

Distant Prediction of Paroxysmal Atrial Fibrillation Using HRV Data Analysis

YV Chesnokov¹, AV Holden², H Zhang²

¹Kuban State University, Krasnodar, Russia

²University of Manchester, Manchester, UK

Abstract

The idea of this research is to determine how long in advance can we predict the onset of paroxysmal atrial fibrillation (PAF) from the HRV data. Having established such methods with great confidence it is possible to avert the PAF onset using pacing techniques, thus relieving patient pains.

We used power spectral density as an estimate on the 30 minute HRV segment. Having analyzed AFPDB records immediately before PAF and distant from it we found statistically significant increase over entire 0.0 – 0.5Hz frequency range for the data immediately before PAF. We trained neural network classifier to distinguish between immediately before PAF and distant HRV segments on the AFPDB database.

For testing of distant prediction we used AFDB database records from 16 patients and achieved mean time of fibrillation prediction 75 ± 43 minutes for 14 patients.

1. Introduction

Paroxysmal atrial fibrillation (PAF) is the most common abnormal heart rhythm encountered in clinical practice, and has serious associated morbidity and mortality as a sudden stroke. Distant prediction of PAF onset can help to avoid it using pacing techniques. Advanced pacing modes turned on only when PAF onset is imminent for the longevity of battery driven implantable devices would be advantageous [1].

2. Methods

The data for analysis was taken from Physionet databases. We used atrial fibrillation prediction database (AFPDB), consisted of 30 minute ECG segments from patients immediately before PAF onset and distant ones, and MIT-BIH AF database (AFDB), consisted of 10 hour recordings with PAF and non-PAF heart rhythms.

We annotated each ECG record using our developed algorithm [2] and extracted HRV data. We used entire 30 minute segment for analysis from AFPDB database. 10

hour AFDB database records were divided to consecutive overlapping 30 minute segments.

Next HRV data was processed with spectral analysis and further automatic classification with artificial neural networks (ANN), which we developed in C++. Statistical hypothesis testing was implemented in Matlab (Statistics Toolbox).

Obtained HRV segments were interpolated to 2Hz and analyzed with power spectral density (PSD) estimation (2.1) in the 0.0 – 0.5Hz frequency range averaging over 0.01Hz frequency span, resulting in the total number of 50 consecutive bins.

$$h(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r(k) \exp(-i\omega k) \quad (2.1)$$

where $r(k)$ is the autocovariance function (2.2).

$$r(k) = E\{(x(t) - \mu)(x(t+k) - \mu)\} \quad (2.2)$$
$$k = 0, \pm 1, \pm 2, \dots$$

where $x(t)$ is the stationary process defined at discrete values $t = 0, \pm 1, \pm 2, \dots$

In order to obtain automatic classification we applied artificial neural networks (feed-forward full-connectionist, with sigmoid activation rule) on the PSD spectrum.

We used backpropagation algorithm with momentum for ANN classifier training. The output y of the single ANN layer is calculated as:

$$y = f(Wx + b), \quad (2.3)$$

where W is the matrix of the layer neurons weights, x – input vector, b – bias weight, f – activation function.

We used sigmoid function as the activation rule:

$$f(x) = \frac{1}{1 + \exp(-x)}, \quad (2.4)$$

The backpropagation algorithm iteration weights update for single layer neurons weights matrix W is defined as:

$$\Delta w_{ij}(t) = (1 - \alpha)\eta x \delta + \alpha \Delta w_{ij}(t-1), \quad (2.5)$$

where α is the momentum, η – learning rule, δ – neuron error.

Input data fed to ANN classifier was normalized with z-score formula (zero mean and unit variance):

$$x_i = \frac{x_i - \mu}{\sigma}, \quad (2.6)$$

where μ is the mean and σ is dispersion of the PSD spectrum calculated from training set (these values were used in the ANN input layer for test data).

We used Sensitivity (Se) and Specificity (Sp) as a classification results evaluation formulas. The Se is defined as:

$$Se = \frac{TP}{TP + FN}, \quad (2.7)$$

where TP (true positives) is the number of correct classifications for positive cases (HRV segments leading to PAF onset correctly classified), FN (false negatives) is the number of misclassifications for positive case being incorrectly classified as negative (HRV segments leading to PAF incorrectly classified as distant from it).

$$Sp = \frac{TN}{TN + FP}, \quad (2.8)$$

where TN (true negatives) is the number of correct classifications for negative cases (HRV segments distant from PAF correctly classified), FP (false positives) is the number of misclassifications for negative case being incorrectly classified as positive (HRV segments distant from PAF incorrectly classified as leading to it).

It is imperative not to miss leading to PAF segments, for our case to predict PAF onset, thus we need as high Se as possible for our method. However, low Sp is tolerable since we would rather be interested to predict PAF onset than to miss prediction.

During cross-validation process of ANN classifier training we used geometric mean metric, which allows obtaining both high Sensitivity and Specificity of the classifier.

$$gm = \sqrt{Se * Sp}, \quad (2.9)$$

3. Results

For the HRV annotation we selected records from AFPDB database without to much corruption with noise. We used both ECG leads from the records with the names of the form p^* and t^* . Obtained HRV data was carefully inspected for the quality of annotation. Total number of 30 minute HRV segments distant from PAF we annotated is equal to 67, the number of HRV segments immediately before PAF is equal to 85. The PSD spectrum of the HRV data from AFPDB database is shown in the fig. 1.

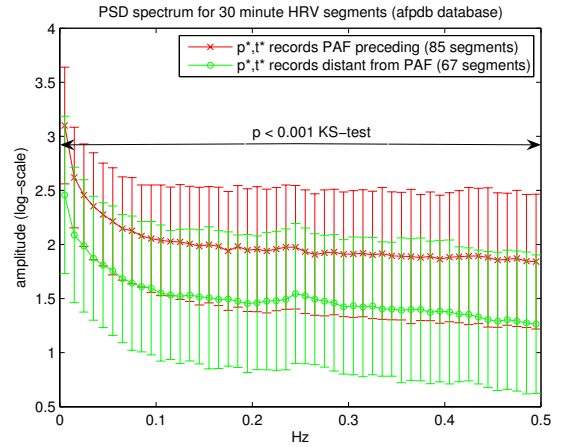


Fig. 1. PSD spectrum for 30 minute HRV segments from AFPDB database. p^* , t^* records – immediately preceding PAF onset. p^* , t^* records – distant from PAF. Statistically significant PSD energy increase ($p < 0.001$, KS-test) for HRV data immediately before PAF. (error bars – mean \pm std).

We can see that there is statistically significant increase ($p < 0.001$, KS-test) over the entire frequency range for the PSD estimate of HRV data immediately before PAF event compared to distant from PAF HRV segments.

First we trained ANN classifier on the HRV data from AFPDB database to distinguish between segments immediately before PAF and distant ones. Then we used that classifier on the HRV data from AFPDB database to predict PAF onset on the long-term data excerpts before fibrillation onset.

We used ANN classifier consisted of 5 layers (50, 15, 10, 5, 1 neurons correspondingly) to distinguish between immediately preceding PAF 30 minute segments and distant ones. We used z-score normalization input layer and geometric mean as the validation metric.

The AFPDB data was randomly split to half for training and half for validation and testing to prevent overfitting. The Sensitivity and Specificity we achieved

for training half Se: 90.4%, Sp: 100%. Validation set Se: 90.4%, Sp: 93.7% and test set Se: 72.7%, Sp: 88.2%.

For the distant prediction of PAF we used AFDB database records from 16 patients. We split their HRV data to segments to 30 minute length with overlapping window of 5 minute stride, computed PSD spectrum for every segment and classified them with trained on the AFPDB data ANN classifier to determine how long in advance the classifier will start to produce positive output as PAF event is imminent. From 16 patients we achieved prediction on 14 with mean time of prediction 75 ± 43 minutes (table 1).

Patient	Times of data before PAF	Prediction time minutes in advance
04043	15:25:04-17:51:20	60 min
04048	12:48:15-14:22:12	40 min
	16:36:30-20:31:05	all positives
04126	11:56:28-19:41:08	positives after PAF
04098	15:00:00-16:20:47	all negatives
	17:08:40-18:45:37	all negatives
	19:07:18-00:31:44	90 min
05091	14:11:02-15:49:00	30 min
05121	16:00:00-18:01:52	35 min
05261	12:00:00-14:23:09	120 min
	14:23:13-16:29:30	50 min
	16:52:20-18:05:27	all positives
06453	20:00:00-20:50:54	all positives
	21:46:02-04:07:23	165 min
06955	14:03:15-16:11:36	all positives
	16:11:51-18:57:16	positives after PAF
07879	22:00:00-23:57:51	50 min
07910	15:49:33-22:08:23	30 min
08215	16:00:00-17:57:17	30 min
08219	20:00:00-23:24:10	175 min
08405	00:00:00-02:48:39	85 min
08434	20:00:00-22:28:07	70 min
08455	00:00:00-03:06:12	70 min
		Mean: 75 ± 43 min

Table 1. HRV excerpts before PAF onset and prediction times. (positives after PAF – only beginning of the HRV excerpt after previous PAF event was positively classified; all positives – every segment classified as leading to PAF; all negatives – every segment classified as not leading to PAF).

The rest of the HRV data segments were classified as negative. However in several records there was positive ANN output immediately after PAF event for several 30 minute segments, when the PAF preceding data excerpt started after previous PAF event. This could be explained that HRV activity with increased PSD energy remains after PAF event and was spotted by ANN classifier. The

corresponding data times of HRV segments before PAF for every patient and prediction time are shown in the table 1.

For 08405 patient, the 30 minute data immediately before fibrillation is present at the AFPDB database (p38.dat record, which was used during the training session). On this patient we achieved fibrillation prediction 85 minutes in advance (fig. 2).

4. Discussion and conclusions

Distant prediction of PAF onset is presented using PSD estimate of the HRV data. Our results well agree with the recent research reported in the literature, however we propose the usage of AI classification methods for distant prognosis of PAF onset, which is suitable for implantable devices.

Computers in Cardiology challenge 2001 participants reported results on immediate (30 minutes before) PAF prediction using AFPDB database with accuracy of 80% using PAC number and P wave variability as predictors.

Independent research also presents interesting results for distant prediction of PAF which could be classified to two categories: complexity and spectral measures on HRV data.

Shin et al [3] used approximate entropy (ApEn) and sample entropy (SmEn) complexity metrics on HRV data. They report statistically significant decrease ($p < 0.05$) both in ApEn and SmEn during 1 hour period before PAF onset. The same ApEn metric along with short-term scaling exponent $\alpha(1)$ are reported in [4] applied to HRV data. They found significant decrease ($p < 0.001$) in these metrics on the HRV data during 2 hour period before PAF.

Spectral analysis of HRV data reported in the literature reveals 3 types of spectral changes before PAF: increase in the LF band (0.04–0.15Hz) only, increase in the HF band (0.15–0.4Hz) only and energy increase in both bands simultaneously. Bettoni [5] found significant HF band increase ($p < 0.001$) during 60 and 20 minutes HRVs before PAF. Amar et al [6] compared 2 hour HRV data before AF with matched controls distant from it. They found both LF and HF energy increase ($p < 0.001$) for the data 2 hour before AF. Tomita et al [7] found gradual increase in the LF and HF bands started 60 minutes before PAF. Tomoda et al [8] also found gradual increase in the LF band of the HRV data started 30 minutes before PAF, but reported no significant changes in the HF band. Similar findings of increased HF energy started 40 minutes before PAF are reported in [9]. Vincenti et al [10] reported vagal prevalence in the 5 minute preceding the onset of PAF both in time and frequency domains. Fioranelli et al [11] provides their findings for the 3 types of spectral changes in the HRV dynamics before PAF onset.

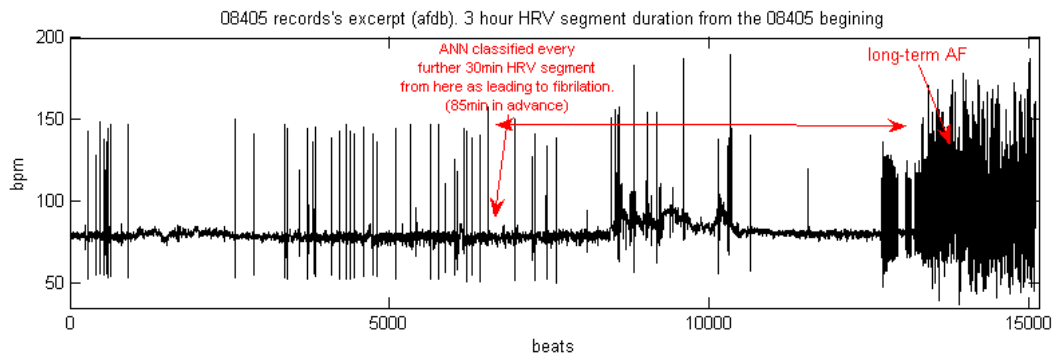


Fig. 2. 08405 patient 3 hour HRV excerpt. 30 minute segment from this record immediately before AF is present at the AFPDB database (p38.dat record, which was used for training ANN classifier). We achieved 85 min prediction in advance for this long-term PAF event.

They compared 60 minute and 5 minute distant times from PAF and found energy increase in the LF band and decrease in the HF band ($p=0.004$) type A, decrease in the LF band and increase in the HF band ($p<0.001$) type B, and in some patients they found simultaneous increase in both LF and HF bands. Herweg et al [12] analyzed HRV samples taken 5, 10 and 20 minutes before fibrillation and compared them. They found increase in the HF band compared 5 to 10 minutes and 10 to 20 minutes before fibrillation onset ($p<0.0001$).

Comparing our own findings in PSD distribution for the HRV data from AFPDB database we conclude that patients from that database correspond to the type with simultaneous energy increase in LF and HF bands. It is possible that missed predictions from AFDB could be attributed to the presence of type A and type B patients. Having included such datasets from both types to the training of ANN classifier, along with complexity metrics, prediction rate would be much better. For the further investigation shorter time segments of 5 – 15 minutes could be examined for automatic classification.

Acknowledgements

I'd like to express my gratitude to my former colleagues Robert Glen, Dmitry Nerukh (Cambridge University), Ian Wilkinson, Carmel McEniery (Addenbrooks Hospital) and Medical Research Council who provided support for this research during my work in Cambridge University and my PhD supervisor at the Kuban State University.

References

- [1] Thong T, et al. Prediction of paroxysmal atrial fibrillation by analysis of atrial premature complexes. *IEEE Trans Biomed Eng* 2004;51(4);561-9.
- [2] Chesnokov YV, Nerukh D, Glen RC. Individually adaptable automatic QT detector. *Computers in Cardiology* 2006;33.
- [3] Shin DG, et al. Prediction of paroxysmal atrial fibrillation

using nonlinear analysis of the R-R interval dynamics before the spontaneous onset of atrial fibrillation. *Circ J* 2006;70(1);94-9.

- [4] Vikman S, et al. Altered complexity and correlation properties of R-R interval dynamics before the spontaneous onset of paroxysmal atrial fibrillation. *Circulation* 1999;100(20);2079-84.
- [5] Bettoni M, Zimmermann M. Autonomic tone variations before the onset of paroxysmal atrial fibrillation. *Circulation* 2002;105(23);2753-9.
- [6] Amar D et al. Competing autonomic mechanisms precede the onset of postoperative atrial fibrillation. *J Am Coll Cardiol* 2003;42(7);1262-8.
- [7] Tomita T et al. Role of autonomic tone in the initiation and termination of paroxysmal atrial fibrillation in patients without structural heart disease. *J Cardiovasc Electrophysiol* 2003;14(6);559-64.
- [8] Tomoda Y et al. Assessment of autonomic nervous activity before the onset of paroxysmal atrial fibrillation. *J Cardiol* 1998;31(1);11-7.
- [9] Huang JL et al. Changes of autonomic tone before the onset of paroxysmal atrial fibrillation. *Int J Cardiol* 1998;66(3);275-83.
- [10] Vincenti A et al. Onset mechanism of paroxysmal atrial fibrillation detected by ambulatory Holter monitoring. *Eurospace* 2006;8(3);204-10.
- [11] Fioranelli M et al. Analysis of heart rate variability five minutes before the onset of paroxysmal atrial fibrillation. *Pacing Clin Electrophysiol* 1999;22(5);743-9.
- [12] Herweg B et al. Power spectral analysis of heart period variability of preceding sinus rhythm before initiation of paroxysmal atrial fibrillation. *Am J Cardiol* 1998;82(7);869-74.

Address for correspondence

Yuriy V Chesnokov
Kuban State University, Krasnodar, Russia
chesnokov.yuriy@gmail.com
chesnokov_yuriy@mail.ru