**PROBAST**

Study:

Monitoring patients with diabetes using wearable sensors: predicting glycaemias using ECG and respiration rate

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?**

Predict hypo- and hyperglycemia

Step 3: Assess risk of bias

**Domain 1: Participants**

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

The study encompassed 30 patients with type II diabetes from Poland and 22 patients with type I diabetes from Italy.

**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 further exclusion criteria mentioned.

**Domain 2: Predictors**

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

Each patient was equipped with Zephyr BioHarness [5] that records ECG, respiration rate, and acceleration, with a GlucoTel [6] glucose monitor.

**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**

Independent of outcome and applicable and assessed similarly.

**Domain 3: Outcome**

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

The glucose level measurements contain, apart from the glucose level value itself, also the information whether the measurement was done before or after a meal, before sleep, at night, or other. Patients were left to decide when to take the measurements. Based on the glucose levels, the measurements were sorted into three groups: hypo-, hyperglycemia, normal glycemia

**3.1 Was the outcome determined appropriately?**

N

**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?**

PN

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

PN

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

Y

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

High

**Rationale of bias rating**

High risk as patients can decide when to take the measurements. Not standardized over all patients.

**Domain 4: Analysis**

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

52 patients. 566 hours of continuous measurement data. Unclear how many different hypo- and hyperglycemic events

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

Various different approaches are described. Logistic regression for T1D

**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)**

10-fold CV

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

Accuracy

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

None/not described

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

To obtain clean data we first processed the ECG and the respiration rate measurements using filters which removed the noisy and unreadable parts, but nevertheless retained the signal morphology. After filtering, we were left with approximately 566 hours of clean ECG and respiration data.

**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?**

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?**

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**

H

**Rationale of bias rating**

Unclear how many outcomes there are. No evaluation metric for imbalanced results.

**Overall Risk of bias**

High