on, or that PCCn+1 is iteratively being computed by sequentially including extra PD participants.

8. I have some concerns about the design of the MLP classification experiment. The results (“The six edges demonstrated good predictive capabilities”) and the Figure 3 caption (“Prediction of Long-Term Improvement Groups in Patients Using 6 Edge Features”) indicate that the edges found to be most correlated with DBS improvement were used as features in the classification analysis. If the correlations were calculated in the full dataset (which includes the data that were later split into training & testing sets), then data leakage occurred (enriching for features which are associated with the outcome in the testing set). The authors should clarify which features are being used in the methods. If they are applying any sort of feature selection, it
needs to be completed without consultation of the testing set at all; they should specifically note this design if this is already the case.

9. Furthermore, the detail is relatively limited for describing the MLP methodology. What is the architecture of the MLP? What hyperparameters are used? What activation function? What features are used? Additionally, the authors mention the generation of an ROC for “each fold” and an “average ROC curve”, but they do not otherwise indicate that they are using cross-validation. They must provide more detail about this if so (How many folds? Nested or not?).

10. More detail should also be included in the XGBoost section. What hyperparameters are tuned in the grid search? What values are tested? Is cross-validation being used to tune hyperparameters? How are feature importance values calculated?

11. Do the authors intend to share code for their analyses? This would help make their work more reproducible, which is a concern given some missing details in the methodology.

RESULTS

12. The results mention a Random Forest classifier (page 12). This is not mentioned in the methods.

13. The last paragraph of the Results mentions Figure 6B, but only 5 figures are included. I’m assuming the authors mean 5B?

DISCUSSION

14. The phrase in the results “exhibited higher nodal local efficiency and degree centrality, resulting in improved long-term outcomes” evokes causality, as if the network metrics in these regions caused the improved outcome. The design is essentially a retrospective case-control study, so causality cannot be inferred. The authors should rephrase to not imply that the differences in brain structure “result[ed]” in the observed clinical improvement.

FIGURES & TABLES

15. Related to my comment #7, Figure 1 could be improved to help describe the method. Figure 1 of Liu et al does a better job of showing how an additional diseased subject is included with the control group to compute PCCn+1. The “k” symbol is still used here, despite not being very helpful.

16. Bar plots in Figure 3A are missing y-axis labels, and scatter plots in 5A are missing both y-xais and x-axis labels. These make the graphs very difficult to interpret.
17. The Figure 4 caption contains the phrases “Degree centrality of nodes that influence long-term motor prognosis” and “Local efficiency of nodes that influence long-term motor prognosis.” Related to my comment #14, these also imply causality and should be changed.

CONCLUSION

No comment.

REFERENCES

No comment.