**PROBAST**

Study:

Feature-Based Machine Learning Model for Real-Time Hypoglycemia Prediction

Step 2: Type of prediction study

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

Prognostic

**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?**

Hypoglycemia

Step 3: Assess risk of bias

**Domain 1: Participants**

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

Patients with T1D.

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

**U**

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

U

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

Unclear

**Rationale of bias rating**

Patient population not described

**Domain 2: Predictors**

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

CGM features, demographic features, time features

**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**

Only past and present values were used for forecasting

**Domain 3: Outcome**

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

Hypoglycemic events at different prediction horizons. <70mg/dL

**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**

Standard outcome for hypoglycemia forecasting.

**Domain 4: Analysis**

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

The CGM datasets were obtained from 112 patients using Dexcom G6 CGM devices over a range of 90 days consisting of over 1 639 921 CGM values under normal living conditions.

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

Feature selection for LR was performed by adding a Least Absolute Shrinkage and Selection Operator (LASSO) penalty.

Two approaches were considered for prediction: (i) Logistic Regression (LR) and (ii) Random Forests (RF).

**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)**

Seventy percent training and 30% testing partition were randomly repeated 10 times and performance results averaged across these 10 replications to generate robust estimates for sensitivity and specificity.

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

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**

Not described

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

Y

**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?**

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?**

PY

**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**

Unclear

**Rationale of bias rating**

The missing data approach is not described. Also it is not clear how many patients or how many datapoints were removed due to missing data. There could also be more metrics such as AUC.

**Overall Risk of bias**

Unclear