Patient-Specific Connectomic Models Correlate With, But Do Not Predict, Outcomes in Deep Brain Stimulation for Obsessive-Compulsive Disorder

Alik S. Widge[1,+], Fan Zhang[2], Aishwarya Gosai[3], George Papadimitrou[3], Peter Wilson-Braun[3], Magdalini Tsintou[3], Senthil Palanivelu[3], Angela M. Noecker[4], Cameron C. McIntyre[4], Lauren O'Donnell[2], Nicole C.R. McLaughlin[5], Benjamin D. Greenberg[5,6], Nikolaos Makris[3], Darin D. Dougherty[3,*], Yogesh Rathi[2,3,*]

1: Department of Psychiatry, University of Minnesota, Minneapolis, MN

2: Department of Radiology, Brigham and Womens Hospital, Boston, MA

3: Department of Psychiatry, Massachusetts General Hospital, Boston, MA

4: Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH

5: Department of Psychiatry and Human Behavior, Butler Hospital and Brown University Medical School, Providence, RI

6: Center for Neurorestoration and Neurotechnology, Providence VA Medical Center, Providence, RI

*: Equal contribution as senior authors

+: Correspondence to: MTRF 3-208, 2001 6th St SE, Minneapolis, MN 55455 ; awidge@umn.edu

Running Title: Patient-Specific Models in DBS for OCD

Keywords: Diffusion magnetic resonance imaging (dMRI); neurostimulation ; machine learning ; internal capsule ; electric field modeling

Abstract

Background: Deep brain stimulation (DBS) of the ventral internal capsule/ventral striatum (VCVS) is an emerging treatment for obsessive-compulsive disorder (OCD). Recently, multiple studies using normative connectomes have correlated DBS outcomes to stimulation of specific white matter tracts. Those studies did not test whether these correlations are clinically predictive, and did not apply cross-validation approaches that are necessary for biomarker development. Further, they did not account for the possibility of systematic differences between DBS patients and the non-diagnosed controls used in normative connectomes.

Methods: We performed patient-specific diffusion imaging in 8 patients who underwent VCVS DBS for OCD. We delineated tracts connecting thalamus and subthalamic nucleus (STN) to prefrontal cortex via VCVS. We then calculated which tracts were likely activated by individual patients' DBS settings. We fit multiple statistical models to predict both OCD and depression outcomes from tract activation. We further attempted to predict hypomania, a VCVS DBS complication. We assessed all models' performance on held-out test sets.

Results: No model predicted OCD response, depression response, or hypomania above chance. Coefficient inspection partly supported prior reports, in that capture of tracts projecting to cingulate cortex was associated with both YBOCS and MADRS response. In contrast to prior reports, however, tracts connected to STN were not reliably correlated with response.

Conclusions: Patient-specific imaging and a guideline-adherent analysis were unable to identify a tractographic target with sufficient effect size to drive clinical decision-making or predict individual outcomes. These findings suggest caution in interpreting the results of normative connectome studies.

Introduction

Deep brain stimulation (DBS) is an emerging approach to treatment-resistant mental disorders (1–3), but response rates in formal clinical trials are mixed (1,4–7). More reliable outcomes might be achieved by improving anatomic targeting. As psychiatric disorders are increasingly understood as network disorders (8,9), psychiatric DBS is moving away from using a single nucleus/structure as the target and towards attempts at affecting networks (1,10–12). There is particular enthusiasm for identifying target networks through diffusion tractography, which may enable DBS electrode placement to be customized to individual patients' anatomy. Although there is controversy over how accurately tractography reconstructs white matter anatomy (13,14), remarkable early results have been reported from DBS placement based on that imaging (10). Further, there are multiple tools available to model the interaction of DBS electric fields and targeted tracts (15–17). These tools could replace trial-and-error DBS programming with a mathematically optimal approach to activating desired pathways while minimizing off-target effects (18). That could overcome the difficulty of correctly programming stimulation, a likely driver of inconsistent clinical outcomes (1,4,19).

To realize that promise, we need to know which tracts should/should not be stimulated. For DBS of the subcallosal white matter for depression, multiple groups have settled on a specific white matter confluence and are studying it prospectively (with varying clinical outcomes (10,20)). For obsessive-compulsive disorder (OCD), a consensus may also be emerging. A theory linking OCD to dysfunction in cortico-striato-thalamic connectivity (21,22) has led to a focus on white matter tracts linking prefrontal cortex (PFC) to striatum, basal ganglia, and thalamus. Retrospective studies from multiple institutions have implicated tracts to/from dorsolateral PFC (23,24), ventrolateral PFC (12,25,26), and anterior cingulate (12,24) as potentially important in response. Recent analyses of patients implanted at two different targets correlated OCD response with a tract linking the ventral internal capsule/striatum (VCVS) and the subthalamic nucleus (STN) with the medial PFC (12,26–28). One study further suggested that capture of tracts from orbitofrontal cortex (OFC) (23) led to non-response, although a

qualitative synthesis (29) suggests that effective DBS tends to activate OFC-related fibers, and OFC-directed circuits can drive compulsive behaviors in animal models (30–32).

Although promising, these prior tractographic analyses are also limited. Many used standard atlases or connectomes derived from healthy controls, comparing these maps against electric fields from patient-specific DBS placements (12,23,27,28). Individual patients, however, show dramatic variation in their white matter topography compared to atlas standards (33). Targeting maps computed using "normative" connectomes differ from those computed from patient-specific DTI images (24). Other studies used simple isotropic field models (25), or distance between electrodes and a target tract (34) which may not accurately capture the DBS-induced electric field (16,35).

Most importantly, these analyses focused on tracts that correlate with clinical response. A variable may correlate strongly with an outcome but not be able to reliably predict that outcome, e.g. if the means are separate but the tails of two distributions overlap (36–38). Best practices in biomarker research suggest explicitly building predictive models, testing those models on held-out data, and reporting predictive performance in addition to correlation (36,37,39,40). Prediction-oriented analyses might better answer the question of whether a tractographic finding can be used as a programming target, i.e. whether it has strong predictive accuracy at the single-patient level (41).

Here, we address these limitations through an explicit attempt to predict single-patient response to DBS for OCD at the VCVS target, based on more precise field modeling approaches and using patient-specific tractography. We replicate in part prior studies' findings that cingulate, medial PFC, and lateral PFC tracts are correlated with clinical response, but we show that these correlations do not provide strong clinical predictive power, and in some cases we identify correlations that contradict earlier reports.

Methods

Study Population and Clinical Treatment

Participants were 6 patients who enrolled in a clinical trial (NCT00640133) of VCVS DBS for OCD (42), plus 2 who received VCVS DBS for OCD under a Humanitarian Device Exemption. All patients received Medtronic model 3387 DBS leads, with the most ventral contact targeted to the ventral striatal grey matter. The Institutional Review Boards of Massachusetts General Hospital and Butler Hospital approved the protocols and provided ethical oversight. All participants gave informed consent, explicitly including separate consent for DBS and for neuroimaging. We report here all patients who agreed to undergo imaging. We analyzed both the Yale-Brown Obsessive-Compulsive Scale (YBOCS) and Montgomery-Asberg Depression Rating Scale (MADRS), collected at visits approximately 2-4 weeks apart by a trained rater.

Imaging and Patient-Specific Tractography

Pre-operative MRI data were acquired on a 3T Siemens TimTrio scanner. Diffusion MRI (dMRI) scans had a spatial resolution of 2 mm (isotropic) with 10 non-diffusion weighted volumes and 60 diffusion weighted volumes, with gradient directions spread uniformly on the sphere with a b-value of 700 s/mm^2. dMRI data were registered to pre-operative T1- and T2-weighted MRI images and post-operative CT scans using a published pipeline (43) available at https://github.com/pnlbwh/. We then performed whole brain tractography from the dMRI data, using a multi-tensor unscented Kalman filter (UKF) (44,45). The UKF fits a mixture model of two tensors to the dMRI data, providing a highly sensitive fiber tracking ability in the presence of crossing fibers (46–49). The UKF method guides each fiber's current tracking estimate by the previous one. This recursive estimation helps stabilize model fitting, making tracking more robust to imaging artifact/noise. Another benefit of UKF is that fiber tracking orientation is controlled by a probabilistic prior about the rate of change of fiber orientation, producing more accurate tracking than the hard limits on curvature used in typical tractography algorithms. We combined the UKF with a fiber clustering algorithm to create an anatomically curated and annotated white matter

atlas (48). The clustering method groups the streamlines from each patient using a spectral embedding algorithm. Each fiber cluster is matched to a tract from an *a priori* labelled atlas of the white matter derived from known connections in monkey and human brains. Fiber clustering was performed only on streamlines longer than 40 mm to annotate medium and long range tracts.

Tract Activation Modeling

For each clinical DBS setting used in each patient, we calculated the volume of tissue activated (VTA) using a modified version of StimVision (15). Briefly, the VTAs were calculated using artificial neural network predictor functions, which were based on the response of multi-compartment cable models of axons coupled to finite element models of the DBS electric field (50). The VTAs used in this study were designed to estimate the spatial extent of activation for large diameter (5.7 μm) myelinated axons near the DBS electrode (51).

Based on theories that VCVS DBS acts by modulating circuits that run primarily in the internal capsule (14,22,29), we estimated activation of pathways linking thalamus with anterior cingulate and pericingulate cortex (ACC-PAC), dorsolateral PFC (dlPFC), ventrolateral PFC (vlPFC), dorsomedial PFC (DMPFC), medial orbitofrontal cortex (MOFC) and lateral OFC (LOFC). Pericingulate cortex includes rostral pre-cingulate cortex, but not the dorsal prefrontal cortex (such as the supplementary motor area). The atlas-guided fiber clustering algorithm (48) and a fiber clustering pipeline (52,53) guided manual delineation of fiber bundles connecting these regions to thalamus. All pathway labelings were performed by an expert neuroanatomist (Dr. Makris). Examples of the traced bundles and their intersections with DBS VTAs are shown in Figure 1A. Recent reports found that a tract connecting subthalamic nucleus (STN) to medial prefrontal cortex was strongly associated with clinical response to DBS in OCD (12,26–28). Therefore, we manually segmented the STN in each subject and extracted all fiber tracts connecting the STN with the prefrontal cortex (Figure 1B).

[Figure 1 about here]

Data Analysis - Independent/Predictor Variables

It is unclear whether the important “dose” of DBS is activation of a sufficient number of fibers (“total fiber” model), vs. the degree to which a sub-circuit is influenced (i.e., the fraction of the overall streamlines in a tract that are within the VTA, or a “percentage” model). We calculated both and fit them as two separate models for each dependent clinical outcome (see below). We also considered the possibility that DBS response is not determined by any individual tract/pathway, but instead requires capture of multiple pathways simultaneously. We therefore added a “total activation” variable to each prediction model. For total fiber models, this variable represented the total number of streamlines activated for all tracts. For percentage models, it represented the mean percentage activation across all reconstructed tracts. We standardized all input variables to the 0-1 interval to ensure that regression coefficients were comparable between independent variables.

All models were fit and evaluated using scikit-learn (0.24.1) in Python (3.8.5). With the exception of a necessary condition analysis described below, variables were coded at the single-visit level. That is, we predicted the clinical outcome at visit T from the DBS settings programmed at visit T-1.

Data Analysis - OCD Response

White matter pathway activation might relate tightly to the degree of clinical improvement (YBOCS as a continuous variable) or to patients’ overall well being (dichotomous responder/non-responder analysis). We thus modeled each separately. We analyzed continuous YBOCS as percentage decrease from baseline. Distribution fitting via the ‘fitdist’ package verified that YBOCS values were most compatible with a gamma distribution. We therefore predicted YBOCS improvement via an L1-regularized generalized linear regression (gamma distribution with identity link, Python package ‘pyglmnet’) and via

a random forest regression with 100 trees. The dependent variable was percentage improvement in YBOCS. We compared these two approaches to assess whether conclusions might be sensitive to the model formulation. Regularized regression emphasizes selection of a small number of highly leveraged variables, which may be more helpful in defining clinical decision rules. Random forests can outperform generalized linear regression in at least some cases (54), particularly where there are nonlinearities better captured by thresholding.

We further analyzed categorical (non)response, defined as a 35% or greater YBOCS decrease from baseline (42). For these, we compared an L1-regularized logistic regression and a random forest classifier with 100 trees. A minority of visits represented clinical response (29 visits out of 165, although 5 of 8 patients were in clinical response during at least one visit). To compensate for this imbalance, we applied the Synthetic Minority Oversampling Technique (SMOTE, (55)) with 3 nearest-neighbor examples. We chose L1 regularization for both regressions because dominant models of OCD argue that dysfunction in specific cortico-striatal loops leads to symptoms (21,22) and/or that a relatively small number of fiber bundles can explain response (12,26–28). This should be reflected in clinical response being driven a small subset of tracts.

Data Analysis - Depression Response

VCVS may have more effects on mood than on compulsivity (56), which would be reflected in better prediction of mood (MADRS) than of YBOCS. We applied the modeling pipeline used for categorical YBOCS response to categorical MADRS response, defined as a 50% or greater MADRS decrease from pre-surgical baseline. 7 out of the 165 visits met MADRS response criteria, although this again represented 5 of 8 patients.

We further assessed tractographic models' prediction of hypomania, a known and voltage-dependent complication of VCVS DBS (57,68); details are in the Supplement.

Data Analysis - Model Evaluation

All categorical data sets were unbalanced, and the outcome of clinical interest was always the minority class. We therefore report balanced accuracy and recall (performance for the minority class) for the categorical dependent variables. Further, we report the area under the receiver operator curve (AUC), which is suggested to be the best summary of a categorical biomarker's performance (36,39). For continuous YBOCS prediction, we report the fraction of variance explained and the coefficient of determination (R^2). We emphasize that R^2 here is not the square of a correlation coefficient (36).

All metrics were calculated on a held-out test set (36,37,39,40). For each model, we held out 2 random patients from the dataset (effectively 4-fold cross-validation with resampling). This improves over leave-one-out approaches, which can overstate predictive performance (58). We left out 25% of patients, rather than visits, because data were highly autocorrelated visit-to-visit, which also falsely inflates performance (36). We then fit the predictive model on the remaining 6 patients, and we report the performance on the visit-level data from the held out patients. To prevent data leakage, the SMOTE upsampling was performed on the training set only, after the split. We obtained confidence intervals for all metrics by repeating this process over all 28 possible leave-two-out combinations, then calculating the range of performance falling within 2 standard deviations of the median performance.

We fit 16 models (4 outcomes x 2 types of model x 2 ways of expressing activation), cross-validating within each model. We interpreted the outcomes using an uncorrected 95% confidence interval to maximize power.

Data Analysis - Predictor Importance

To detect potentially relevant tracts, we performed importance scoring on all models, regardless of whether they correctly predicted the clinical outcomes. For regression models, we computed the median and standard deviation of the regression coefficient for each tract, across all the train-test splits. For random forests, we applied permutation importance as implemented in scikit-learn. We permuted each independent variable 5 times for each of the train-test splits.

Data Analysis - Alternative Univariate Approach

Recent papers (12,26–28) used a different approach, based on comparison of VTAs to population-scale tractography. As an additional analysis (not pre planned), we attempted a similar approach on this dataset. We calculated all linear correlations between YBOCS improvement (continuous variable) and the activation of each individual tract (either as a total fiber or percentage activation). These correlations were performed on the training set after holding out 2 random patients, consistent with (12). To test whether this approach produced more generalizable predictors of DBS response, we used the same data to fit a univariate linear regression for each independent variable, then evaluated the model performance (coefficient of determination, R^2) on the 2 held out patients.

In a further exploratory analysis (see Supplement), we considered whether DBS outcomes depended not on the tracts activated, but the integrity of those tracts.

Results

Clinical Outcomes - YBOCS

The mean YBOCS improvement (considering each patient's best time point) was 46.6%, and 5 of the 8 patients (62.5%) were clinical responders (≥35% YBOCS drop) for at least one visit.

No tract reliably predicted continuous YBOCS improvement. By all metrics, model performance was worse than chance on the held-out test set (Table 1), for both total-activation and percentage-activation models. Consistent with this, no coefficients in the regression models were above zero (i.e., the dataset mean was more reliable than any tractographic predictor). In the random forest models, the highest importance was percentage activation of fibers connecting thalamus to left OFC, but this was at chance level (change in R^2 across models: mean 0.09, SD 0.24).

[Table 1 about here]

Similarly, no model exceeded chance for response/nonresponse prediction (Table 2). In the logistic regression, highly weighted features across models were the number (but not percentage) of activated streamlines connecting thalamus to left cingulate, lateral OFC, medial OFC, and vlPFC. Cingulate and lateral OFC streamline activation were positively associated with response, whereas medial OFC and vlPFC activation were negatively associated (Figure 2). For all of these tracts, the confidence interval for the coefficient estimated across all train-test splits included 0. These findings were sensitive to the modeling approach; the same tracts did not show median importance scores different from 0 in the random forest models. The ACC-PAC findings were corroborated by a Necessary Condition Analysis on white matter integrity (Supplementary Results).

[Table 2 and Figure 2 about here]

The alternate mass-univariate approach also did not reliably predict response on the held-out test sets (Table 3). It was concordant with the categorical response analysis in that it identified streamlines connecting the left cingulate to thalamus as correlated with response, and similarly streamlines from bilateral vlPFC as correlated with non-response. There was more discordance than similarity, however. The medial OFC tracts identified by regression were not selected in the mass univariate approach, and conversely, the mass univariate approach predicted nonresponse if tracts projecting to dlPFC were

within the VTA. Further, the mass univariate approach emphasized percentage capture, while the logistic regression emphasized total fibers within a VTA. We note that tracts from STN to PFC were negatively correlated with clinical outcomes, whereas prior reports identify them as positively correlated (12,27,28).

Clinical Outcomes - MADRS

The mean MADRS improvement (considering each patient's best time point) was 55.69%, and 5 of the 8 (62.5%) were responders (≥50% MADRS drop) at some point. Mood and OCD response were not linked (r=0.13 for correlation between response status on YBOCS and MADRS). Consistent with other reports (56), there were more observations of MADRS response without YBOCS than of YBOCS response without MADRS (22 vs. 4).

No model reliably predicted MADRS response above chance (Table 4). For comparison with the YBOCS analysis, we further examined the non-zero coefficients of the total-fiber regression. Capture of streamlines between right cingulate and thalamus was correlated with MADRS response, and the confidence interval for this coefficient excluded zero (Figure 3). This was not true of any other tract. Left vlPFC was associated with non-response (as it was in the categorical YBOCS analysis), but the distribution of coefficients across analyses included zero. Random forest importance scores were centered around zero.

[Table 4 and Figure 3 about here]

Discussion

Our results are both concordant and discordant with prior efforts to predict clinical OCD DBS response from tractographic modeling of cortico-striatal and cortico-basal circuits. Critically, we implemented

multiple analytic steps beyond prior studies: individualized, patient-specific tracts registered to individual lead placements, activation volume calculation beyond simple electric field assumptions, consideration of multiple clinical timepoints for each patient, and formal evaluation of predictive power (as compared to measurement of correlations between activation and response or group mean differences). With this more guideline-adherent approach, we found that no tract could reliably predict clinical response or complications, whether those were considered in a continuous or categorical approach. This is likely not a surprise – we and others have highlighted that group-level significant correlations/separations often do not have clinical predictive power (36–39). In this sense, our results support calls for caution regarding the clinical role of tractography (16,41). We also showed that outcomes can be sensitive to the analytic approach – our random forest and regularized regression approaches produced very different results, even though both are commonly used approaches to prediction and variable selection.

Model inspection may offer some insight into variables for further investigation, even if pathway activation modeling approaches are not yet able to strongly predict response. Numerically, predictive power was greater (more non-zero regression coefficients after regularization) when predicting categorical rather than continuous outcomes. This may be because categorical outcomes effectively smooth out small fluctuations in continuous rating scales, fluctuations that may be primarily due to inter-rater variability or disease-unrelated variables rather than to DBS settings. The YBOCS in particular shows non-linear behavior at high scores that may exacerbate this (59). We obtained non-zero regression coefficients for models using activated fiber counts, but not for percentage-activated models, implying that it is more important to get at least a portion of a key tract within the VTA. These results also make sense in the context of our finding that the integrity (traceability) of these tracts varies greatly between patients with OCD -- a tract where response depends on tract integrity will have a large coefficient in a total-fibers model, but not in a percentage-activation model.

Our results in part support and in part diverge from a series of recent papers implicating pathways between PFC and basal ganglia as critical for OCD DBS (12,26–28). Consistent with that work, the ACC/PAC to thalamus tracts were implicated in both YBOCS and MADRS response, and were the most positively weighted in our mass-univariate approach. Our white matter integrity analysis identified the same tracts as having the largest effect size (necessity). Also similar to that prior work, we found that activation of connections to medial OFC produced numerically worse outcomes. Inconsistent with the prior work (12,26–28), we found negative correlations (in the mass univariate analysis) or null effects (in the predictive models) specifically for tracts connecting PFC to STN or vlPFC to thalamus. This again may reflect the importance of patient-specific imaging. Given that we have previously shown these tracts to have substantial inter-individual variability in their position within the internal capsule (33), and that here we note them to have similar variability in their overall integrity, a normative connectomic analysis may not reflect the actual fibers being successfully modulated in DBS cases. Alternatively, our results may highlight programming and surgical differences. These patients were implanted and programmed following the approach in (60), which emphasizes an initial search for a positive affective response. Other centers have reported very different programming algorithms (61), based more on standard anatomic positions. If response correlates with, e.g., the quality of concomitant therapy (26,62) or general clinical expertise (63), those factors will likely be strongly correlated with the programming clinician, and thus will spuriously load onto the tracts and implant locations that clinician happens to prefer. Most importantly, our results highlight the importance of applying analyses designed specifically to identify clinical predictors (36). Interestingly, we found that OFC engagement predicted worse OCD clinical response. OFC-originating components of cortico-striato-thalamic circuits are heavily emphasized in theoretical (21,22,29) and animal (30,32,64) models of OCD, and these findings may contribute to an ongoing debate over those models.

These results are tempered by three limitations. First, our sample size is small, consistent with the rarity of these patients (65). Second, imaging was not performed on a connectome-optimized scanner, and scanning at 7 Tesla (as has now become more common (66)) might identify more tracts. Third, we used

relatively simple models of DBS activation. All of these add noise, reducing our ability to detect subtle correlations, particularly given DTI's susceptibility to false positives (14). Practically, however, these limitations may not affect the clinical importance of our findings. We mitigated the lower resolution of these scans by use of an algorithm that is specifically designed to perform well in the presence of noise (45) and ensuring that our extracted tracts matched known, anatomically verified fiber bundles (48). Further, small sample sizes tend to inflate effect sizes and bias towards positive conclusions (67), not the negative result we report. Most importantly, for a tractographic result to be sufficiently reliable to inform clinical targeting/programming, it would need to have a large and clear influence on outcomes, with robustness to minor variations in analytic or clinical technique. Such a large effect would be clearly detectable and consistent across studies even at small sample sizes, like the clinical effect of VCVS DBS, which shows consistent 60-70% response rates across many small to medium cohorts (56,68–71). In that context, failure to identify a significant predictor in this small sample is relevant to both clinical practice and future study design.

Overall, our results support a growing argument that circuits linking ACC to thalamus and basal ganglia are important to VCVS DBS response. They dovetail with other work linking modulation of those circuits to increased cognitive control (72,73), a construct that is thought to be deficient in OCD (74,75). At the same time, they highlight that the current level of tractographic understanding does not have strong clinical predictive power, and that multiple confounds remain to be controlled/addressed. With multiple technologies emerging to better verify target engagement and address patient heterogeneity (1,16), that understanding will likely grow in coming years.

Data/Code Availability

De-identified data tables and analysis code used to produce all exhibits in this manuscript will be available at the time of publication at https://github.com/tne-lab .

Acknowledgements

This work was supported by NIH grants R03MH111320 (ASW, CCM, DDD), UH3NS100548 (ASW, DDD), R01MH111917 (ASW, DDD, NMM, YR), U01U01MH076179 (BDG), P20GM130452 (NCM, BDG), K23MH100607 (NCM), and P50MH106435 (NCM, BDG). ASW further acknowledges support from the MnDRIVE Brain Conditions and Minnesota Medical Discovery Team – Addictions initiatives. All opinions and conclusions herein are those of the authors. They do not represent the views or official policy of any public or private funding body.

Disclosures

ASW and DDD have received research funding, device donations, and honoraria from Medtronic, which manufactured the DBS devices used in patients' clinical care. Medtronic had no financial or technical role in this study. ASW and CCM have multiple patents and patent filings in the area of deep brain stimulation, including methods for optimizing/customizing stimulation parameters. CCM is a paid consultant for Boston Scientific Neuromodulation, receives royalties from Hologram Consultants, Neuros Medical, Qr8 Health, and is a shareholder in the following companies: Hologram Consultants, Surgical Information Sciences, CereGate, Autonomic Technologies, Cardionomic, and Enspire DBS. All other authors affirm no related financial interests.

References

1. Sullivan CRP, Olsen S, Widge AS (2021): Deep brain stimulation for psychiatric disorders: From focal brain targets to cognitive networks. *NeuroImage* 225: 117515.
2. Bari AA, Mikell CB, Abosch A, Ben-Haim S, Buchanan RJ, Burton AW, *et al.* (2018): Charting the road forward in psychiatric neurosurgery: proceedings of the 2016 American Society for Stereotactic and Functional Neurosurgery workshop on neuromodulation for psychiatric disorders. *J Neurol Neurosurg Psychiatry* jnnp-2017-317082.

3. Ramirez-Zamora A, Giordano J, Gunduz A, Alcantara J, Cagle JN, Cernera S, *et al.* (2020): Proceedings of the Seventh Annual Deep Brain Stimulation Think Tank: Advances in Neurophysiology, Adaptive DBS, Virtual Reality, Neuroethics and Technology. *Front Hum Neurosci* 14: 54.

4. Widge AS, Malone DAJ, Dougherty DD (2018): Closing the loop on deep brain stimulation for treatment-resistant depression. *Front Neurosci* 12: 175.

5. Luyten L, Hendrickx S, Raymaekers S, Gabriëls L, Nuttin B (2016): Electrical stimulation in the bed nucleus of the stria terminalis alleviates severe obsessive-compulsive disorder. *Mol Psychiatry* 21: 1272–1280.

6. Holtzheimer PE, Husain MM, Lisanby SH, Taylor SF, Whitworth LA, McClintock S, *et al.* (2017): Subcallosal cingulate deep brain stimulation for treatment-resistant depression: a multisite, randomised, sham-controlled trial. *Lancet Psychiatry* 4: 839–849.

7. Dougherty DD, Rezai AR, Carpenter LL, Howland RH, Bhati MT, O'Reardon JP, *et al.* (2015): A randomized sham-controlled trial of deep brain stimulation of the ventral capsule/ventral striatum for chronic treatment-resistant depression. *Biol Psychiatry* 78: 240–248.

8. Williams LM (2016): Defining biotypes for depression and anxiety based on large-scale circuit dysfunction: a theoretical review of the evidence and future directions for clinical translation. *Depress Anxiety* 34: 9–24.

9. Avena-Koenigsberger A, Misic B, Sporns O (2018): Communication dynamics in complex brain networks. *Nat Rev Neurosci* 19: 17.

10. Riva-Posse P, Choi KS, Holtzheimer PE, Crowell AL, Garlow SJ, Rajendra JK, *et al.* (2018): A connectomic approach for subcallosal cingulate deep brain stimulation surgery: prospective targeting in treatment-resistant depression. *Mol Psychiatry* 23: 843–849.

11. Karas PJ, Lee S, Jimenez-Shahed J, Goodman WK, Viswanathan A, Sheth SA (2019): Deep brain stimulation for obsessive compulsive disorder: evolution of surgical stimulation target parallels changing model of dysfunctional brain circuits. *Front Neurosci* 12: 998.

12. Li N, Baldermann JC, Kibleur A, Treu S, Akram H, Elias GJB, *et al.* (2020): A unified connectomic target for deep brain stimulation in obsessive-compulsive disorder [no. 1]. *Nat Commun* 11: 3364.
13. Maier-Hein KH, Neher PF, Houde J-C, Côté M-A, Garyfallidis E, Zhong J, *et al.* (2017): The challenge of mapping the human connectome based on diffusion tractography [no. 1]. *Nat Commun* 8: 1349.
14. Haber SN, Tang W, Choi EY, Yendiki A, Liu H, Jbabdi S, *et al.* (2020): Circuits, networks, and neuropsychiatric disease: transitioning from anatomy to imaging. *Biol Psychiatry* 87: 318–327.
15. Noecker AM, Choi KS, Riva-Posse P, Gross RE, Mayberg HS, McIntyre CC (2018): StimVision software: examples and applications in subcallosal cingulate deep brain stimulation for depression. *Neuromodulation Technol Neural Interface* 21: 191–196.
16. Noecker AM, Frankemolle-Gilbert AM, Howell B, Petersen MV, Beylergil SB, Shaikh AG, McIntyre CC (2021): StimVision v2: examples and applications in subthalamic deep brain stimulation for parkinson's disease. *Neuromodulation Technol Neural Interface* 24: 248–258.
17. Horn A, Kühn AA (2015): Lead-DBS: A toolbox for deep brain stimulation electrode localizations and visualizations. *NeuroImage* 107: 127–135.
18. Peña E, Zhang S, Patriat R, Aman JE, Vitek JL, Harel N, Johnson MD (2018): Multi-objective particle swarm optimization for postoperative deep brain stimulation targeting of subthalamic nucleus pathways. *J Neural Eng* 15: 066020.
19. Provenza NR, Matteson ER, Allawala AB, Barrios-Anderson A, Sheth SA, Viswanathan A, *et al.* (2019): The case for adaptive neuromodulation to treat severe intractable mental disorders. *Front Neurosci* 13. https://doi.org/10.3389/fnins.2019.00152
20. Ramasubbu R, Clark DL, Golding S, Dobson KS, Mackie A, Haffenden A, Kiss ZH (2020): Long versus short pulse width subcallosal cingulate stimulation for treatment-resistant depression: a randomised, double-blind, crossover trial. *Lancet Psychiatry* 7: 29–40.
21. Robbins TW, Vaghi MM, Banca P (2019): Obsessive-compulsive disorder: puzzles and prospects. *Neuron* 102: 27–47.

22. Dougherty DD, Brennan B, Stewart SE, Wilhelm S, Widge AS, Rauch SL (2018): Neuroscientifically informed formulation and treatment planning for patients with obsessive-compulsive disorder: a review. *JAMA Psychiatry* 75: 1081–1087.

23. Hartmann CJ, Lujan JL, Chaturvedi A, Goodman WK, Okun MS, McIntyre CC, Haq IU (2016): Tractography activation patterns in dorsolateral prefrontal cortex suggest better clinical responses in OCD DBS. *Neuroprosthetics* 519.

24. Baldermann JC, Melzer C, Zapf A, Kohl S, Timmermann L, Tittgemeyer M, *et al.* (2019): Connectivity profile predictive of effective deep brain stimulation in obsessive-compulsive disorder. *Biol Psychiatry* 85: 735–743.

25. Barcia JA, Avecillas-Chasín JM, Nombela C, Arza R, Albea JG-, Pineda J-Á, *et al.* (2018): Personalized striatal targets for deep brain stimulation in obsessive-compulsive disorder. *Brain Stimulat*. https://doi.org/10.1016/j.brs.2018.12.226

26. Mosley PE, Windels F, Morris J, Coyne T, Marsh R, Giorni A, *et al.* (2021): A randomised, double-blind, sham-controlled trial of deep brain stimulation of the bed nucleus of the stria terminalis for treatment-resistant obsessive-compulsive disorder [no. 1]. *Transl Psychiatry* 11: 1–17.

27. Smith AH, Choi KS, Waters AC, Aloysi A, Mayberg HS, Kopell BH, Figee M (2021): Replicable effects of deep brain stimulation for obsessive-compulsive disorder. *Brain Stimulat* 14: 1–3.

28. Bouwens van der Vlis TAM, Ackermans L, Mulders AEP, Vrij CA, Schruers K, Temel Y, *et al.* (n.d.): Ventral capsule/ventral striatum stimulation in obsessive-compulsive disorder: toward a unified connectomic target for deep brain stimulation? *Neuromodulation Technol Neural Interface* n/a. https://doi.org/10.1111/ner.13339

29. Haber SN, Yendiki A, Jbabdi S (2020): Four deep brain stimulation targets for obsessive-compulsive disorder: Are they different? *Biol Psychiatry* S000632232031773X.

30. Wood J, Ahmari SE (2015): A framework for understanding the emerging role of corticolimbic-ventral striatal networks in OCD-associated repetitive behaviors. *Front Syst Neurosci* 171.

31. Burguière E, Monteiro P, Feng G, Graybiel AM (2013): Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science* 340: 1243–1246.

32. Heilbronner SR, Rodriguez-Romaguera J, Quirk GJ, Groenewegen HJ, Haber SN (2016): Circuit based cortico-striatal homologies between rat and primate. *Biol Psychiatry* 80: 509–521.

33. Makris N, Rathi Y, Mouradian P, Bonmassar G, Papadimitriou G, Ing WI, *et al.* (2016): Variability and anatomical specificity of the orbitofrontothalamic fibers of passage in the ventral capsule/ventral striatum (VC/VS): precision care for patient-specific tractography-guided targeting of deep brain stimulation (DBS) in obsessive compulsive disorder (OCD). *Brain Imaging Behav* 10: 1054–1067.

34. Liebrand LC, Caan MWA, Schuurman PR, van den Munckhof P, Figee M, Denys D, van Wingen GA (2018): Individual white matter bundle trajectories are associated with deep brain stimulation response in obsessive-compulsive disorder. *Brain Stimulat*. https://doi.org/10.1016/j.brs.2018.11.014

35. Howell B, McIntyre CC (2016): Analyzing the tradeoff between electrical complexity and accuracy in patient-specific computational models of deep brain stimulation. *J Neural Eng* 13: 036023.

36. Poldrack RA, Huckins G, Varoquaux G (2020): Establishment of best practices for evidence for prediction: a review. *JAMA Psychiatry* 77: 534–540.

37. Grzenda A, Kraguljac NV, McDonald WM, Nemeroff CB, Torous J, Alpert JE, *et al.* (Accepted): Evaluating the machine learning literature: a primer and user's guide for psychiatrists. *Am J Psychiatry*.

38. Lo A, Chernoff H, Zheng T, Lo S-H (2015): Why significant variables aren't automatically good predictors. *Proc Natl Acad Sci* 112: 13892–13897.

39. Widge AS, Bilge MT, Montana R, Chang W, Rodriguez CI, Deckersbach T, *et al.* (2019): Electroencephalographic biomarkers for treatment response prediction in major depressive illness: a meta-analysis. *Am J Psychiatry* 176: 44–56.

40. Woo C-W, Chang LJ, Lindquist MA, Wager TD (2017): Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci* 20: 365–377.

41. Coenen VA, Schlaepfer TE, Varkuti B, Schuurman PR, Reinacher PC, Voges J, *et al.* (2019): Surgical decision making for deep brain stimulation should not be based on aggregated normative data mining. *Brain Stimul Basic Transl Clin Res Neuromodulation* 12: 1345–1348.

42. McLaughlin N, Dougherty DD, Eskandar EN, Ward HE, Foote KD, Malone Jr. DA, *et al.* (n.d.): Double blind randomized controlled trial of deep brain stimulation for obsessive-compulsive disorder: clinical trial design. *Contemp Clin Trials*.

43. reckbo, Tashrif Billah, Isaiah Norton (2019): *Pnlbwh/Pnlpipe: Easy Install and Multiprocessing*. Zenodo. https://doi.org/10.5281/zenodo.3270927

44. Malcolm JG, Shenton ME, Rathi Y (2010): Filtered multitensor tractography [no. 9]. *IEEE Trans Med Imaging* 29: 1664–1675.

45. Reddy CP, Rathi Y (2016): Joint Multi-Fiber NODDI Parameter Estimation and Tractography Using the Unscented Information Filter. *Front Neurosci* 10. https://doi.org/10.3389/fnins.2016.00166

46. Gong S, Zhang F, Norton I, Essayed WI, Unadkat P, Rigolo L, *et al.* (2018): Free water modeling of peritumoral edema using multi-fiber tractography: Application to tracking the arcuate fasciculus for neurosurgical planning [no. 5]. *PLoS ONE* 13. https://doi.org/10.1371/journal.pone.0197056

47. Liao R, Ning L, Chen Z, Rigolo L, Gong S, Pasternak O, *et al.* (2017): Performance of unscented Kalman filter tractography in edema: Analysis of the two-tensor model. *NeuroImage Clin* 15: 819–831.

48. Zhang F, Wu Y, Norton I, Rigolo L, Rathi Y, Makris N, O'Donnell LJ (2018): An anatomically curated fiber clustering white matter atlas for consistent white matter tract parcellation across the lifespan. *NeuroImage* 179: 429–447.

49. Chen Z, Tie Y, Olubiyi O, Rigolo L, Mehrtash A, Norton I, *et al.* (2015): Reconstruction of the arcuate fasciculus for surgical planning in the setting of peritumoral edema using two-tensor unscented Kalman filter tractography. *NeuroImage Clin* 7: 815–822.

50. Chaturvedi A, Luján JL, McIntyre CC (2013): Artificial neural network based characterization of the volume of tissue activated during deep brain stimulation. *J Neural Eng* 10: 056023.

51. Butson CR, McIntyre CC (2006): Role of electrode design on the volume of tissue activated during deep brain stimulation. *J Neural Eng* 3: 1–8.

52. O'Donnell LJ, Westin C-F (2007): Automatic tractography segmentation using a high-dimensional white matter atlas [no. 11]. *IEEE Trans Med Imaging* 26: 1562–1575.

53. O'Donnell LJ, Wells WM, Golby AJ, Westin C-F (2012): Unbiased groupwise registration of white matter tractography. In: Ayache N, Delingette H, Golland P, Mori K, editors. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, vol. 7512. Berlin, Heidelberg: Springer Berlin Heidelberg, pp 123–130.

54. Couronné R, Probst P, Boulesteix A-L (2018): Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinformatics* 19: 270.

55. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002): SMOTE: Synthetic Minority Over-sampling Technique. *J Artif Intell Res* 16: 321–357.

56. Tyagi H, Apergis-Schoute AM, Akram H, Foltynie T, Limousin P, Drummond LM, *et al.* (2019): A randomised trial directly comparing ventral capsule and anteromedial subthalamic nucleus stimulation in obsessive compulsive disorder: clinical and imaging evidence for dissociable effects. *Biol Psychiatry* 85: 726–734.

57. Widge AS, Licon E, Zorowitz S, Corse A, Arulpragasam AR, Camprodon JA, *et al.* (2015): Predictors of hypomania during ventral capsule/ventral striatum deep brain stimulation. *J Neuropsychiatry Clin Neurosci* 28: 38–44.

58. Smith GCS, Seaman SR, Wood AM, Royston P, White IR (2014): Correcting for optimistic prediction in small data sets. *Am J Epidemiol* 180: 318–324.

59. Storch EA, Rasmussen SA, Price LH, Larson MJ, Murphy TK, Goodman WK (2010): Development and psychometric evaluation of the Yale–Brown Obsessive-Compulsive Scale—Second Edition. *Psychol Assess* 22: 223.

60. Widge AS, Dougherty DD (2015): Managing patients with psychiatric disorders with deep brain stimulation. In: Marks Jr. WJ, editor. *Deep Brain Stimulation Management*, 2nd ed. Cambridge : New York: Cambridge University Press.

61. van Westen M, Rietveld E, Bergfeld IO, Koning P de, Vullink N, Ooms P, *et al.* (2021): Optimizing deep brain stimulation parameters in obsessive–compulsive disorder. *Neuromodulation Technol Neural Interface* 24: 307–315.
62. Mantione M, Nieman DH, Figee M, Denys D (2014): Cognitive–behavioural therapy augments the effects of deep brain stimulation in obsessive–compulsive disorder. *Psychol Med* 44: 3515–3522.
63. van Westen M, Rietveld E, Denys D (2019): Effective deep brain stimulation for obsessive-compulsive disorder requires clinical expertise. *Front Psychol* 10: 2294.
64. Burguière E, Monteiro P, Mallet L, Feng G, Graybiel AM (2015): Striatal circuits, habits, and implications for obsessive–compulsive disorder. *Curr Opin Neurobiol* 30: 59–65.
65. Garnaat SL, Greenberg BD, Sibrava NJ, Goodman WK, Mancebo MC, Eisen JL, Rasmussen SA (2014): Who qualifies for deep brain stimulation for OCD? Data from a naturalistic clinical sample. *J Neuropsychiatry Clin Neurosci* 26: 81–86.
66. Duchin Y, Shamir RR, Patriat R, Kim J, Vitek JL, Sapiro G, Harel N (2018): Patient-specific anatomical model for deep brain stimulation based on 7 Tesla MRI. *PloS One* 13: e0201469.
67. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR (2013): Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 14: 365–376.
68. Denys D, Graat I, Mocking R, de Koning P, Vulink N, Figee M, *et al.* (2020): Efficacy of deep brain stimulation of the ventral anterior limb of the internal capsule for refractory obsessive-compulsive disorder: a clinical cohort of 70 patients. *Am J Psychiatry* appi.ajp.2019.19060656.
69. Menchón JM, Real E, Alonso P, Aparicio MA, Segalas C, Plans G, *et al.* (2021): A prospective international multi-center study on safety and efficacy of deep brain stimulation for resistant obsessive-compulsive disorder [no. 4]. *Mol Psychiatry* 26: 1234–1247.
70. Greenberg B, Gabriels L, Malone D, Rezai A, Friehs G, Okun M, *et al.* (2010): Deep brain stimulation of the ventral internal capsule/ventral striatum for obsessive-compulsive disorder: worldwide experience. *Mol Psychiatry* 15: 64–79.

71. Nuttin B, Cosyns P, Demeulemeester H, Gybels J, Meyerson B (1999): Electrical stimulation in anterior limbs of internal capsules in patients with obsessive-compulsive disorder. *The Lancet* 354: 1526.

72. Basu I, Yousefi A, Crocker B, Zelmann R, Paulk AC, Peled N, *et al.* (Accepted in principle): Closed loop enhancement and neural decoding of human cognitive control. *Nat Biomed Eng*. https://doi.org/10.1101/2020.04.24.059964

73. Widge AS, Zorowitz S, Basu I, Paulk AC, Cash SS, Eskandar EN, *et al.* (2019): Deep brain stimulation of the internal capsule enhances human cognitive control and prefrontal cortex function. *Nat Commun* 10: 1536.

74. Gruner P, Pittenger C (2017): Cognitive inflexibility in Obsessive-Compulsive Disorder. *Neuroscience* 345: 243–255.

75. Abramovitch A, Abramowitz JS, Mittelman A (2013): The neuropsychology of adult obsessive–compulsive disorder: A meta-analysis. *Clin Psychol Rev* 33: 1163–1171.

76. Brennan BP, Jacoby RJ, Widge AS (2018): A case of severe intractable contamination-based obsessive-compulsive disorder. *JAMA Psychiatry* 75: 1088–1089.

77. Smith SM (2002): Fast robust automated brain extraction [no. 3]. *Hum Brain Mapp* 17: 143–155.

78. Jenkinson M, Pechaud M, Smith S (2005): BET2 - MR-Based Estimation of Brain, Skull and Scalp Surfaces. presented at the Eleventh Annual Meeting of the Organization for Human Brain Mapping. Retrieved from http://mickaelpechaud.free.fr/these/HBM05.pdf

79. *Advanced Normalization Tools (ANTs)* (2021): Advanced Normalization Tools Ecosystem. Retrieved April 1, 2021, from https://github.com/ANTsX/ANTs

80. Howell B, Choi KS, Gunalan K, Rajendra J, Mayberg HS, McIntyre CC (2019): Quantifying the axonal pathways directly stimulated in therapeutic subcallosal cingulate deep brain stimulation. *Hum Brain Mapp* 40: 889–903.

81. Dul J (2016): Necessary Condition Analysis (NCA): logic and methodology of “necessary but not sufficient” causality. *Organ Res Methods* 19: 10–52.

82. Dul J (2021): *NCA: Necessary Condition Analysis*, version 3.1.0. Retrieved April 1, 2021, from https://CRAN.R-project.org/package=NCA

2. Bari AA, Mikell CB, Abosch A, Ben-Haim S, Buchanan RJ, Burton AW, *et al.* (2018): Charting the road forward in psychiatric neurosurgery: proceedings of the 2016 American Society for Stereotactic and Functional Neurosurgery workshop on neuromodulation for psychiatric disorders. *J Neurol Neurosurg Psychiatry* jnnp-2017-317082.

3. Ramirez-Zamora A, Giordano J, Gunduz A, Alcantara J, Cagle JN, Cernera S, *et al.* (2020): Proceedings of the Seventh Annual Deep Brain Stimulation Think Tank: Advances in Neurophysiology, Adaptive DBS, Virtual Reality, Neuroethics and Technology. *Front Hum Neurosci* 14: 54.

4. Widge AS, Malone DAJ, Dougherty DD (2018): Closing the loop on deep brain stimulation for treatment-resistant depression. *Front Neurosci* 12: 175.

5. reckbo, Tashrif Billah, Isaiah Norton (2019): *Pnlbwh/Pnlpipe: Easy Install and Multiprocessing*. Zenodo. https://doi.org/10.5281/zenodo.3270927

6. Luyten L, Hendrickx S, Raymaekers S, Gabriëls L, Nuttin B (2016): Electrical stimulation in the bed nucleus of the stria terminalis alleviates severe obsessive-compulsive disorder. *Mol Psychiatry* 21: 1272–1280.

7. Widge AS, Licon E, Zorowitz S, Corse A, Arulpragasam AR, Camprodon JA, *et al.* (2015): Predictors of hypomania during ventral capsule/ventral striatum deep brain stimulation. *J Neuropsychiatry Clin Neurosci* 28: 38–44.

8. Holtzheimer PE, Husain MM, Lisanby SH, Taylor SF, Whitworth LA, McClintock S, *et al.* (2017): Subcallosal cingulate deep brain stimulation for treatment-resistant depression: a multisite, randomised, sham-controlled trial. *Lancet Psychiatry* 4: 839–849.

9. Denys D, Graat I, Mocking R, de Koning P, Vulink N, Figee M, *et al.* (2020): Efficacy of deep brain stimulation of the ventral anterior limb of the internal capsule for refractory obsessive-compulsive disorder: a clinical cohort of 70 patients. *Am J Psychiatry* appi.ajp.2019.19060656.

10. Dougherty DD, Rezai AR, Carpenter LL, Howland RH, Bhati MT, O'Reardon JP, *et al.* (2015): A randomized sham-controlled trial of deep brain stimulation of the ventral capsule/ventral striatum for chronic treatment-resistant depression. *Biol Psychiatry* 78: 240–248.

11. Williams LM (2016): Defining biotypes for depression and anxiety based on large-scale circuit dysfunction: a theoretical review of the evidence and future directions for clinical translation. *Depress Anxiety* 34: 9–24.

12. Avena-Koenigsberger A, Misic B, Sporns O (2018): Communication dynamics in complex brain networks. *Nat Rev Neurosci* 19: 17.

13. Riva-Posse P, Choi KS, Holtzheimer PE, Crowell AL, Garlow SJ, Rajendra JK, *et al.* (2018): A connectomic approach for subcallosal cingulate deep brain stimulation surgery: prospective targeting in treatment-resistant depression. *Mol Psychiatry* 23: 843–849.

14. Karas PJ, Lee S, Jimenez-Shahed J, Goodman WK, Viswanathan A, Sheth SA (2019): Deep brain stimulation for obsessive compulsive disorder: evolution of surgical stimulation target parallels changing model of dysfunctional brain circuits. *Front Neurosci* 12: 998.

15. Li N, Baldermann JC, Kibleur A, Treu S, Akram H, Elias GJB, *et al.* (2020): A unified connectomic target for deep brain stimulation in obsessive-compulsive disorder [no. 1]. *Nat Commun* 11: 3364.

16. Maier-Hein KH, Neher PF, Houde J-C, Côté M-A, Garyfallidis E, Zhong J, *et al.* (2017): The challenge of mapping the human connectome based on diffusion tractography [no. 1]. *Nat Commun* 8: 1349.

17. Haber SN, Tang W, Choi EY, Yendiki A, Liu H, Jbabdi S, *et al.* (2020): Circuits, networks, and neuropsychiatric disease: transitioning from anatomy to imaging. *Biol Psychiatry* 87: 318–327.

18. Noecker AM, Choi KS, Riva-Posse P, Gross RE, Mayberg HS, McIntyre CC (2018): StimVision software: examples and applications in subcallosal cingulate deep brain stimulation for depression. *Neuromodulation Technol Neural Interface* 21: 191–196.

19. Noecker AM, Frankemolle-Gilbert AM, Howell B, Petersen MV, Beylergil SB, Shaikh AG, McIntyre CC (2021): StimVision v2: examples and applications in subthalamic deep brain stimulation for parkinson's disease. *Neuromodulation Technol Neural Interface* 24: 248–258.

20. Horn A, Kühn AA (2015): Lead-DBS: A toolbox for deep brain stimulation electrode localizations and visualizations. *NeuroImage* 107: 127–135.

21. Peña E, Zhang S, Patriat R, Aman JE, Vitek JL, Harel N, Johnson MD (2018): Multi-objective particle swarm optimization for postoperative deep brain stimulation targeting of subthalamic nucleus pathways. *J Neural Eng* 15: 066020.

22. Provenza NR, Matteson ER, Allawala AB, Barrios-Anderson A, Sheth SA, Viswanathan A, *et al.* (2019): The case for adaptive neuromodulation to treat severe intractable mental disorders. *Front Neurosci* 13. https://doi.org/10.3389/fnins.2019.00152

23. Ramasubbu R, Clark DL, Golding S, Dobson KS, Mackie A, Haffenden A, Kiss ZH (2020): Long versus short pulse width subcallosal cingulate stimulation for treatment-resistant depression: a randomised, double-blind, crossover trial. *Lancet Psychiatry* 7: 29–40.

24. Robbins TW, Vaghi MM, Banca P (2019): Obsessive-compulsive disorder: puzzles and prospects. *Neuron* 102: 27–47.

25. Dougherty DD, Brennan B, Stewart SE, Wilhelm S, Widge AS, Rauch SL (2018): Neuroscientifically informed formulation and treatment planning for patients with obsessive-compulsive disorder: a review. *JAMA Psychiatry* 75: 1081–1087.

26. Hartmann CJ, Lujan JL, Chaturvedi A, Goodman WK, Okun MS, McIntyre CC, Haq IU (2016): Tractography activation patterns in dorsolateral prefrontal cortex suggest better clinical responses in OCD DBS. *Neuroprosthetics* 519.

27. Baldermann JC, Melzer C, Zapf A, Kohl S, Timmermann L, Tittgemeyer M, *et al.* (2019): Connectivity profile predictive of effective deep brain stimulation in obsessive-compulsive disorder. *Biol Psychiatry* 85: 735–743.

28. Barcia JA, Avecillas-Chasín JM, Nombela C, Arza R, Albea JG-, Pineda J-Á, *et al.* (2018): Personalized striatal targets for deep brain stimulation in obsessive-compulsive disorder. *Brain Stimulat*. https://doi.org/10.1016/j.brs.2018.12.226

29. Mosley PE, Windels F, Morris J, Coyne T, Marsh R, Giorni A, *et al.* (2021): A randomised, double-blind, sham-controlled trial of deep brain stimulation of the bed nucleus of the stria terminalis for treatment-resistant obsessive-compulsive disorder [no. 1]. *Transl Psychiatry* 11: 1–17.

30. Smith AH, Choi KS, Waters AC, Aloysi A, Mayberg HS, Kopell BH, Figee M (2021): Replicable effects of deep brain stimulation for obsessive-compulsive disorder. *Brain Stimulat* 14: 1–3.

31. Bouwens van der Vlis TAM, Ackermans L, Mulders AEP, Vrij CA, Schruers K, Temel Y, *et al.* (n.d.): Ventral capsule/ventral striatum stimulation in obsessive-compulsive disorder: toward a unified connectomic target for deep brain stimulation? *Neuromodulation Technol Neural Interface* n/a. https://doi.org/10.1111/ner.13339

32. Haber SN, Yendiki A, Jbabdi S (2020): Four deep brain stimulation targets for obsessive-compulsive disorder: Are they different? *Biol Psychiatry* S000632232031773X.

33. Wood J, Ahmari SE (2015): A framework for understanding the emerging role of corticolimbic-ventral striatal networks in OCD-associated repetitive behaviors. *Front Syst Neurosci* 171.

34. Burguière E, Monteiro P, Feng G, Graybiel AM (2013): Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science* 340: 1243–1246.

35. Heilbronner SR, Rodriguez-Romaguera J, Quirk GJ, Groenewegen HJ, Haber SN (2016): Circuit based cortico-striatal homologies between rat and primate. *Biol Psychiatry* 80: 509–521.

36. Makris N, Rathi Y, Mouradian P, Bonmassar G, Papadimitriou G, Ing WI, *et al.* (2016): Variability and anatomical specificity of the orbitofrontothalamic fibers of passage in the ventral capsule/ventral striatum (VC/VS): precision care for patient-specific tractography-guided targeting of deep brain stimulation (DBS) in obsessive compulsive disorder (OCD). *Brain Imaging Behav* 10: 1054–1067.

37. Liebrand LC, Caan MWA, Schuurman PR, van den Munckhof P, Figee M, Denys D, van Wingen GA (2018): Individual white matter bundle trajectories are associated with deep brain stimulation

response in obsessive-compulsive disorder. *Brain Stimulat*. https://doi.org/10.1016/j.brs.2018.11.014

38. Howell B, McIntyre CC (2016): Analyzing the tradeoff between electrical complexity and accuracy in patient-specific computational models of deep brain stimulation. *J Neural Eng* 13: 036023.

39. Poldrack RA, Huckins G, Varoquaux G (2020): Establishment of best practices for evidence for prediction: a review. *JAMA Psychiatry* 77: 534–540.

40. Grzenda A, Kraguljac NV, McDonald WM, Nemeroff CB, Torous J, Alpert JE, *et al.* (Accepted): Evaluating the machine learning literature: a primer and user's guide for psychiatrists. *Am J Psychiatry*.

41. Lo A, Chernoff H, Zheng T, Lo S-H (2015): Why significant variables aren't automatically good predictors. *Proc Natl Acad Sci* 112: 13892–13897.

42. Widge AS, Bilge MT, Montana R, Chang W, Rodriguez CI, Deckersbach T, *et al.* (2019): Electroencephalographic biomarkers for treatment response prediction in major depressive illness: a meta-analysis. *Am J Psychiatry* 176: 44–56.

43. Woo C-W, Chang LJ, Lindquist MA, Wager TD (2017): Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci* 20: 365–377.

44. Coenen VA, Schlaepfer TE, Varkuti B, Schuurman PR, Reinacher PC, Voges J, *et al.* (2019): Surgical decision making for deep brain stimulation should not be based on aggregated normative data mining. *Brain Stimul Basic Transl Clin Res Neuromodulation* 12: 1345–1348.

45. McLaughlin N, Dougherty DD, Eskandar EN, Ward HE, Foote KD, Malone Jr. DA, *et al.* (n.d.): Double blind randomized controlled trial of deep brain stimulation for obsessive-compulsive disorder: clinical trial design. *Contemp Clin Trials*.

46. Malcolm JG, Shenton ME, Rathi Y (2010): Filtered multitensor tractography [no. 9]. *IEEE Trans Med Imaging* 29: 1664–1675.

47. Reddy CP, Rathi Y (2016): Joint Multi-Fiber NODDI Parameter Estimation and Tractography Using the Unscented Information Filter. *Front Neurosci* 10. https://doi.org/10.3389/fnins.2016.00166

744 48. Gong S, Zhang F, Norton I, Essayed WI, Unadkat P, Rigolo L, *et al.* (2018): Free water modeling of
745 peritumoral edema using multi-fiber tractography: Application to tracking the arcuate fasciculus
746 for neurosurgical planning [no. 5]. *PLoS ONE* 13. https://doi.org/10.1371/journal.pone.0197056
747 49. Liao R, Ning L, Chen Z, Rigolo L, Gong S, Pasternak O, *et al.* (2017): Performance of unscented
748 Kalman filter tractography in edema: Analysis of the two-tensor model. *NeuroImage Clin* 15:
749 819–831.
750 50. Zhang F, Wu Y, Norton I, Rigolo L, Rathi Y, Makris N, O'Donnell LJ (2018): An anatomically curated
751 fiber clustering white matter atlas for consistent white matter tract parcellation across the
752 lifespan. *NeuroImage* 179: 429–447.
753 51. Chen Z, Tie Y, Olubiyi O, Rigolo L, Mehrtash A, Norton I, *et al.* (2015): Reconstruction of the
754 arcuate fasciculus for surgical planning in the setting of peritumoral edema using two-tensor
755 unscented Kalman filter tractography. *NeuroImage Clin* 7: 815–822.
756 52. Chaturvedi A, Luján JL, McIntyre CC (2013): Artificial neural network based characterization of the
757 volume of tissue activated during deep brain stimulation. *J Neural Eng* 10: 056023.
758 53. Butson CR, McIntyre CC (2006): Role of electrode design on the volume of tissue activated during
759 deep brain stimulation. *J Neural Eng* 3: 1–8.
760 54. O'Donnell LJ, Westin C-F (2007): Automatic tractography segmentation using a high-dimensional
761 white matter atlas [no. 11]. *IEEE Trans Med Imaging* 26: 1562–1575.
762 55. O'Donnell LJ, Wells WM, Golby AJ, Westin C-F (2012): Unbiased groupwise registration of white
763 matter tractography. In: Ayache N, Delingette H, Golland P, Mori K, editors. *Medical Image*
764 *Computing and Computer-Assisted Intervention – MICCAI 2012*, vol. 7512. Berlin, Heidelberg:
765 Springer Berlin Heidelberg, pp 123–130.
766 56. Couronné R, Probst P, Boulesteix A-L (2018): Random forest versus logistic regression: a large-
767 scale benchmark experiment. *BMC Bioinformatics* 19: 270.
768 57. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002): SMOTE: Synthetic Minority Over-
769 sampling Technique. *J Artif Intell Res* 16: 321–357.

58. Tyagi H, Apergis-Schoute AM, Akram H, Foltynie T, Limousin P, Drummond LM, *et al.* (2019): A randomised trial directly comparing ventral capsule and anteromedial subthalamic nucleus stimulation in obsessive compulsive disorder: clinical and imaging evidence for dissociable effects. *Biol Psychiatry* 85: 726–734.

59. Smith GCS, Seaman SR, Wood AM, Royston P, White IR (2014): Correcting for optimistic prediction in small data sets. *Am J Epidemiol* 180: 318–324.

60. Widge AS, Dougherty DD (2015): Managing patients with psychiatric disorders with deep brain stimulation. In: Marks Jr. WJ, editor. *Deep Brain Stimulation Management*, 2nd ed. Cambridge : New York: Cambridge University Press.

61. van Westen M, Rietveld E, Bergfeld IO, Koning P de, Vullink N, Ooms P, *et al.* (2021): Optimizing deep brain stimulation parameters in obsessive–compulsive disorder. *Neuromodulation Technol Neural Interface* 24: 307–315.

62. Mantione M, Nieman DH, Figee M, Denys D (2014): Cognitive–behavioural therapy augments the effects of deep brain stimulation in obsessive–compulsive disorder. *Psychol Med* 44: 3515–3522.

63. van Westen M, Rietveld E, Denys D (2019): Effective deep brain stimulation for obsessive-compulsive disorder requires clinical expertise. *Front Psychol* 10: 2294.

64. Burguière E, Monteiro P, Mallet L, Feng G, Graybiel AM (2015): Striatal circuits, habits, and implications for obsessive–compulsive disorder. *Curr Opin Neurobiol* 30: 59–65.

65. Garnaat SL, Greenberg BD, Sibrava NJ, Goodman WK, Mancebo MC, Eisen JL, Rasmussen SA (2014): Who qualifies for deep brain stimulation for OCD? Data from a naturalistic clinical sample. *J Neuropsychiatry Clin Neurosci* 26: 81–86.

66. Duchin Y, Shamir RR, Patriat R, Kim J, Vitek JL, Sapiro G, Harel N (2018): Patient-specific anatomical model for deep brain stimulation based on 7 Tesla MRI. *PloS One* 13: e0201469.

67. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR (2013): Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 14: 365–376.

68. Menchón JM, Real E, Alonso P, Aparicio MA, Segalas C, Plans G, *et al.* (2021): A prospective international multi-center study on safety and efficacy of deep brain stimulation for resistant obsessive-compulsive disorder [no. 4]. *Mol Psychiatry* 26: 1234–1247.

69. Greenberg B, Gabriels L, Malone D, Rezai A, Friehs G, Okun M, *et al.* (2010): Deep brain stimulation of the ventral internal capsule/ventral striatum for obsessive-compulsive disorder: worldwide experience. *Mol Psychiatry* 15: 64–79.

70. Nuttin B, Cosyns P, Demeulemeester H, Gybels J, Meyerson B (1999): Electrical stimulation in anterior limbs of internal capsules in patients with obsessive-compulsive disorder. *The Lancet* 354: 1526.

71. Basu I, Yousefi A, Crocker B, Zelmann R, Paulk AC, Peled N, *et al.* (Accepted in principle): Closed loop enhancement and neural decoding of human cognitive control. *Nat Biomed Eng*. https://doi.org/10.1101/2020.04.24.059964

72. Widge AS, Zorowitz S, Basu I, Paulk AC, Cash SS, Eskandar EN, *et al.* (2019): Deep brain stimulation of the internal capsule enhances human cognitive control and prefrontal cortex function. *Nat Commun* 10: 1536.

73. Gruner P, Pittenger C (2017): Cognitive inflexibility in Obsessive-Compulsive Disorder. *Neuroscience* 345: 243–255.

74. Abramovitch A, Abramowitz JS, Mittelman A (2013): The neuropsychology of adult obsessive–compulsive disorder: A meta-analysis. *Clin Psychol Rev* 33: 1163–1171.

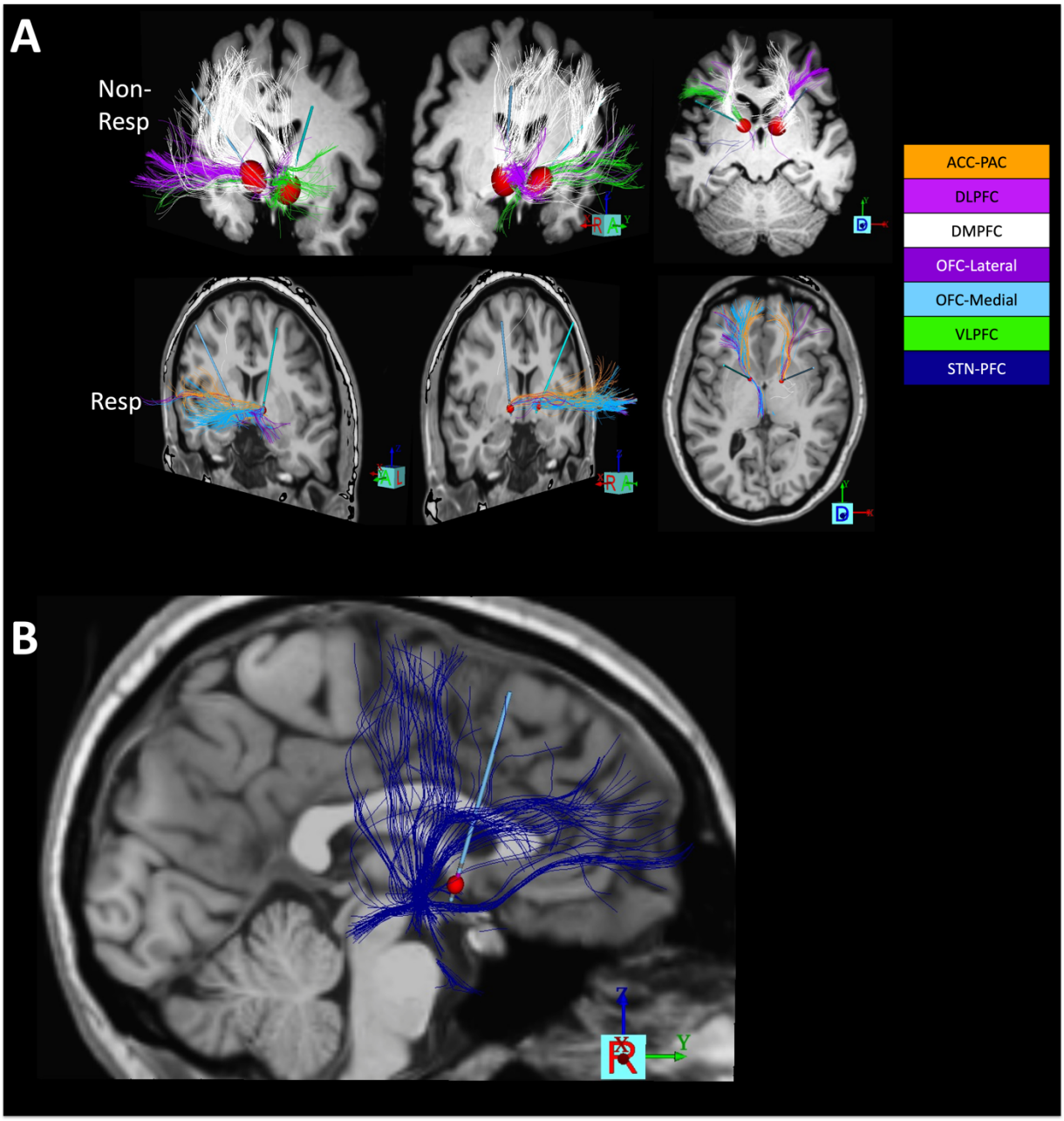


Figure 1: Patient-specific tractographic mapping of OCD DBS response. (A), Tract tracing and activation modeling examples. Shown are left/right oblique and axial views from one non-responder and one responder, with cortico-thalamic and cortico-STN tracts indicated by different colors. DBS leads are shown in teal and VTAs in red. In this panel, we show only tracts intersecting the VTAs for clarity. (B), Tracing of tracts between STN and frontal cortex, in the same responder as (A). To ensure capture of the tract reported in (12), we broadly traced all streamlines originating in a seed around STN and extending anterior to the central sulcus. This includes fibers coursing dorsally to motor regions, and

826 tracts as in (12) connecting STN to ACC and medial PFC. Very few of these intersect the VTA in this
827 patient, despite the good clinical response (YBOCS drop of 61% from baseline). To emphasize that
828 point, this panel shows all fibers traced from the STN seed in this patient, regardless of VTA
829 intersection.

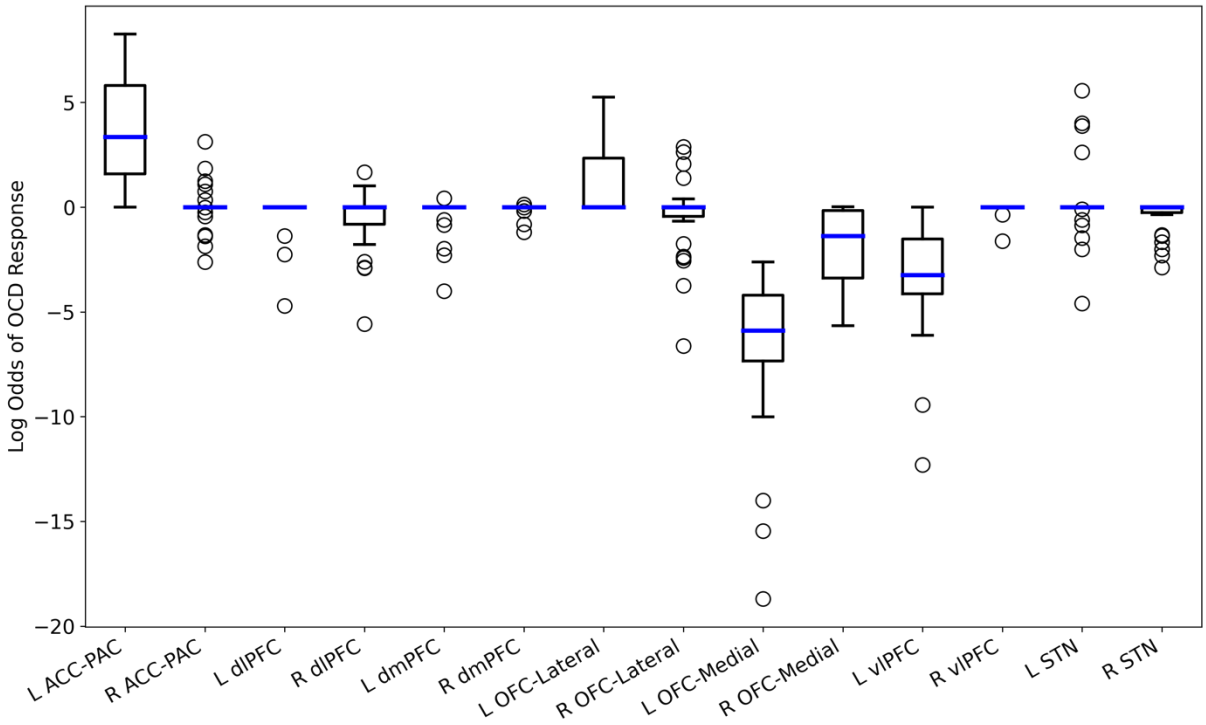
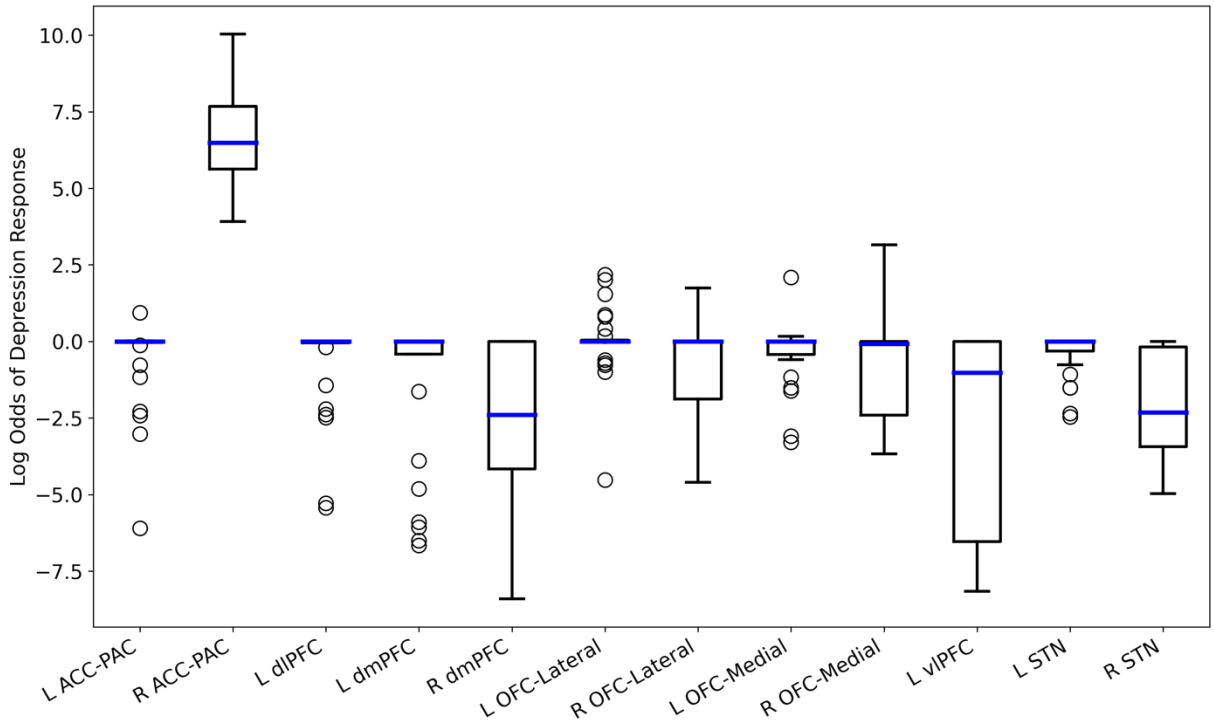


Figure 2: Non-zero regression coefficients across exhaustive leave-two-out cross-validation of regularized logistic regression to predict YBOCS response. All confidence intervals include 0, with left medial OFC (non-response) and left ACC (response) coming closest to significance. All reported results are for total fiber capture; percentage capture did not have non-zero coefficients in this analysis. Data are coded such that positive regression coefficients represent clinical improvement.



838

839 **Figure 3:** Non-zero regression coefficients across exhaustive leave-two-out cross-validation of

840 regularized logistic regression to predict MADRS response. All confidence intervals include 0, except

841 for the right cingulate cortex. All reported results are for total fiber capture; percentage capture did not

842 have non-zero coefficients in this analysis.

843

	R^2			Explained Variance		
	Median	CI Lower Bound	CI Upper Bound	Median	CI Lower Bound	CI Upper Bound
L1 Regression (Percentage)	-0.194	-1.743	1.355	0	0	0
L1 Regression (Total Fibers)	-0.196	-1.747	1.356	0	0	0
Random Forest (Percentage)	-0.792	-3.794	2.21	-0.023	-0.902	0.857
Random Forest (Total Fibers)	-1.389	-4.934	2.156	-0.222	-1.249	0.805

Table 1: Modeling outcomes for YBOCS improvement as a continuous variable. All confidence intervals include 0. Negative coefficients of determination (R^2) imply a model that performs worse than chance.

	Balanced Accuracy			Recall			AUC		
	Median	CI Lower Bound	CI Upper Bound	Median	CI Lower Bound	CI Upper Bound	Median	CI Lower Bound	CI Upper Bound
L1 Logistic (Percentage)	0.5	0.5	0.5	0	0	0	0.5	0.5	0.5
L1 Logistic (Total Fibers)	0.5	0.093	0.907	0.571	-0.24	1.383	0.572	0.281	0.864
Random Forest (Percentage)	0.46	0.192	0.728	0	-0.582	0.582	0.58	0.368	0.793
Random Forest (Total Fibers)	0.421	0.193	0.65	0	-0.527	0.527	0.588	0.374	0.802

Table 2: Modeling outcomes for YBOCS improvement as a categorical response. All confidence intervals include chance (0.5 for Balanced Accuracy and AUC, 0 for Recall of the minority class).

	Training Set R			Test Set R^2		
Tract	**Median**	**CI Lower Bound**	**CI Upper Bound**	**Median**	**CI Lower Bound**	**CI Upper Bound**
L dlPFC (Percentage)	-0.405	-0.607	-0.203	-0.139	-1.842	1.564
L vlPFC (Total Fibers)	-0.389	-0.591	-0.187	-0.268	-2.813	2.277
R vlPFC (Percentage)	-0.384	-0.504	-0.263	-0.162	-1.622	1.297
L STN (Percentage)	-0.373	-0.62	-0.125	-0.301	-3.189	2.587
L vlPFC (Percentage)	-0.366	-0.603	-0.129	-0.411	-27.119	26.297
R dmPFC (Percentage)	-0.359	-0.481	-0.237	-0.275	-2.182	1.633
R dlPFC (Percentage)	-0.356	-0.542	-0.17	-0.14	-1.706	1.426
All Regions (Percentage)	-0.347	-0.555	-0.138	-0.35	-1.598	0.897
L dlPFC (Total Fibers)	-0.332	-0.594	-0.07	-0.275	-2.149	1.6
R STN (Percentage)	-0.321	-0.525	-0.118	-0.315	-2.149	1.519
L dmPFC (Percentage)	-0.315	-0.567	-0.063	-0.434	-22.196	21.329
All Regions (Total Fibers)	-0.311	-0.553	-0.069	-0.421	-1.868	1.026
R dmPFC (Total Fibers)	-0.266	-0.471	-0.062	-0.35	-1.832	1.132
R vlPFC (Total Fibers)	-0.266	-0.461	-0.071	-0.182	-2.991	2.626
L ACC-PAC (Total Fibers)	0.268	0.067	0.469	-0.33	-1.959	1.298
R OFC-Lateral (Percentage)	0.319	0.07	0.567	-0.315	-1.771	1.141

Table 3: Correlations between individual fiber tracts and YBOCS response, in the style of (12), filtered to tracts whose confidence interval excludes 0 on the training sets. No such tract has clinical predictive power on held-out test sets (all R^2 values less than 0).

	Balanced Accuracy			Recall			AUC		
	Median	CI Lower Bound	CI Upper Bound	Median	CI Lower Bound	CI Upper Bound	Median	CI Lower Bound	CI Upper Bound
L1 Logistic (Percentage)	0.5	0.5	0.5	0	0	0	0.5	0.5	0.5
L1 Logistic (Total Fibers)	0.5	0.093	0.907	0.571	0	1	0.572	0.281	0.864
Random Forest (Percentage)	0.46	0.192	0.728	0	0	0.582	0.58	0.368	0.793
Random Forest (Total Fibers)	0.421	0.193	0.65	0	0	0.527	0.588	0.374	0.802

Table 4: Leave-two-out prediction outcomes for categorical depression response (MADRS). No model exceeded chance accuracy on the test set (all confidence intervals include a balanced accuracy/AUC of 0.5 or recall of 0).

Patient-Specific Connectomic Models Correlate With, But Do Not Predict, Outcomes in Deep Brain Stimulation for Obsessive-Compulsive Disorder

Supplemental Methods

Study Population and Clinical Treatment

In all cases, DBS implantation and programming followed the protocols described in (60). 2 patients were implanted and followed at Butler Hospital/Brown University Medical School; the other 6 were implanted and followed at Massachusetts General Hospital. All patients received Medtronic model 3387 DBS leads, with the most ventral contact targeted to the ventral striatal grey matter. All participants gave informed consent after multiple meetings with the site study teams, which explicitly included separate consent for neuroimaging.

One patient, reported in (76), declined rating visits with the study team after her first several months of treatment, due to being substantially improved and not desiring further programming. She had a series of telephone notes captured in the electronic medical record over the course of a year documenting that she was doing well with no DBS setting changes from her last programming visit. We carried her YBOCS and MADRS forward from the last available rating, despite suggestions in the clinical record that she was doing better than these scores reflect.

Imaging Details

Pre-operative MRI data were acquired on a 3T Siemens TimTrio scanner. T1-weighted and T2-weighted images were acquired with a 1.2 mm isotropic voxel size; diffusion MRI (dMRI) scans had a spatial resolution of 2 mm (isotropic) with 10 non-diffusion weighted volumes and 60 diffusion weighted volumes, with gradient directions spread uniformly on the sphere with a b-value of 700 s/mm^2. After

DBS implantation, a postoperative CT scan was acquired at a spatial resolution of 0.43 x 0.43 x 0.63 mm^3.

All MRI data was processed using a published pipeline (43) available at https://github.com/pnlbwh/. We applied axis alignment to the T1-weighted and T2-weighted images, and performed eddy current and motion correction for the dMRI images. T1-weighted images were skull-stripped using the Brain Extraction Toolkit (BET) (77,78). These masks were then manually checked and edited to ensure accurate brain extraction. In a similar manner, the CT scans were also masked to only retain the brain. The ANTS registration software (79) was then used to coregister the CT images to the T1-weighted images as well as the dMRI scans. Further, Freesurfer (v6.1) was run on the T1-weighted images to parcellate the brain (both cortex and subcortical structures) using the Desikan-Killiany atlas. Next, we registered the Freesurfer segmentation to the diffusion weighted images using pipeline scripts. All registrations (CT to T1-weighted, CT to dMRI, Freesurfer) were manually checked for accuracy.

During the tract tracing and annotation, fibers terminating in ACC vs. PAC could not be reliably distinguished across all subjects. We thus treated these adjacent regions as a single terminus.

We defined the STN-PFC tracts as all streamlines connecting the STN with the following Freesurfer segmented regions: caudal anterior cingulate, caudal middle frontal, lateral orbitofrontal, medial orbitofrontal, parsopercularis, parsorbitalis, rostral anterior cingulate, rostral middle frontal, superior frontal, and frontal pole. To ensure that we captured the relevant bundles, we dilated the STN by 1 voxel (~ 2mm) in all directions (from our hand drawn segmentations) to include more tracts that might be close to the STN. These streamlines were analyzed identically to the cortico-thalamic tracts.

Data Analysis - Hypomania

We considered the possibility that pathway modeling might be more useful for minimizing off-target effects than for determining response (23,80). The most common complication of VCVS DBS is a hypomanic-impulsive syndrome that may affect up to 50% of patients (57,68). (We refer to this as "hypomania" for conciseness, but acknowledge the substantial debate (61) over the naming and nature of this complication.) We ascertained hypomania by chart review of visit notes and case report forms. 3 patients experienced at least one hypomanic episode, over 50 total visits. We again compared L1-regularized logistic regression and random forest classification on this restricted dataset. It was not possible to perform this analysis on the full dataset, because the rarity of the complication meant that cross-validation strategies would produce training/test sets without sufficient examples. Further, it would likely be possible to achieve above-chance classification by identifying which patient contributed a given data point, which would be subtly reflected in the overall pattern of activation (i.e., there is a strong potential for data leakage).

Data Analysis - Model Evaluation

We altered the cross-validation process for the hypomania model, because restricting the dataset to the 3 patients who experienced at least one hypomanic event made leave-2-out infeasible. We therefore cross-validated at the level of individual visits. We split the dataset with 80% of the data points as training and 20% as test set data, stratifying the split so that hypomanic events occurred in each dataset. We repeated this splitting process 1,000 times to obtain confidence intervals, and again performed all oversampling after the split.

Data Analysis - Overall White Matter Integrity

In a further exploration, we noted that some tracts had very few streamlines in non-responders. We considered that white matter integrity might affect our results, such that patients could only respond if a given tract were sufficiently intact. For every tract in the previous analyses, we defined its integrity as

the total number of streamlines traced, divided by the overall intracranial volume (to ensure that we did not simply trace more fibers in larger brains). Following the total/percentage fiber analyses above, we also created a predictor from the mean integrity across all tracked bundles. We could not correlate integrity against response at the visit level, since tract integrity does not change with DBS settings. Instead, we classified patients as overall responders if a plurality (40% or more) of their clinical visits had YBOCS scores below the 35% response threshold; 3 of the 8 met this criterion. We did not consider the first three months of the clinical course in determining this response, in order to emphasize response from DBS as opposed to lesion effect. We did not SMOTE this analysis, as the dataset was too small.

To test for a “threshold” level of integrity for DBS efficacy, we performed a Necessary Condition Analysis (NCA, (81)), as implemented in R package “NCA” (82), using the free disposal hull (CE-FDH) ceiling calculation. We separately tested the necessity of each tract and of overall (mean) integrity. We assessed significance of the resulting NCA effect sizes by 1000-fold permutation (shuffling responder/non-responder labels), with Benjamini-Hochberg false discovery rate correction.

Supplemental Results

Hypomania

3 patients (37.5%) experienced hypomania (6 total episodes) in this dataset. This is somewhat below the rate we reported from a different VCVS dataset (57), but consistent with the largest published cohort of this target in OCD (68).

No model was able to predict hypomania with better than chance performance (Table S5). Further, all fitted models had 0 median recall, i.e. predictive power was generally achieved by always predicting the majority (non-hypomanic) outcome. The percentage-capture regression model assigned all coefficients to 0, whereas the total-fiber model had some nonzero coefficients (Figure S4). Specifically, the left-sided PFC-STN connection was protective against hypomania. The left DLFPC-thalamus fibers had a similar but non-significant effect. Activation of right lateral OFC and STN fibers, on the other hand, predisposed towards hypomania, consistent with a prior analysis that found an association between right-sided monopolar stimulation and hypomania (57). The random forest model had no importance scores that systematically differed from zero.

	Balanced Accuracy			Recall			AUC		
	Median	**CI Lower Bound**	**CI Upper Bound**	**Median**	**CI Lower Bound**	**CI Upper Bound**	**Median**	**CI Lower Bound**	**CI Upper Bound**
L1 Logistic (Percentage)	0.5	0.5	0.5	0	0	0	0.5	0.5	0.5
L1 Logistic (Total Fibers)	0.4	0	0.84	0	0	0.899	0.7	0.486	0.914
Random Forest (Percentage)	0.5	0.043	0.957	0	0	0.912	0.55	0.152	0.948
Random Forest (Total Fibers)	0.5	0.003	0.997	0	0	0.988	0.55	0.109	0.991

Table S5: Prediction outcomes for hypomania, on a test set composed of 20% of visits from the 3 patients who had hypomanic episodes. No model consistently achieves recall > 0 or better than chance accuracy/AUC.

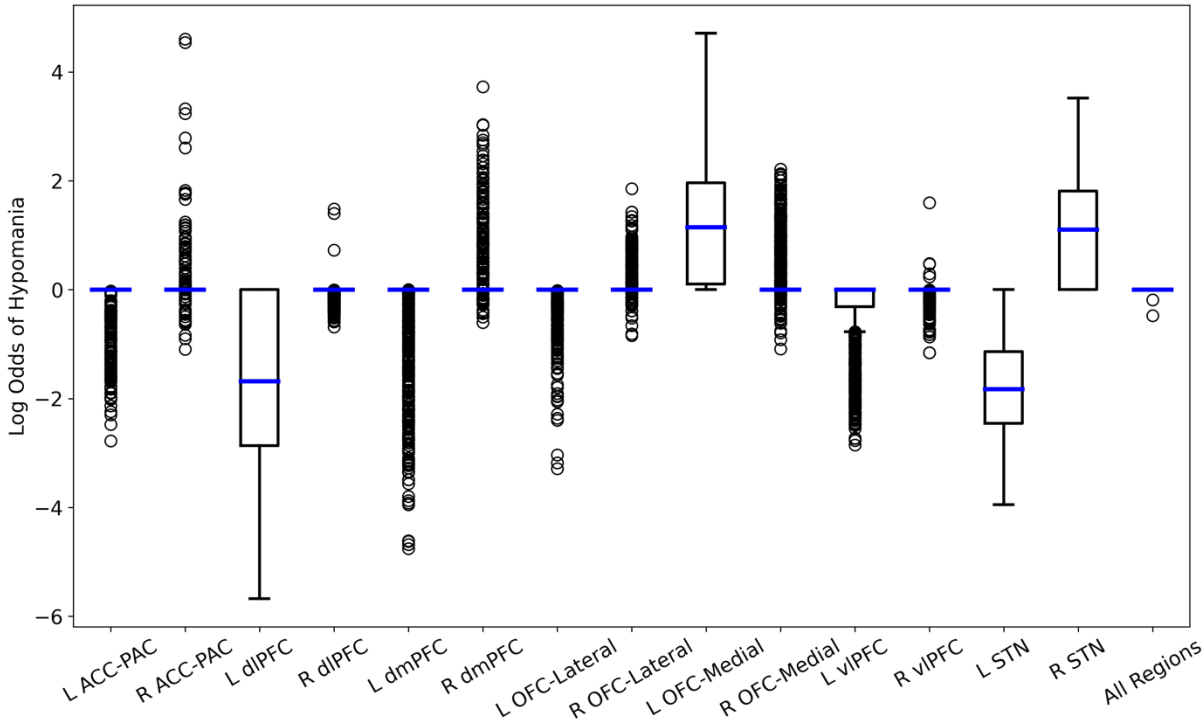


Figure S4: Non-zero regression coefficients for regularized logistic regression to predict hypomania. The only tract whose confidence interval excludes 0 connected left PFC to STN, and was protective against hypomania. All reported results are for total fiber capture; percentage capture did not have non-zero coefficients in this analysis. Here, positive coefficients signify greater risk of hypomania.

White Matter Integrity

There was substantial variability in white matter integrity among individual patients and tracts (Table S2). Overall (mean) white matter integrity had a moderate effect size in the NCA (0.62). Consistent with the regression analyses, the left and right cingulo-thalamic tracts had the highest effect size of any individual tract. None of these effects reached significance even at the uncorrected level (Table S3).

Subj	L ACC-PAC	L dlPFC	L dmPFC	L OFC-Lateral	L OFC-Medial	L vlPFC	L STN	R ACC-PAC	R dlPFC	R dmPFC	R OFC-Lateral	R OFC-Medial	R vlPFC	R STN	Mean Integrity	Responder
A	0.064	0.280	0.306	0.017	0.029	0.414	0.220	0.155	0.331	1.083	0.016	0.010	0.691	0.115	0.278	N
B	0.187	0.506	1.076	0.012	0.055	0.279	0.025	0.160	0.319	1.718	0.013	0.033	0.713	0.025	0.392	Y
C	0.003	0.063	0.117	0.000	0.002	0.075	0.005	0.002	0.167	0.239	0.000	0.000	0.041	0.005	0.055	N
D	0.280	0.364	0.669	0.081	0.080	0.473	0.093	0.210	0.342	0.588	0.057	0.023	0.169	0.102	0.264	Y
E	0.148	0.470	0.834	0.043	0.089	0.057	0.105	0.296	1.079	0.993	0.008	0.011	0.418	0.104	0.350	N
F	0.194	0.398	0.871	0.038	0.092	0.226	0.054	0.263	0.302	1.206	0.051	0.126	0.688	0.074	0.347	N
G	0.311	0.655	1.327	0.018	0.042	0.145	0.567	0.156	0.413	0.589	0.016	0.056	0.149	0.166	0.342	Y
H	0.098	0.309	0.806	0.100	0.308	0.084	0.103	0.128	0.257	1.283	0.043	0.116	0.198	0.176	0.295	N

Table S6: White matter integrity. Each cell represents the number of traced streamlines within a given tract, divided by each subject's intracranial volume. Responder/nonresponder status is determined by the fraction of clinical visits where the YBOCS had improved 35% or more from baseline.

Name	Effect Size	p	p(FDR)
Mean Integrity	0.620	0.635	0.635
L ACC-PAC	0.595	0.077	0.475
R ACC-PAC	0.526	0.190	0.475
L dlPFC	0.509	0.190	0.475
L dmPFC	0.456	0.373	0.622
R dmPFC	0.236	0.635	0.635
R OFC-Lateral	0.225	0.335	0.622
L vlPFC	0.211	0.170	0.475
R OFC-Medial	0.182	0.178	0.475
R dlPFC	0.167	0.187	0.475
R vlPFC	0.160	0.635	0.635
L OFC-Medial	0.130	0.373	0.622
L OFC-Lateral	0.123	0.635	0.635
R STN	0.113	0.635	0.635
L STN	0.036	0.635	0.635

Table S3: Necessary Condition Analysis for white matter integrity predicting YBOCS response, with p-values from 1000-fold bootstrap resampling, both raw and False Discovery Rate corrected. No variable reaches corrected or uncorrected significance.