**PROBAST**

Study:

Unveiling the Unpredictable in Parkinson's Disease: Sensor-Based Monitoring of Dyskinesias and Freezing of Gait in Daily Life

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?**

Logistic regression

**What is the outcome of interest?**

Detect FOG and dyskinesia in PD patients

Step 3: Assess risk of bias

**Domain 1: Participants**

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

Patients with PD were longitudinally enrolled at the movement disorder outpatient clinics of the Sapienza University of Rome and the University of Turin (Italy) between January 2023 and February 2024. The inclusion criteria for the study encompassed diagnosis of idiopathic PD based on current consensus criteria [1]; Hoehn and Yahr (H&Y)between 1.5 and 4; PD duration of at least 3 years; and chronic therapy with L-Dopa ±other dopaminergic drugs. Exclusion criteria included diagnosis of possible or probable atypical parkinsonism; inability to walk independently; and comorbidities potentially affecting gait(e.g., neurological conditions other than PD, orthopedic and/or rheumatologic issues).

**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

**Domain 2: Predictors**

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

Each patient underwent long-term gait monitoring for a minimum of 5 days, at least 8 h per day, during daily life activities using a single wearable device (STAT-ONTM, Sense4Care, Barcelona, Spain) on the left side of the waist through an elastic belt (Figure 1). The device was positioned so that the x, y, and z axes of the embedded sensors represented the anterior, vertical (upward), and lateral

**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**

Accelerometer data are applicable and assessed the same way

**Domain 3: Outcome**

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

All participants were regularly followed up by an expert in movement disorders and classified as patients with (PD-Dys) and without dyskinesias (PD-nDys) based on the use of the Unified Parkinson’s Disease Dyskinesia Rating Scale (UDysRS) (score ≥1 at item 1) and the direct observation of dyskinesias on at least one visit. Likely, the classification of participants as patients with (PD-FOG) and without FOG (PD-nFOG) was based on the clinical diagnosis of FOG through the direct observation of the disorder on at least one visit and using the FOG-Questionnaire.

A binary classification approach was also employed to differentiate patients experiencing both dyskinesias and FOG from those who were unaffected by each of the two disorders based on both dyskinesias and FOG measures.

**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**

Gold standard tests were taken to diagnose patients with dyskinesia and FoG

**Domain 4: Analysis**

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

71 patients

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

Then, the performance of a logistic regression model in leave-one-subject-out vali-dation was evaluated in terms of sensitivity, specificity, positive and negative predictive values (PPV and NPV, respectively), accuracy, and area under the curve (AUC).

**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)**

leave-one-subject-out validation

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

Then, the performance of a logistic regression model in leave-one-subject-out vali-dation was evaluated in terms of sensitivity, specificity, positive and negative predictive values (PPV and NPV, respectively), accuracy, and area under the curve (AUC).

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

All enrolled patients meeting the study’s eligibility criteria completed the experimental procedures without any dropouts.

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

No missing data were encountered in the dataset

**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?**

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?**

U

**Risk of bias introduced by the analysis**

H

**Rationale of bias rating**

71 patients is very little for Machine learning.

**Overall Risk of bias**

High