Computer Methods and Programs in Biomedicine 151 (2017) 111–121

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

# Beat-to-beat P-wave morphology as a predictor of paroxysmal atrial fibrillation

Dimitrios Filos[a,*], Ioanna Chouvarda[a], Dimitris Tachmatzidis[b], Vassilios Vassilikos[b], Nicos Maglaveras[a]

[a] Laboratory of Computing and Medical Informatics, Aristotle University of Thessaloniki, Box 323, 54124, Thessaloniki, Greece
[b] 3rd Cardiology Department, Aristotle University of Thessaloniki, Greece

ARTICLE INFO

*Article history:*
Received 14 October 2016
Revised 11 August 2017
Accepted 21 August 2017

*Keywords:*
Atrial fibrillation
P-wave morphology
SVM
Wavelet Transform

ABSTRACT

*Background and objectives:* Atrial Fibrillation (AF) is the most common cardiac arrhythmia. The initiation and the perpetuation of AF is linked with phenomena of atrial remodeling, referring to the modification of the electrical and structural characteristics of the atrium. P-wave morphology analysis can reveal information regarding the propagation of the electrical activity on the atrial substrate. The purpose of this study is to investigate patterns on the P-wave morphology that may occur in patients with Paroxysmal AF (PAF) and which can be the basis for distinguishing between PAF and healthy subjects.
*Methods:* Vectorcardiographic signals in the three orthogonal axes (X, Y and Z), of 3–5 min duration, were analyzed during SR. In total 29 PAF patients and 34 healthy volunteers were included in the analysis. These data were divided into two distinct datasets, one for the training and one for the testing of the proposed approach. The method is based on the identification of the dominant and the secondary P-wave morphology by combining adaptive k-means clustering of morphologies and a beat-to-beat cross correlation technique. The P-waves of the dominant morphology were further analyzed using wavelet transform whereas time domain characteristics were also extracted. Following a feature selection step, a SVM classifier was trained, for the discrimination of the PAF patients from the healthy subjects, while its accuracy was tested using the independent testing dataset.
*Results:* In the cohort study, in both groups, the majority of the P-waves matched a main and a secondary morphology, while other morphologies were also present. The percentage of P-waves which simultaneously matched the main morphology in all three leads was lower in PAF patients ($90.4 \pm 7.8\%$) than in healthy subjects ($95.5 \pm 3.4\%$, $p = 0.019$). Three optimal scale bands were found and wavelet parameters were extracted which presented statistically significant differences between the two groups. Classification between the two groups was based on a feature selection process which highlighted 7 features, while an SVM classifier resulted a balanced accuracy equal to 93.75%. The results show the virtue of beat-to-beat analysis for PAF prediction.
*Conclusion:* The difference in the percentage of the main P-wave-morphology and in the P-wave time-frequency characteristics suggests a higher electrical instability of the atrial substrate in patients with PAF and different conduction patterns in the atria.



## 1. Introduction

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia, occurring in 1–2% of the general population. Over 6 million Europeans suffer from this arrhythmia, and its prevalence is estimated to at least double in the next 50 years as the population ages [1]. While AF itself is not a life-threatening condition, its incidence leads to increased risk of stroke [2], and high rates of mortality, while the economic burden of AF is enormous [3]. "Paroxysmal" AF (PAF) is a type of AF [1] which may last up to 48 hours and is usually self-terminated. This may later evolve to a more persistent condition.

AF is characterized by an irregular and chaotic electrical activation of the atria resulting in the modification of the ventricular activity and increased risk of clot formation [4,5]. AF causes electrical and structural remodeling. Electrical remodeling favors the occurrence of focal triggering mechanisms such as multiple wavelets

* Corresponding author.
*E-mail addresses:* dimfilos@auth.gr (D. Filos), ioanna@med.auth.gr (I. Chouvarda), tachmatzidis@gmail.com (D. Tachmatzidis), vvassil@med.auth.gr (V. Vassilikos), nicmag@med.auth.gr (N. Maglaveras).

http://dx.doi.org/10.1016/j.cmpb.2017.08.016
0169-2607/

re-entry, increasing the susceptibility to AF [6]. Nevertheless, it is considered reversible in contrast to the structural remodeling which may be the result of prolonged AF duration or presence of structural heart disease. Structural remodeling refers to the conduction heterogeneities caused by modifications of the connectivity tissue and fibrosis [1]. There are usually no distinct morphological deformities in the atria, detectable by the usual clinical tools in patients with PAF [4].

Given the absence of remarkable structural remodeling in patients with PAF, the detection of possible changes in the surface electrocardiogram (ECG) is challenging. The atrial depolarization is represented on the surface ECG as the P-wave. The first part of this wave corresponds mainly to right atrial depolarization, while the second part to the left atrial depolarization [7–9]. P-wave closely correlates with the conduction in specific parts of the atria [10,11] and thus any deflection of the P-wave characteristics implies regional changes in atrial activation time and conduction route.

The main factors which influence the P-wave morphology include: (1) the origin of the sinus rhythm (SR), (2) the left atrial breakthrough and (3) both the shape and the size of the atrial chambers which affect the time required for the depolarization process completion [12], while any structural abnormalities, such as fibrosis and fatty infiltration, affect the atrial conduction routes towards the conduction slowing and variable propagation of the electrical activation, influencing the P-wave morphology [13,14].

The standard approaches for time domain P-wave analysis focus on the signal-averaged P-wave. Its prolongation has been associated with a history of AF [15,16] as well as a marker of the atrial remodeling extend, outside the pulmonary veins, and it has been used as marker of PVI success [17]. On the contrary, over the recent years there seems to be robust evidence that short P-waves can also be associated with the incidence of AF [18]. P-wave dispersion, the difference between the longest and the shortest P-wave duration in multi lead ECG [19], the duration of the high frequency P-wave and the root mean square amplitude of the last 20msec of the filtered P-wave (RMS2040) [20], P-wave variance and PR interval [21] were also used for the prediction of AF prone patients, denoting statistical significant differences with healthy subjects. The main weakness of these methods involves the extraction of a signal-averaged P-wave, considered to represent all the P-waves appearing in the ECG. These approaches prerequisite the accurate identification of the P-wave onset and offset. Several methods were proposed for the automatic delineation of the ECG fiducial points [22], however, a possible shifting of the ECG fiducial points may affect negatively the calculation of the average P-wave and thus the reliability of the results [23,24].

The idea behind beat-to-beat (b2b) analysis is that mild to moderate remodeling may not constantly change the P-wave morphology. Increased b2b variability of time domain and other morphological parameters, such as P-wave velocity, arc and area, were associated with the approximation of PAF episodes [25,26], suggesting a distributed conduction in the atrial substrate. Other studies combine the b2b analysis of the P-wave with correlation techniques in order to explore the P-wave morphology. In [27] *k*-means clustering algorithm was combined with time and morphological parameters and each beat was classified under the strong assumption that in each signal two P-wave morphology clusters exist, while in [13] each beat was classified, according to its morphology, into one of five pre-defined classes and the percentage of beats not matching the most common P-wave morphology was found to be a predictor of circumferential pulmonary vein isolation success. In [28] the P-wave morphology of the average P-wave is modeled by fitting a combination of Gaussian functions. The results denote that the morphological parameters extracted can effectively identify patients of different risks for AF development. Nevertheless, none of the studies considered the existence of a distinct secondary P-wave morphology in a sinus node originated ECG signal. In the current study, a machine learning technique was applied for the adaptive identification of possible clusters of P-wave morphologies from sinus node originated beats and it was supervised by the clinical experts. Time domain characteristics alone may not reveal the atrial conduction disturbances and the development of PAF in patients without structural heart disease [29] as well as any detected difference may be attributed to the anatomy of the atrium and not to the pathology itself [30].

The P-waves preceding the onset of the PAF episodes contain high frequency information compared to the normal beats [31] which is associated to the intracardiac high frequency sources in several regions in the atrial myocardium [32]. Spectral turbulence was used for assessing the frequency content of the averaged P-wave and predicting the recurrence of AF [33,34]. However, the transition from SR to AF is associated with the gradual progression to AF and with the aforementioned methods the spectral variability of the P-waves over time is masked. The P-wave frequency content variability in PAF patients was studied by the analysis of the energy in the bands between 80 and 150 Hz [31]. However, when the frequency derived information was combined with the time-domain features, was the classification increased.

Wavelet transform (WT) is a widely used method for time-frequency transformations and it was adopted by several studies in order to capture the time specific frequency content appearing in the biomedical signals such as ECG [35,36]. In [37] the continuous WT was applied on each P-waves of an ECG signal as well as on the average one in order to discriminate healthy signals from those with atrial conduction pathologies. The wavelet energy, on 30 randomly selected, consecutive P-waves demonstrated to have predictive value towards the recurrence of PAF episodes within a year, for specific frequency zones, and could be used for the early identification of such patients [38]. The b2b analysis of the P-waves is more complete and unbiased than analyzing a randomly selected beat or an average one, which smooths the local characteristics of the P-waves resulting in the loss of valuable information [39,40].

The present work identified a gap in the analysis of the P-wave characteristics for the discrimination of the signals from patients prone to PAF during SR and it proposes an end-to-end P-wave analysis in all domains in a b2b basis. In this respect, a features set is constructed including the P-wave wavelet characteristics along with morphological characteristics and the standard time domain parameters. The most informative features among them are employed in a classification procedure, to separate normal from PAF. The article is structured as follows: the introduction section provides a description of the problem of AF and the methods used for the P-wave analysis, the methodology section describes the methods used in the current study while the results are presented in the related section. The discussion section summarizes the main finding of the study.

## 2. Methodology

### 2.1. Data description and pre-processing

The current study emphasized on patients characterized by paroxysmal atrial fibrillation as their main clinical complaint, creating a homogeneous patient group comparable to healthy volunteers as much as possible. In this respect many patients with comorbidities, such as previous cardiovascular surgery, previous cardiac ablation, heart failure NYHA class III-IV, severe valvular heart disease, prosthetic valves, reduced life expectancy, age $> 75$ years, atrioventricular block, presence of implanted pacemaker or cardiac defibrillator, moderate/severe renal or hepatic impairment,