

# Automatic Real Time Detection of Atrial Fibrillation

S. DASH,<sup>1</sup> K. H. CHON,<sup>1</sup> S. LU,<sup>1</sup> and E. A. RAEDER<sup>2</sup>

<sup>1</sup>Department of Biomedical Engineering, SUNY at Stony Brook, HSC T18, Rm. 030, Stony Brook, NY 11794, USA; and  
<sup>2</sup>Department of Medicine, Division of Cardiology, SUNY at Stony Brook, Stony Brook, NY, USA

(Received 14 December 2008; accepted 6 June 2009; published online 17 June 2009)

**Abstract**—Atrial fibrillation (AF) is the most common sustained arrhythmia and is associated with significant morbidity and mortality. Timely diagnosis of the arrhythmia, particularly transient episodes, can be difficult since patients may be asymptomatic. In this study, we describe a robust algorithm for automatic detection of AF based on the randomness, variability and complexity of the heart beat interval (RR) time series. Specifically, we employ a new statistic, the Turning Points Ratio, in combination with the Root Mean Square of Successive RR Differences and Shannon Entropy to characterize this arrhythmia. The detection algorithm was tested on two databases, namely the MIT-BIH Atrial Fibrillation Database and the MIT-BIH Arrhythmia Database. These databases contain several long RR interval series from a multitude of patients with and without AF and some of the data contain various forms of ectopic beats. Using thresholds and data segment lengths determined by Receiver Operating Characteristic (ROC) curves we achieved a high sensitivity and specificity (94.4% and 95.1%, respectively, for the MIT-BIH Atrial Fibrillation Database). The algorithm performed well even when tested against AF mixed with several other potentially confounding arrhythmias in the MIT-BIH Arrhythmia Database (Sensitivity = 90.2%, Specificity = 91.2%).

**Keywords**—Cardiac arrhythmia, Turning points ratio, RMSSD, Shannon entropy, ROC.

## INTRODUCTION

Atrial fibrillation (AF) is the most common clinical arrhythmia affecting approximately 3 million Americans. It has a prevalence of 17.8% and an incidence of 20.7/1000 patient years in individuals older than 85. At age 55 the lifetime risk of developing AF is approximately 23%.<sup>11</sup> Atrial fibrillation is an independent risk factor for death (relative risk in men 1.5 and in women 1.9)<sup>2</sup> and a major cause of ischemic stroke whose impact increases with age, reaching 23.5% in patients older than 80 years.<sup>28</sup> Accurate

detection of AF is crucial since treatment options such as chronic anticoagulation, antiarrhythmic therapy and radiofrequency ablation offer significant benefits but also carry potentially serious risks. Despite the ubiquity of this arrhythmia, its diagnosis rests largely on the presence of symptoms (e.g., rapid and irregular heart rate) and on serendipity. Unfortunately, since patients are sometimes unaware of their irregular pulse<sup>13,19</sup> the diagnosis may only be established during a fortuitous doctor visit. The challenge is even greater when episodes of AF are asymptomatic and intermittent.

The prevalence of asymptomatic AF found incidentally on clinical examination is ~20%,<sup>12,14</sup> even higher with Holter or event recorders.<sup>15,23</sup> One study of patients with implantable pacemakers for AF detection found an incidence of 50% of asymptomatic AF.<sup>4</sup> Given the significant risk of mortality and morbidity and the fact that asymptomatic AF is not detected unless specifically looked for, there is a strong impetus for ambulatory monitoring. With the greater need for ambulatory monitoring, accurate and automated detection of asymptomatic AF becomes an important task. It is impractical for a trained technician to sift through ~100,000 beats of data in order to identify the presence of AF on a daily basis.

Several algorithms have been developed to detect AF which either rely on the absence of P-waves<sup>1,3,5,7,8,16,18,20,22</sup> or are based on RR variability.<sup>6,17,22,24,26,27</sup> Since there is no uniform depolarization of the atria during AF and consequently no discernible P-waves in the ECG, their absence has been utilized in the detection of AF. However, locating the P-wave fiducial point is very difficult because the low amplitude of the P-wave makes it susceptible to corruption by noise. The methods in the second category are based on RR interval dynamics and do not require identification of the P-wave. However, few algorithms in this category show high predictive value for clinical application.<sup>6,24,26,27</sup> Notable exceptions include Duverney *et al.*,<sup>6</sup> Sarkar *et al.*<sup>24</sup> and Tateno

Address correspondence to K. H. Chon, Department of Biomedical Engineering, SUNY at Stony Brook, HSC T18, Rm. 030, Stony Brook, NY 11794, USA. Electronic mail: ki.chon@sunysb.edu

and Glass.<sup>26,27</sup> Duverney *et al.*<sup>6</sup> used wavelet transform of the RR time series while the latter used the Kolmogorov–Smirnov test to compare the density histogram of the test RR (and  $\Delta$ RR) segment with previously compiled standard density histograms of RR (and  $\Delta$ RR) segments during AF. Sarkar *et al.*<sup>24</sup> used the Lorenz distribution of a time series of RR intervals for AF and tachycardia detection for its use in a chronic implantable monitor. Tateno and Glass reported a sensitivity of 94.4% and specificity of 97.2% for the MIT BIH Atrial Fibrillation database.<sup>9,10</sup> Although the accuracy of the study by Duverney *et al.* was high, their results were based on a small database.<sup>6</sup> The main drawback of these algorithms is that they are dependent on the robustness of the training data. For example, if the characteristics of AF are different from those learned in the training data, the accuracy of AF detection is compromised.

In the current study, we use a combination of three different statistical methods capable of detecting the presence of randomness in a signal. Our expectation is that using such an approach will minimize the need for extensive storage capacity (as in the case of histogram comparisons) while preserving the accuracy of the detection. A beat-by-beat analysis of the detection results is presented and accuracy is shown through a ROC curve analysis. In addition we also used an ectopic beat filtering scheme to prevent misdetection of ectopic rhythms as AF. The ROC analysis revealed that the optimal segment length is 128 RR intervals with at least 50% AF to ensure correct classification of the segment as AF.

## METHODS

The approach we present here is based on the generally accepted characteristic of AF as a random sequence of heart beat intervals with markedly increased beat-to-beat variability and complexity. We have developed an algorithm combining three statistical techniques to exploit these characteristics, namely the Root Mean Square of Successive RR Differences to quantify variability (RMSSD), the Turning Points Ratio to test for randomness of the time series (TPR) and Shannon entropy to characterize its complexity (SE). In addition, in contrast to the Tateno–Glass method,<sup>26,27</sup> which relies on training data histograms, the current method is purely statistical in nature and thus less dependent on the diversity of training data. Once the thresholds for maximum sensitivity and specificity are determined for each of the three algorithms using the ROC curve analysis, no further heuristic tuning of the threshold values is required.

### Turning Point Ratio (TPR)

To determine whether a RR time series is random we apply a nonparametric statistical test<sup>29</sup> comparing the value of each RR relative to its neighbors. Given three random numbers  $a_1, a_2, a_3$ , where  $a_1 > a_2 > a_3$ , there are six possible combinations to generate a series. Among them,  $(a_1a_3a_2)$ ,  $(a_2a_3a_1)$ ,  $(a_2a_1a_3)$  and  $(a_3a_1a_2)$  include turning points while  $(a_1a_2a_3)$  and  $(a_3a_2a_1)$  do not. Thus if the time series is random the probability of a RR being surrounded by either two longer or two shorter intervals (“Turning Point”) is equal to  $2/3$ . In a random series of length  $l$  the expected number of turning points is found to be  $\frac{2l-4}{3}$ , and the standard deviation is  $\sqrt{\frac{16l-29}{90}}$ . Hence, the expected TPR of a random series is given by  $\frac{2l-4}{3} \pm \sqrt{\frac{16l-29}{90}}$ .<sup>29</sup> One can also define confidence limits of this ratio around the mean and standard deviation just described (assuming the TPR values for all  $l$ -length random segments to be normally distributed) to estimate the boundaries of randomness. A series with a ratio below and above the 95% confidence interval exhibits periodicity (e.g. sinus rhythm) whereas a TPR within the 95% confidence limit signifies random characteristics. Of course, one can optimize the confidence interval threshold to achieve best results with regard to sensitivity and specificity calculations, as explained under “Detector Optimization”. The confidence interval threshold is denoted by  $TprThresh$ .

### Root Mean Square of Successive Differences (RMSSD)

The second component of the algorithm, beat-to-beat variability, is estimated by the RMSSD. Since AF exhibits higher variability than regular rhythms such as sinus rhythm, the RMSSD is expected to be higher. For a given segment  $a(i)$  of RR intervals of some length  $l$ , the RMSSD is given by:

$$RMSSD = \sqrt{\frac{1}{(l-1)} \sum_{j=1}^{l-1} (a(j+1) - a(j))^2} \quad (1)$$

During optimization, the selection of the threshold value for RMSSD was compared against the mean value of the RR interval (MeanRR) rather than the individual RR interval series. This strategy is used to compensate against possible outliers (e.g. premature ventricular contraction) which can lead to false detection of AF. Consequently, the numerical value of the RMSSD threshold changes with each segment, but the threshold for the ratio (RMSSD/MeanRR) remains constant. This threshold is denoted by  $RmsThresh$ .

### Shannon Entropy (*SE*)

Shannon entropy provides a quantitative measure of uncertainty for a random variable. Specifically, the SE quantifies the likelihood that runs of patterns exhibiting regularity over some duration of data also exhibit similar regular patterns over the next incremental duration of data. For example, a random white noise signal (data are independent) is expected to have the highest SE value (1.0) since it shows maximum uncertainty in predicting the pattern of the signal whereas a simple sinusoidal signal (data are not independent) has a very low SE value approaching 0. Thus, the SE of normal sinus rhythm is expected to be significantly lower than in AF. To calculate SE of the RR time series, we first construct a histogram of the segment considered. The eight longest and eight shortest RR values in the segment are considered outliers and are removed from consideration. The remaining RRs are sorted into equally spaced bins whose limits are defined by the shortest and longest R after removing outliers. To obtain a reasonably accurate measure of the SE, at least 16 such bins are required. The probability distribution is computed for each bin as the number of beats in that bin divided by the total number of beats in the segment (after removing outliers), i.e.,

$$p(i) = \frac{N_{\text{bin}(i)}}{l - N_{\text{outliers}}} \quad (2)$$

where  $N_{\text{bin}(i)}$  is the number of beats in the  $i$ th bin,  $l$  is the total number of beats in the segment and  $N_{\text{outliers}}$  is the number of outliers in that segment (16 in this case).

Finally, the Shannon Entropy is calculated as follows:

$$\text{SE} = - \sum_{i=1}^{16} p(i) \frac{\log(p(i))}{\log(\frac{1}{16})} \quad (3)$$

For optimization purposes, we compare sensitivity and specificity of the algorithm by varying the threshold of the SE (denoted by  $\text{SeThresh}$ ).

### Filtering of Ectopic Beats

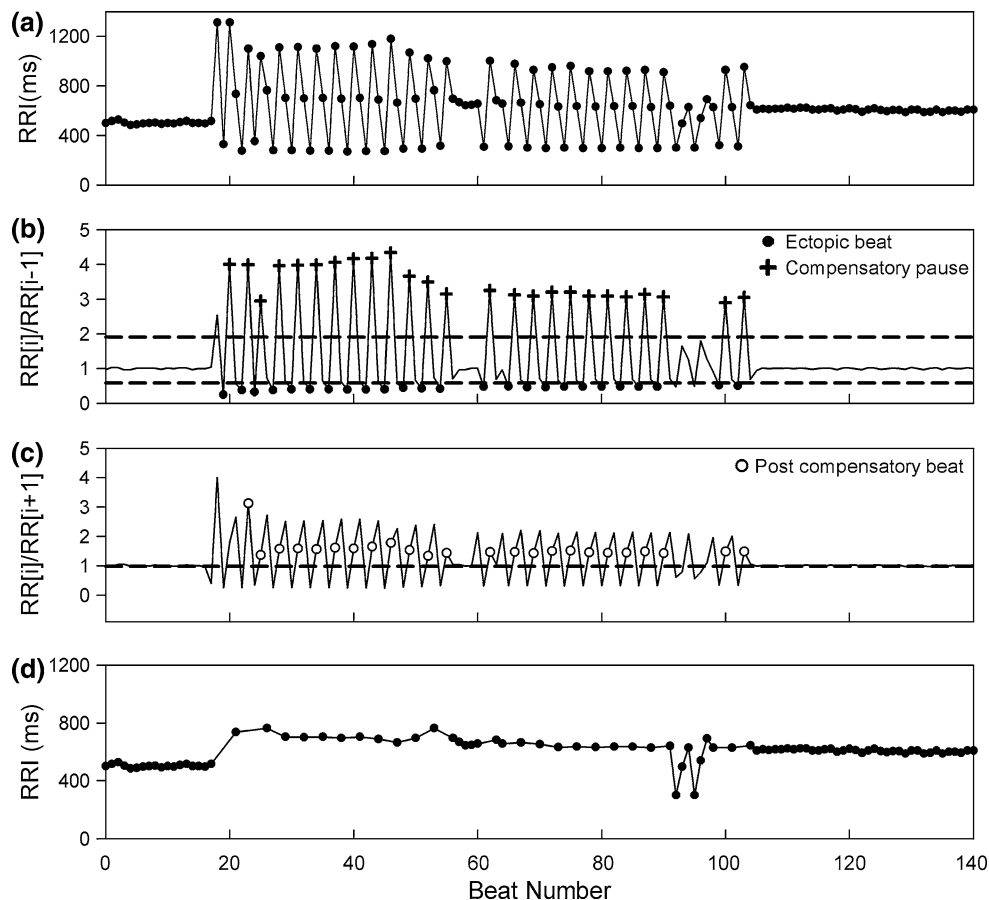
Ectopic beats during regular sinus rhythm are a potential cause of false detection of AF since they confound all components of the algorithm. Premature beats can be recognized by their signature short-long RR sequence (ectopic coupling interval and compensatory pause, respectively) wedged between normal RRs. Thus, if  $\text{RR}[i]$  is premature and followed by the compensatory pause  $\text{RR}[i + 1]$ , then  $\text{RR}[i - 1] > \text{RR}[i] < \text{RR}[i + 1]$  and  $\text{RR}[i] < \text{RR}[i + 1] > \text{RR}[i + 2]$

yielding at least two additional turning points and three if  $\text{RR}[i + 1] > \text{RR}[i + 2] < \text{RR}[i + 3]$ . In order to recognize the characteristic short-long pattern we compute the ratio  $\text{RR}[i]/\text{RR}[i - 1]$  for each RR in the time series. During sinus rhythm this ratio is close to unity with small fluctuations reflecting physiologic variability. In the case of an ectopic beat the sequence of ratios approximates  $\text{RR}[i]/\text{RR}[i - 1] \leq 0.8$ ,  $\text{RR}[i + 1]/\text{RR}[i] \geq 1.3$ , and  $\text{RR}[i + 2]/\text{RR}[i + 1] \leq 0.9$  depending on the type of ectopic beat (e.g. supraventricular vs. ventricular and unifocal vs. multifocal). Hence, rather than relying on these arbitrary fixed ratios we identify diverse ectopic beats with varying coupling intervals by searching for RR sequences which satisfy the conditions  $\text{RR}[i]/\text{RR}[i - 1] < \text{Perc1}$  and  $\text{RR}[i + 1]/\text{RR}[i] > \text{Perc99}$  and  $\text{RR}[i + 1]/\text{RR}[i + 2] > \text{Perc25}$  where  $\text{Perc1}$ ,  $\text{Perc99}$  and  $\text{Perc25}$  denote the first, 99th and 25th percentiles, respectively. Thus, we label a beat as ectopic when the preceding RR interval ratio belongs to the shortest 1% and the RR ratio following it belongs to the longest 1% of all RR interval ratios. When an ectopic beat is encountered, it is excluded from further analysis along with its compensatory pause, thereby creating a ‘‘clean’’ time series largely devoid of ectopic beats. In addition, ambulatory recordings often contain undetected R waves causing dropouts in the beat series. We use the above approach to eliminate these artifacts. Specifically, if  $\text{RR}[i]$  comprises one or more undetected RR, then  $\text{RR}[i]/\text{RR}[i - 1] > \text{Perc99}$  and  $\text{RR}[i + 1]/\text{RR}[i] < \text{Perc1}$ . The percentile threshold values, though chosen empirically, were found to be fairly robust across both databases. Figure 1 illustrates an example of the filtering scheme, where panel (a) shows the original RR series contaminated with ectopic beats, panel (b) and (c) show the sequence of ratios as defined above along with the thresholds, and panel (d) shows the RR sequence after removal of the ectopic beats.

### Detector Optimization

After removal of ectopic beats, the complete RR interval series is linked together yielding a continuous RR interval sequence free of ectopic beats on which the analysis is carried out. It must be noted that the true temporal location of each segment is preserved, so that after completion of the detection process the time of onset of AF and non-AF is not distorted. The condition for AF classification of an ectopy-free RR segment of length  $l$  is now given by a simple logical AND condition:

- if  $((\text{RMS}/\text{MeanRR} > \text{RmsThresh}) \text{ AND } (\text{TPR} \text{ within } \text{TprThresh} \text{ confidence interval}) \text{ AND } (\text{SE} > \text{SeThresh}))$  then classify segment as AF
- else classify segment as non-AF.



**FIGURE 1.** (a) A sample beat sequence from file 8219 in MIT BIH AFIB database containing normal sinus rhythm (NSR) punctuated by ectopy; (b) the ratio of each RR interval to the one preceding it. Lower and upper dashed lines correspond to Perc1 and Perc99 thresholds respectively. Ectopic beat candidates were those beats for which this ratio was  $< \text{Perc1}$  and the ratio value at the subsequent location was  $> \text{Perc99}$ . (c) The ratio of each RR interval to the one occurring just after it. For the ectopy candidates in (b) this ratio was calculated for beats occurring 2 beats after the candidate. If it exceeded the Perc25 threshold (dashed line: value of 0.9), the beat was confirmed as an ectopic beat. (d) The final clean RR segment from which all such ectopic beats and compensatory pauses were deleted.

Our algorithm is based on statistics calculated from  $l$ -beat segments, and consequently it would not be fair to compare the result of this algorithm with beat-to-beat annotations, because of the inherent  $l$ -beat uncertainty in the calculation. A more useful technique would be to convert the original beat-to-beat annotations to  $l$ -beat resolution such that a particular  $l$ -beat segment is classified as a true AF segment only if the number of true AF beats (as annotated in the database) is more than a minimum threshold number or percentage. We denote this minimum threshold as PercThresh. The detection result for each  $l$ -beat segment can then be compared to the new annotation. Note that the objective here is not to classify whether *each* beat is AF, but whether *every*  $l$ -beat RR segment is AF or not, and hence this conversion of annotations is valid. The ROC analyses are used to find  $l$  and PercThresh for optimal results. A similar annotation conversion procedure has been employed in Sarkar *et al.*<sup>24</sup>

In addition to  $l$  and PercThresh, the thresholds for the three statistics we have used are also tuned for optimum sensitivity and specificity. In summary, the algorithm parameters considered for this optimization problem are:

1. RR interval segment length  $l$ , varied from 32 (minimum number of beats must be 16 for outlier conditions) to 480 at intervals of 32 beats.
2. PercThresh (varied from 0 to 100%) at intervals of 10%.
3.  $\frac{\text{RMSSD}}{\text{MeanRR}}$  threshold (RmsThresh) which is varied from 0 to 1 at intervals of 0.02.
4. TPR threshold (TprThresh), varied from a 99.99% confidence interval to 50% confidence interval at intervals of 0.01%.
5. SE threshold (SeThresh), varied from 0 to 1 at intervals of 0.01.



We can now define a 5-element vector of algorithm parameters  $\alpha$  as

$$\alpha = [l, \text{PercThresh}, \text{RmsThresh}, \text{TprThresh}, \text{SeThresh}]^T \quad (4)$$

This vector parameter can now be varied according to the ranges defined above. For each particular value of the vector  $\alpha_k$ , we find the number of True Positives ( $\text{TP}_k$ ), True Negatives ( $\text{TN}_k$ ), False Positives ( $\text{FP}_k$ ) and False Negatives ( $\text{FN}_k$ ) in the detection results using  $\alpha_k$  as the threshold. We use the sensitivity ( $\text{TP}_k/(\text{TP}_k + \text{FN}_k)$ ) and specificity ( $\text{TN}_k/(\text{TN}_k + \text{FP}_k)$ ) metrics in order to quantify accuracy of the detection for the vector parameter  $\alpha_k$ . Sensitivity and specificity thus quantify the ability of the algorithm to correctly detect AF and to correctly identify non-AF segments. We used the MIT BIH<sup>9,10</sup> Atrial Fibrillation database to optimize the thresholds for the algorithm, after which the detection algorithm was tested on the MIT BIH Arrhythmia database.

#### *MIT BIH Atrial Fibrillation Database*

This database contains 25 fully annotated ECG recordings containing a total of 299 AF episodes. Each ECG recording is approximately 10 h long and is sampled at 250 Hz. It also contains some cases of atrial flutter and Normal Sinus Rhythm. Data sets 4936 and 5091 were excluded from our study because some RR intervals were incorrectly annotated.

#### *MIT BIH Arrhythmia Database*

This database consists of 48 half-hour annotated ECG recordings sampled at 360 Hz. Of these, 23 are in the 100 series and the rest are in the 200 series. The recordings in the 100 series contain sinus rhythm and arrhythmias but no AF episodes. The 200 series contains AF, various arrhythmias and sinus rhythm.

An example calculation of the aforementioned statistics along with the final detection using the corresponding thresholds for a sample recording from the MIT BIH Atrial Fibrillation database is shown in Fig. 2.

## RESULTS

The  $\alpha$  value for optimum sensitivity and specificity was found to be (compare to Eq. (4)):

$$\alpha_{\text{opt}} = [128, 50\%, 0.1, 99.9\%, 0.7]^T \quad (5)$$

These threshold values have also been summarized in Table 1. For  $\alpha = \alpha_{\text{opt}}$ , we obtained a sensitivity = 94.4%, specificity = 95.1% for the MIT-BIH Atrial

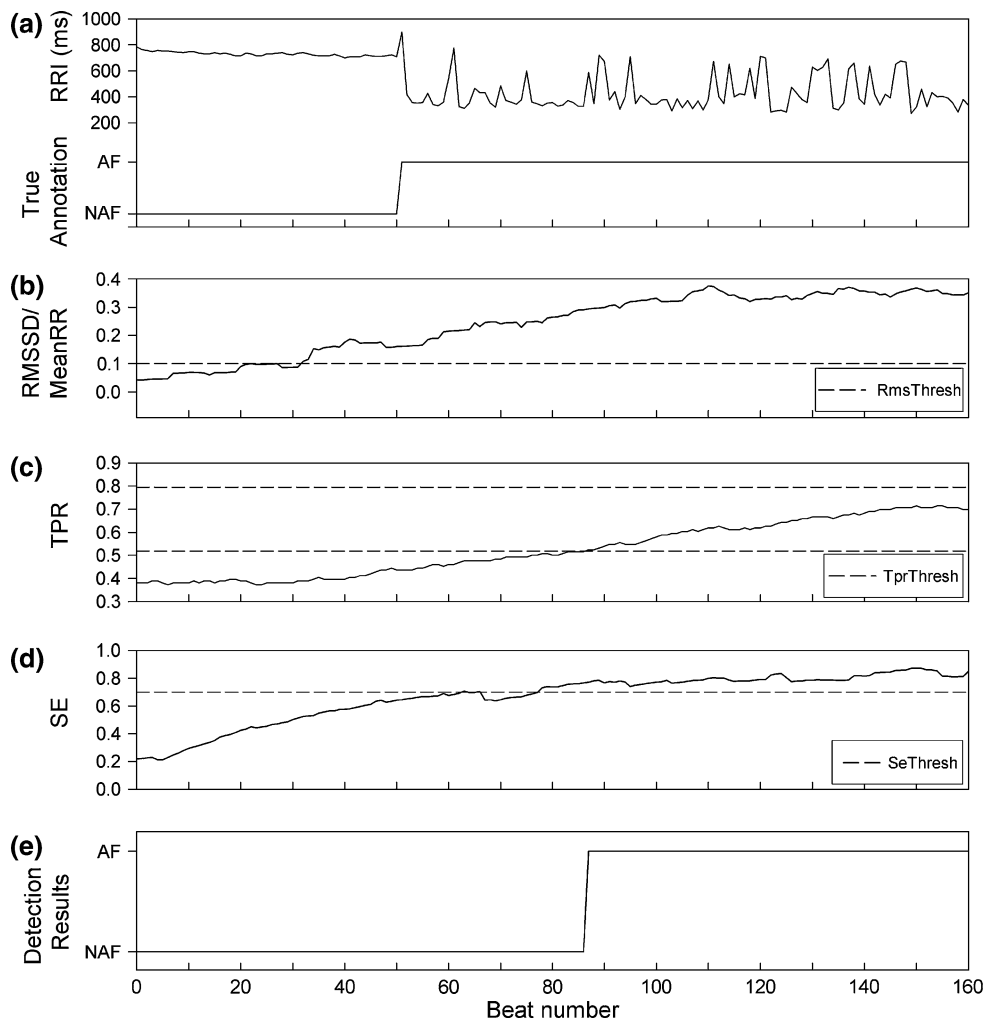
Fibrillation database and sensitivity = 90.2%, specificity = 91.2% for the MIT BIH Arrhythmia database (200 series).

For the 100 series in the same database, the specificity was 99.52%, and since this series contains no true AF beats, the sensitivity cannot be quantified. These results are tabulated in Table 2.

The optimization parameter  $\alpha$  is a 5-dimensional vector, thus, it is difficult to plot the accuracy metrics as a function of all the elements of the vector. For ease of visualization, we have shown only the ROC curves (sensitivity vs. specificity) obtained by varying SeThresh while keeping the other elements of the parameter  $\alpha$  constant. Figure 3a shows these ROC curves obtained for different segment lengths and Fig. 3b shows similar ROC curves obtained for different values of PercThresh. The values for RmsThresh and TprThresh were kept equal to the optimum values obtained. It should be noted that the actual optimization was performed for the entire range of each parameter, and not just for the SeThresh range.

## DISCUSSION

There is an extensive body of literature on the characteristics of AF and whether its RR sequence is deterministic<sup>21</sup> or random.<sup>25</sup> The prevailing view considers AF to be random<sup>25</sup> and this assumption underlies our algorithm for AF detection which employs a nonparametric statistic to test for randomness of the RR time series.<sup>29</sup> In order to enhance its robustness, the algorithm also analyzes RR variability and complexity. Our ability to detect AF with high sensitivity and specificity in two large well documented databases is consistent with the notion that AF is indeed random. Prior studies have differentiated AF from other rhythms by examining the distribution of RR intervals, successive RR differences,<sup>26,27</sup> or ratios of successive RR. Compared with the results reported by Tateno and Glass<sup>26,27</sup> in the MIT-BIH Atrial Fibrillation Database our algorithm was equally sensitive (94.4% vs. 94.4%) but slightly less specific (95.1% vs. 97.2%). For the 200 series of the MIT-BIH Arrhythmia Database, both sensitivity (90.2% vs. 88.2% obtained by Tateno and Glass) and specificity (91.2% vs. 87.2%) were higher, most likely due to the elimination of ectopic beats by our filtering scheme. The MIT-BIH AF database results by Sarkar *et al.*<sup>24</sup> are promising. However, their results on MIT-BIH Arrhythmia Database have not been reported. It should be noted that the accuracy of the algorithm by Tateno and Glass,<sup>27</sup> and Sarkar *et al.*<sup>24</sup> depends on the richness of the training data. If a new AF dataset has characteristics that are different from those of the



**FIGURE 2.** (a) Original RR interval time series from a section of file 4048 of the MIT BIH Atrial Fibrillation database along with the true annotation. Onset of AF can be seen. Calculation of the (b) RMSSD/MeanRR, (c) TPR, (d) Shannon Entropy is shown along with the optimum thresholds (dashed lines). (e) The final detection results based on whether these statistics cross their respective thresholds.

**TABLE 1.** Parameters used to classify an *l*-beat RR interval segment as AF.

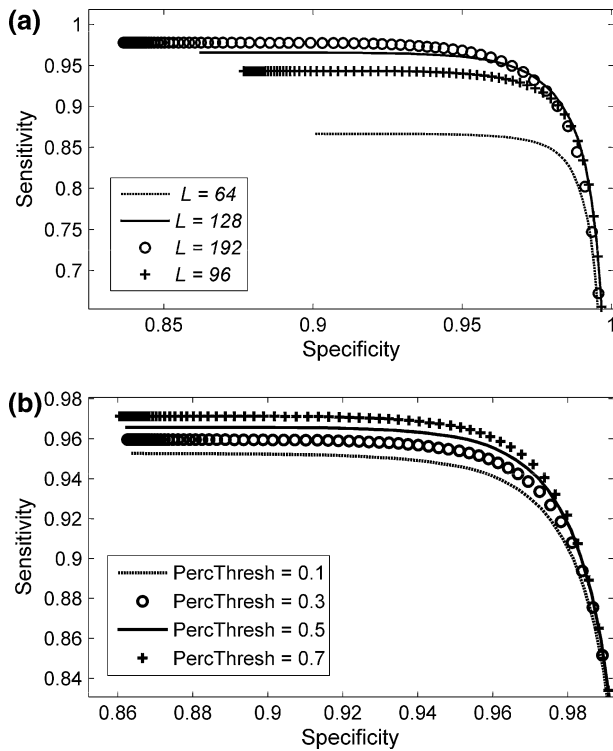
Parameter	Meaning	Range of parameter values for classifying segment as AF
RMSSD	Root mean square of successive differences	$RMSSD > 0.1 * \text{meanRR}$
TPR	Turning point ratio	$0.54 < TPR < 0.77$
SE	Shannon entropy	$SE > 0.7$
<i>l</i>	Segment length in beats	128

MIT-BIH<sup>9,10</sup> database, their accuracy is likely to be diminished. While our method’s threshold values were also trained on the MIT-BIH Database, it is less dependent on the diversity of the training data since the three algorithms are well-established statistical

**TABLE 2.** Sensitivity and specificity values for all databases tested, for the optimal parameter vector  $\alpha = \alpha_{opt} = [128, 50\%, 0.1, 99.9\%, 0.7]^T$ .

Name of database	AF beats		AF episodes Accuracy
	Sensitivity	Specificity	
MIT BIH atrial fibrillation database	94.4%	95.1%	99.1%
MIT BIH arrhythmia (100-series) database			
With ectopy	NA	96.2%	NA
Without ectopy	NA	99.5%	NA
MIT BIH arrhythmia (200-series) database			
With ectopy	96.5%	69.4%	100%
Without ectopy	90.2%	91.2%	100%

Results for before and after removal of all PVCs have been shown in the case of the Arrhythmia databases. The Arrhythmia 100-series database did not have any true AF beats and so only specificity can be calculated for this database.



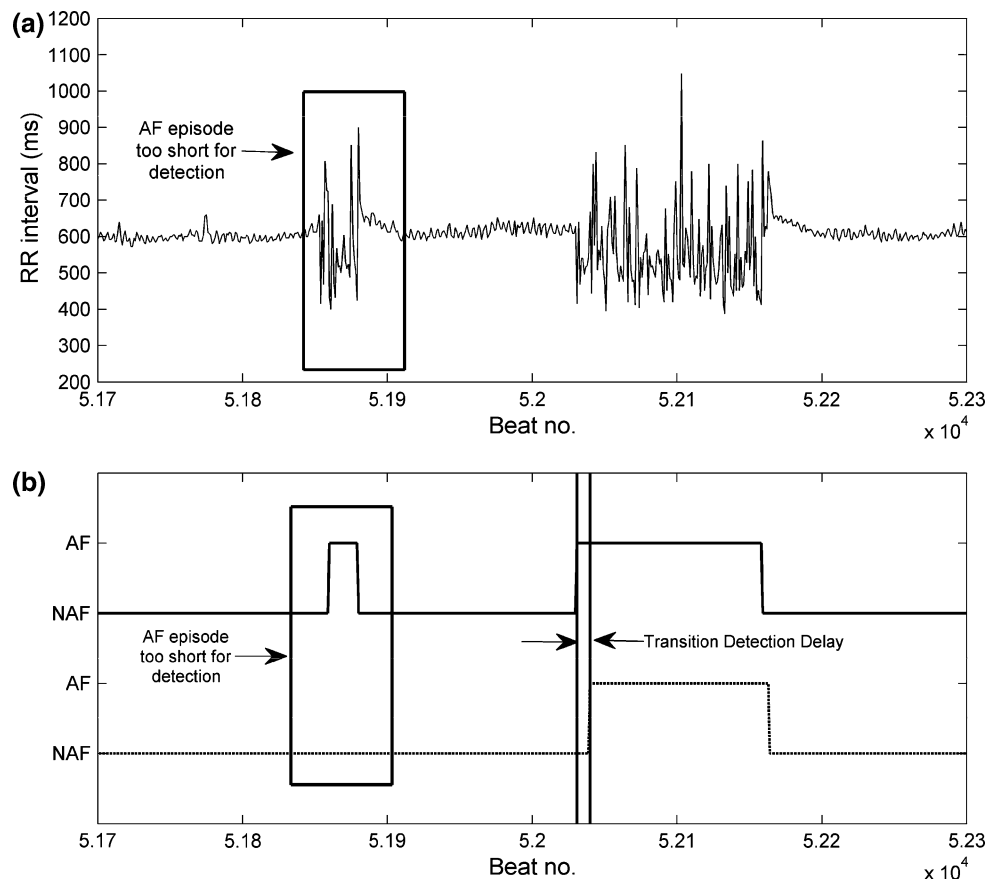
**FIGURE 3. Receiver Operating Characteristic (ROC) curves with  $RmsThresh = 0.1$ ,  $TprThresh = 99.9$ th percentile and (a)  $PercThresh = 0.5$ ,  $l$  varying,  $SeThresh$  varying; (b)  $l = 128$ ,  $PercThresh$  and  $SeThresh$  varying.**

methods. The algorithms by Tateno and Glass,<sup>27</sup> and Sarkar *et al.*<sup>24</sup> also require a memory bank of histogram characteristics of AF while our method only requires three threshold values. Furthermore, the computational speed of our method is faster than Tateno and Glass. The computation time of a 128-beat data segment is of the order of 2 ms, whereas the Tateno–Glass method takes around 30 ms for a 100-beat segment (programs run on MATLAB). Hence, automatic real time detection of AF in a clinical setting would be computationally less expensive. The computation time for each 128-beat segment can be further lessened since we are dealing with a logical AND condition, which means it is not necessary to calculate all three statistics for each and every RR segment. If even one of the statistics (e.g., RMSSD since it is the most computationally efficient) fails to cross the corresponding threshold, one can simply mark that segment as non-AF and skip to the next RR segment.

The segment length of 128 beats blurs the transition between non-AF and AF, causing a transition detection delay. For the MIT-BIH Atrial Fibrillation Database, the average delay was 17 beats (non-AF to AF) and 19 beats (AF to non-AF) and for the MIT-BIH Arrhythmia Database, 13 and 17 beats,

respectively (Fig. 4). Hence, selection of this 128-beat window delays the detection of AF by 10–15 s on average. However, from a clinical perspective it is important to note that AF beats do not exist in isolation but only as part of AF episodes. Accordingly we have converted the original beat-to-beat annotations to  $l$ -beat resolution annotations as explained in “Detector Optimization”. A cardiologist observing real-time telemetry needs to observe a minimum number of RR intervals in order to diagnose AF. Moreover, the accuracy of this diagnosis increases with the number of RR intervals. Thus, the window size (128 beats in this study) represents a trade-off between accuracy and speed of diagnosis. In the MIT-BIH Atrial Fibrillation Database (excluding files 4936 and 5091), the total number of true AF episodes is 254 of which 224 were correctly detected by our algorithm. The remaining 30 episodes were less than 75 beats in length. In the MIT-BIH Arrhythmia Database, our algorithm detected 96/107 episodes in the 200 series; the remaining 11 episodes were less than 11 beats long which suggests that AF detection is robust except for very short episodes. In keeping with the optimization results, we choose the value of 64 beats ( $PercThresh = 0.5$  for  $l = 128$ ) as the shortest AF episode that the algorithm can accurately detect. It should be recognized that for clinical applications, the most relevant objective is to detect the presence of AF in a given recording, i.e. not necessarily every single AF beat. Using this latter criterion and excluding episodes that are less than 64 beats long, we achieved an episode-detection accuracy of 99.1% and 100% for the MIT-BIH AF and Arrhythmia (200 series) databases, respectively (see Table 2).

Another potential issue is the possible decrease in the detection accuracy because of the linking of previously temporally separated RR segments because of ectopic beat removal. While it is plausible that such a “linking” process may produce some statistical artifacts, we have compensated for it by “de-linking” the segments after detection was done on the “linked” series. In other words, the temporal location of the detected AF segments is preserved throughout the algorithm. Hence the only significant distortions would be at the boundaries of the two “de-linked” segments. For example, consider an RR series in which a short NSR segment is followed by some ectopic rhythm and then by an AF segment. In this case, removal of the ectopic segment would cause the sinus rhythm and AF segments to be temporally linked. The detection algorithm will now operate on a series consisting of NSR followed by AF rhythms and it is possible that some of the NSR beats might be incorrectly marked as AF. It must however be noted



**FIGURE 4.** (a) RR interval series segment from file 4043 of the MIT BIH Atrial Fibrillation database showing a very short AF episode (around 20 beats) and a normal AF episode. (b) The solid line shows the true annotations of the RR intervals series and the dotted line shows the detection using the current protocol. The figure shows the delay in detection of the normal episode which is denoted as the Transition Detection Delay. Also shown is the failure to detect AF episodes that are too short.

that exactly the same rationale would have been valid had we not removed the ectopic segment. In that case, some of the (non-AF) ectopic beats (instead of the NSR beats) would have been incorrectly marked as AF. However, the number of incorrect AF markings would be greater in the case when the ectopy is not removed because in general, ectopic rhythms exhibit more randomness than NSR. The purpose of the ectopic beat filtering scheme is simply to reduce the number of “blurry” transitions from non-AF to AF segments arising because of ectopic beats and their compensatory pauses. After the detection is performed on such “linked” series, the results are then temporally “de-linked” so that they correspond correctly to the actual time of occurrence.

Note that the MIT-BIH database annotations for cardiac rhythm were performed by experts but they do have a component of subjectivity. While the accuracy of annotation by experts depends on the fidelity of RR intervals of the MIT-BIH database,<sup>9,10</sup> we are fortunate to have such a databank to test the

performance of our algorithms in comparison to other methods<sup>24,27</sup> which have also used the same database.

## FUTURE DEVELOPMENT

The use of ROC curves allows tuning of thresholds to suit specific applications, modifying the sensitivity and specificity of the algorithm. For example, in ambulatory monitoring a high sensitivity may be more important than specificity when detection of rare AF episodes is of paramount importance. On the other hand, unnecessary alarms are undesirable in hospital settings, favoring an emphasis on specificity. Future development may lead to improvements of AF detection from RR sequences and decrease the transition detection delay. The ultimate goal is to implement the algorithm in portable non-invasive sensors which can be used for screening of patients at risk for atrial fibrillation.



## REFERENCES

- <sup>1</sup>Aytemir, K., S. Aksoyek, A. Yildirim, N. Ozer, and A. Oto. Prediction of atrial fibrillation recurrence after cardioversion by P wave signal-averaged electrocardiography. *Int. J. Cardiol.* 70:15–21, 1999.
- <sup>2</sup>Benjamin, E. J., P. A. Wolf, R. B. D'Agostino, H. Silbershatz, W. B. Kannel, and D. Levy. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. *Circulation* 98:946–952, 1998.
- <sup>3</sup>Clavier, L., J. M. Boucher, R. Lepage, J. J. Blanc, and J. C. Cornily. Automatic P-wave analysis of patients prone to atrial fibrillation. *Med. Biol. Eng. Comput.* 40:63–71, 2002.
- <sup>4</sup>Defaye, P., F. Dournaux, and E. Mouton. Prevalence of supraventricular arrhythmias from the automated analysis of data stored in the DDD pacemakers of 617 patients: the AIDA study. The AIDA Multicenter Study Group. Automatic Interpretation for Diagnosis Assistance. *Pacing Clin. Electrophysiol.* 21:250–255, 1998.
- <sup>5</sup>Dotsinsky, I. Atrial wave detection algorithm for discovery of some rhythm abnormalities. *Physiol. Meas.* 28:595–610, 2007.
- <sup>6</sup>Duverney, D., J. Gaspoz, V. Pichot, F. Roche, R. Brion, A. Antoniadis, and J. Barthélémey. High accuracy of automatic detection of atrial fibrillation using wavelet transform of heart rate intervals. *Pacing Clin. Electrophysiol.* 25:457–462, 2002.
- <sup>7</sup>Ehrlich, J. R., K. Schadow, K. Steul, G. Q. Zhang, C. W. Israel, and S. H. Hohnloser. Prediction of early recurrence of atrial fibrillation after external cardioversion by means of P wave signal-averaged electrocardiogram. *Z. Kardiol.* 92:540–546, 2003.
- <sup>8</sup>Fukunami, M., T. Yamada, M. Ohmori, K. Kumagai, K. Umemoto, A. Sakai, N. Kondoh, T. Minamino, and N. Hoki. Detection of patients at risk for paroxysmal atrial fibrillation during sinus rhythm by P wave-triggered signal-averaged electrocardiogram. *Circulation* 83:162–169, 1991.
- <sup>9</sup>Goldberg, R. J. To the framingham data, turn, turn, turn. *Circulation* 119:1189–1191, 2009.
- <sup>10</sup>Goldberger, A. L., L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101:E215–E220, 2000.
- <sup>11</sup>Heeringa, J., D. A. van der Kuip, A. Hofman, J. A. Kors, G. van Herpen, B. H. Stricker, T. Stijnen, G. Y. Lip, and J. C. Witteman. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. *Eur. Heart J.* 27: 949–953, 2006.
- <sup>12</sup>Humphries, K. H., C. R. Kerr, S. J. Connolly, G. Klein, J. A. Boone, M. Green, R. Sheldon, M. Talajic, P. Dorian, and D. Newman. New-onset atrial fibrillation: sex differences in presentation, treatment, and outcome. *Circulation* 103:2365–2370, 2001.
- <sup>13</sup>Israel, C. W., G. Gronefeld, J. R. Ehrlich, Y.-G. Li, and S. H. Hohnloser. Long-term risk of recurrent atrial fibrillation as documented by an implantable monitoring device: implications for optimal patient care. *J. Am. Coll. Cardiol.* 43:47–52, 2004.
- <sup>14</sup>Kerr, C., J. Boone, S. Connolly, M. Greene, G. Klein, R. Sheldon, and M. Talajic. Follow-up of atrial fibrillation: the initial experience of the Canadian Registry of Atrial Fibrillation. *Eur. Heart J.* 17(Suppl C):48–51, 1996.
- <sup>15</sup>Kinlay, S., J. W. Leitch, A. Neil, B. L. Chapman, D. B. Hardy, and P. J. Fletcher. Cardiac event recorders yield more diagnoses and are more cost-effective than 48-hour Holter monitoring in patients with palpitations. A controlled clinical trial. *Ann. Intern. Med.* 124:16–20, 1996.
- <sup>16</sup>Michalkiewicz, D., M. Dziuk, G. Kaminski, R. Olszewski, M. Cholewa, A. Cwetsch, and L. Markuszewski. Detection of patients at risk for paroxysmal atrial fibrillation (PAF) by signal averaged P wave, standard ECG and echocardiography. *Pol. Merkur. Lekarski* 20:69–72, 2006.
- <sup>17</sup>Murgatroyd, F. D., B. Xie, X. Copie, I. Blankoff, A. J. Camm, and M. Malik. Identification of atrial fibrillation episodes in ambulatory electrocardiographic recordings: validation of a method for obtaining labeled r-r interval files. *Pacing Clin. Electrophysiol.* 18:1315–1320, 1995.
- <sup>18</sup>Opolski, G., P. Scislo, J. Stanislawska, G. Aleksander, R. Steckiewicz, and A. Torbicki. Detection of patients at risk for recurrence of atrial fibrillation after successful electrical cardioversion by signal-averaged P-wave ECG. *Int. J. Cardiol.* 60:181–185, 1997.
- <sup>19</sup>Page, R. L., W. E. Wilkinson, W. K. Clair, E. A. McCarthy, and E. L. Pritchett. Asymptomatic arrhythmias in patients with symptomatic paroxysmal atrial fibrillation and paroxysmal supraventricular tachycardia. *Circulation* 89:224–227, 1994.
- <sup>20</sup>Portet, F. P wave detector with PP rhythm tracking: evaluation in different arrhythmia contexts. *Physiol. Meas.* 29:141–155, 2008.
- <sup>21</sup>Rawles, J. M., and E. Rowland. Is the pulse in atrial fibrillation irregularly irregular? *Br. Heart J.* 56:4–11, 1986.
- <sup>22</sup>Reddy, S., B. Young, Q. Xue, B. Taha, D. Brodnick, and J. Steinberg. Review of methods to predict and detect atrial fibrillation in post-cardiac surgery patients. *J. Electrocardiol.* 32(Suppl):23–28, 1999.
- <sup>23</sup>Roche, F., J. M. Gaspoz, A. Da Costa, K. Isaaz, D. Duverney, V. Pichot, F. Costes, J. R. Lacour, and J. C. Barthélémey. Frequent and prolonged asymptomatic episodes of paroxysmal atrial fibrillation revealed by automatic long-term event recorders in patients with a negative 24-hour Holter. *Pacing Clin. Electrophysiol.* 25:1587–1593, 2002.
- <sup>24</sup>Sarkar, S., D. Ritscher, and R. Mehra. A detector for a chronic implantable atrial tachyarrhythmia monitor. *IEEE Trans. Biomed. Eng.* 55:1219–1224, 2008.
- <sup>25</sup>Stein, K. M., J. Walden, N. Lippman, and B. B. Lerman. Ventricular response in atrial fibrillation: random or deterministic? *Am. J. Physiol.* 277:H452–H458, 1999.
- <sup>26</sup>Tateno, K., and L. Glass. A method for detection of atrial fibrillation using RR intervals. *Comput. Cardiol.* 27:391–394, 2000.
- <sup>27</sup>Tateno, K., and L. Glass. Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and  $\Delta$ RR intervals. *Med. Biol. Eng. Comput.* 39:664–671, 2001.
- <sup>28</sup>Wolf, P. A., R. D. Abbott, and W. B. Kannel. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 22:983–988, 1991.
- <sup>29</sup>Zar, J. H. Serial randomness of measurements: nonparametric testing. In: *Biostatistical Analysis*, 4th edn. Prentice-Hall, 1999, pp. 587–588.