**PROBAST**

Study:

Machine learning for prediction of ventricular arrhythmia episodes from intracardiac electrograms of automatic implantable cardioverter-defibrillators.

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?**

Convolutional neural network

**What is the outcome of interest?**

Prediction of future ventricular arrythmias

Step 3: Assess risk of bias

**Domain 1: Participants**

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

The study included patients from Biotronik’s CERTITUDE registry, which is a postmarket, real-world database consisting of remote cardiac device monitoring data from qualifying patients who have been registered to Biotronik Home Monitoring® (Biotronik SE & Co KG, Berlin, Germany) and have provided authorization for the use of their device’s data for research purposes. This study collected data from CERTITUDE patients with im-planted Biotronik ICDs between January 1, 2010, and December 31, 2020.

**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 eligibility criteria given.

**Domain 2: Predictors**

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

This study collected data from CERTITUDE patients with im-planted Biotronik ICDs between January 1, 2010, and December 31, 2020. The IEGM data were collected from periodically transmitted recordings, and the device detected ventricular and atrial episode events. Transmitted IEGM data are lossy com-pressed with antialiasing, high-pass filtering, and downsampling.

The focus of data extraction and analysis was IEGMs. The stored data included scheduled IEGMs recorded periodically by the ICD, mainly at nighttime in preprogrammed intervals between 0 and 100 days. Most recordings were recorded within 90-day intervals (every 3 months remote monitoring per standard practice and reimbursement). The unscheduled recordings were triggered whenever an episode of sustained VA was detected. The stored IEGM during the episode included the signals right before (5 seconds), during, and after the arrhythmia was resolved by antitachycardia pacing or shock therapies.

**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**

IEGM recordings were used to predict future events so no dependence on outcome. All patients had Biotronik devices so similar assessment.

**Domain 3: Outcome**

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

The recordings were accompanied by related data information, including anonymized patient identification, device name and model, arrhythmia type, time from implantation, time of day, and specified rhythm markers (such as ventricular sensing, arrhythmia onset, VT, and VF).

**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**

Outcome definition was done by an internal device algorithm from Medtronic. It was done the same way for every patient and was independent of prior recordings.

**Domain 4: Analysis**

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

The study included 13,516 patients. Of301,647 IEGM recordings collected from these participants,27,845 episodes of sustained VT or VF were observed in4467 patients (33.0%)

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

The models used for the prediction of the different scenarios were convolutional neural network classifiers whose architecture was inspired by ResNet architecture.10,14 Thei nput was 1280 samples (10 seconds) in the long-range and mid-range scenarios. The training stage network was fed with continuous 8 seconds chosen randomly for each cycle from the 10 seconds. This jittering process aims to make the network phase invariant and to regulate the model when the test signal is presented. Random jittering could not be done for the short-range predictions in which only 2 seconds were presented to the network.

**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)**

Data were randomly divided into 3 distinct sets by patient identification number: training (70%), tuning and validation(10%), and testing (20%).

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

AUC

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

None

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

The following data were excluded: the IEGMs representing ventricular paced rhythm in patients with ventricular pacemaker dependency and cardiac resynchronization to avoid the influence of paced signals on data; devices with episodes for which either metadata or IEGM data were missing;

**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?**

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?**

PN

**4.8 Were model overfitting and optimism in model performance accounted for?**

N

**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**

High

**Rationale of bias rating**

No proper validation method applied. Only AUC score reported.

**Overall Risk of bias**

High