

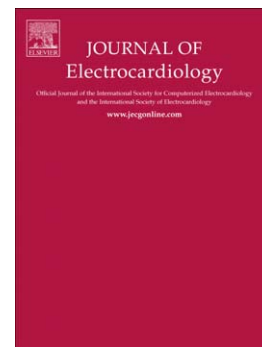
## Accepted Manuscript

Diagnostic decision support systems for atrial fibrillation based on a novel electrocardiogram approach

Jonathan Araujo Queiroz, Alfredo Junior, Fausto Lucena, Allan Kardec Barros

PII: S0022-0736(17)30425-9  
DOI: doi: [10.1016/j.jelectrocard.2017.10.014](https://doi.org/10.1016/j.jelectrocard.2017.10.014)  
Reference: YJELC 52529

To appear in: *Journal of Electrocardiology*



Please cite this article as: Queiroz Jonathan Araujo, Junior Alfredo, Lucena Fausto, Barros Allan Kardec, Diagnostic decision support systems for atrial fibrillation based on a novel electrocardiogram approach, *Journal of Electrocardiology* (2017), doi: [10.1016/j.jelectrocard.2017.10.014](https://doi.org/10.1016/j.jelectrocard.2017.10.014)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Diagnostic decision support systems for atrial fibrillation based on a novel electrocardiogram approach

Jonathan Araujo Queiroz<sup>a</sup>, Alfredo Junior<sup>a</sup>, Fausto Lucena<sup>b</sup>, Allan Kardec Barros<sup>a</sup>

<sup>a</sup>Department of Electrical Engineering, Federal University of Maranhão, São Luís, Brazil

<sup>b</sup>University of CEUMA, Imperatriz - MA, Brazil

---

## Abstract

Atrial fibrillation (AF) occurs when fast and irregular electrical signals cause the atria does not contract, just fibrillate. The electrocardiogram (ECG) is one of the most non-invasive techniques to give support to the AF diagnosis. Several authors use the temporal difference between two consecutive R waves, a method known as RR interval, to perform the AF diagnosis. However, RR interval-based analysis does not detect distortions on the other ECG waves. Thus, the present work proposes a diagnostic decision support systems for atrial fibrillation based on higher order spectrum analysis of voltage variation on the ECG. The proposed method was used aiming AF classifying. The classifier is composed by two screening stages: one based on the average and an-other on the average deviation of kurtosis of the ECG signals. Heartbeat obtained from the MIT-BIH atrial fibrillation and MIT-BIH normal were used. ECG signal featured by kurtosis outperforms second order statistics based metrics in up to 476 times, and up to 110 times above the RR interval. The screening methods obtained sensitivity equal to 100% and specificity is up to 84.04%. The two screening methods combined provided an AF classifier with a accuracy rate at diagnosis of 100%. The results presented take into account windows of up to five heartbeats and a 99.73% confidence interval. Therefore, the results obtained by the proposed method can be used to sup-port decision-making in clinical practices with a diagnostic accuracy rate of 90.04% to 100%.

Keywords: Kurtosis; Statistical Analysis; RR Interval.

---

## 1. Introduction

Atrial fibrillation (AF) is an arrhythmia where there is a complete disorder on the atrial electrical activity, causing a non-contraction and a fibrillation of the atria, hampering the correct blood flow to the ventricles not generating atrial systoles [1, 2].

Naccarelli et al. estimate that the number of AF cases in the USA will increase from 3.03 million in 2005 to 7.56 million in 2050 [3]. In more recent studies, as one developed by Collila et al., those numbers are far more alarming, going from 5.2 million in 2010 to 12.1 million in 2030 [4]. The prevision for the European Union is that the number of AF cases will double from 2010 to 2060 [5].

One of the most used non-invasive techniques to help the AF diagnosis is the electrocardiogram (ECG), which registers the heart muscle electrical activity [6, 7, 8]. Methods based on the temporal difference of the R waves (RR intervals), which are the higher waves of the ECG signal, were proposed [9, 10]. Nevertheless, the analysis of the RR intervals is not able to measure changes on other ECG waves. For instance, such as the distortions on P-wave for AF and with an inverted sawtooth F-wave pattern in the atrial flutter. Also, for temporal arrhythmias similar to AF (such as atrial flutter), the temporal analysis is inefficacious [11, 12, 13, 14, 15]. Thus, developing methods that evaluate the distortions on the other ECG waves is a challenge to be overcome. The present work proposes to investigate the voltage (mV) variation occurring at each heartbeat interval using kurtosis. Kurtosis is indicated for measuring small variation distributions, and sparse signals such as ECG [16, 17] (Fig. 1)

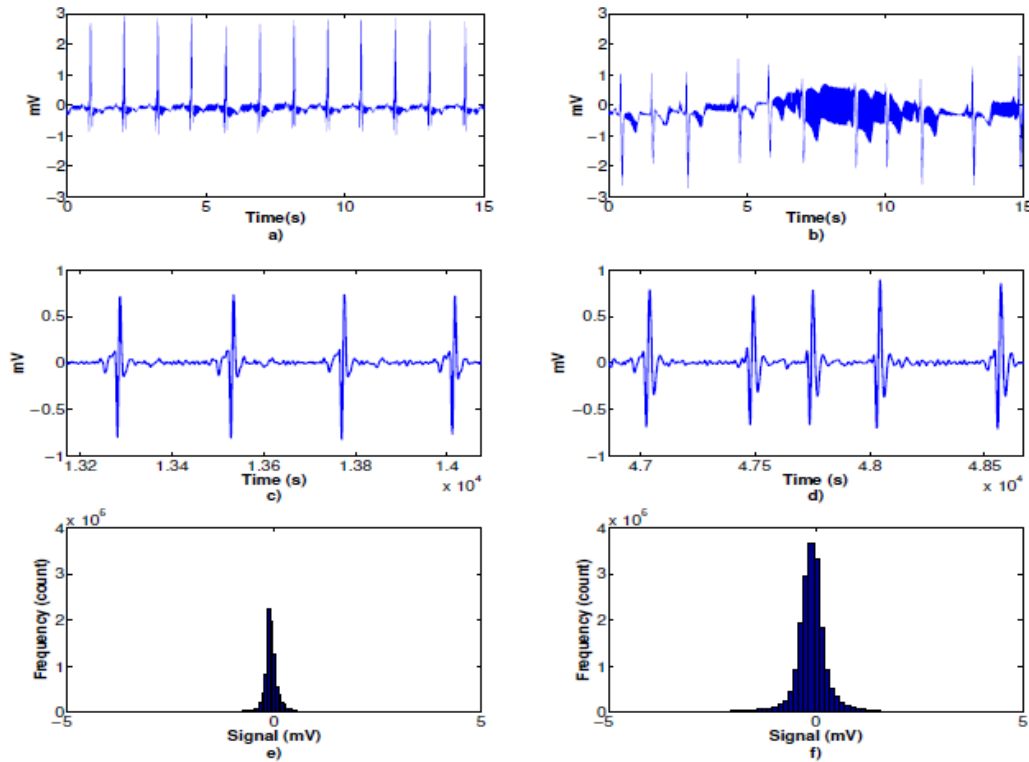


Figure 1: Comparison between healthy and with AF ECG signals. a) Example of healthy ECG record. b) Example of with AF ECG record. c) Example of zoom healthy ECG record. d) Example of with zoom AF ECG record. e) Histogram of database MIT-BIH healthy. f) Histogram of database MIT-BIH AF.

The results of the proposed methodology are evaluated in terms of specificity (how efficient the method is for diagnosing healthy patients), sensibility (how efficient the method is for diagnosing patients with AF), and accuracy (how efficient is the method regarding the diagnosis).

## 2. Materials and Methods

### 2.1. Materials

The ECG signals were obtained from the databases (DB) MIT-BIH Normal Sinus Rhythm (NRS) [18] and MIT-BIH AF [19], being 18 healthy patients, and 25 patients with AF. The sampling frequency of the healthy group is 128Hz and the AF group is 256Hz. For each dataset, two or more cardiologists independently annotated each record; disagreements were resolved to obtain the computer-readable reference annotations for each beat (approximately 110,000 annotations in all) included with the database.

#### 2.1.1. Preprocessing

From the records of the MIT-BIH NSR and MIT-BIH AF bases, as described in the 2.1 section, 50000 healthy heartbeat and 50000 heartbeat of people with AF were withdrawn. Less than 1%, at the beginning and at the end, of the ECG signals were excluded because of measurement error. It was used the algorithm of Pan and Tompkins for identifying the R-peaks from the QRS complex. This algorithm was chosen because it does not require large computational effort and has a 99.3% success rate of the QRS complexes for the MIT-BIH Arrhythmia database [20].

### 2.2. Method

#### 2.3. Usual feature extraction

Cardiac variability is usually used to provide temporal information of the ECG signal. Cardiac variability is obtained by the RR interval that is defined as

$$Interval_{RR} = R_m - R_{m-1}, \quad (1)$$

where  $R_m$  is the time instant of the m-th R-peak.

Based on the RR interval, the ECG can be defined as

$$ECG_{RR} = (RR_1, RR_2, \dots, RR_n), \quad (2)$$

where  $RR_j$  represents the j-th RR interval.

#### 2.4. Proposed feature extraction

The extraction of morphological information of the ECG signal proposed in this work is carried out by analyzing the voltage variation on each heartbeat. Heartbeat feature extraction proposed in this work is defined as

$$B_t = (x_1, x_2, \dots, x_n), \quad (3)$$

where  $B_t$  is the  $t$ -th beat,  $x_n$  is the  $n$ -th sample of  $B_t$  and  $L_i \leq n \leq L_s$ , such that  $L_s$  is the upper limit and  $L_i$  is the lower limit. . The upper ( $L_s$ ) and lower ( $L_i$ ) limits are given by

$$L_s = P_R + F_s \lambda, \quad (4)$$

and

$$L_i = P_R - F_s \theta, \quad (5)$$

where  $P_R$  is the position of the peak R,  $\lambda$  and  $\theta$  are weights of the proportions of the ECG signal to be used, being  $\lambda + \theta \leq 1$  and  $F_s$  is the sampling frequency of the ECG signal.  $L_s$  and  $L_i$  are limits of position and not of variation of voltage of  $B_t$ . The Fig. 1 illustrates the proposed method for ECG feature extraction.

Figure 2 illustrates the proposed method for ECG feature extraction.

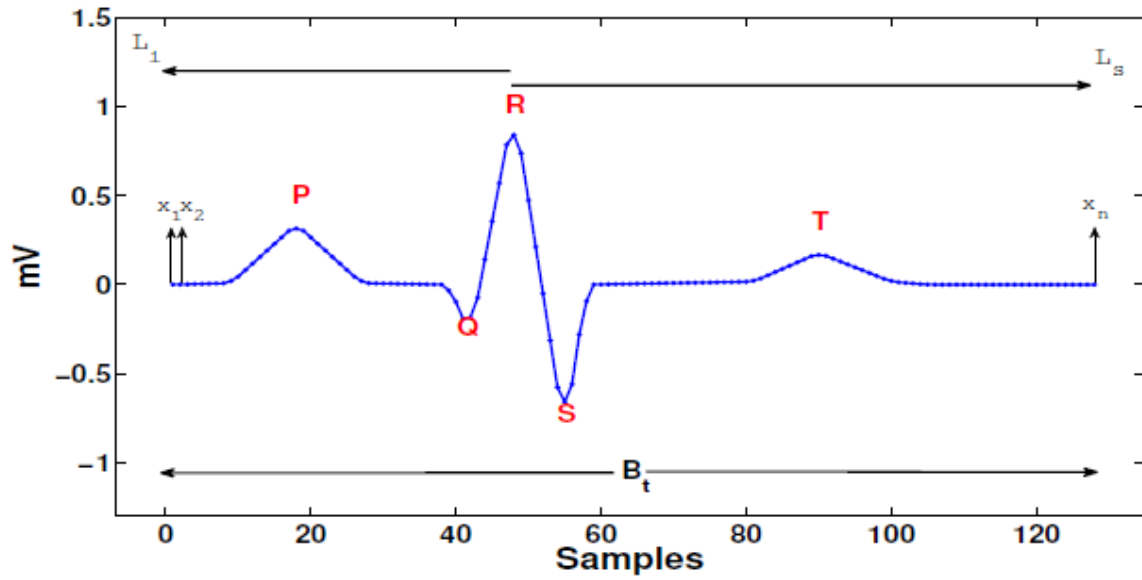


Figure 2: Proposed method of extracting ECG feature.  $B_t$  is the vector that represents each heartbeat,  $x_n$  is the  $n$ -th sample of a heartbeat.  $L_s$  and  $L_i$  are the upper and lower position limits, respectively.

Based on Eq. 3 the ECG can be redefined as

$$ECGB_t = ((x_1, x_2, x_3, \dots, x_n), (y_1, y_2, y_3, \dots, y_n), \dots, (z_1, z_2, z_3, \dots, z_n)), \quad (6)$$

where  $x$ ,  $y$  and  $z$  are different heartbeat.

The technique used in this work to analyze and characterize each beat will be the Higher Order Specter (HOS) that is given by

$$M_k(B_t) = E(x_i - \mu)^k \quad (7)$$

where  $\mu$  is the average of the  $x_n$  and  $k$  represents the  $k$ -th moment.

Statistical moments used in this work are the 1-th to 4-th order and are receptively the avarange, variance, skewness and kurtosis.

To evaluate the relevance of this novel ECG method, we propose two screening methods and a classification method for AF in the following subsections.

### 3. Screening Methods and Classification of AF

In order to perform the AF diagnosis, it is necessary to carry out some steps. Among them, is the development of screening methods which will help the clinical diagnosis. Therefore this work proposes the use of two screening methods as stages of the AF classifier.

In order to guarantee the results and the generalization of the method, besides the definition of the decision thresholds, we propose to establish a confidence interval (CI) of 99.73% [21], i.e., three average standard deviation for each threshold. The decision thresholds proposed in the literature are empirically attributed [9, 10]. Therefore, the challenge is to develop adaptive decision thresholds not needing to be adjusted.

#### 3.1. Screening method based on the average ( $S_A$ )

The adaptive threshold based on the average is given by

$$S_A = \varphi \pm 3\sigma_{AF}, \quad (8)$$

where  $\varphi$  is the average of  $M_K(B_t)$  and  $s$  is the standard deviation of  $M_K(B_t)$ , both from the group with AF.

#### 3.2. Screening method based on the average deviation ( $S_{Ad}$ )

The adaptive threshold based on the average deviation is given by

$$S_{Ad} = [M_K(B_t)_{AF} - \varphi] \pm 3\sigma_{AF}, \quad (9)$$

#### 3.3. AF Classifier (CAF)

The screening methods are good indicators for backing up the AF diagnosis. How-ever, the ideal is to develop a method capable of classifying and not only carrying out the screening of patients with AF. Thus, the following decision function will be used as a classification method:

$$C_{Ad} = \begin{cases} AF, & M_K(B_t) \in S_A \wedge M_K(B_t) \in S_A \\ Healthy, & otherwise. \end{cases}$$

Therefore, the patient is diagnosed with AF, if only if, satisfies the two method of screening. Otherwise, the patient is classified as healthy.

### 3.4. Evaluation Metrics

The results are evaluated in four categories, as follows: True positive (TP), in which a record with AF is classified with AF; True negative (TN), wherein a normal record is classified as normal; False positive (FP) a normal record is classified with AF; False negative (FN), for this category, a record with AF is classified as normal.

The metric for evaluating the classification stage, we use the sensitivity (SENS), which reflects how the methodology adopted is effective in correctly identifying the TP, specificity (SPEC), reflecting how the methodology adopted is effective in correctly identifying TN and accuracy (ACC) indicating how the method is effective in correctly performing the diagnosis.

The sensitivity and specificity are defined, respectively, given by

$$SENS = \frac{TP}{TP+FN} \times 100, \quad (11)$$

and

$$SPEC = \frac{TN}{TN+FP} \times 100. \quad (12)$$

And the accuracy is given by

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \times 100. \quad (13)$$

## 4. Simulation and Results

### 4.1. Simulation

The records of the databases MIT-BIH NSR and MIT-BIH AF were used as described in the section 2.1. To select the waveform of the ECG signal that undergoes the deformations in the AF episode (that is, the P wave) the parameters of the ECG signal ratio weights used for the upper ( $\lambda$ ) and lower ( $\theta$ ) are  $\lambda=0$  and  $\theta=0.4$ .

#### 4.2. Results

The results are presented for a set of 100,000 preprocessed heartbeats and the diagnosis evaluated in relation to sensitivity, specificity and accuracy. Figure 3 illustrates the 50,000 healthy heartbeats and 50,000 heartbeat with AF before and after the proposed method was applied.

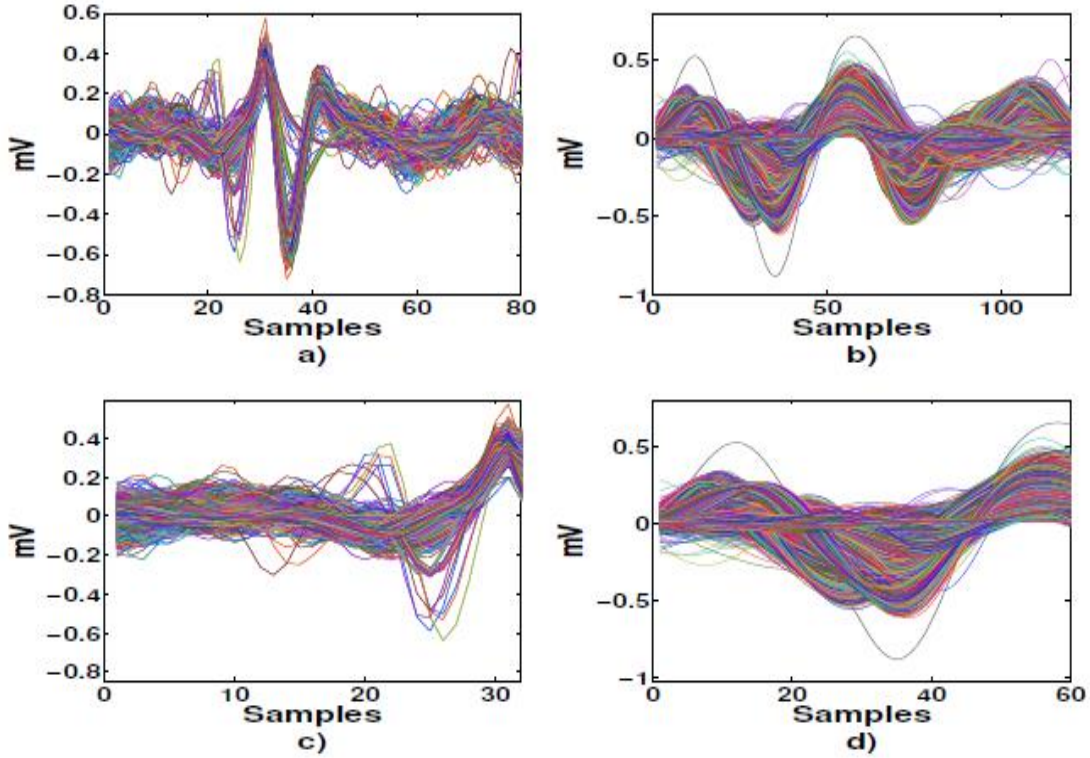


Figure 3: 100,000 preprocessed heartbeats. a) Illustrates 50,000 preprocessed heartbeats with weighting parameter based on  $\lambda = 0:6$  and  $\theta = 0:4$  for healthy patients. Fig. b) Illustrates 50,000 preprocessed heartbeats with weighting parameter based on  $\lambda = 0:6$  and  $\theta = 0:4$  for patients with AF. c) Illustrates 50,000 preprocessed heartbeats with weighting parameter based on  $\lambda = 0$  and  $\theta = 0:4$  for healthy patients. d) Illustrates 50,000 preprocessed heartbeats with weighting parameter based on  $\lambda = 0$  and  $\theta = 0:4$  for patients with AF.

#### 4.3. Statistical analysis for $M_K(B_t)$

The criterion for choosing the  $M_K(B_t)$  that separate the groups is represented in Tab. 1.



Table 1: Dispersion analysis of AF and healthy groups based on different statistical metrics.  $\Delta$  is the difference between the medians of the groups.

Analysis		Healthy	AF	$ \Delta $
RR Interval	Upper Limit	2.0625	2.36	0.104
	Median	1.312	1.416	
	Lower Limit	0.6000	0.3640	
Variance ( $M_2(B_t)$ )	Upper Limit	0.2054	1.3018	0.024
	Median	0.062	0.086	
	Lower Limit	0.022	0.007	
Skewness ( $M_3(B_t)$ )	Upper Limit	4.9496	1.6949	2.7776
	Median	2.3241	-0.4535	
	Lower Limit	-0.9574	-2.3424	
kurtosis ( $M_4(B_t)$ )	Upper Limit	36.9601	10.91	11.4294
	Median	18.6701	7.2407	
	Lower Limit	8.1533	5.6508	

Median is a position measure that divides the dataset in half, that is, the farther the medians of the AF and healthy groups, more distant those groups will be. Therefore,  $\Delta$ , which is the difference between the medians of groups AF and the healthy one, is used as criteria for choosing the statistics that characterizes the ECG signal. Then, the greater  $\Delta$  is, farther away the groups will be. Thus, based on Tab. 1, kurtosis  $M_4(B_t)$  is the statistics that characterizes the ECG signal.

The comparison between the RR interval and  $M_K(B_t)$  (variance, skewness, and kurtosis) is shown in Tab. 1, and illustrated in Fig. 4.

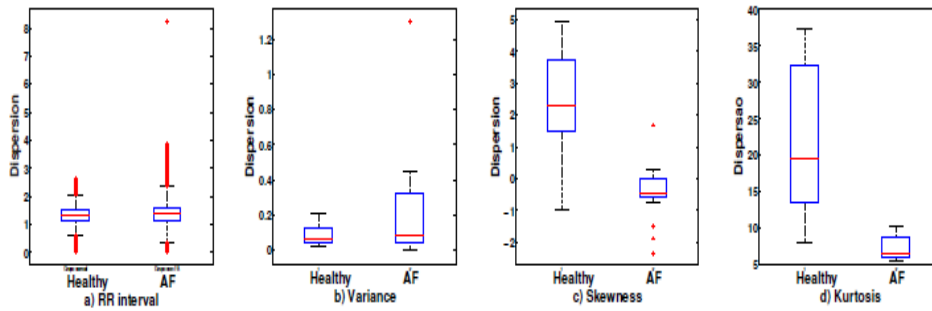


Figure 4: Boxplot of sampling distributions. a) Dispersion of the RR interval, b) Dispersion of the  $M_2(B_t)$  (variance), c) Dispersion of the  $M_3(B_t)$  (skewness), d) Dispersion of the  $M_4(B_t)$  (kurtosis).

#### 4.4. Analysis of the ECG windowing signal for $M_4(B_t)$

Some studies show that the size at which the ECG signals is windowed influences its analysis [22, 23]. So, in this work the analysis of the windowing for classifies AF is presented in Tab. 2.

Table 2: Windowing analysis of the ECG signal for AF and healthy groups. Where,  $\Delta$  is the difference between the medians between healthy and AF group.

Groups		Heartbeat										
		100	90	80	70	60	50	40	30	20	10	5
Healthy	Upper Limit	36.96	37.00	37.00	37.01	36.96	36.96	36.96	36.72	36.54	36.39	35.03
	Median	18.45	19.03	18.81	18.87	18.67	19.09	19.13	18.78	19.15	18.97	18.01
	Lower Limit	8.37	8.87	8.87	8.86	8.84	8.86	8.88	8.31	8.13	7.77	7.05
AF	Upper Limit	10.77	10.88	10.79	10.91	10.75	10.84	10.84	10.69	11.24	11.02	10.57
	Median	8.60	8.12	8.02	7.79	7.24	7.26	7.08	6.75	6.72	7.03	6.14
	Lower Limit	5.83	5.73	5.75	5.78	5.78	5.75	5.75	5.81	5.50	4.82	5.83
$\Delta$		9.85	10.91	10.79	11.08	11.43	11.83	12.05	12.03	12.43	11.93	11.87

#### 4.5. AF Classification

For the screening method (Eq. 8 and Eq. 9), as well as for AF classifier (Eq. 10), the ECG signal is characterized by kurtosis according to Tab. 1, and windowed at 20-5 heartbeat according to Tab. 2. The results are shown in Tab. 3, and evaluated with respect to sensibility, specificity, and accuracy.

Table 3: Performance analysis of sensibility, specificity, and accuracy for screening methods ( $S_A$  and  $S_{Ad}$ ), as well as the classifier ( $C_{AF}$ ), in different sizes of windows (20-5 heartbeats).

Proposed methods	Performance (%)								
	20 heartbeat			10 heartbeat			5 heartbeat		
	SENS	SPEC	ACC	SENS	SPEC	ACC	SENS	SPEC	ACC
$S_A$	100	84.04	92.85	100	84.04	92.85	100	77.77	90.04
$S_{Ad}$	100	77.77	88.88	100	66.66	85.71	100	66.66	85.71
$C_{AF}$	100	100	100	100	100	100	100	100	100
Screening method based on the average ( $S_A$ ); Screening method based on the average deviation ( $S_{Ad}$ ); AF Classifier ( $C_{AF}$ ); Sensitivity (SENS); Specificity (SPEC); Accuracy (ACC).									

Figure 5 illustrates the performance analysis among sensibility, specificity, and accuracy for different window sizes of the ECG according to the proposed method.

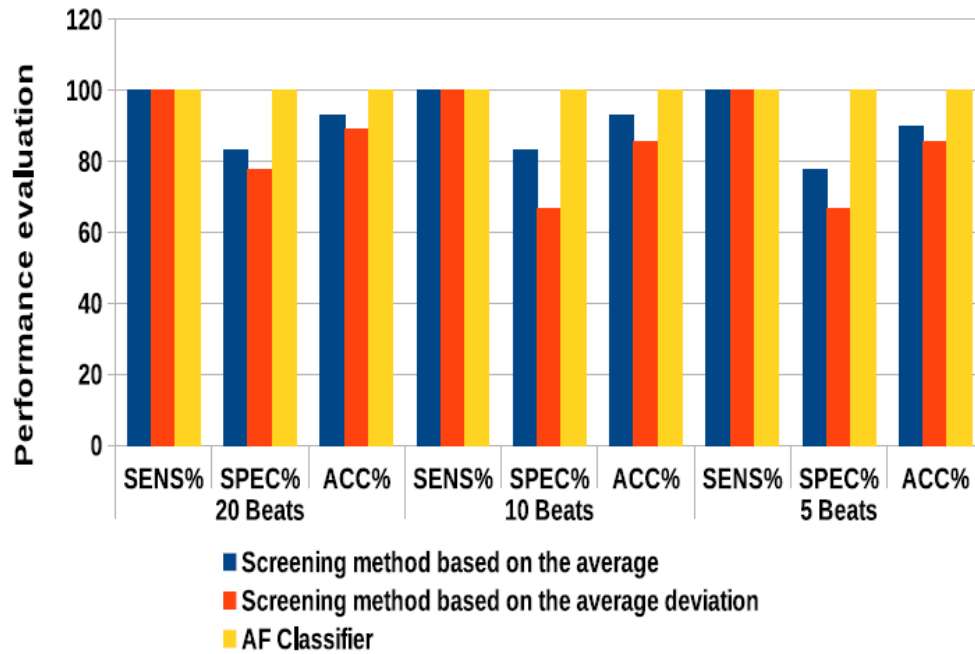


Figure 5: Performance analysis of the proposed method in different window sizes (20-5 heartbeats).

Figure 6 illustrates the acceptance region for the 99.73% CI, for the ( $S_A$  and  $S_{Ad}$ ) method screening, and for the AF ( $C_{AF}$ ) classifier.

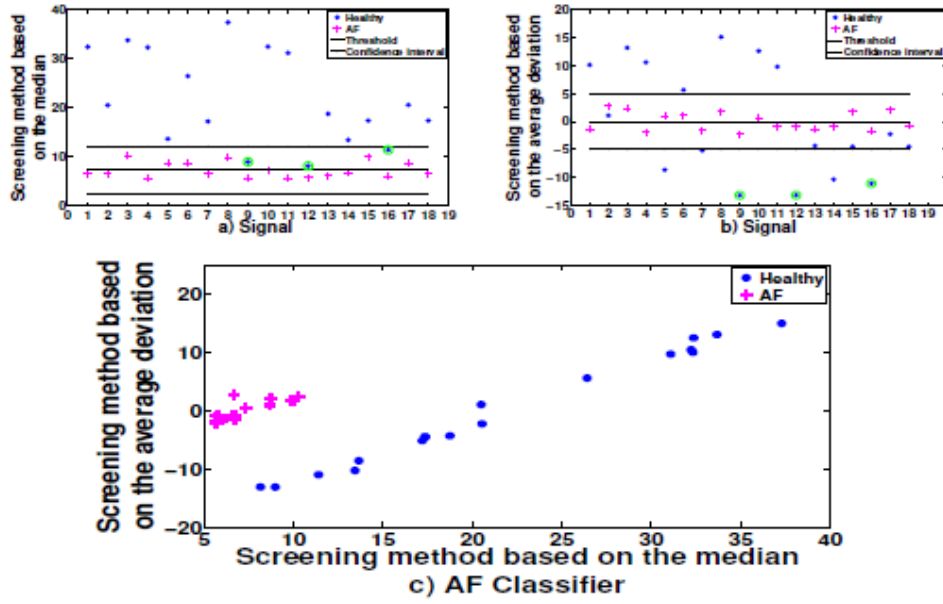


Figure 6: AF classification method composed by two screening methods. a) Screening method based on the median. b) Screening method based on the average deviation. c) AF classifier. For the classifier, every patient will be diagnosed with AF if and only if they are identified in both screening methods. Such being the case, the healthy patients 9, 12 and 16 are false positive in Fig. 6.a but they are true negative in Fig. 6.b, consequently classified as healthy in Fig 6.c.

Investigations using the RR interval were also performed in order to verify the separability of the healthy and AF groups (Fig. 7).

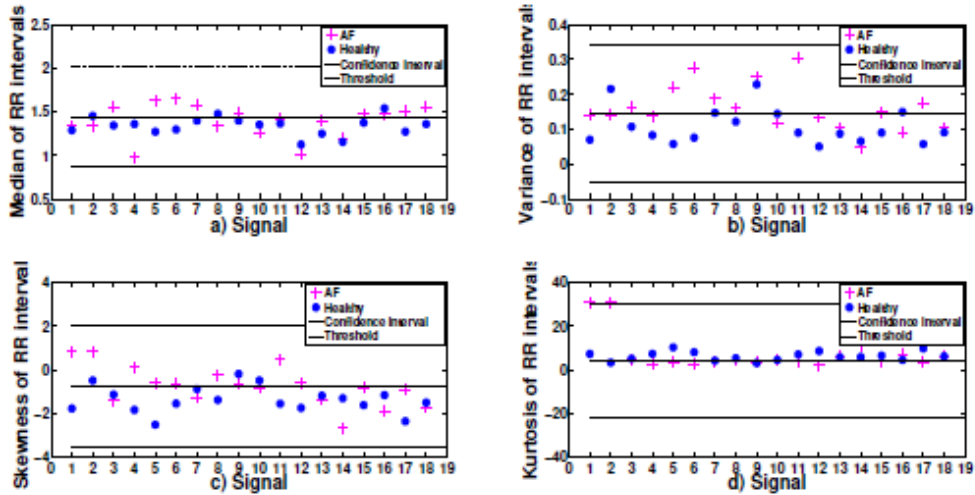


Figure 7: Methods screening based on the RR interval. a) Screening method using AF and healthy groups medians. b) Screening method using variance of group AF and healthy. c) Screening method using the skewness of healthy and AF groups. d) Screening method using kurtosis for group healthy and AF.

## 5. Discussion

The present work proposes a novel approach for ECG aiming AF classifying. The proposed method allows the selection of specific parts of ECG signal, in this case the wave P (Fig.3). This novel approach, was able to separate AF and healthy ECG signals using HOS than the RR interval (Tab. 1). Among the evaluated statistics, the one that separates the groups was kurtosis, reaffirming studies by Martis et al. and Lucena et al., which point out that kurtosis could be an appropriate approach to measure sparse signals such as ECG [16, 17]. In absolute value  $M_4(B_t)$  is up to 476 times upper to other evaluated statistics ( $M_2(B_t)$ ), and up to 110 times upper to the RR interval. The  $M_k(B_t)$  analysis is only more efficient than the RR interval when HOS is used. The  $M_k(B_t)$  dispersion presented by AF and healthy groups is upper to the RR interval for all statistics (Fig.4).

The investigations show that the best windowings for  $M_4(B_t)$  were the 20-40 heart-beat (Tab. 2). Such behavior was expected for two reasons: a clinical and an analytical aspect. The former deals with the recovery period of the cardiac frequency, which were  $30 \pm 12$  heartbeat [24, 25, 26, 27]. The other is the analytical aspect which takes into account the quantity of windows where the ECG signal is shown, since the smaller is the windowing, more parts of the ECG signal are obtained. According to the central theorem, leaving the sampling set closer to the AF patients [28]. However, even for lower windows (5 heartbeat) the result of the proposed classifier was not changed.

For all window sizes tested, the screening methods obtained sensitivity equal to 100% and specificity is up to 84.04%. The two screening methods combined provided an AF classifier with an accuracy rate at diagnosis of 100% (Tab. 3 and Fig. 5). The results presented take into account a 99.73% CI.

The first method of screening serves to identify the true positives and the true negatives. In this first screening method, all true positives were sorted correctly, although, not all true negatives were sorted correctly (Fig. 6 a). Therefore, the second method of screening serves primarily to identify false positives of the first screening method (Fig. 6 b) that is, the patient will be diagnosed with AF if and only if the diagnosis satisfies the two screening methods (Fig. 6 c). Otherwise, the patient is classified as healthy.

The screening methods as well as the AF classifier proposed in this work were evaluated using the RR interval. However, the RR interval did not present significant results in the separation of the groups AF and healthy (Fig. 7).

Table 4 compares the proposed methodology with the best performance in the literature with respect to sensitivity, specificity and accuracy.

Table 4: Performance comparison of proposed methodology with the literature with respect to sensitivity, specificity and accuracy. The symbol (-) represent values of sensibility, specificity, and accuracy are not specified in the works.

Author (Year)	Database	Features	Window	Classifiers	SENS%	SPEC%	ACC%
Screening1 (2017)	AFDB	P wave morphology	5 heartbeat	ATD+CI	100	77.77	90.04
			10 heartbeat		100	84.04	92.85
			20 heartbeat		100	77.77	90.04
Screening2 (2017)	AFDB	P wave morphology	5 heartbeat	ATD+CI	100	66.66	85.75
			10 heartbeat		100	66.66	85.75
			20 heartbeat		100	77.77	88.88
Classifier (2017)	AFDB	P wave morphology	5 heartbeat	ATD+CI	100	100	100
			10 heartbeat		100	100	100
			20 heartbeat		100	100	100
Kennedy et al. (2016)[29]	AFDB	CoSEn+ CV+RMSSD +MAD	30 RR	RF	92.80	98.30	-
Orchard et al. (2016)[12]	44 patients	Absence of P wave	30 s	New algorithm	95.00	99.00	-
Huo et al. (2015)[13]	70 patients	Absence of P wave	250 ms	MLR	41.50	94.10	-
Petrenas et al. (2015)[9]	AFDB	RR	8 heartbeat	TD	97.10	98.30	-
Zhou et al. (2014)[10]	AFDB	SD+SE	17 RR	TD	97.53	98.26	98.16
Coefficient of Sample Entropy (CoSEn); Coefficient of Variance (CV); Root Mean Square of the Successive Differences (RMSSD); Median Absolute Deviation (MAD); Symbolic Dynamics (SD); Shannon Entropy (SE); RR intervals (RR); Second (S); Milliseconds (Ms); Adaptive Threshold Detection (ATD); Threshold (TD); Confidence Interval (CI); Random Forests (RF); Multivariate Logistic Regression (MLR).							

When compared with more complex techniques for extraction of feature such as those used by Kennedy et al. [29], which were coefficient of sample entropy (CoSEn), coefficient of variance (CV), root mean square of successive differences (RMSSD), and median absolute deviation (MAD) or in Zhou et al. [10] using Symbolic Dynamics (SD) and Shannon Entropy (SE), the proposed method outperforms the sensitivity rate in all performance comparisons. The specificity of the screening methods separately, is lower in all performance comparisons. However, when the two screening methods are used together it is possible to achieve a 100 % rate in the AF diagnosis (Tab. 4).

Still according to the Tab.4, even in methods similar to that proposed in this paper, as in Orchard et al. [12] and Huo et al. [13] which are based on the absence of the P wave, the results are lower than the proposed AF classifier. This result is a consequence of the higher amount of information that the proposed method can obtain in each heartbeat, providing an AF classifier that

measures the deformations in the P wave and not only verifies its absence. Even when compared to Petrenas et al. [9], which classifies the AF with a high hit rate (97.10%) using simply the irregularities of the RR interval, it is advantageous to use the proposed method since it uses threshold of adaptive decision and CI (Eq. 8 and Eq. 9), which makes possible the application of the proposed method in other populations with AF.

The implementation of the proposed method is relevant, especially for countries with wide territorial space such as Brazil, where the concentration of the number of physicians is 2.09 physician per 1000 inhabitants, and in certain States, such as Maranhão, it is about 0.79 per 1000 inhabitants. The statistics get worse when it comes to specialist physicians such as cardiologists. There are only 13.420 cardiologists out of 419.224 registered physicians for a population of 201.032.714 inhabitants [30].

In at least two points the proposed method is deficient. First of all, to kurtosis is sensitive to outliers which may alter the AF diagnosis. The second concerns the amount of healthy heartbeats that exist in each window; if the healthy beats are 95% higher than the AF beats the correct AF diagnosis can be impaired.

## 6. Conclusions

The present work proposes investigate the voltage variation (mV) at each heartbeat using kurtosis, to AF classify. The proposed AF classifier is composed of two screening steps, which may be used in epidemiological studies and decision making in medical practice, with 91.67% diagnosis rate accuracy. However, for more rigorous epidemiological studies, the classifier is the most indicated, with a 100% accuracy at diagnosis. Furthermore the simplicity of the proposed method and the larger quantity information, utilized for classifying AF patients allows for application in embedded systems similar to the 24-hour Holter. Still, such a system will be able not only to record the heart electrical activities and its variations but will also provide a prognosis for AF. Implementing and evaluating this AF classifier in real time is our next challenge.

## Acknowledgements

We would like to thank PPGEE-UFMA, CNPq, FAPEMA and CAPES for infrastructure development and financial support

- [1] B. Freedman, J. Camm, H. Calkins, J. S. Healey, M. Rosenqvist, J. Wang, C. M. Albert, C. S. Anderson, S. Antoniou, E. J. Benjamin, G. Boriani, J. Brachmann, A. Brandes, T.-F. Chao, D. Conen, J. Engdahl, L. Fauchier, D. A. Fitzmaurice, L. Friberg, B. J. Gersh, D. J. Gladstone, T. V. Glotzer, K. Gwynne, G. J. Hankey, J. Harbison, G. S. Hillis, M. T. Hills, H. Kamel, P. Kirchhof, P. R. Kowey, D. Krieger, V. W. Y. Lee, L.-Å. Levin, G. Y. H. Lip, T. Lobban, N. Lowres, G. H. Mairesse, C. Martinez, L. Neubeck, J. Orchard, J. P. Piccini, K. Poppe, T. S. Potpara, H. Puererfellner, M. Rienstra, R. K. Sandhu, R. B. Schnabel, C.-W. Siu, S. Steinhubl, J. H. Svendsen, E. Svennberg, S. Themistoclakis, R. G. Tieleman, M. P. Turakhia, A. Tveit, S. B. Uittenbogaart, I. C. Van Gelder, A. Verma, R. Wachter, B. P. Yan, Screening for atrial fibrillation, *Circulation* 135 (19) (2017) 1851–1867. arXiv:<http://circ.ahajournals.org/content/135/19/1851.full.pdf>, doi:10.1161/CIRCULATIONAHA.116.026693.URL <http://circ.ahajournals.org/content/135/19/1851>
- [2] C. T. January, L. S. Wann, J. S. Alpert, H. Calkins, J. E. Cigarroa, J. C. Cleveland, J. B. Conti, P. T. Ellinor, M. D. Ezekowitz, M. E. Field, K. T. Murray, R. L. Sacco, W. G. Stevenson, P. J. Tchou, C. M. Tracy, C. W. Yancy, 2014aha/acc/hrs guideline for the management of patients with atrial fibrillation, *Circulation* 130 (23) (2014) e199–e267. arXiv:<http://circ.ahajournals.org/content/130/23/e199.full.pdf>, doi:10.1161/CIR.0000000000000041. URL <http://circ.ahajournals.org/content/130/23/e199>
- [3] G. V. Naccarelli, H. Varker, J. Lin, K. L. Schulman, Increasing prevalence of atrial fibrillation and flutter in the united states, *The American Journal of Cardiology* 104 (11) (2009) 1534 – 1539. doi:<http://dx.doi.org/10.1016/j.amjcard.2009.07.022>.URL <http://www.sciencedirect.com/science/article/pii/S0002914909013885>
- [4] S. Colilla, A. Crow, W. Petkun, D. E. Singer, T. Simon, X. Liu, Estimates of current and future incidence and prevalence of atrial fibrillation in the u.s. adult population, *The American Journal of Cardiology* 112 (8) (2013) 1142 – 1147. doi:<http://dx.doi.org/10.1016/j.amjcard.2013.05.063>. URL <http://www.sciencedirect.com/science/article/pii/S0002914913012885>
- [5] B. P. Krijthe, A. Kunst, E. J. Benjamin, G. Y. Lip, O. H. Franco, A. Hofman, J. C. Witteman, B. H. Stricker, J. Heeringa, Projections on the number of individuals with atrial fibrillation in the european union, from 2000 to 2060, *European Heart Journal* 34 (35) (2013) 2746–2751. arXiv:<http://eurheartj.oxfordjournals.org/content/34/35/2746.full.pdf>, doi:10.1093/eurheartj/eh280. URL <http://eurheartj.oxfordjournals.org/content/34/35/2746>
- [6] M. L. Meyer, A. Jaensch, U. Mons, L. P. Breitling, H. Hahmann, W. Koenig, H. Brenner, D. Rothenbacher, Atrial fibrillation and long-term prognosis of patients with stable coronary heart disease: Relevance of routine electro-cardiogram, *International Journal of Cardiology* 203 (2016)



1014 – 1015. doi:<http://dx.doi.org/10.1016/j.ijcard.2015.11.111>.URL  
<http://www.sciencedirect.com/science/article/pii/S0167527315309219>

[7] W. Shuai, X. xing Wang, K. Hong, Q. Peng, J. xiang Li, P. Li, J. Chen, X. shu Cheng, H. Su, Is 10-second electrocardiogram recording enough for accurately estimating heart rate in atrial fibrillation, *International Journal of Cardiology* 215 (2016) 175 – 178. doi:<http://dx.doi.org/10.1016/j.ijcard.2016.04.139>.

URL <http://www.sciencedirect.com/science/article/pii/S0167527316308403>

[8] T. Lankveld, C. B. de Vos, I. Limantoro, S. Zeemering, E. Dudink, H. J. Crijns, U. Schotten, Systematic analysis of {ECG} predictors of sinus rhythm maintenance after electrical cardioversion for persistent atrial fibrillation, *Heart Rhythm* 13 (5) (2016) 1020 – 1027. doi:<http://dx.doi.org/10.1016/j.hrthm.2016.01.004>. URL  
<http://www.sciencedirect.com/science/article/pii/S1547527116000060>

[9] L. S. Andrius Petrėnas, Marozas Vaidotas, Low-complexity detection of atrial fib-rillation in continuous long-term monitoring, *Computers in Biology and Medicine* (2015) 1–8doi:[10.1016/j.combiomed.2015.01.019](https://doi.org/10.1016/j.combiomed.2015.01.019).

[10] X. Zhou, H. Ding, B. Ung, E. Pickwell-MacPherson, Y. Zhang, et al., Automatic online detection of atrial fibrillation based on symbolic dynamics and shannon entropy, *Biomedical engineering online* 13 (1) (2014) 18.

[11] A. Maan, M. Mansour, J. N. Ruskin, E. K. Heist, Impact of catheter ablation on p-wave parameters on 12-lead electrocardiogram in patients with atrial fibrillation, *Journal of Electrocardiology* 47 (5) (2014) 725 – 733. doi:<http://dx.doi.org/10.1016/j.jelectrocard.2014.04.010>. URL <http://www.sciencedirect.com/science/article/pii/S0022073614001332>

[12] J. Orchard, N. Lowres, S. B. Freedman, L. Ladak, W. Lee, N. Zwar, D. Peiris, Y. Kamaladasa, J. Li, L. Neubeck, Screening for atrial fibrillation during influenza vaccinations by primary care nurses using a smartphone electrocardiograph (iecg): A feasibility study, *European Journal of Preventive Cardiology* 23 (2\_suppl) (2016) 13–20, pMID: 27892421. arXiv:<http://dx.doi.org/10.1177/2047487316670255>, doi:[10.1177/2047487316670255](https://doi.org/10.1177/2047487316670255).URL  
<http://dx.doi.org/10.1177/2047487316670255>

[13] Y. Huo, F. Holmqvist, J. Carlson, T. Gaspar, G. Hindricks, C. Piorkowski, A. Bollmann, P. G. Platonov, Variability of p-wave morphology predicts the outcome of circumferential pulmonary vein isolation in patients with recur-rent atrial fibrillation, *Journal of Electrocardiology* 48 (2) (2015) 218 –

225.doi:<http://dx.doi.org/10.1016/j.jelectrocard.2014.11.011>.URL

<http://www.sciencedirect.com/science/article/pii/S0022073614004865>

[14] F. Rahman, N. Wang, X. Yin, P. T. Ellinor, S. A. Lubitz, P. A. LeLorier, D. D. McManus, L. M. Sullivan, S. Seshadri, R. S. Vasan, E. J. Benjamin, J. W. Magnani, Atrial flutter: Clinical risk factors and adverse outcomes in the framingham heart study, *Heart Rhythm* 13 (1) (2016) 233 – 240. doi:<http://dx.doi.org/10.1016/j.hrthm.2015.07.031>. URL

<http://www.sciencedirect.com/science/article/pii/S1547527115009443>

[15] K. Wahbi, F. A. Sebag, N. Lellouche, A. Lazarus, H.-M. Bécane, G. Bassez, T. Stojkovic, A. Fayssoil, P. Laforêt, A. Béhin, C. Meune, B. Eymard, D. Duboc, Atrial flutter in myotonic dystrophy type 1: Patient characteristics and clinical outcome, *Neuromuscular Disorders* 26 (3) (2016) 227 – 233. doi:<http://dx.doi.org/10.1016/j.nmd.2016.01.005>.URL

<http://www.sciencedirect.com/science/article/pii/S0960896615301231>

[16] R. J. Martis, U. R. Acharya, C. M. Lim, K. M. Mandana, A. K. Ray, C. Chakraborty, Application of higher order cumulant features for cardiac health diagnosis using ecg signals., *Int. J. Neural Syst.* 23 (4). URL <http://dblp.uni-trier.de/db/journals/ijns/ijns23.html#MartisALMRC13>

[17] F. Lucena, A. K. Barros, J. C. Príncipe, N. Ohnishi, Statistical coding and decoding of heartbeat intervals, *PloS one* 6 (6) (2011) e20227.

[18] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, H. E. Stanley, Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals, *Circulation* 101 (23) (2000) 215–220.

[19] G. B. Moody, R. G. Mark, A new method for detecting atrial fibrillation using R-R intervals (1983).

[20] J. Pan, W. J. Tompkins, A real-time QRS detection algorithm., *IEEE transactions on bio-medical engineering* 32 (3) (1985) 230–236. doi:10.1109/TBME.1985.325532.

[21] C. Dytham, Choosing and using statistics: a biologist's guide, John Wiley & Sons, 2011.

[22] R. Chen, Y. Huang, J. Wu, Multi-window detection for p-wave in electrocardiograms based on bilateral accumulative area, *Computers in Biology and Medicine* 78 (2016) 65 – 75.

doi:<http://dx.doi.org/10.1016/j.compbimed.2016>.

URL

<http://www.sciencedirect.com/science/article/pii/S0010482516302384>

[23] I. Beraza, I. Romero, Comparative study of algorithms for ecg segmentation, *Biomedical Signal Processing and Control* 34 (2017) 166 – 173. doi:<http://dx.doi.org/10.1016/j.bspc.2017.01.013>.

URL <http://www.sciencedirect.com/science/article/pii/S1746809417300216>

[24] I. Antelmi, E. Y. Chuang, C. J. A. Grupi, M. d. R. A. D. O. Latorre, A. J. A. Mansur, Heart rate recovery after treadmill electrocardiographic exercise stress test and 24-hour heart rate variability in healthy individuals, *Arquivos Brasileiros de Cardiologia* 90 (2008) 413 – 418. URL [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0066-782X2008000600005&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0066-782X2008000600005&nrm=iso)

[25] M. Gayda, M. G. Bourassa, J.-C. Tardif, A. Fortier, M. Juneau, A. Nigam, Heart rate recovery after exercise and long-term prognosis in patients with coronary artery disease, *Canadian Journal of Cardiology* 28 (2) (2012) 201 – 207. doi:<http://dx.doi.org/10.1016/j.cjca.2011.12.004>. URL <http://www.sciencedirect.com/science/article/pii/S0828282X11014346>

[26] H. Itoh, R. Ajisaka, A. Koike, S. Makita, K. Omiya, Y. Kato, H. Adachi, M. Nagayama, T. Maeda, A. Tajima, N. Harada, K. Taniguchi, Heart rate and blood pressure response to ramp exercise and exercise capacity in relation to age, gen-der, and mode of exercise in a healthy population, *Journal of Cardiology* 61 (1) (2013) 71 – 78. doi:<http://dx.doi.org/10.1016/j.jjcc.2012.09.010>. URL <http://www.sciencedirect.com/science/article/pii/S0914508712002687>

[27] P. J. Kannankeril, F. K. Le, A. H. Kadish, J. J. Goldberger, Parasympathetic effects on heart rate recovery after exercise, *Journal of Investigative Medicine* 52 (6) (2016) 394–401. arXiv:<http://jim.bmj.com/content/52/6/394.full.pdf>, doi:10.1136/jim-52-06-34. URL <http://jim.bmj.com/content/52/6/394>

[28] N. Breinegaard, a. S. A. Rabe-Hesketh, S, The transition model test for serial dependence in mixed-effects models for binary data, *Statistical Methods in Medical Research* 0 (0) (2015) 1–21.

[29] A. Kennedy, D. D. Finlay, D. Guldenring, R. R. Bond, K. Moran, J. McLaughlin, Automated detection of atrial fibrillation using r-r intervals and multivariate- based classification, *Journal of Electrocardiology* 49 (6) (2016) 871 – 876. doi:<http://dx.doi.org/10.1016/j.jelectrocard.2016.07.033>. URL <http://www.sciencedirect.com/science/article/pii/S002207361630111X>

[30] M. e. a. Scheffer, Demografia médica no brasil 2015, *Arquivos Brasileiros de Cardiologia*.

## Highlights

The present work proposes to investigate the voltage (mV) variation occurring at each heartbeats interval using kurtosis. This new method of feature extraction was allow a beat-to-beat analysis, unlike the RR interval in which each beat is associated with a single real number, the proposed method associates each beat to a set of points, that is, to a vector. In addition, we propose adaptive decision threshold aiming AF classifying. The results obtained by the proposed method can be used to support decision-making in clinical practices with a diagnostic accuracy rate of 90.04% to 100%.