New Bayesian Discriminator for Detection of Atrial Tachyarrhythmias

Weichao Xu, BE, MSEE; Hung-Fat Tse, MD; Francis H.Y. Chan, PhD; Peter Chin Wan Fung, PhD; Kathy Lai-Fun Lee, MB; Chu-Pak Lau, MD

Background—Accurate, rapid detection of atrial tachyarrhythmias has important implications in the use of implantable devices for treatment of cardiac arrhythmia. Currently available detection algorithms for atrial tachyarrhythmias, which use the single-index method, have limited sensitivity and specificity.

Methods and Results—In this study, we evaluated the performance of a new Bayesian discriminator algorithm in the detection of atrial fibrillation (AF), atrial flutter (AFL), and sinus rhythm (SR). Bipolar recording of 364 rhythms (AF=156, AFL=88, SR=120) at the high right atrium were collected from 20 patients who underwent electrophysiological procedures. After initial signal processing, a column vector of 5 features for each rhythm were established, based on the regularity, rate, energy distribution, percent time of quiet interval, and baseline reaching of the rectified autocorrelation coefficient functions. Rhythm identification was obtained by use of Bayes decision rule and assumption of Gaussian distribution. For the new Bayesian discriminator, the overall sensitivity for detection of SR, AF, and AFL was 97%, 97%, and 94%, respectively; and the overall specificity for detection of SR, AF, and AFL was 98%, 98%, and 99%, respectively. The overall accuracy of detection of SR, AF, and AFL was 98%, 97% and 98%, respectively. Furthermore, sensitivity, specificity, and accuracy of this algorithm were not affected by a range of white Gaussian noises with different intensities.

Conclusions—This new Bayesian discriminator algorithm, based on Bayes decision of multiple features of atrial electrograms, allows rapid on-line and accurate (98%) detection of AF with robust anti-noise performance. (Circulation. 2002;105:1472-1479.)

Key Words: tachyarrhythmias, intervals, fibrillation, pacemakers, atrium

Atrial fibrillation (AF) is the most common supraventricular arrhythmia associated with a considerable risk of morbidity and mortality.1–3 Theoretical analyses and high-density mapping studies have suggested that the most common mechanism of AF is the presence of multiple wavefronts or “wavelets” circulating irregularly throughout the atrial tissue.4–6 In dual-chamber pacemakers, accurate AF detection is critically important to avoid rapid ventricular pacing by activating automatic mode switching. In an implantable cardioverter-defibrillator, accurate recognition of AF can avoid false discharges. Furthermore, the recent development of automatic implantable atrial defibrillators has created a critical need for speedy and reliable discrimination of AF from other types of intra-atrial electrograms.7–9 Previous studies have used time-domain, frequency-domain, or time-frequency analyses to differentiate fibrillatory from nonfibrillatory rhythms. However, most of these techniques are essentially single-index methods and have some limitations: (1) long process time, (2) lack of robustness to noise or far-field events, (3) poor performance to discriminate atrial flutter (AFL) from sinus rhythm (SR), and, (4) relatively low sensitivity and specificity.

Motivated by the need to overcome the aforementioned problems, we developed a new method for AF detection based on multivariate Bayes decision (Appendix A), which combine 5 different features of the intra-atrial electrogram. The purpose of this study was (1) to compare the sensitivity, specificity, and accuracy for detection of AF, AFL, and SR by using different numbers of features; (2) to evaluate the sensitivity, specificity, and accuracy of this new Bayesian discriminator algorithm for detection of AF, AFL, and SR; and (3) to test the robustness of this detection algorithm against different ranges of noise.

Methods

Data Acquisition

Bipolar intra-atrial electrograms at the high anterolateral right atrium (with a 1-cm interelectrode distance) from 20 patients in AF, AFL, and SR were amplified and recorded (CardioLab 4.11, Pruka Engineering, Inc) during electrophysiological procedures. These consecutive patients were seen in our electrophysiology.
laboratory for internal cardioversion of AF, electrophysiology study, and/or radiofrequency ablation procedures for their underlying arrhythmias. Up to 220 seconds (mean, 190 ± 20 seconds; range, 180 to 220 seconds) of simultaneous unfiltered (band pass, 0.04 to 5000 Hz) recordings from each patient were digitized at 1000 Hz. Data were then split into 1 (AF and AFL) or 2 seconds (SR) segments for analysis so that at least 2 atrial events were recorded during SR. To have an unbiased data set, nearly the same numbers of episodes were randomly collected from each patient. Computer processing was performed with the use of a Matlab 5.3 computer program (The Mathwork, Inc).

Signal Manipulation
Before extracting the features of the signal, each rhythm episode was processed with the following manipulations: (1) third-order Butterworth bandpass filtering (40 to 250 Hz), (2) absolute valuing, (3) low-pass filtering (0 to 20 Hz), (4) autocorrelation, and (5) rectification (Figure 1). Steps 1 to 3 output a flattened signal proportional to the high-frequency energy contained in the input episode.10,11 The autocorrelation process is performed to avoid drastic fluctuation of the amplitude of atrial electrograms with time.12,13 Then, the rectification process removes all the negative parts of the processed signal to facilitate the mathematical treatment during feature extraction.

Feature Extraction Procedure
Five relevant feature parameters were extracted from the final processed signal by a feature extraction procedure (Figure 1). The first feature (f1) is defined as the first peak, occurring at time (t), which is positively related to the regularity of the input. The second feature (f2) is defined as \( f_2 = t/1000 \) and is proportional to the input atrial rate. Feature 3 (f3) is defined as the percentage of energy contained in the two time bands \( [E_1/E, E_2/E] \), where \( E_1 \) is the energy within 0 to 100 ms, \( E_2 \) is the energy within 500 ms to 1000 ms, and \( E \) is the total energy within 0 to 1000 ms. The typical sinus rate is 60 to 120 beats per minute, that is, the corresponding peak-to-peak interval is 500 to 1000 ms. In SR, the energy is mainly distributed in the above-mentioned two time bands. Therefore, \( f_1 \) is helpful in distinguishing SR signals from the other two classes of rhythm (AF or AFL). \( f_2 \) is very close to 1 for SR and smaller for AF or AFL. Feature four (f4) measures the percent time interval corresponding to zero amplitude signal (percent quiet interval). It is calculated by the sum of time intervals with zero value over the total duration of rectified autocorrelation function. Feature five (f5) measures the number of components that reach baseline in 1 second (baseline reaching). Both \( f_2 \) and \( f_5 \) reflect the chaotic extent or randomness of the input signals and hence are supposed to be sensitive to fibrillatory rhythm (AF). The whole group of parameters (\( f_1, f_2, f_3, f_4, \) and \( f_5 \)) form a vector in 5 dimensions, which only can be determined if all the values of these 5 variables are known.

Training Process
We randomly selected 60% of the collected rhythm episodes as our initial training data set of this new algorithm. Then, the values of \( f_1, f_2, f_3, f_4, \) and \( f_5 \) and the corresponding feature vector for these 3 classes of rhythm signals were obtained. The distribution of each of 5 features has been found to follow approximately the normal distribution; therefore, the corresponding feature vectors of each class of rhythm also satisfy approximately a 5-dimensional normal distribution. Similar to one variate normal distribution, multivariate normal distributions are also determined by two parameters, the mean and the so-called covariant matrix; both could be estimated by the feature vectors of the training data set (also close data set). These two parameters are necessary for the discrimination procedure depicted in the next subsection (see Appendix B for details).

Discrimination Procedure
To reach the optimal detection performance, we used the multivariate Bayes decision theory, which has found wide use in the pattern recognition area (Appendix A). By using Bayes theorem, we can get the posterior probability, which is the chance that a feature vector of any episode should belong to any of the three classes of rhythm. Then, we have obtained a so-called “discrim-
In order to train our algorithm and discriminate between rhythms, we implemented three discrimination functions, \( g \), \( g_{AF} \), and \( g_{AFL} \), which correspond to the probabilities of the episode belonging to SR, AF, and AFL, respectively. The decision for each rhythm episode is simply determined by the absolute value that is the largest of the above three (Figure 2).

### Anti-Noise Performance

To test the anti-noise performance of our algorithm, we intentionally added Gaussian white noises with different signal-to-noise ratio (SNR) to each episode of the closed test data set.

### Statistical Analysis

Continuous variables are expressed as mean\(\pm\)SD. Statistical comparisons were performed by \(\chi^2\) analysis and Student’s t test as appropriate. To test the performance of the new algorithm, the sensitivity, specificity, and accuracy for detection of SR, AF, and AFL are calculated.\(^{14}\) Probability values <0.05 were considered statistically significant.

### Results

#### Patient Characteristics

This study consisted of 20 patients (17 men and 3 women; mean age, 55\(\pm\)16 years, \(\pm\)SD). Mean left ventricular ejection fraction was 56\%\(\pm\)10\% and mean left atrial diameter was 4.6\(\pm\)1.7 cm, as measured by echocardiography. Clinical characteristics are summarized in Table 1.

A total of 364 bipolar recordings were collected from the patients. All rhythm episodes were assessed blindly and classified into AF, AFL, or SR by 2 experienced electrophysiologists. Of these recordings, 156 episodes were AF, 88 episodes were AFL (mean atrial cycle length, 320\(\pm\)40 ms; range, 290 to 345 ms), and 120 episodes were SR, including 50 episodes of sinus tachycardia during isoprenaline infusion (mean sinus cycle length, 535\(\pm\)30 ms; range, 505 to 570 ms). Each patient contributed nearly the same number of episodes to the data set (18 to 22 episodes per patient). We randomly selected 219 (60\%) and 145 (40\%) rhythms as the closed-test data set and the open-test data set, respectively.

#### Signal Features

The results of 5 extracted features for the closed and open data sets are presented in Figure 3. The values of each of 5 features were significantly different between AF, AFL, and SR for both closed and open data sets. However, there were also significant overlaps between the values among the three types of rhythm for each feature. There were no significant differences in the values of each of 5 features for AF, AFL, and SR between closed and open data sets, which suggests that the two data sets were very similar.

#### Justification of Feature Selections

Figure 4 shows, respectively, the sensitivity, specificity, and accuracy of rhythm detection versus the increase of
features. When the number of feature(s) used increases from 1 to 5, the performance increases significantly ($P < 0.01$) from $\approx 80\%$ to $>95\%$. This result also indicates the advantage of multifeature detection over single-feature detection.

### Discrimination Performances

The performances of the new Bayesian discriminator algorithm for the closed and open data sets are summarized in Table 2. A total of 3 episodes (4%) of false-positive AF detection occurred in 2 patients during SR as the result of

---

**Figure 3.** Comparison of features for open and/or closed data sets. White, black, and shadowed bars represent SR, AFL, and AF, respectively. Each of 5 features is significantly different between AF, AFL, and SR for both open and closed data sets. There are no significant differences in the values of each of 5 features for AF, AFL, and SR between closed and open data sets.
the presence of far-field R-wave sensing. All 50 episodes of sinus tachycardia were correctly identified as SR. The sensitivity, specificity, and accuracy of rhythm detection for both closed and open data sets were similar. The overall sensitivity for detection of SR, AF, and AFL was 97%, 97%, and 94%, respectively; and the overall specificity for detection of SR, AF, and AFL was 98%, 98%, and 99%, respectively. The overall accuracy of detection of SR, AF, and AFL was 98%, 97%, and 98%, respectively (Table 2).

Anti-Noise Performances
The effects of increasing the SNR on the performances of the new Bayesian discriminator algorithm are presented in Figure 5. With a decrease in SNR, the sensitivity for detection of more regular rhythms such as SR and AFL decreases accordingly, whereas the sensitivity for AF detection is maintained at a high level. However, the specificity for AF detection decreases with the reduction of SNR, whereas the specificity for SR and AFL detection is maintained at a high level. As a result, the overall accuracy for detection of SR, AFL, and AF are similar at different SNR. When the SNR is >10 dB, this algorithm has an accuracy of ~95% in the detection of SR, AFL, and AF (Figure 5).

Discussion
Main Findings
The results of this study demonstrate that the features of intracardiac atrial electrograms, which include the regularity, rate, energy distribution, percent time of quiet interval, and numbers of baseline reaching, are significantly different during SR, AFL, and AF. However, detection algorithms that used only one or few of these features have only limited sensitivity, specificity, and accuracy for detection of SR, AFL, and AF. The new Bayesian discriminator algorithm, which is based on the Bayes decision rule

<table>
<thead>
<tr>
<th>Algorithm Rhythm Decision</th>
<th>Algorithm Performances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR</td>
</tr>
<tr>
<td>Close data set: actual rhythm</td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>70</td>
</tr>
<tr>
<td>AF</td>
<td>1</td>
</tr>
<tr>
<td>AFL</td>
<td>1</td>
</tr>
<tr>
<td>Open data set: actual rhythm</td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>47</td>
</tr>
<tr>
<td>AF</td>
<td>1</td>
</tr>
<tr>
<td>AFL</td>
<td>1</td>
</tr>
</tbody>
</table>
on all of these 5 features of atrial electrograms, allows rapid on-line and accurate (98%) detection of SR, AFL, and AF, with robust anti-noise performance.

**Previous Studies**

Previously proposed techniques for detecting AF fall roughly into 4 categories: (1) methods based on time-domain features\(^{10,11,15-20}\); (2) algorithms that use frequency-domain properties\(^{21-24}\); (3) techniques making use of time and frequency analysis\(^{25,26}\); and (4) miscellaneous.\(^{27}\) Botteron and Smith\(^{10,11}\) developed an algorithm based on the cross-correlation of two preprocessed bipolar intra-atrial signals, of which an active space constant was extracted. Tse et al\(^{15}\) depicted a 2-phase AF detection method that directly processed the time-domain signals. More recently, Sih et al\(^{16}\) proposed an approach that uses the mean square error in the linear prediction between two unipolar epicardial electrograms. Swerdlow et al\(^{17}\) used a technique that combined the median cycle length and an atrial tachyarrhythmias evidence counter that used the number of sensed atrial electrograms in consecutive R-R intervals. Chen et al\(^{19,20}\) proposed a modified sequential algorithm-based technique. Instead of measuring the rate, they used blanking variability to measure the temporal irregularity and thus improved the detection accuracy.

Except for the time-domain measures mentioned above, there exist some methods rooted in spectral analysis, such as the coherence spectrum method and frequency analysis with the use of the surface ECG.\(^{21,22}\) Recently, Chen et al\(^{23,24}\) proposed a 2-stage arrhythmia discrimination method that uses a damped exponential modeling algorithm, which gives higher frequency resolution than simple fast Fourier transform methods. Slocum et al\(^{25}\) designed an algorithm that took into account both the morphological information (atrial rate and amplitude probability function) and frequency-domain features (power spectrum analysis). Lovett and Ropella\(^{26}\) have analyzed atrial rhythms through the use of a time-frequency distribution of coherence. From the viewpoint of dynamic systems, Zhang et al\(^{27}\) proposed a complexity-based approach for the discrimination of ventricular tachycardias and fibrillation. Their method cannot simply be categorized into the time- and frequency-domain approaches but appears to have a few advantages over the conventional detection techniques.\(^{24}\) However, these methods need rather long episodes (>5 seconds) to get a satisfactory performance.

In this study, we provide a new and yet simple algorithm to treat intracardiac signals after a brief training period. This algorithm requires a very short computing time, as 250 ms is needed to make a decision for a rhythm episode of 1000 ms, on a computer of 500 MHz CPU, 64 M RAM. As shown in this study, the use of multiple feature discrimination provides a much higher sensitivity, specificity, and accuracy (all >94%) for rhythm detection than single or double features methods, as described above.

**Anti-Noise Performance**

Sometimes the intracardiac signals may be corrupted by noises introduced by external electromagnetic interference and myopotential sensing. It is important for the method to be robust when processing noisy episodes. As shown in this study, the SNR has significant effects on the performance of the detection algorithm. A decrease in SNR reduces the sensitivity for detection of regular rhythms such as SR and AFL. This phenomenon is due to the “noisy nature” of AF signals. The additive noises increase the randomness of all three classes of signals, which makes the algorithm liable to judge all episodes as AF and hence favors the AF class. As a result, the specificity for detection of AF also decreases as the SNR is reduced. This
new Bayesian discriminator algorithm has a satisfactory performance (>95%) for detection of SR, AFL, and AF when the SNR is ≥10 dB.

**Limitations**

The main limitation of this new detection algorithm is the presence of far-field R-wave interference; there is a tendency for misclassification of SR as AF, according to our algorithm. This problem also exists in other algorithms and cannot be resolved yet, although appropriate cross-chamber blanking and careful positioning of the atrial lead to avoid far-field R-waves may minimize this problem.

**Clinical Implications**

As device therapies for atrial tachyarrhythmias become more sophisticated in their ability to deliver several modes of therapy, such as antitachycardiac pacing and defibrillation, more accurate detection of therapy, such as antitachycardiac pacing and defibrillation, will be critical. Furthermore, accurate detection of SR from AFL and AF can also prevent inappropriate device therapy. The new Bayesian discriminator algorithm described in this study, which is based on multiple feature detection, can be easily implemented in the implantable device and provides rapid (~250 ms) and accurate (>97%) detection of AF, with robust anti-noise performance.

**Appendix A**

Bayes Decision Rule

Bayes decision theory assumes that the decision problem (whether an observed episode belongs to one class or another) is posed in probabilistic terms and that all of the relevant probabilities are known. $P(w_i)$ is denoted to be the prior probability that a certain episode belongs to one class or another) is posed according to Bayes rule. Assume that $p(\mathbf{v})$ is normal, we then

## Equation (1)

$$f(\mathbf{v}) = \frac{1}{(2\pi)^{d/2}} \sum_{i=1}^{n} \exp \left[ -\frac{1}{2} (\mathbf{v} - \mathbf{\mu}_i)^T \mathbf{\Sigma}_i^{-1} (\mathbf{v} - \mathbf{\mu}_i) \right].$$

where $\mathbf{\mu}_i$ is the vector of features, $\mathbf{\Sigma}_i$ is the covariance matrix generated by the vector $\mathbf{v}$, $\mathbf{\mu}$ denotes transpose, and $-1$ denotes inverse of a matrix.

## Equation (2)

$g_i(\mathbf{v}) = \log P(\mathbf{v} | w_i) + \log P(w_i)$

Substituting (1) into (2), we obtain a convenient form for the “discrimination function” $g_i(\mathbf{v})$:

## Equation (3)

$$g_i(\mathbf{v}) = \mathbf{w}_i \cdot \mathbf{v} + \frac{1}{2} \mathbf{v}^T \mathbf{\Sigma}_i^{-1} \mathbf{v} + \log P(w_i).$$

After calculating the three values of $g_i(\mathbf{v})$ ($i=1,2,3$), the i value corresponding to the maximum $g_i$ is chosen according to the Bayes decision rule.

**Appendix B**

Training Process

The objective of the training process is to estimate the prior probability $P(w_i)$ and the class distribution $p(\mathbf{v} | w_i)$. These two items are necessary to get the posterior probability $P(w_i | \mathbf{v})$, which is critical for the discrimination procedure. In practice, $P(w_i)$ can be approximated by $n_i / \sum_{i=1}^{n} n_i$, where $n_i$ is the total episode number of the $i$th class. $P(\mathbf{v} | w_i)$ can be calculated by $p(\mathbf{v} | w_i)P(w_i)$ according to Bayes rule. Assume that $p(\mathbf{v} | w_i)$ is normal, we then arrive at

$$p(\mathbf{v} | w_i) = \frac{1}{(2\pi)^{d/2}} \sum_{i=1}^{n} \exp \left[ -\frac{1}{2} (\mathbf{v} - \mathbf{\mu}_i)^T \mathbf{\Sigma}_i^{-1} (\mathbf{v} - \mathbf{\mu}_i) \right],$$

where $\mathbf{\mu}_i = \mathbf{E}[\mathbf{v} | w_i]$ is the mean of $\mathbf{v}_i$, and $\Sigma = \mathbf{E} [(\mathbf{v} - \mathbf{\mu})(\mathbf{v} - \mathbf{\mu})^T]$ is the covariance matrix generated by the vector $(\mathbf{v} - \mathbf{\mu})$; $t$ denotes transpose and $-1$ denotes inverse of a matrix.

**Appendix C**

Detection Process

Theoretically, the detection process is used to calculate the posterior probabilities $P(w_i | \mathbf{v}) = p(\mathbf{v} | w_i)P(w_i)$ of all 3 classes, given one unknown episode; however, because normal distribution has exponential terms, which are time consuming to calculate, for computation efficiency, taking the logarithm on both sides of the above equation, we have

$$g_i(\mathbf{v}) = \log P(\mathbf{v} | w_i) + \log P(w_i).$$

Substituting (1) into (2), we obtain a convenient form for the “discrimination function” $g_i(\mathbf{v})$:

## Equation (3)

$$g_i(\mathbf{v}) = \mathbf{w}_i \cdot \mathbf{v} + \frac{1}{2} \mathbf{v}^T \mathbf{\Sigma}_i^{-1} \mathbf{v} + \log P(w_i).$$

After calculating the three values of $g_i(\mathbf{v})$ ($i=1,2,3$), the i value corresponding to the maximum $g_i$ is chosen according to the Bayes decision rule.

**Acknowledgments**

The research of Dr Lau was supported by CRCG grant, University of Hong Kong. The research of Dr Fung was supported by the same CRCG grant, as well as the Central Allocation Fund, University of Hong Kong. W. Xu is a recipient of a Hong Kong University postgraduate scholarship.

**References**


New Bayesian Discriminator for Detection of Atrial Tachyarrhythmias
Weichao Xu, Hung-Fat Tse, Francis H.Y. Chan, Peter Chin Wan Fung, Kathy Lai-Fun Lee and Chu-Pak Lau

_Circulation._ 2002;105:1472-1479; originally published online March 11, 2002;
doi: 10.1161/01.CIR.000012349.14270.54
_Circulation_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2002 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://circ.ahajournals.org/content/105/12/1472

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation_ is online at:
http://circ.ahajournals.org/subscriptions/