**PROBAST**

Study:

Wearable-Enabled Algorithms for the Estimation of Parkinson's Symptoms Evaluated in a Continuous Home Monitoring Setting Using Inertial Sensors.

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?**

Convolutional neural network

**What is the outcome of interest?**

Detection of PD motor symptoms

Step 3: Assess risk of bias

**Domain 1: Participants**

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

Two sets of data were constructed for this study, which incorporated both in-clinic and at-home data collections.

No eligibility criteria mentioned.

**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**

No eligibility criteria mentioned except for PD diagnosis.

**Domain 2: Predictors**

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

Accelerometer data from wearables.

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

Y

**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**

Low

**Rationale of bias rating**

Predictors are assessed the same way for every patient.

**Domain 3: Outcome**

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

They simultaneously filled out an adapted version of the Veterans Affairs Patient Motor Diary (VA Patient Motor Diary) [81] every half an hour (available in the supplementary material – Table A1).

**3.1 Was the outcome determined appropriately?**

Y

**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**

Patient reported outcomes with a validated questionnaire. Assessed the same way for every patient.

**Domain 4: Analysis**

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

Additionally, 20 PD subjects (seventeen from the above trial and three new patients) took part in the at-home data collection.

Unclear how many outcomes.

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

A convolutional neural network (CNN) was considered fora tremor classification model

**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)**

These data were split into a training and test set (approx. 80%/20%)with different subjects in each set such that no data from a participant in the training set would appear in the test set and vice versa. The split was made randomly while attempting to preserve the percentage of samples for each class as much as possible given this constraint of non-overlapping subject data between the training and test sets. Samples created by data augmentation were removed from the test set before prediction. For reporting purposes, this procedure was repeated five times and the mean and standard deviation (SD) recorded in the results

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

BACC, REC, PRE, AUC

**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?**

U

**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?**

Y

**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**

U

**Rationale of bias rating**

It is unclear how many outcomes (dyskinesia episodes etc.) there are in total over the 20 patients. Therefore we cannot conclude whether the dataset size is big enough.

**Overall Risk of bias**

Unclear