

# Quantification of not-dipolar components of atrial depolarization by principal component analysis of the P-wave

Federica Censi<sup>(a)</sup>, Giovanni Calcagnini<sup>(a)</sup>, Pietro Bartolini<sup>(a)</sup>,  
Renato Pietro Ricci<sup>(b)</sup> and Massimo Santini<sup>(b)</sup>

<sup>(a)</sup>Dipartimento di Tecnologie e Salute, Istituto Superiore di Sanità, Rome, Italy

<sup>(b)</sup>Dipartimento di Cardiologia, Ospedale San Filippo Neri, Rome, Italy

**Abstract.** *Background.* Principal component analysis (PCA) of the T-wave has been demonstrated to quantify the dipolar and not-dipolar components of the ventricular activation, the latter reflecting repolarization heterogeneity. Accordingly, the PCA of the P-wave could help in analyzing the heterogeneous propagation of sinus impulses in the atria, which seems to predispose to fibrillation. *Aim.* The aim of this study is to perform the PCA of the P-wave in patients prone to atrial fibrillation (AF). *Methods.* PCA is performed on P-waves extracted by averaging technique from ECG recordings acquired using a 32-lead mapping system (2048 Hz, 24 bit, 0-400 Hz bandwidth). We extracted PCA parameters related to the dipolar and not dipolar components of the P-wave using the first 3 eigenvalues and the cumulative percent of variance explained by the first 3 PCs (explained variance EV). *Results and conclusions.* We found that the EV associated to the low risk patients is higher than that associated to the high risk patients, and that, correspondingly, the first eigenvalue is significantly lower while the second one is significantly higher in the high risk patients respect to the low risk group. Factor loadings showed that on average all leads contribute to the first principal component.

*Key words:* principal component analysis, atrial fibrillation, P-wave, not dipolar components.

**Riassunto** (*Quantificazione delle componenti non dipolari della depolarizzazione atriale mediante l'analisi in componenti principali dell'onda P*). *Introduzione.* L'analisi in componenti principali dell'onda T è stata usata per quantificare le componenti dipolari e non dipolari dell'attivazione ventricolare, associate all'eterogeneità della ripolarizzazione. Allo stesso modo, l'analisi in componenti principali dell'onda P può essere utile nell'analizzare la propagazione del segnale elettrico negli atri, la cui eterogeneità sembra essere un fattore predisponente per la fibrillazione atriale. *Obiettivo.* Lo scopo dello studio è di effettuare l'analisi in componenti principali dell'onda P in pazienti suscettibili alla fibrillazione atriale. *Metodi.* L'analisi in componenti principali è effettuata su onde P estratte effettuando la media sincronizzata su registrazioni ECG acquisite mediante un sistema di mappaggio ECG a 32 derivazioni (2048 Hz, 24 bit, 0-400 Hz). *Risultati e conclusioni.* I parametri estratti per quantificare le componenti dipolari e non dipolari dell'onda P sono stati i primi 3 autovalori e la percentuale di varianza cumulata spiegata dalle prime 3 componenti principali. Quest'ultima è risultata maggiore nel gruppo di pazienti a basso rischio di sviluppare fibrillazione atriale rispetto al gruppo ad alto rischio. Inoltre, corrispondentemente, il primo auto valore è significativamente più basso mentre il secondo è significativamente più alto nei pazienti ad alto rischio rispetto a quelli a basso rischio. L'analisi dei fattori ha mostrato che in media tutte le derivazioni contribuiscono alla prima componente principale.

*Parole chiave:* fibrillazione atriale, componenti principali, onda P, componente non dipolare.

## INTRODUCTION

Atrial fibrillation (AF) is the most commonly encountered arrhythmia in clinical practice. It is defined by the absence of coordinated atrial systole, since it results from multiple re-entrant electrical wavelets that move randomly around the atria. Although it is not a lethal disease, AF may increase mortality up to 2-fold, primarily due to embolic stroke.

Indeed, the lack of coordinated atrial contraction leads to unusual fluid flow states through the atrium that could favour the formation of thrombus at risk to embolize, especially after return to normal sinus rhythm.

The incidence of atrial fibrillation increases significantly with advancing age. When a patient spontaneously alternates between AF and normal rhythm,

the condition is known as paroxysmal AF. When a patient continues with AF as the dominant cardiac rhythm without reversion to the normal rhythm, the condition is known as chronic AF. Two main electrophysiological conditions are indicated for AF initiation and perpetuation [1]: slower conduction velocity in some atrial areas and heterogeneity of cell refractory periods. This heterogeneity of structural and electrophysiological properties leads to a longer and more fragmented P-wave [2-4].

Thus, many studies focused on the analysis of the P-wave to extract parameters to recognize a patient with paroxysmal AF as well as to predict the development of AF [4-8]. Given the technical difficulties to analyze the P-wave and the different acquisition and processing systems used, these studies often lead to diverse and not-comparable results in terms of cut-off values.

On the other hand, the analysis of the T-wave, corresponding to the ventricular repolarization, has been extensively used to quantify repolarization inhomogeneity that may create an arrhythmogenic ventricular substrate. Promising results have been obtained by measuring the QT interval (QT dispersion) and by performing the principal component analysis of the T-wave [9-14].

The former analysis has been analogously applied to the P-wave: P-wave dispersion (which is the difference between the maximum and the minimum P-wave duration recorded from the 12 standard leads) has been shown to distinguish patients with paroxysmal AF [4-8]. Principal component analysis (PCA) of the T-wave has been extensively used to quantify both the complexity and the not dipolar components of the T-wave [9-12]: particularly, if the ECG would be completely explained by a single electrical dipole, the three largest principal components (PCs), and their corresponding orthogonal eigenvectors, would span the real three dimensional space (dipolar components), while the remaining PCs (not dipolar components) would be zero [12]. For the T-wave it has been demonstrated that the not dipolar components, quantified by the PCA, are not zero, and reflect local repolarization heterogeneity [12]. Accordingly, the PCA of the P-wave could help in analyzing the heterogeneous propagation of sinus impulses in the atria, which seems to predispose to fibrillation. The aim of this study is to perform the PCA of the P-wave in patients prone to AF, during sinus and paced rhythm. For the best of our knowledge, principal component analysis has never been applied to investigate the characteristics of the P-wave.

## MATERIALS AND METHODS

### Study population

Nineteen patients with paroxysmal atrial fibrillation and permanent dual chamber pacemakers (AT500-Medtronic Inc., Minneapolis, MN, USA) were recruited from Ospedale San Filippo Neri, Rome, Italy. The atrial pacing leads were positioned in the right atrium.

The AT500 device combines atrial sensing and detection algorithms for monitoring and diagnostics, and atrial therapy delivery functions.

The system can store up to 35 episodes of atrial tachycardia/flutter with electrograms and up to 128 episodes text summaries, without electrograms.

This pacemaker allows for accurate classification of atrial fibrillation episodes, with detailed information about episode instant of occurrence and duration; it also features three distinct programmable pacing algorithms that suppress atrial tachyarrhythmia trigger mechanisms.

When an episode occurs, the device is also programmed for arrhythmia termination. Three atrial pace-termination algorithms can recognise treatable atrial tachycardias and deliver antitachycardia paces-therapies to restore sinus rhythm.

The study population consisted of 9 female and 10 men, aged  $72 \pm 10$ . Two classes of risk have been defined, according to the number of AF episodes recorded by dual chamber pacemakers in the last 6 months preceding the study: class LR (low risk) without AF episodes and class HR (high risk) with at least one AF episode.

### Experimental protocols

Five-minute ECG recording was performed for each subject, with the pacemaker programmed in VVI mode, *i.e.* in single-chamber ventricular pacing mode set to a rate of 40 beats/min, so that to have spontaneous rhythm. Recordings were made using a multi-lead mapping system for high-resolution biopotential measurement (ActiveTwo, Biosemi, The Netherlands).

The system is made of a battery powered isolated AD box that digitises the signals and transfers them to a PCI receiver on computer through a fibre-optic connection.

The signals were digitised at a sampling rate of 2048 Hz and a resolution of 24 bits with a frequency response in the full DC-400 Hz range.

No further filtering was applied to the data. Thirty-two leads were positioned on the thorax (*Figure 1*), to allow accurate recordings of atrial signals.

ECG recordings were acquired as single-ended signals, with respect to a common reference position.

Before starting the acquisition, signals were visualised on a computer screen to check for good electrode contact.

### P-wave pre-processing

Every lead signal was pre-processed and analysed to extract the average P-wave characteristic.

The first step is to isolate the P-waves from the acquired signals: after detecting the R-wave (using an algorithm similar to that proposed by Pan and Tompkins [15]), P-waves are extracted in a 200 ms-long window (410 samples) starting 300 ms before the R-wave (*Figure 1*). This is a crucial point since the not correct identification of the P-wave beginning leads to the impossibility to construct the P-wave template.

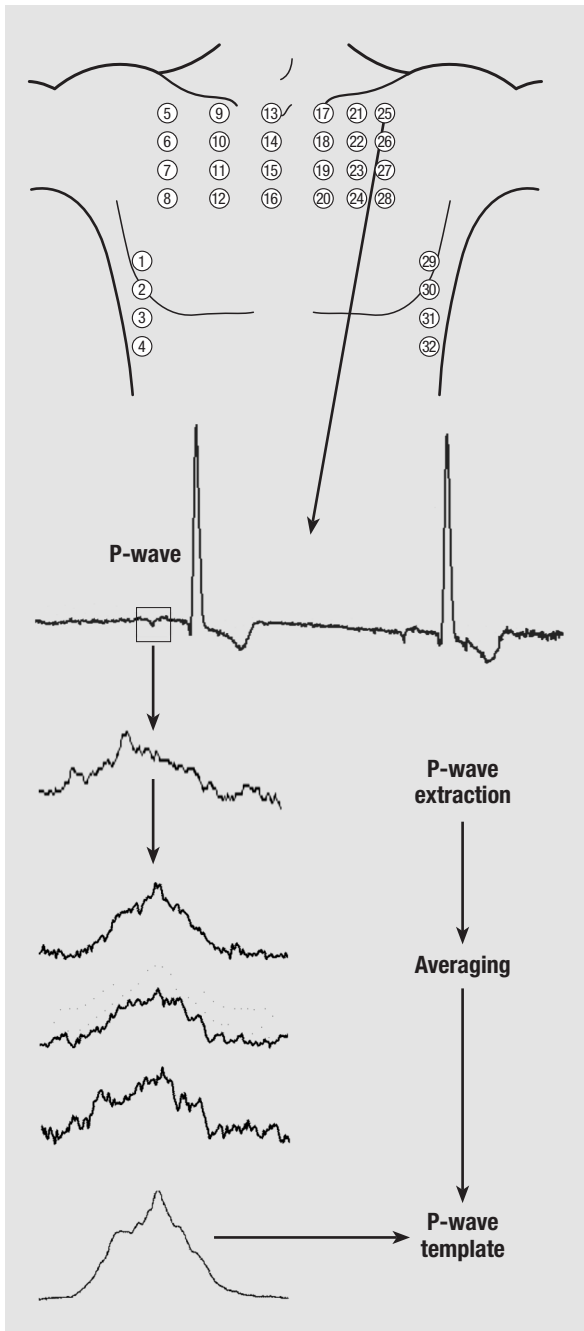


Fig. 1 | Scheme of the electrodes positioning and of the P-wave pre-processing procedure.

Secondly, a beat-by-beat linear piecewise interpolation was used to remove baseline wander, on each P-wave. Fiducial points for linear interpolation were taken from TP and PQ tracks of each beat.

Third, a P-wave template is constructed (Figure 1) by averaging each extracted P-wave having a cross-correlation coefficient with the current template higher than 0.9.

In order to take into account the variations in PR interval and/or the inaccuracy in R-wave detection

before averaging, P-waves were aligned according to the lag at which the cross-correlation function between the current averaged P-wave and each single P-wave shows its maximum (coherent averaging procedure).

The coherent averaging procedure went on until 200 beats were included. If the residual noise level (measured in the isoelectric TP track) remained at more than 1  $\mu\text{V}$  even after averaging of 200 beats, averaging procedure continued until the noise level reached a value lower than 1  $\mu\text{V}$ . If it was impossible, the lead was excluded from the study.

**Principal component analysis**

For each patient, PCA of the 32 averaged P-waves extracted from the 32 leads has been performed. Data from 32 leads were organized in a data matrix of 32 rows and 410 columns: each row is the 410-samples-long P-wave template of each lead. Singular value decomposition (SVD) of the data matrix has been performed. Since PCA transforms the measured P-wave to virtual parameters that are mutually independent (orthogonal), the 3 largest PCs would contain all the information in the P-wave stemming from the vectorial concept of a single electrical dipole. Following an approach already applied to the T-wave [10, 12], the other principal components (in this case from the 4<sup>th</sup> to the 32<sup>nd</sup>) represent the not dipolar components of the atrial depolarization.

The quantification of the dipolar and of the not-dipolar component of the heart electrical dipole, was based on the extraction of the first 3 eigenvalues (namely L1, L2 and L3) and of the cumulative percent of variance explained by the first 3 PCs (explained variance, EV). EV gives a direct indication of the amount of variability explained by the dipolar component of the electrical dipole. Of course the residual variance (100%-EV) is that associated to the not dipolar component.

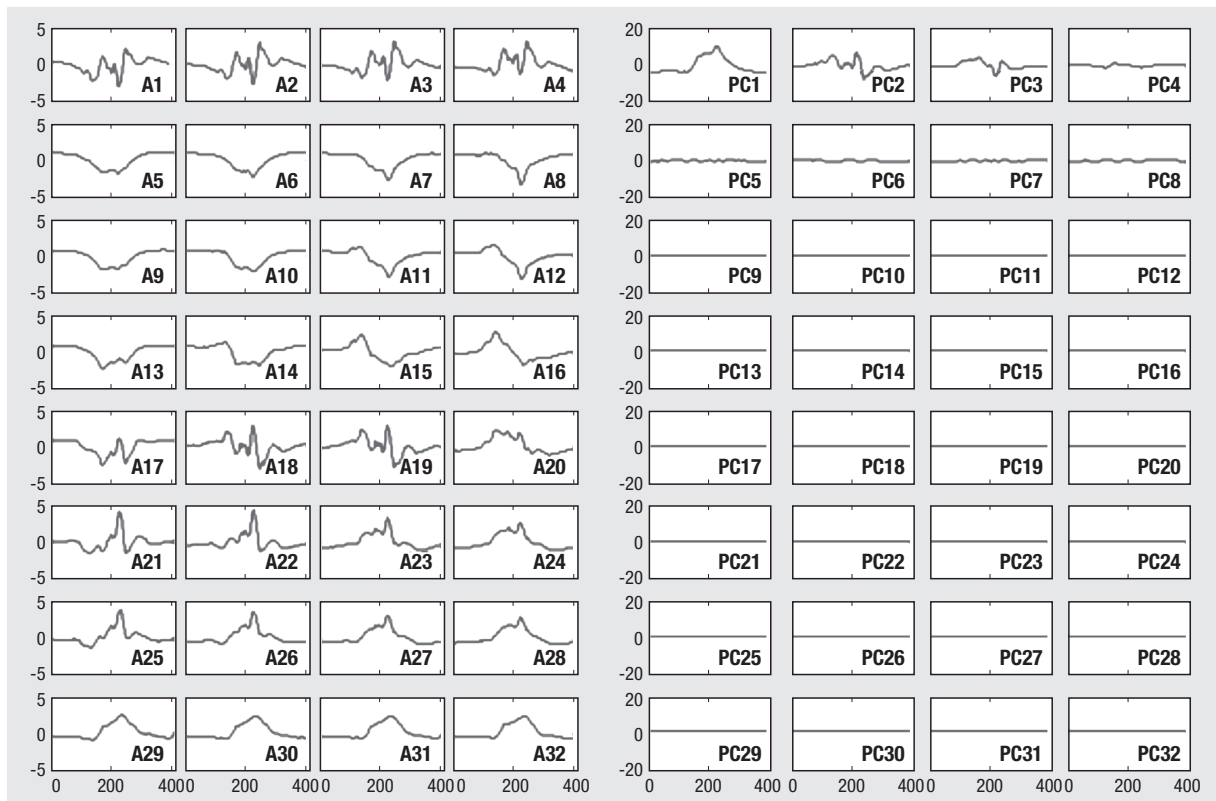
Each PCs is a linear combination of the original 32 signals. In order to estimate to which extent each lead contributes to the first principal component, factor loadings have been calculated. Analogous to Pearson's coefficient, the squared factor loading is the percent of variance in that variable explained by that PC (i.e. the degree of correlation between the original data and the first principal component expressed in percentage).

**RESULTS**

According to the definition of LR and HR risk patients, 7 patients belonged to the LR class and 12 patients to HR class.

Figure 2 shows the 32 P-wave templates and the results of the PCA for one patient. Table 1 summarizes the results obtained by the PCA parameters, L1, L2, L3 and EV.

The dipolar component of atrial activation, as expressed by the EV, of the LR patients resulted to be



**Fig. 2** | Example of the 32 P-wave templates and of the results of the principal component analysis for one patient.

**Table 1** | PCA-derived parameters for the P-wave extracted for low risk and high risk patients

Low risk patients	L1	L2	L3	EV (%)
1	22.4	6.6	0.48	97.9
2	22.8	6.4	1.7	98.6
3	21.7	7.6	0.5	98.5
4	22.23	6.5	0.78	98.4
5	21.8	7.02	0.68	96.5
6	20	8.1	0.73	98.4
7	22.3	5.7	0.3	99.01
<b>Mean ± standard deviation</b>	<b>21.9 ± 0.9</b>	<b>6.8 ± 0.8</b>	<b>0.7 ± 0.4</b>	<b>98.2 ± 0.8</b>
High risk patients	L1	L2	L3	EV (%)
1	19.9	10.04	1.13	97.2
2	19.2	10.2	1.3	98.9
3	15.18	12.82	0.85	98.5
4	19.7	8.03	2.44	98.2
5	19.53	10.42	1.7	98.7
6	17	8.8	2.11	96.5
7	18.1	13.1	0.6	98.9
8	18.9	10.1	2.2	98.8
9	14.8	14.2	0.8	98.1
10	16.3	9.2	1.4	97.5
11	16.8	10.2	2.3	98.15
12	15.9	7.5	0.85	95.75
<b>Mean ± standard deviation</b>	<b>17.6 ± 1.8</b>	<b>10.4 ± 2.0</b>	<b>1.5 ± 0.6</b>	<b>97.9 ± 1.0</b>
<b>Mann Whitney U-test (p)</b> <b>Low risk vs high risk patients</b>	p < 0.001	p < 0.001	p = 0.02	p = 0.6

PCA: principal component analysis; EV: explained variance.

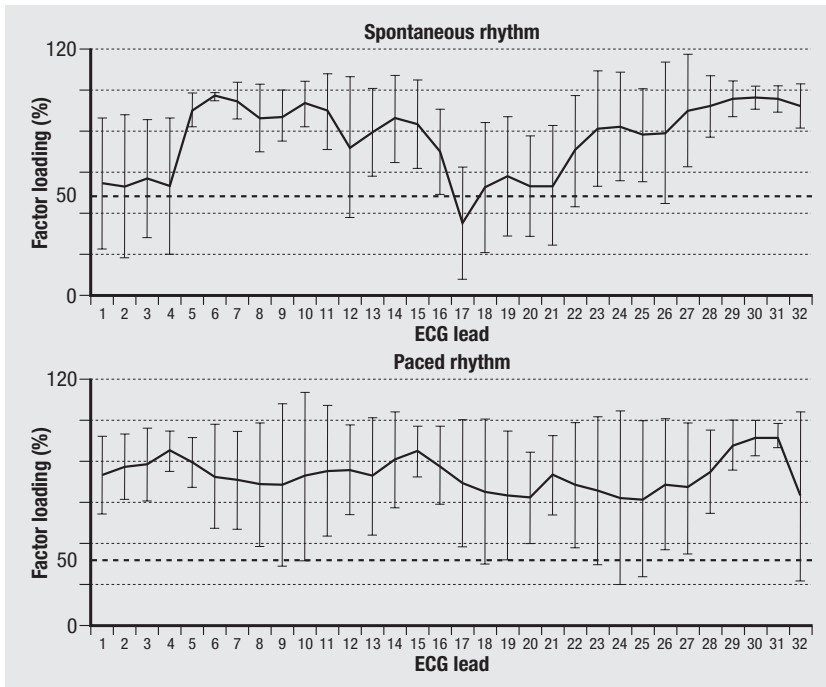


Fig. 3 | Factor loadings obtained during spontaneous rhythm and during pacing.

higher than that computed for HR patients, even if such a difference does not reach a statistical significance ( $98.2 \pm 0.8$  vs  $97.9 \pm 1.0$ ). However, when the eigenvalues of the first 3 components are analyzed, we find that while the first one (L1) is significantly higher for LR patients respect to HR patients, L2 and L3 are lower in LR group respect to the HR one (L2 in a significant way).

The computation of the factor loadings shows that on average all leads contribute to the first principal component. Figure 3 shows the factor loadings averaged (in absolute values) all over the population. Each lead but one in spontaneous rhythm (lead A17) correlates with the first principal component. No differences have been found between LR and HR patients.

Table 2 reports the correlation coefficients (R-square) between PCA parameters and P-wave duration parameters, previously published [16-18] (maximum and minimum P-wave duration in any of the 32 leads ( $P_{max}$ ,  $P_{min}$ ) as well as P wave dispersion ( $P_{disp} = P_{max} - P_{min}$ )). We find weak correlations, except a slightly higher correlation between EV and Pmax and Pdisp.

**DISCUSSION**

Patients with paroxymal AF are thought to have atrial depolarization heterogeneities, which have been extensively analyzed in terms of electrocardiographic indexes extracted from the P-wave to recognize or to predict the development and the perpetuation of AF [4-8].

However, the technical difficulties to acquire and process the P-wave have so far limited its clinical use.

Quantitative analysis of the T-wave, instead, has been carried out extensively, and is now used in the clinical practice. The PCA of the T-wave led to prom-

ising results, in terms of quantification of ventricular repolarization inhomogeneity that may create an arrhythmogenic ventricular substrate [9-12]. PCA can be used to estimate the dipolar and the not dipolar components of the P-wave, and thus its use could led to interesting results in terms of analysis of atrial path from surface ECG.

We hereby used an 32-lead ECG acquisition system particularly suitable for P-wave analysis, having 24 bit resolution and being DC-coupled. We performed the PCA of the P-wave in patients prone to AF. PCA has been applied to the average P-wave extracted in any of the 32 leads.

For each patient we extracted the first 3 eigenvalues (L1, L2 and L3) and the cumulative percent of variance explained by the first 3 PCs (EV). Such parameters are different from those employed for the T-wave analysis [10-12], even if, as for the T-wave, the chosen parameters are related to the first three PCs associated to the dipolar component of the P-wave, and to the remaining PCs associated to the not dipolar component of the P-wave (100%-EV).

To our knowledge this is the first time the PCA is performed on the P-wave, thus physiological interpretation and critical discussion can be related only to

Table 2 | Correlation coefficients between principal component analysis parameters and P-wave duration parameters

	L1	L2	L3	EV
$P_{min}$	0.37	0.31	0.52	0.34
$P_{max}$	0.39	0.29	0.36	0.55
$P_{disp}$	0.17	0.13	0.35	0.60

previous experimental evidences of ventricular conduction disturbance (PCA of the T-wave, [10, 12]).

The first important result is that the EV associated to the low risk patients is higher than that associated to the high risk patients, even if this difference does not reach a statistical significance. However, we found that, correspondingly, the first eigenvalue is significantly lower while the second one is significantly higher in the high risk patients respect to the low risk group. It means that while the not dipolar component does not change significantly, the distribution of the dipolar component over the three orthogonal axis changes. These results are consistent with those obtained by analyzing the dipolar component of the P-wave derived from orthogonal ECG leads [17, 18]: differences in orthogonal P-wave morphology were found between healthy individuals and patients suffering from paroxysmal atrial fibrillation group.

The second important results is that, on average, all the 32 leads contributes to the first PC, having a significant correlation coefficient with almost all variables.

Since any leads systematically show a significant correlation with first PC, each lead seems to contribute to a similar extent to the dipolar component. However, we found an inter-patient variability for the factor loadings – some patients had not significant factor loadings in some leads. This result suggests that maps of the correlation with the first PC (or of the average correlation with the first 3 PCs) could help in identifying those leads (*i.e.* body surface zones) which mainly contribute to the dipolar component of the atrial depolarization.

When performing PCA of the P-waves, attention must be paid to the system used to record and ac-

quire the ECG and to the processing algorithm used to extract the P-wave [19, 20]. Indeed, the analysis presented in this paper (PCA) has been performed on ECG recordings obtained by a system particularly suitable for the analysis of the P-wave, given its high resolution (24 bit, corresponding to 31 nV). Such a resolution guarantees a detail level of the P-wave which allows to obtain important information on the subtle changes in P-wave morphology, which are in turn due to the atrial conduction path. Whether such analysis can be performed even by commonly used commercial ECG system has to be investigated. In addition, it has to be mentioned that PCA-derived parameters of the T-wave have been investigated in terms of their reproducibility, which has been considered somewhat suspect, because of different pre-processing algorithms used to extract the T-wave (10-second median beat, single beat, averaging over 200 beats) [21]. Thus, if PCA is applied to analyze the P-wave, care must be paid even to the pre-processing algorithm used to extract the signal; although a standardized procedure is desirable, signal averages consisting of several complexes are required to produce reproducible and reliable values for P-wave parameters obtained by PCA.

#### Conflict of interest statement

There are no potential conflicts of interest or any financial or personal relationships with other people or organizations that could inappropriately bias conduct and findings of this study.

Received on 13 October 2011.

Accepted on 18 December 2011.

#### References

- Clavier L, Boucher JM, Lepage R, Blanc JJ, Cornily JC. Automatic P-wave analysis of patients prone to atrial fibrillation. *Med Biol Eng Comput* 2002;40(1):63-71. <http://dx.doi.org/10.1007/BF02347697>
- Davies LG, Ross IP. Abnormal P-waves and paroxysmal tachycardia. *Br Heart J* 1963;25:570-4. <http://dx.doi.org/10.1136/hrt.25.5.570>
- Kawano S, Hiraoka M, Sawanobori T. Electrocardiographic features of P waves from patients with transient atrial fibrillation. *Jpn Heart J* 1988;29(1):57-67. <http://dx.doi.org/10.1536/ihj.29.57>
- Dilaveris PE, Gialafos JE, Sideris S. Simple electrocardiographic markers for the prediction of paroxysmal idiopathic atrial fibrillation. *Am Heart J* 1998;135:73-8. [http://dx.doi.org/10.1016/S0002-8703\(98\)70030-4](http://dx.doi.org/10.1016/S0002-8703(98)70030-4)
- Jordaens L, Tavernier R, Gorgov N, 'T Kindt H, Dimmer C, Clement D. Signal-averaged P-wave: predictor of atrial fibrillation. *J Cardiovasc Electrophysiol* 1998;9:S30-4.
- Dilaveris PE, Gialafos JE. P-wave dispersion: a novel predictor of paroxysmal atrial fibrillation. *ANE* 2001;6:159-65. <http://dx.doi.org/10.1111/j.1542-474X.2001.tb00101.x>
- Darbar D, Jahangir A, Hammill SC, Gersh BJ. P-wave signal-averaged electrocardiography to identify the risk for atrial fibrillation. *Pace* 2002;25:1447-53. <http://dx.doi.org/10.1046/j.1460-9592.2002.01447.x>
- Dilaveris P, Gialafos J. Future concept in P-wave morphological analyses. *Card Electrophysiol Rev* 2002;6:221-4. <http://dx.doi.org/10.1023/A:1016320807103>
- De Ambroggi L, Aimè E, Ceriotti C, Rovida M, Negroni S. Mapping of ventricular repolarization potentials in patients with arrhythmogenic right ventricular dysplasia: principal component analysis of the ST-T waves. *Circulation* 1997;96(12):4314-8. <http://dx.doi.org/10.1161/01.CIR.96.12.4314>
- Acar B, Yi G, Hnatkova K, Malik M. Spatial, temporal and wavefront direction characteristics of 12-lead T-wave morphology. *Med Biol Eng Comput* 1999;37(5):574-84. <http://dx.doi.org/10.1007/BF02513351>
- Malik M, Acar B, Gang Y, Yap YG, Hnatkova K, Camm AJ. QT dispersion does not represent electrocardiographic interlead heterogeneity of ventricular repolarization. *J Cardiovasc Electrophysiol* 2000;11(8):835-43. <http://dx.doi.org/10.1111/j.1540-8167.2000.tb00061.x>
- Kesek M, Jernberg T, Lindahl B, Xue J, Englund A. Principal component analysis of the T wave in patients with chest pain and conduction disturbances. *Pacing Clin Electrophysiol* 2004; 27(10):1378-87. <http://dx.doi.org/10.1111/j.1540-8159.2004.00642.x>
- Porthan K, Viitasalo M, Jula A, Reunanen A, Rapola J, Väinänen H, Nieminen MS, Toivonen L, Salomaa V, Oikarinen L. Predictive value of electrocardiographic QT

- interval and T-wave morphology parameters for all-cause and cardiovascular mortality in a general population sample. *Heart Rhythm* 2009;6(8):1202-8.  
<http://dx.doi.org/10.1016/j.hrthm.2009.05.006>
14. Okin PM, Devereux RB, Fabsitz RR, Lee ET, Galloway JM, Howard BV. Principal component analysis of the T wave and prediction of cardiovascular mortality in American Indians: the Strong Heart Study. *Circulation* 2002;105(6):714-9.  
<http://dx.doi.org/10.1161/hc0602.103585>
  15. Pan J, Tompkins WJ. A real-time QRS detection algorithm. *IEEE Trans Biomed Eng* 1985;32(3):230-6.  
<http://dx.doi.org/10.1109/TBME.1985.325532>
  16. Keane D, Stafford P, Baker S, Lewis S, Jackson G, Vincent R. Signal-averaged electrocardiography of the sinus and paced P-wave in sinus node disease. *Pacing Clin Electrophysiol* 1995; 18(7):1346-53.  
<http://dx.doi.org/10.1111/j.1540-8159.1995.tb02595.x>
  17. Holmqvist F, Platonov PG, Havmøller R, Carlson J. Signal-averaged P wave analysis for delineation of interatrial conduction - further validation of the method. *BMC Cardiovasc Disord* 2007;7:29.  
<http://dx.doi.org/10.1186/1471-2261-7-29>
  18. Carlson J, Havmøller R, Herreros A, Platonov P, Johansson R, Olsson B. Can orthogonal lead indicators of propensity to atrial fibrillation be accurately assessed from the 12-lead ECG? *Europace* 2005;7(Suppl. 2):39-48.  
<http://dx.doi.org/10.1016/j.eupc.2005.04.012>
  19. Censi F, Ricci C, Calcagnini G, Triventi M, Ricci RP, Santini M, Bartolini P. Time-domain and morphological analysis of the P-wave. Part I. Technical aspects for automatic quantification of P-wave features. *Pacing Clin Electrophysiol* 2008;31(7):874-83.  
<http://dx.doi.org/10.1111/j.1540-8159.2008.01102.x>
  20. Censi F, Ricci C, Calcagnini G, Triventi M, Ricci RP, Santini M, Grammatico A, Bartolini P. Time-domain and morphological analysis of the P wave. Part II: effects of atrial pacing on P-wave features. *Pacing Clin Electrophysiol* 2008;31(8):935-42.  
<http://dx.doi.org/10.1111/j.1540-8159.2008.01119.x>
  21. Batdorf BH, Feiveson AH, Schlegel TT. The effect of signal averaging on the reproducibility and reliability of measures of T-wave morphology. *J Electrocardiol* 2006;39(3):266-70. Epub 2006 Mar 10.  
<http://dx.doi.org/10.1016/j.jelectrocard.2005.11.004>