**PROBAST**

Study:

Motor network gamma oscillations in chronic home recordings predict dyskinesia in Parkinson's disease.

Step 2: Type of prediction study

**Is the study a diagnostic or a prognostic study?**

Diagnostic

**Is the study a development only, development and validation or validation only study?**

Development only

**What is the model of interest?**

Linear regression

**What is the outcome of interest?**

Estimate dyskinesia score in PD

Step 3: Assess risk of bias

**Domain 1: Participants**

**Describe the sources of data and criteria for participant selection**

We recruited participants from the movement disorders surgery clinic at the University of California San Francisco in accordance with the Declaration of Helsinki. Inclusion criteria were a diagnosis of idiopathic Parkinson’s disease by a movement disorders neurologist and standard clinical indications for subthalamic nucleus or globus pallidus deep brain stimulation. Exclusion criteria were significant cognitive impairment as determined by a Montreal Cognitive Assessment score of 20 or lower or an untreated mood disorder as evaluated by a neuropsychologist.

**1.1 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?**

Y

**1.2 Were all inclusions and exclusions of participants appropriate?**

Y

**Risk of bias introduced by selection of participants:**

Low

**Rationale of bias rating**

Reasonable exclusion criteria

**Domain 2: Predictors**

**List and describe predictors included in the final model, e.g. definition and timing of assessment**

We analysed intracranial sensorimotor cortical and basal ganglia field potentials from subjects with chronically implanted cortical paddles and depth leads in the home environment while engaging in activities of daily living. Recordings were completed more than 1 week after implantation, before initiating therapeutic deep brain stimulation 1 month after implantation, and for recording durations of at least 8 h per hemisphere. For cortical recordings in Subjects 1–5, we configured overlapping bipolar recording channels (contacts 1–3 and 2–4) and configured all subsequent subjects with non- overlapping bipolar recording channels (contacts 1–2 and 3–4). For the non-overlapping recording channels from the cortical site, the recording configuration from contacts 1–2 sampled the precentral gyrus, while the recording configuration from contacts 3–4 sampled the postcentral gyrus.

Subjects streamed neural data at home while wearing Global Kinetics Pty Ltd. Personal KinetiGraph® (PKG®) monitors on both wrists during normal daily activities and while on their preoperative dose of antiparkinsonian medication (Table 1). This wearable device pro-vides continuous tremor, bradykinesia and dyskinesia scores in 2-min intervals using a three-axis accelerometer and a proprietary, validated commercial algorithm.

**2.1 Were predictors defined and assessed in a similar way for all participants?**

N

**2.2 Were predictor assessments made without knowledge of outcome data?**

Y

**2.3 Are all predictors available at the time the model intended to be used?**

Y

**Risk of bias introduced by predictors or their assessment**

High

**Rationale of bias rating**

Implantation of intracranial sensors differed between the patients

**Domain 3: Outcome**

**Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:**

Subjects streamed neural data at home while wearing Global Kinetics Pty Ltd. Personal KinetiGraph® (PKG®) monitors on both wrists during normal daily activities and while on their preoperative dose of antiparkinsonian medication (Table 1). This wearable device pro-vides continuous tremor, bradykinesia and dyskinesia scores in 2-min intervals using a three-axis accelerometer and a proprietary, validated commercial algorithm.

We transformed the wearable dyskinesia numerical output to have discrete ordinal labels. We binned the dyskinesia scores from bins 0–3 in an exponential manner, such that each consecutive bin contains half the amount of values as the previous bin, to mimic the duration of time subjects experience dyskinesia theoretically. Bin 4 contains all the remaining values. The ordinal bin with a label of 0 corresponds to the lowest 50th percentile of dyskinesia scores, with each consecutive label containing the remaining lowest 50th percentile of scores such that bin 0 contains per-centiles 0–5

**3.1 Was the outcome determined appropriately?**

PN

**3.2 Was a pre-specified or standard outcome definition used?**

Y

**3.3 Were predictors excluded from the outcome definition?**

Y

**3.4 Was the outcome defined and determined in a similar way for all participants?**

Y

**3.5 Was the outcome determined without knowledge of predictor information?**

Y

**3.6 Was the time interval between predictor assessment and outcome determination appropriate?**

Y

**Risk of bias introduced by the outcome or its determination**

Low

**Rationale of bias rating**

Some risk due to ordinal binning of outcome, however this risk might be rather small.

**Domain 4: Analysis**

**Describe number of participants, number of candidate predictors, outcome events and events per candidate predictor**

We implanted 30 hemispheres from 16 subjects with Parkinson’s disease with cortical and basal ganglia leads (2/16 female, mean age at surgery 57 ± 12 years). Subject demographics are in Table 1. Twelve subjects had a clinical history of dyskinesia (UPDRS IVa dyskinesia score >0), 10 of whom had implants in the subthalamic nucleus. In the OFF-medication state, 15 subjects had mild to moderate tremor scores in at least one limb (UPDRS III 15–17), and all subjects had non-zero bradykinesia scores in all limbs (UPDRS III 4–9, 14, recorded in the OFF-medication state).

**Describe how the model was developed, predictor selection and risk group definition**

For linear regression analysis, we used the sklearn package. We split data into an 80/20 ratio of training and testing subsets. Using the training subset, we trained a linear regression model and calculated the correlation coefficient (r) between the predicted and actual dependent variable from the testing subset.

**Describe whether and how the model was validated, either internally (cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants)**

For cross-validation, we regenerated the data split ratio five times across non-overlapping epochs of the data. Then, we averaged all statistics across each split.

**Describe the performance measures of the model, e.g. calibration, discrimination, classification, net benefit, and whether they were adjusted for optimism**

Correlation coefficient

**Describe any participants who were excluded from the analysis**

None

**Describe missing data on predictors and outcomes as well as methods used for missing data**

Not described

**4.1 Were there a reasonable number of participants with the outcome?**

N

**4.2 Were continuous and categorical predictors handled appropriately?**

Y

**4.3 Were all enrolled participants included in the analysis?**

Y

**4.4 Were participants with missing data handled appropriately?**

U

**4.5 Was selection of predictors based on univariable analysis avoided?**

Y

**4.6 Were complexities in the data (e.g. censoring, competing risks, sampling of controls)**

**accounted for appropriately?**

Y

**4.7 Were relevant model performance measures evaluated appropriately?**

N

**4.8 Were model overfitting and optimism in model performance accounted for?**

Y

**4.9 Do predictors and their assigned weights in the final model correspond to the results**

**from multivariable analysis?**

U

**Risk of bias introduced by the analysis**

High

**Rationale of bias rating**

Low amount of outcomes. Only correlation coefficient reported.

**Overall Risk of bias**

High