**PROBAST**

Study:

A Predictive Machine Learning Tool for Asthma Exacerbations: Results from a 12-Week, Open-Label Study Using an Electronic Multi-Dose Dry Powder Inhaler with Integrated Sensors

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?**

XGBoost

**What is the outcome of interest?**

Asthma exacerbations

Step 3: Assess risk of bias

**Domain 1: Participants**

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

This 12-week, open-label study was conducted across 45 study centers in the USA between February 2017 and February 2018. The study consisted of a 2-week screening period and a 12-week intervention period. The study population comprised adult patients with a physician diagnosis of asthma, at least one moderate or severe asthma exacerbation over the 12 months prior to screening, and poorly controlled asthma as defined by an Asthma Control Questionnaire-5 (ACQ-5) score of ≥1.5.

All patients were required to be on moderate or high doses of inhaled corticosteroids, equivalent to at least 440 μg daily of fluticasone propionate, with or without other asthma maintenance medications (long-acting beta2-agonist, leukotriene antagonist, long-acting antimuscarinic agent, biologic, or maintenance oral corticosteroids). Patients were excluded if they had any confounding underlying lung disorder other than asthma or had used any investigational drugs within five half-lives of discontinuation.

**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?**

PY

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

Low

**Rationale of bias rating**

Somewhat limiting eligibility criteria but still adequate

**Domain 2: Predictors**

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

The primary measure for the predictive analysis was albuterol use, and parameters of interest included total number of inhalations in the days preceding an exacerbation peak (defined as the day on which the patient began using SCS), the number of days prior to exacerbation peak during which albuterol use increased, and the amount of albuterol use in the 24 hours prior to a moderate or severe exacerbation.

**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 independent and applicable

**Domain 3: Outcome**

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

Moderate exacerbations were those involving worsening asthma and requiring administration of systemic corticosteroids (SCS) above baseline for at least 3 days, or an unscheduled HCP visit (eg, doctor’s office or emergency care) associated with an increase in asthma therapy. Severe exacerbations were those requiring both administration of SCS as above and an unscheduled HCP visit.

The target of the predictive model was defined as the prediction of an exacerbation within the following 5 days.

The study had several limitations. Digihaler data were only downloaded and analyzed at the end of the study, and exacerbations were identified by monthly patient phone calls, potentially resulting in recall bias and unidentified exacerbations, which may have affected the precision of the model.

**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?**

N

**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**

High

**Rationale of bias rating**

Multiple options are given for the outcome which may be of different nature. Still adequate for asthma. However as written in the Discussion there may be recall bias.

**Domain 4: Analysis**

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

Patients who completed the study with at least one valid inhalation using the Digihaler (n=360 [91%]) were eligible for inclusion in the predictive analysis dataset.

Of the 360 patients who made ≥1 valid inhalation and completed the study, and so were eligible for inclusion in the predictive analysis, 64 (18%) experienced a total of 78 moderate/severe exacerbations.

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

To develop the predictive model, machine learning techniques were applied to a combination of case report form data taken on study Day 1 (age, body mass index, blood pressure, previous exacerbations, and the number of exacerbations and hospitalizations in the previous 12 months), data from the Digihaler prior to (and including) the day of the prediction, and patient baseline characteristics from the Digihaler (timestamp of inhalation, inhalation status, PIF, inhalation volume, time to PIF, and inhalation duration). The number of inhalations and mean (standard deviation [SD]) of each inhalation parameter during the first 10 days of the study were considered as baseline features for the predictive model. A feature engineering process was conducted to determine the most relevant features for the 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)**

Patients were randomly divided into three groups to train the model (“training set”), test and optimize the model (“test set”), and validate the chosen model (“validation set”). A 4-fold cross validation technique was used to compare the predictive performance metrics of the algorithms.

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

AUC score

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

Patients who completed the study with at least one valid inhalation using the Digihaler (n=360 [91%]) were eligible for inclusion in the predictive analysis dataset. Excluded from the predictive analysis dataset were patients who experienced an exacerbation during the first 10 days of the study (n=6), and patients who made no inhalations from the Digihaler during this period (n=47), during the period after the first 10 days (n=2) or during the 4-day period preceding an exacerbation (n=7). The predictive analysis population, therefore, comprised 298 patients (Figure 3).

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

**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?**

PY

**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?**

PN

**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**

Only 64 exacerbations. Missing data handling not described. Only AUC score reported

**Overall Risk of bias**

High