Screening Patients with Paroxysmal Atrial Fibrillation (PAF) from Non-PAF Heart Rhythm Using HRV Data Analysis

YV Chesnokov[1], AV Holden[2], H Zhang[2]

[1]Kuban State University, Krasnodar, Russia
[2]University of Manchester, Manchester, UK

Abstract

The idea of this research is to determine can we tell from the HRV data without paroxysmal atrial fibrillation present at the recording if the patient suffers from this arrhythmia. The benefit is we can provide time and cost effective preliminary screening procedure during short time visit to the clinic.

To achieve this goal we used Fourier analysis of the 30 minute HRV segment duration. We found statistically significant increase in the energy above 0.1Hz for the patients with documented PAF history. This suggests that people with this arrhythmia has increased parasympathetic activity.

For automatic classification of the patient we trained artificial neural networks on the HRV Fourier spectrum of AFPDB database. Testing on the AFDB (66.5 hours of HRV data from PAF patients) and NSRDB (352 hours of HRV data from healthy ones) databases produced Se 94.5% and Sp 96.5%.

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. As PAF occurrence usually hard to catch using conventional ECG recording during short visit to a clinic, screening if a patient is prone to PAF from non-PAF heart rhythm would facilitate diagnosis. The screening is especially valuable for patients with heart diseases as hypertophic cardiomyopathy (HCM) and abnormal conditions that could lead to development of PAF: hypertension, hyperthyroidism, etc. To achieve this goal we studied non-PAF heart rhythms from PAF documented patients and patients without that disease.

2. Methods

The data for analysis was taken from Physionet databases. We used atrial fibrillation prediction database (AFPDB), consisted of 30 minute non-PAF ECG segments from PAF patients, healthy controls and diseased patients without PAF, MIT-BIH AF database (AFDB), consisted of 10 hour recordings from PAF suffering patients with PAF and non-PAF rhythms and normal sinus rhythm database (NSRDB), consisted of 24 hour recordings.

We annotated each ECG record using our own developed algorithm [1] and extracted HRV data. We used entire 30 minute segment for analysis from AFPDB database. Long-term records from AFDB and NSRDB databases 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 processed with Fourier analysis (FFT) estimation (2.1) in the 0.01 – 0.5Hz frequency range averaging over 0.01Hz frequency span, resulting in the total number of 49 consecutive bins.

$$X(\omega) = \sum_{t=1}^{m} x(t) \exp(-i\omega t), \quad \pi \le \omega \le -\pi \qquad (2.1)$$

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

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

$$y = f(Wx + b), \qquad (2.2)$$

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

We used sigmoid function as the activation rule:

$$f(x) = \frac{1}{1+\exp(-x)}, \qquad (2.3)$$

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

$$\Delta w_{ij}(t) = (1-\alpha)\eta x\delta + \alpha\Delta w_{ij}(t-1), \qquad (2.4)$$

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):

$$x_i = \frac{x_i - \mu}{\sigma}, \qquad (2.5)$$

where μ is the mean and σ is dispersion of the FFT spectrum calculated from the training set (these values were used as the preprocessing in the ANN input layer).

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

$$Se = \frac{TP}{TP+FN}, \qquad (2.6)$$

where TP (true positives) is the number of correct classifications for positive cases (HRV segments from unhealthy patients correctly classified), FN (false negatives) is the number of misclassifications for the positive case being incorrectly classified as negative (HRV segments from unhealthy patients incorrectly classified as healthy).

$$Sp = \frac{TN}{TN+FP}, \qquad (2.7)$$

where TN (true negatives) is the number of correct classifications for negative cases (HRV segments from patients without PAF history correctly classified), FP (false positives) is the number of misclassifications for negative case being incorrectly classified as positive (HRV segments from patients without PAF history incorrectly classified as belonging to the patients with PAF history).

In medical diagnosis it is imperative not to miss unhealthy patients, for our case to identify patients with probable PAF, thus we need as high Se as possible for our method. However, low Sp is tolerable, and the suspect patients could be investigated with additional methods as ultrasound, long-term ECG recording etc.

During cross-validation process of ANN classifier training we used geometric mean metric, which allows obtaining both high Sensitivity and Specificity of the classifier in the case of biased training data distribution, when we have limited number of patients with PAF history and more patients without that arrhythmia.

$$gm = \sqrt{Se * Sp}, \qquad (2.8)$$

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 n*, p* and t*. Obtained HRV data was carefully inspected for the quality of annotation. Total number of 30 minute HRV segments (free from PAF rhythm) from PAF patients we annotated is equal to 136, the number of HRV segments from the patients not suffering from PAF is equal to 118. The FFT spectrum of the HRV data from AFPDB database is shown in the fig. 1.

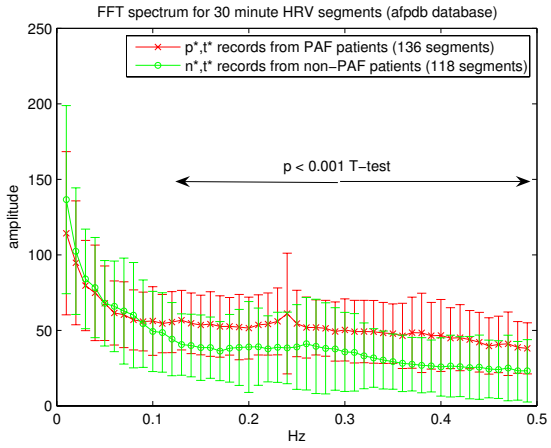

Fig. 1. FFT spectrum for 30 minute HRV segments from AFPDB database. p*, t* records from the patients with documented PAF (136 segments) and n*, t* records from the patients without PAF (118 segments). There is statistically significant ($p<0.001$, T-test) increase in the 0.1 – 0.5Hz frequency range for the patients with documented PAF history. (error bars – mean ± std).

We can see that there is statistically significant increase ($p<0.001$, T-test) in the frequency range 0.1 – 0.5Hz for the patients with documented PAF history compared to the ones without this arrhythmia. Below 0.1Hz there is no statistically significant ($p>0.05$, T-test)

difference.

We compared HRV FFT spectra from AFDB and NSRDB databases to the ones from AFPDB. From AFDB database we annotated 16 subjects (table 1) with the total of 66.5 hours (667 overlapping 30 minute segments) of non-PAF rhythm. As the duration of the non-PAF rhythm not restricted to 30 minute length as in AFPDB, we used overlapping window with 5 minute stride. From the NSRDB we used also 16 subjects (table 2, overlapping window with 10 minute stride).

The error bar FFT plots are shown in the fig. 2 and fig. 3.

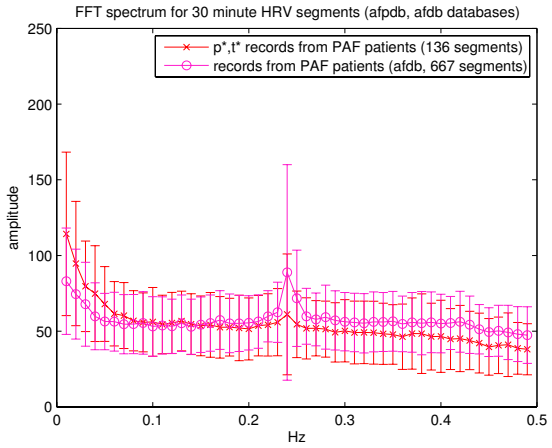

Fig. 2. FFT spectrum for 30 minute HRV segments from AFDB and AFPDB databases. p*, t* records from the patients with documented PAF (136 segments, AFPDB) and records from 16 patients (667 segments) from AFDB. There is close correspondence between two databases spectra.

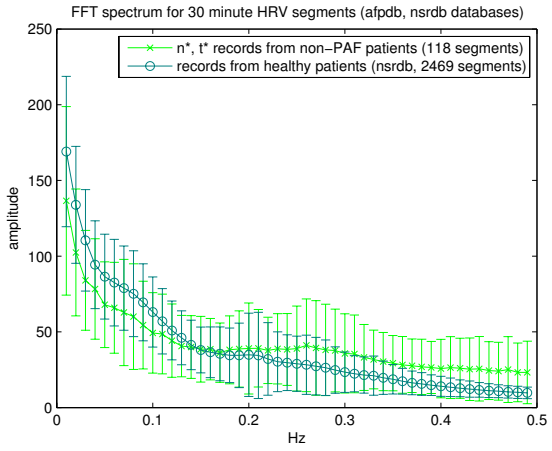

Fig. 3. FFT spectrum for 30 minute HRV segments from NSRDB and AFPDB databases. n*, t* records from the patients without PAF (118 segments, AFPDB) and records from 16 healthy patients (2469 segments) from NSRDB. There is close correspondence between two databases spectra.

We can see that FFT spectra closely resemble the ones from AFPDB database. AFDB non-PAF rhythm HRVs has the same peak around 0.24Hz (fig. 2). Otherwise AFDB has slight elevation above 0.27Hz and small degradation below 0.05Hz compared to AFPDB spectra.

First we trained ANN on the non-PAF HRV data from AFPDB database to distinguish between patients prone to PAF and non-PAF patients. Three healthy patients from NSRDB were also added. Then we applied that ANN model to AFDB (on the non-PAF HRV data) and NSRDB database for testing.

ANN consisted of 5 layers (49, 15, 10, 5, 1 neurons correspondingly). 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 (Se), Specificity (Sp) we achieved for training half Se: 95.5%, Sp: 92.4%. Validation set Se: 100%, Sp: 91.6% and test set Se: 91.6%, Sp: 92.1%.

This trained ANN classifier was then used on the AFDB and NSRDB 16 patients for automatic classification. Mean Sensitivity on the 30 minute per-segment classification for 16 AFDB patients was 94.5% and mean Specificity for 16 NSRDB patients was 96.5%. Classification rates for individual subject from AFDB and NSRDB are shown in the table 1 and table 2 correspondingly.

Patient	Non-PAF rhythm total times	Sensitivity
04043	87 minutes	100%
04048	328 minutes	97.92%
04126	465 minutes	100%
04098	500 minutes	73.49%
05091	97 minutes	53.85%
05121	120 minutes	100%
05261	340 minutes	97.92%
06453	430 minutes	100%
06955	290 minutes	100%
07879	120 minutes	88.89%
07910	380 minutes	100%
08215	120 minutes	100%
08219	200 minutes	100%
08405	170 minutes	100%
08434	150 minutes	100%
08455	190 minutes	100%
	Total: 66.45 hours	Mean: 94.5%

Table 1. Per-segment classification sensitivity for 16 PAF documented patients from AFDB database.

Patient	Non-PAF rhythm total times	Specificity
04043	~20-24 hours	100%
04048	–	98.35%
04126	–	94%
04098	–	100%
05091	–	93.53%
05121	–	81.25%
05261	–	97.1%
06453	–	97.54%
06955	–	93.6%
07879	–	100%
07910	–	92.42%
08215	–	100%
08219	–	100%
08405	–	97.79%
08434	–	99.57%
08455	–	98.73%
	Total: ~352 hours	Mean: 96.5%

Table 2. Per-segment classification specificity for 16 healthy patients from NSRDB database.

4. Discussion and conclusions

HRV data spectral analysis is presented as the simple method for preliminary risk assessment of PAF. Results we achieved on 32 patients from independent test databases are encouraging.

Current research in this field can be divided to time-domain and frequency-domain analysis applied to either ECG data or derived from it HRV and PP indices. The most common parameters with statistically significant differences separating controls and PAFs are: P wave duration, P wave dispersion, left atrial (LA) diameter, root mean square (RMS) voltage of the P wave, atrial early potentials (EP), P wave spectral areas ratios. These parameters are used with SAECG and high resolution ECG for PAF risk assessment of hypertensive patients, HCM, hyperthyroidism patients. Reported results on these markers for the researchers own datasets present Sensitivity in the range of 62–96% and Specificity of 72–93%. Participants of the Computers in Cardiology 2001 reported 80% accuracy on the AFPDB database using PAC number and P wave variability parameters.

Recent research on the same databases and spectral analysis of the 30 minute HRV segments that we used is reported in [2]. Authors also used AFPDB as a training database and NSRDB, AFDB as a large independent test sets for their methods. They applied periodogram estimate of the power spectral density of the 30 minute HRV segments and PAC number as the markers. For the classification purposes whether analyzed segment comes from PAF or non-PAF patient they used Fisher's linear discriminant classifier. They achieved Se 85% and Sp 81% on the training AFPDB database. Per-segment Specificity on the 18 subjects from NSRDB is reported as 98.8% and Sensitivity on the 24 subjects from AFDB is 43%. They attribute bad Sensitivity results on the AFDB to the possibility that training data from AFPDB was immediately before PAF onset and testing data from AFDB was in the long-term non-PAF excerpts which are in majority distant from PAF. Thus they explain distant from PAF HRV data could miss characteristic changes that are present immediately before PAF. However our results of FFT estimate show that spectral distribution is closely similar for the non-PAF segments from AFPDB and AFDB databases. And our automatic classification results with non-linear ANN classifier corroborate that fact with per-segment Sensitivity of 94.5% for 14 patients. We did not use the rest of the patients from AFDB as the quality of the other recordings did not allows us to provide reliable HRV annotation and that could potentially lead to the bias in the results.

In summary we achieved better classification rates for the Physionet databases reported in the literature and very close rates to the best reported results in the literature on the time domain indices. Our method thus can be combined with high resolution P wave indices and provide far more reliable screening procedure.

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] Chesnokov YV, Nerukh D, Glen RC. Individually adaptable automatic QT detector. Computers in Cardiology 2006;33.

[2] Hickey B, Heneghan C, Chazal P. Non-episode-dependent assessment of paroxysmal atrial fibrillation through measurement of RR interval dynamics and atrial premature contractions. Annals of Biomedical Engineering 2004;32(5);677-87.

Address for correspondence

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