Unravelling the black box of machine learning for atrial fibrillation forecasting: role of heart rate variability and premature beat counting

Jean-Marie Grégoire[1,2], Cédric Gilon[1], Stéphane Carlier[2], Hugues Bersini[1]
[1] IRIDIA, Université Libre de Bruxelles, Brussels, Belgium
[2] Faculty of Medecine and Pharmacy, UMONS, Mons, Belgium

Background

The identification of patients still in sinus rhythm who will present one month later an atrial fibrillation (AF) episode is possible using machine learning (ML) techniques. However, these new ML algorithms do not provide any relevant information about the underlying pathophysiology.
Preventive pacing algorithms implemented in pacemakers are mainly based on premature beats (PB) and have given disappointing results.

Purposes

To compare the predictive performance for forecasting AF between a machine learning algorithm and other parameters with known pathophysiological mechanisms triggering arrhythmia (i.e. the count of PB) and HRV.

Materials and method

We conducted a retrospective study from an outpatient clinic. In 10484 Holter ECG recordings screened, 250 analyzable AF onsets were labelled. We developed a deep neural network model (composed of convolutional neural network layers and bidirectional gated recurrent) that was trained for the forecast of paroxysmal AF episodes, using 300 RR intervals windows. This model works like a black box. For comparison purposes, we used a "random forest" (RF) model of ML to obtain forecast results using HRV parameters with (input = 300 RR) and without PB (input = 300 NN). These models allow the evaluation of the forecasting relevance of HRV parameters and PB. We calculated the area under the curve (AUC, see Table 1) of the receiver operating characteristic curves for different time windows counted in RR intervals before the AF onset.

Results

According to our dataset, the forecasting value of the deep neural network model was not statistically superior to the random forest algorithm. Prediction value of both decreased when analyzing RR intervals further away from the onset of AF. Most of the predictive information was found in HRV.

Number of RR intervals before AF	ML AUC (%) (Black box model)	HRV+PB AUC (%) (RF model)	HRV AUC (%) (RF model)	P value (anova)
0-30	0.71 (0,69-0,73)	0.70 (0.69-0.72)	0.69 (0.68-0.71)	NS
31-60	0.66 (0.64-0.68)	0.66 (0.65-0.68)	0.65 (0.64-0.67)	NS
61-300	0.60 (0.58-0.62)	0.60 (0.59-0.62)	0.60 (0.58-0.62)	NS

Table 1: Results of the 3 models
AUC : Area Under the Curve of the Receiver Operating Characteristics curve
NS : non-significant using two-way ANOVA: hypothesis; values in parentheses indicate 95% confidence intervals

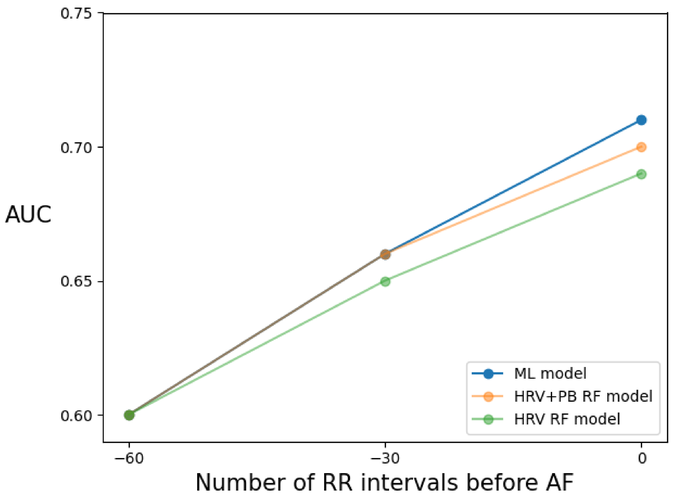

Fig. 1 AUC of evolution of the 3 models

Conclusion

These results suggest that HRV plays a predominant role in triggering AF episodes. Premature beats add limited additional information according to the random forest model. Moreover, the closer the window from the AF onset, the better the accuracy, regardless of the method used. Detection algorithms based on HRV might prove useful to prevent AF by changing pacing sequence while patients would still be in sinus rhythm. However this remains to be demonstrated.

Conflicts of interest: none