**PROBAST**

Study:

Wearable Sensor-Based Assessments for Remotely Screening Early-Stage Parkinson's Disease.

Step 2: Type of prediction study

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

Diagnoistc

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

Development only

**What is the model of interest?**

Random forest

**What is the outcome of interest?**

Distinguish PD patients from healthy controls

Step 3: Assess risk of bias

**Domain 1: Participants**

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

Participants living with Parkinson’s disease (PD) and healthy controls (HC) were recruited to complete the multi-center (n = 17) WATCH-PD (Wearable Assessments in the Clinic and at Home in PD) (NCT03681015) observational study at a designated Parkinson Study Group research site. The WCGTM Institutional Review Board approved the procedures used in the study, and there was full compliance with human experimentation guidelines. Criteria for enrollment into the PD group included: (1) a diagnosis that has been clinically documented by a movement disorder specialist; (2) the participant is older than 30 years at diagnosis; (3) a disease duration of less than two years; (4) a Hoehn and Yahr stage of <3; (5) no baseline use of dopaminergic or other PD medications; and (6) no alternative Parkinsonian diagnosis. Approximately half of the PD group underwent DaTscan screening to confirm their diagnosis. Criteria for enrollment into the HC group included: (1) age-match to the PD group; (2) no previous PD diagnosis; and (3) no other significant neurologic disease.

**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 eligibility criteria for PD and healthy controls

**Domain 2: Predictors**

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

Features were engineered from all the continuous data sources available from each assessment (see Table 1). Our library of features was modeled after previous works, extracting features from cognitive speech and accelerometer-based mobility assessments

**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 were gathered the same way for every participant and are independent of outcome and can be done when model is intended to be used.

**Domain 3: Outcome**

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

Differentiate PD from healthy control. Diagnosis done before recruitment: Criteria for enrollment into the PD group included: (1) a diagnosis that has been clinically documented by a movement disorder specialist;

**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**

No risk of bias as outcome is based on PD diagnosis based on specialist.

**Domain 4: Analysis**

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

PD (n = 82) and HC (n = 50) participants enrolled in the WATCH-PD (NCT03681015)study across 17 Parkinson Study Group research sites

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

To enumerate how many features showed significant distributional differences between the PD and HC participants, univariate linear regression was performed on each feature independently, wherein a given feature was the dependent variable and the participant’s group (PD, HC) was a categorical independent variable (Feature~Group [PD, HC]).

Feature reduction routines were performed to reduce the dimensionality of the featuresused during modeling by eliminating multicollinearity between features.

Nine unique models were developed and evaluated.

**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)**

LOOCV

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

ACC, AUC, SEN, SPE

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

Not described

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

For example, if a participant partially completed some assessments in a single session, the row for that session would be missing some features. Features generated across all sessions—both at home andin the clinic—were averaged together for each participant, producing a 1 ×n row of features for each participant, thereby minimizing the session and temporal variance in the measurements and normalizing features to a normal distribution.

**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?**

U

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

Y

**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?**

Y

**Risk of bias introduced by the analysis**

High

**Rationale of bias rating**

Small amount of patients. Unclear whether some patients were excluded.

**Overall Risk of bias**

High