# Three Different Algorithms for Identifying Patients Suffering from Atrial Fibrillation during Atrial Fibrillation Free Phases of the ECG

N Kikillus, G Hammer, N Lentz, F Stockwald, A Bolz

University of Karlsruhe, Karlsruhe, Gemany

## Abstract

*The paper presents and compares three different methods to detect atrial fibrillation using algorithms. Each of these algorithms can identify patients suffering from atrial fibrillation even if there is no atrial fibrillation visible on the ECG. All methods are based on RR-intervals; thus, only a single channel ECG is required.*

*These algorithms have been tested using the MIT-BIH atrial fibrillation database and the MIT-BIH normal sinus rhythm database. The sensitivity and specificity of method 1 is 91.5% and 96.9% respectively. Method 2 results in a sensitivity of 93.3% and a specificity of 92.8% and method 3 in a sensitivity of 94.1% and a specificity of 93.4%. Even if an atrial fibrillation burden of 0% is assumed, the sensitivity still proves satisfactory (sensitivity of method 1, 2 and 3 is 82.9%, 96.3% and 94.1%, respectively).*

## 1. Introduction

Atrial fibrillation (AF) is quickly becoming a widespread disease. About one percent of the entire population is affected. The most severe complication is an embolism and a stroke in particular. Patients suffering from AF have to be identified in order to prevent ischemic strokes by treating them early.

Unfortunately, AF remains underdiagnosed because of two reasons: Firstly, it is often asymptomatic: up to 30% of all patients with AF are unaware of their diagnosis [1]. Thus, they are highrisk patients but are unaware of the disease. Secondly, AF often occurs intermittently and therefore may not be detected on a standard 12-lead ECG or even on a 24-hour ECG recording (Holter).

## 2. Methods and Material

## 2.1. Algorithms

## 2.1.1. General items

All algorithms presented are based on the parameter of the heart rate variability (HRV); thus only a single channel ECG is required.

## 2.1.2. Signal preprocessing

A Pan-Tompkins QRS-detector [2] assigns all R-peaks in the ECG. Any ectopic beats are removed, as recommended by the Task Force [3], since they have a negative impact on the analysis of HRV parameters. Afterwards, all time differences between two consecutive R peaks, which are called RR intervals, are calculated.

## 2.1.3. Algorithm 1

**Normalization**

Algorithm 1 requires 3500 RR intervals that are normalized to 60 bpm by the following equation:

$$RRInterval_{normalized} = RRInterval_{latest} * \frac{heartrate_{15\,sec}}{60\frac{1}{min}}$$

The latest RR interval is multiplied by the mean heart rate of the last 15 seconds and divided by 60 bpm. This normalization is necessary in order to become independent from pulse variations (due to emotions and movements) and to be able to compare the ECGs of different patients. Figure 1 shows the heart rate over 60 minutes without normalization on the left side and the same segment after normalization on the right side.
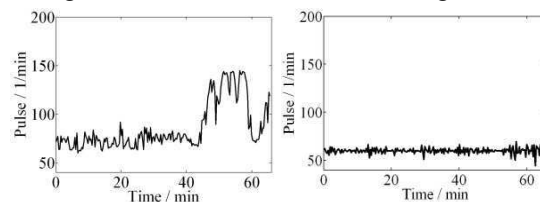


Figure 1. Chronological sequence of the heart rate: heart rate without (left), and with normalization (right).

**Time domain analysis**

The time domain parameter SDSD (standard deviation of the temporal differences of consecutive RR intervals) is determined from all 3500 RR intervals using the following equation:

$$SDSD = SD\Big|_{1}^{N-1}(X_N - X_{N+1})$$

SD is the standard deviation, X represents the difference between two consecutive RR intervals and N is the number of RR intervals. SDSD corresponds to a risk level, which depends on the calculated value. There are up to 4 risk levels.

**Image analysis**

The Poincaré Plot, a non-linear method, is created by plotting all 3500 RR intervals in a two dimensional system in the following way: Two consecutive RR intervals represent one point in the plot. The first RR interval ($RR_N$) represents the x-coordinate, the second interval ($RR_{N+1}$) the y-coordinate. Figure 2 shows a Poincaré Plot of a healthy patient on the left side, rotated with respect to the x-axis and with its center shifted to zero. Its allocation density is displayed on the right side.



Figure 2. Poincaré plot of a patient with a normal sinus rhythm (left side) and the corresponding density allocation (right side).

Poincaré Plots from patients with AF look very different (Figure 3). The contours vary: triangles, circles in various sizes, randomly distributed but often concentrated points and any merged contours.
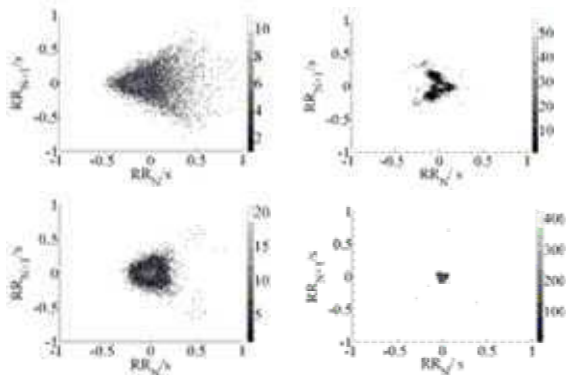


Figure 3. Poincaré plots with density allocation from patients with AF. The contours vary.

The figure shown in Figure 4 is used to analyze these Poincaré Plots. This figure is automatically adapted to the input signal according to the density allocation. The automatic image analysis determines the number of points within the following patterns:
1. Inside the eye
2. Inside the circle
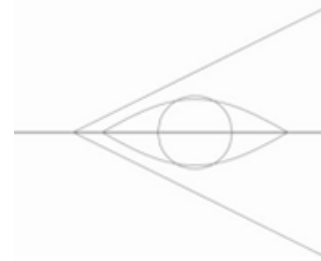3. Inside the eye and outside the circle, but only on the



Figure 4. Figure of the image analysis.

   left side of the eye
4. Inside the triangle, but outside the eye
5. Outside all of the figures

The number of points inside each of these patterns corresponds to a risk level. All risk levels of these five patterns are then summed up. Additionally, a parameter $V_{shape}$, which determines whether the scatter plot is more similar to a circle or to an eye, is calculated using the following equation:

$$V_{shape} = Inside_{eye} - Inside_{circle} - Inside_{triangle}$$

$Inside_{eye}$, $Inside_{circle}$ and $Inside_{triangle}$ are the numbers of calculated points within the various patterns. $V_{shape}$ is also associated with a risk level which is added to the risk level from the image analysis.

**Risk level**

The final risk level is calculated from the risk levels of both the image analysis and the time domain. The risk levels are added using weighted calculation. There are up to 20 risk levels. A patient whose risk level is greater than two is identified as a patient with AF.

## 2.1.4   Algorithm 2

**FFT**

Algorithm 2 consists of three different tests that compare different frequencies. They require 45 minutes of RR intervals. In order to prevent leakage a window is applied to the RR tachogram built from these RR intervals. The windowed signal is spline-interpolated and resampled with 128 Hz and afterwards transformed into the frequency domain (FFT). The following three tests consider only the absolute value of the Fourier coefficients between 0 and 64 Hz.

**Test 1**

Test 1 calculates the ratio of two integrals of adjacent frequency bands. The ratio is compared with a threshold that has been predetermined empirically. If the value of the ratio is greater than the threshold, the decision is "AF", otherwise it is "No AF". This is shown in Figure 5. A total of seven different ratios of integrals of two

adjacent frequency bands are calculated. The frequency bands are always within a range from 0,05 Hz to 0.25 Hz. Each of the seven crosschecks with an associated threshold yields one decision: "AF" or "No AF". The final decision results from all seven single decisions by a majority decision (e.g. 4 x AF and 3 x No AF result in AF).
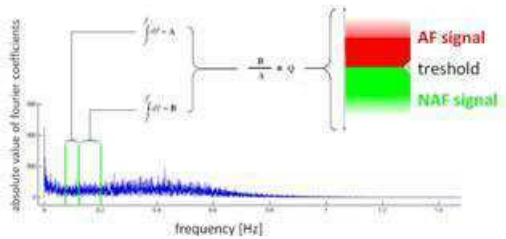


Figure 5. Test 1 calculates the ratio of two adjacent integrals.

## Test 2

The integral from 0 Hz to 64 Hz is computed. After this, the computed area is split in two equal halves to determine the center point. This center point is compared with a threshold that has been predetermined empirically. If the center point value is equal to or greater than 0.22 Hz, the test opts for "AF", otherwise for "No AF". Figure 6 shows a general schema of test 2.
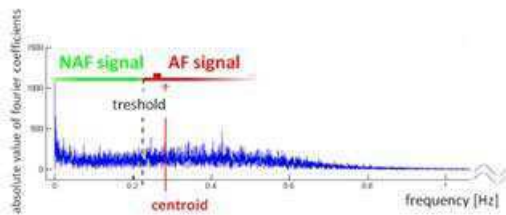


Figure 6. Test 2 determines the center point of the integral and compares it to a threshold.

## Test 3

Test 3 consists of 5 single tests that work in the same way. First, the sector ranging from 0 Hz to the center point calculated in test 2 is divided into x bands. X is different for each of the 5 tests. The integrals of two non-adjacent bands are computed, the ratio is calculated and afterwards compared to a threshold that has been predetermined empirically. If the value is greater than the
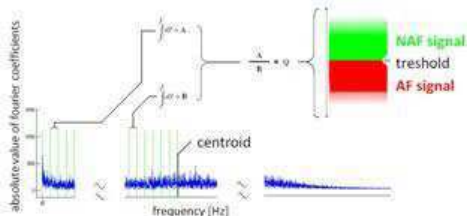


Figure 7. Test 3 compares two integrals from non-adjacent bands

threshold, the test opts for "AF", otherwise for "No AF". As in test 1 the final decision for or against AF results from a majority decision based on the 5 single decisions. Figure 7 shows the schematic illustration.

**The Decision Maker**

For the final result of algorithm 2 the decision maker shown in Figure 8 is used. If test 1 opts for a "No AF"-Signal, tests 2 and 3 will not be taken into account. If test 1 opts for an "AF"-signal but test 2 and 3 for a "No AF"-signal, the result will be a "No AF"-signal. Otherwise it will be an "AF"-signal.
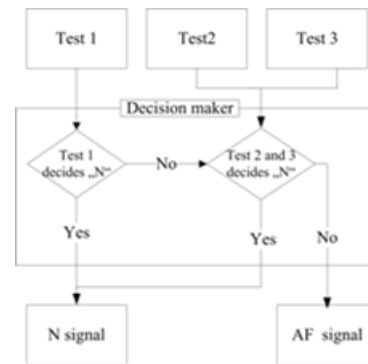


Figure 8. Decision Maker of algorithm 2

### 2.1.5 Algorithm 3

**Differences of RR intervals**

Algorithm 3 consists of 2 tests and requires 3500 differences of adjacent RR intervals (3500 dRR). The second RR interval is always subtracted from the first.

**Test 1: Histogram and basis difference**

The 3500 dRR are subdivided into 7 blocks of 500 dRR each. For each block a histogram with 50 classes (with ranges between 0s and 1s) is generated and a so called basis difference (bd) is calculated. Bd is the upper limit of the least class whose absolute frequency is less than 10 dRR but maximal 0.2 s. The number of dRR within 4 different ranges (2xbd and 3xbd, 3xbd and 4xbd, 4xbd and 5xbd, 5xbd and 1s) is counted. The histogram, the bd and its multiples can be seen in Figure 9.
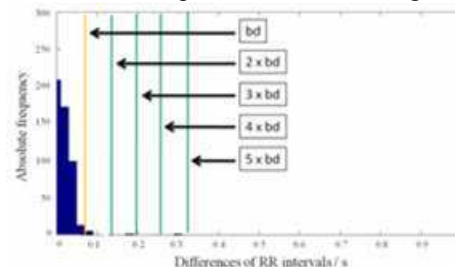


Figure 9. The histogram of dRR forms a normal sinus rhythm ECG histogram.

Depending on the number of dRR within the ranges, points (0, 1 or 3) are assigned. Afterwards the points of the 4 ranges are summed up. If the value is greater than 1, each of the 500 dRR blocks will be designated 1. If it is smaller than or equal to 1, it will be designated 0. This procedure is repeated for each block (thus 7 times). After all single results have been calculated, they are summed up again. If the sum is greater than 2, the value will be set to 2 (therefore results between 0 and 2 are possible).

**Test 2: pNN200**

pNN200 is a modification of the standard parameter pNN50. It describes the percentage of RR intervals that differ at least 200 ms from their antecedent RR interval. pNN200 is calculated from all 3500 dRR. Depending on the percentage computed, points are assigned according to Table 1

Table 1: Assignment of points depending on pNN200

| pNN200 | points |
|--------|--------|
| < 5 % | 0 |
| 5 % < …. < 7 % | 1 |
| > 7 % | 2 |

**Analysis**

The designated points of test 1 and 2 are summed up. If the sum is greater than 1, the algorithm opts for a patient suffering from AF (Figure 10).
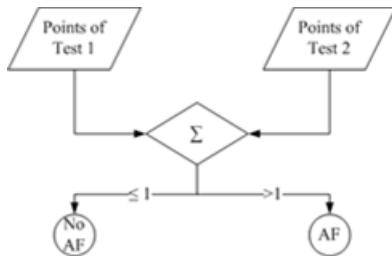


Figure 10. Analysis of the two test of algorithm 3.

## 2.2    ECG Records

All three algorithms were developed and tested with 2 ECG databases from Physiobank [4]: The MIT-BIH Atrial Fibrillation Database (MIT-BIH AF DB) and the MIT-BIH Normal Sinus Rhythm Database (MIT-BIH NSR DB). To be able to test the algorithm in AF free phases of an ECG we also generated some AF free data and added it to the tested ECG files. Hence, the class AF0% is generated only during ECG phases in which AF is not visible on the ECG (AF burden = 0%). The so-called annotations from PhysioBank are used to get the AF free phases of the ECG records and to calculate sensitivity and specificity later. These annotations have

been added by cardiologists who mark characteristic points in ECGs, like R-peaks, the start of a normal sinus rhythm, the start of atrial fibrillation, etc..

## 3.    Results

The sensitivity and specificity of the three algorithms are shown in Table 2. Table 3 shows the sensitivity of the class AF0%.

Table 2. Sensitivity and specificity of the 3 algorithms

|  | sensitivity | specificity |
|--------|--------|--------|
| Algorithm 1 | 91,5 % | 96,9 % |
| Algorithm 2 | 93,3 % | 92,8 % |
| Algorithm 3 | 94,1 % | 93,4 % |

Table 3. Sensitivity of the 3 algorithms for AF0%

|  | sensitivity (AF0%) |
|--------|--------|
| Algorithm 1 | 82,9 % |
| Algorithm 2 | 96,3 % |
| Algorithm 3 | 94,1 % |

## 4.    Discussion and conclusions

It seems to be possible to identify patients with atrial fibrillation even if no atrial fibrillation is visible on the ECG. We developed three different algorithms that yield good results. In future, we will combine all three algorithms to improve the sensitivity and specificity.

It is important to test the algorithms with ECGs of patients suffering from other heart diseases as well since the heart rate variability can also be influenced by many heart diseases. Additionally the algorithms will be tested with more annotated ECG-signals.

## References

[1] Furberg CD, Bsaty BM, Manolio TA, Gardin JM, Smith VE, Rautaharju PM. Prevalence of atrial fibrillation in elderly subjects (the Cardiovascular) Health Study. Am J Cardiol 1994; 74: 236–241.

[2] Pan J, Tompkins W. A Real-Time QRS Detection Algorithm. IEEE Trans Biom Eng 1985; 32(3): 230-236.

[3] Malik M. Heart Rate Variability Standards of Measurement, Physiological Interpretation, and Clinical Use. Circulation 1996; 93: 1043-1065.

[4] Physionet [homepage on the internet]. Camebridge, MA: PhysioNet; [updated 2007 Sept. 01; cited 2007 March 27]. Available from: http://www.physionet.org/

Address for correspondence:

Nicole Kikillus
Hertzstr. 16, 76187 Karlsruhe, Germany
Kikillus@ibt.uka.de