

An Automatic ECG Processing Algorithm to Identify Patients Prone to Paroxysmal Atrial Fibrillation

G Schreier, P Kastner, W Marko

Austrian Research Centers Seibersdorf, Graz, Austria

Abstract

An algorithm to identify patients prone to paroxysmal atrial fibrillation (PAF) has been developed and evaluated using the PAF Prediction Challenge Database. The procedure is based on conventional electrocardiogram (ECG) signal pre-processing techniques for beat detection and classification, a correlation based assessment of the P-wave morphology of both regular and premature heartbeats of supraventricular origin, and a statistical test to calculate the PAF predictive parameter, i.e. the probability that a certain degree of P-wave variability is associated with potential triggers for PAF.

This probability, finally, is used to differentiate between patients with and without PAF (screening) and to find out which of the two recordings of each patient immediately precedes the onset of PAF (prediction), respectively. The obtained diagnostic accuracies of 82% and 84%, respectively, indicate that this concept may be useful in terms of clinical PAF risk stratification.

1. Introduction

The ECG reflects the electrical activity of the heart and is an excellent means to non-invasively assess the cardiac state. A number of cardiac diseases are related with short- (beat-to-beat) or long-term changes of the way and rhythm of cardiac excitation and hence are associated with characteristic changes in the ECG.

Atrial fibrillation (AF) is the most common cardiac arrhythmia, affecting about two million people in the US [1]. AF is characterized by uncoordinated electrical and mechanical activity of the atria and subsequent loss of their pumping function. In general, AF does not present a life threatening condition, can occur without symptoms or, in many cases, may even stay unnoticed for quite some time but, in any case, is associated with a high risk of cardiovascular morbidity and mortality [2].

Chronic AF is often preceded by PAF. Besides using antiarrhythmic drugs, considerable effort is presently underway to develop adequate therapies to suppress the onset of AF based on pacing devices [3]. To administer such therapies efficiently, methods are required to:

- a) assess the basic risk of a certain patient for developing PAF, and
- b) to detect situations which are likely to be followed by a PAF episode, i.e. to detect PAF triggers.

Numerous concepts for AF onset risk assessment based on ECG analysis have already been investigated. A reliable and clinically validated method, however, has not yet been established [4].

Previous studies have indicated that P-wave characteristics (morphology, duration, dispersion, etc.) are related to the onset of AF episodes [5]. Premature atrial complexes, particularly if being associated with an aberrant P-wave morphology in the ECG, have also been identified as possible PAF triggers. Reliable analysis of beat-to-beat P-wave changes, however, is difficult since it requires to differentiate between intrinsic variability, i.e. changes caused by physiological processes like respiration and beat-to-beat changes beyond such effects, i.e. disease specific irregularities. This process can be further complicated by the presence of noise and other artifacts within the signal.

The present paper describes an automatic ECG processing method for robust assessment of single-beat P-wave changes in order to investigate the association of such ECG changes to PAF.

2. Methods

2.1. Data management

The development and evaluation of the algorithms were based on the PhysioNet PAF Prediction Challenge Database, that has been provided specifically to be used for the Computers in Cardiology Challenge 2001 [6,7].

The learning and test set ECGs were downloaded from the Internet to a local computer. The signals were stored within a database which was specifically designed to provide a basis for systematically developing and improving the algorithms. The database allowed to track every change associated with algorithm modifications, to store the extracted parameter values finally to be used for classification, and to provide the manually obtained "expert" classification. An additional purpose of the database was to manage high performance serial ECG processing using multiple computers simultaneously.

2.2. Expert classification

After having consulted a cardiologist to improve the authors knowledge on long-term ECG interpretation, the downloaded ECG recordings were repeatedly visually inspected in order to detect effects and configurations possibly being associated to PAF. This led to an “expert” classification for each of the test set ECGs, either as normal recordings or as recording immediately preceding the onset of PAF, based on the authors experience. Every new classification produced by the automatic algorithm was checked against this expert classification in order to assess the plausibility of the results and the value of the underlying algorithm improvements.

2.3. ECG preprocessing

Prior to the computation of any specific PAF related parameters, conventional ECG preprocessing was performed. The information provided by the annotation files was not used since this classification does not reflect different QRS morphologies (this, however, is essential in discriminating premature beats of ventricular and supraventricular origin). Basic ECG processing comprised of the following major steps:

1. Heartbeat detection and classification was applied to consecutive 60 seconds ECG sections and consisted of detecting all events with super-threshold first derivative values and cross-correlating the detected beats with a set of templates of variable size (low correlation coefficients triggered the generation of additional beat classes). Those beats with the dominant signal morphology were considered to be the ‘regular’ heartbeats.

2. The first derivative of the sequence of the beat coupling intervals was checked for the occurrence of beats with at least a 20% prematurity and the same morphology as the preceding beat. Such beats were assumed to be premature heartbeats of supraventricular origin, following premature atrial contractions.

3. A representative P-wave template was generated by coherent averaging of the respective signal part of all regular heartbeats. Subsequently, all regular and premature heartbeats were compared to this P-wave template using linear regression analysis (Figure 1).

4. The two groups of correlation coefficients corresponding to the regular and the premature heart beats were considered to represent two samples to which the non-parametric U-test was applied in order to assess the probability (p value) that both samples stem from the same distribution. The logarithm of the p value, served as the PAF screening parameter.

2.4. Diagnosis models

1. The decision on whether a particular signal is indicative of PAF (Event 1) was based on a single threshold diagnosis model. A risk for PAF was assumed for a particular patient if at least one of both signals exhibited a p value below the cut-off threshold. The cut-off threshold was determined by computing the receiver operating characteristic of this diagnosis model for the learning record set and searching for the threshold that gave the highest overall number of correctly classified patients (Figure 2).

2. The decision on which of the two recordings is the one immediately preceding PAF onset (Event 2), was

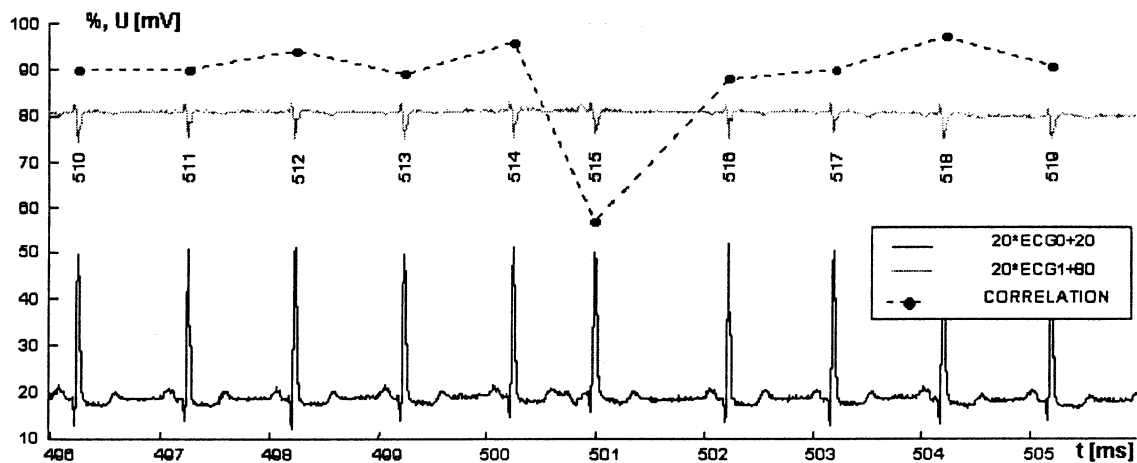


Figure 1. A 10 second section from ECG t85.dat showing both signal channels and the correlation coefficients (dots) as obtained for each heartbeat. Beat number 515 represents an “A” beat, i.e. a premature atrial complex followed by a normal QRS complex (gray shaded area). The P-Wave of this heartbeat shows an inverse signal morphology as compared to the P-waves of the neighboring heartbeats and exhibits a much lower correlation coefficient of about 55%. This indicates that this beat may be a potential trigger for the onset of PAF. The legend displays the ECG channels scaling factors that have been used for displaying purposes.

based on the PAF “trigger moment” PAF_{im} , which was calculated as the sum of 100% minus the actual correlation coefficients cc_i for all n premature heartbeats and weighted according to the following equation:

$$PAF_{im} = \sum_{i=1}^n (100\% - cc_i) \frac{1}{1 - \frac{t_i - t_0}{t_E - t_0}}$$

The PAF_{im} increases with the number of premature beats n , with aberrant P-wave morphology, i.e. lower cc_i and when the respective premature beats occur at time t_i more towards the end t_E of the sequence than at the beginning t_0 .

The recordings with the higher PAF_{im} values were assumed to be the ones followed by the onset of PAF.

3. Results

Overall, 10 prospective classification attempts with subsequent algorithm improvements were performed. Figure 2 displays the application of the diagnosis model to the learning data set, resulting in an overall diagnostic accuracy of 84% (42 out of 50 patients correctly classified, sensitivity = 84%, specificity = 84%) at the optimum cut-off threshold of -1.85 . Using this threshold, finally, the algorithm was able to correctly classify 41 out of 50 patients from the test data set.

The comparison of the PAF_{im} values as obtained from the learning data set indicated that a higher value in the recordings preceding the onset of PAF was present in 19 of 25 cases of PAF positive patients. By choosing the test set recordings with the higher PAF_{im} values in each particular patient, the recordings preceding the onset of PAF were correctly identified in 42 out of 50 cases (including the cases of PAF negative patients which are always considered correctly classified).

These figures correspond to a diagnostic accuracy for PAF screening and PAF prediction of 82% and 84%, respectively.

4. Conclusions

The obtained numbers of correctly classified cases indicate the potential of the proposed approach and confirms that premature and aberrant P-waves may herald or even trigger PAF. If further studies confirm the feasibility of the present concept in a larger patient population, the method may present a first step towards the strongly desired tool for a reliable PAF risk assessment and, eventually, be used to guide PAF suppression therapies.

Acknowledgements

We gratefully acknowledge the cardiological advise of Dr. Brigitte Rotman from the University Hospital in Graz, Department of Internal Medicine/Cardiology.

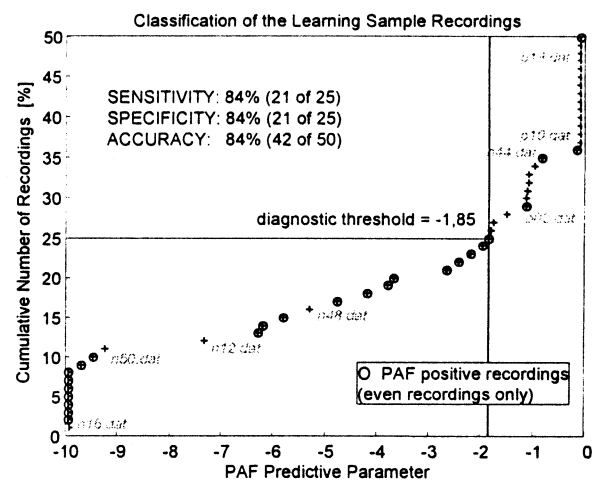


Figure 2. Cumulative distribution function of the PAF Predictive Parameter as obtained from the 50 even numbered recordings of the learning data set. Each patient is represented by a “+”, the encircled cases correspond to the p-recordings, i.e. to patients with PAF. Using the diagnostic threshold as shown, a total of 42 out of the 50 cases were classified correctly, and the depicted diagnostic indices were obtained. This particular threshold has been chosen because it gives the highest overall diagnostic accuracy in terms of the overall number of correct classifications. Incorrectly classified cases are marked by the respective recording name.

References

- [1] American Heart Association. 2001 Heart and Stroke Statistical Update. Dallas, Texas, American Heart Association 2000.
- [2] Savelieva I, Camm AJ. Silent atrial fibrillation - another Pandora's box. PACE 2000;23:145-8.
- [3] Hesselson AB, Gallagher PL, et al. Pacing suppression of spontaneous atrial ectopy in patients with focal atrial fibrillation. PACE 2001;24:554.
- [4] <http://www.physionet.org/physiobank/database/afpdb>
- [5] Ciaroni S, Cuenoud L, Bloch A. Clinical study to investigate the predictive parameters for the onset of atrial fibrillation in patients with essential hypertension. Am Heart J 2000;139:814-9.
- [6] Moody GB; Goldberger AL, McClennen S, Swiryn SP. Predicting the onset of paroxysmal atrial fibrillation: The Computers in Cardiology Challenge 2001. This issue.
- [7] Moody GB, Mark RG, Goldberger AL. PhysioNet: A Web-based resource for the study of physiologic signals. IEEE Engineering in Medicine and Biology 2001;20:70-5.

Address for correspondence:

Dr. Günter Schreier
Austrian Research Centers Seibersdorf – Graz Office
Grottenhofstraße 3
A-8053 Graz
Austria
guenter.schreier@arcs.ac.at