Analysis of Atrial Fibrillation after CABG using Waveletes

N Maglaveras[1], I Chouvarda[1], G Dakos[2], V Vasilikos[2], S Mochlas[2], G Louridas[2]

[1]Lab of Medical Informatics, Medical School, Aristotelian University of Thessaloniki, Greece
[2]1st Cardiology Dept, AHEPA Hospital, Aristotelian University of Thessaloniki, Greece

Abstract

In the present study, Wavelet analysis of P wave for the prediction of Atrial Fibrillation after CABG is evaluated.

Continuous Wavelet Transform is applied to ECG and Mean/Max parameters are calculated within the P window for different frequency bands. Thus, 24 parameters are available, which, along with the 4 window lengths (corresponding to P wave length of X, Y, Z, V signals), make a pool of available parameters to be used for classification. Linear regression is used for the classification of the two groups and bootstrapping is applied in order to enhance statistical robustness. The features to be used in the regression model are selected from the pool of available parameters by use of an iterative procedure. The outcome of the feature selection procedure shows that X and Z-axis features as well as vector-magnitude features are the most important ones for the prediction of Atrial Fibrillation after CABG.

1. Introduction

Postoperative Atrial Fibrillation (AF) is associated with significant morbidity, longer hospital stay, and higher related costs [1]. AF usually occurs during or before the third postoperative day. Identification of patients vulnerable for arrhythmia will allow targeting of those most likely to benefit from prophylactic therapy. Although the etiologic mechanism of postoperative AF and its optimum method of prophylaxis or management are not well defined, progress has been made during the past decade [2].

Within this context, various studies have taken place in order to evaluate clinical characteristics and laboratory methods, like the duration of P-wave, which can be used as predictors of the development of this arrhythmia. [3], [4], [5], [6], [7]. It has been widely documented that prolonged Signal Averaged P-wave Duration, advanced age, and male sex are independent predictors that identify patients at high risk for development of AF after CABG. However, P wave dispersion and amplitude did not provide significant information in the prediction of postoperative AF. The P-wave duration method has been tested in classic leads as well as in orthogonal leads.

Other methods have been investigated as well. RR variability characteristics using wavelet and fractal analysis has reported high sensitivity and specificity for automatic AF detection [8], and can be used in screening for AF in large populations at risk. In another work [9], raised cosine wavelet transform (RCWT) has been applied for Atrial Fibrillation determination, and a low frequency band in the range of 0-5 Hz has been found to be sensitive to AF.

The purpose of the current work is to evaluate the predictive value of Continuous Wavelet Transform in AF after CABG. Specifically, it is investigated whether Wavelet-based features of the P-wave can help distinct AF from non-AF cases. Therefore, various parameters describing the morphology of CWT have been determined, and their value for the classification of AF has been assessed.

2. Methods

The method applied consists of two parts. First, ECG processing is applied, consisting of P-wave marking followed by wavelet analysis and calculation of parameters based on Wavelet Transform. Classification of the two groups of patients (AF and non-AF) is done, based on these parameters. Customised software has been used for Wavelet Transform, while further analysis and classification took place in MATLAB.

2.1. Data

ECG data were acquired from AHEPA Cardiology Unit of Aristotle University of Thessaloniki. The sample used for the analysis consisted of 37 subjects, 24 without Atrial Fibrillation and 13 with Atrial Fibrillation, during the first days after CABG. Vectorcardiogram at 1000Hz is available for each patient, X, Y and Z leads as well as the vector magnitude. P wave is semi-automatically located, by the physician using an interactive graphical tool. In order to reassure a more robust statistic and override the problems related to single beat processing, three consecutive beats are stored and analyzed for each patient.

2.2. Feature extraction

Morlet mother function is used for Wavelet analysis in the frequency range 50-200Hz. This range is divided in three bands (band1=150-200Hz, band2 =100-150 Hz and band3=50-100 Hz). In each of these time-frequency areas defined by the wavelet transform and distribution (DWVD), the following parameters are calculated:

$$Mean_i = \sqrt{\frac{\sum_{t=t_i}^{t_{i+1}} \sum_{f=f_j}^{f_{j+1}} \left|DWVD(t,f)\right|^2}{num}}$$

$$Max_i = \max\{DWVD_{abs}(t,f)\}, \quad (1)$$

$$t_i \le t < t_{i+1},\ f_i \le f < f_{i+1}$$

The Mean and Max stated in (1) correspond to the total energy and the maximum value of a time-frequency area. Eighteen parameters are eventually extracted from each DWVD, corresponding to basic features of the time-frequency map. Three consecutive beats are analysed and an average parameter is calculated.

The pool of available parameters consists of:

- 4 window lengths corresponding to P wave length (of X,Y,Z,V signals), semi-automatically defined by the clinician
- mean/max parameters for 3 frequency bands for 4 signals (X,Y,Z,V), totally 24 available wavelet parameters

Feature enumeration is explained in Table 1.

Table 1. The pool of available features from which the feature vector can be selected for classification.

	X	Y	Z	V
Width	3	4	5	6
Mean1	7	13	19	25
Max1	8	14	20	26
Mean2	9	15	21	27
Max2	10	16	22	28
Mean3	11	17	23	29
Max3	12	18	24	30

2.3. Classification method

In the present work, the available sample of patients is apriori classified in two groups, according to the presence of AF after CABG, a distinction based on clinical evidence. Our hypothesis is that a linear combination of the selected features (linear discriminant finction) will classify cases in two classes reflecting the two groups of AF.

$$f = \sum_i a_i \cdot parameter_i$$

$$f = 10 \quad \text{for group1} \qquad (2)$$

$$f = -10 \quad \text{for group2}$$

The time-frequency features discussed in section 2.2 are used as variables in the linear expression of (2) and the weights a_i of the discriminant function for each parameter i are estimated (the weights a_i in (2)). Finally, their contribution to the discrimination of the two groups is investigated.

The selected features to be used in the classification process have a physical meaning. However, it has to be investigated whether this set of features is the optimal or a subset would lead to the smallest classification error. Instead of an exhaustive search of the optimum feature vector, which would be extremely time-consuming, an iterative method is proposed for feature selection [11].

Specifically, the regression/classification process is repeated and the addition of a feature each time is evaluated based on predefined criteria. The monotonicity property of the criteria is considered as a clue for the necessity of each feature. The procedure starts with the very best features selected by a statistical test and sequentially a feature is added that improves results, according to predefined criteria, until there is no more improvement or all features are added.

A cost function is empirically defined for the optimization algorithm, which includes as criteria not only the minimum classification error on training set but also criteria related to good clustering, for example the distance between the classes, the compactness of the classes etc.

Since the amount of medical data is not large in the present work, a bootstrap technique through resampling of data is applied in order to produce robust statistics from a small sample of measurements [12]. The basic idea is the repetition of the regression algorithm (Eq 2), each time with a different dataset generated via permutation of the initial dataset. After a large number of such repetitions, a statistic is produced for each estimated weight and the appropriate weight values are depicted from the statistic, while confidence intervals are also calculated for each estimate.

3. Results

3.1. Statistical analysis

The first step is the investigation of the statistical characteristics. An initial statistical analysis is performed on the available dataset. Actually, two datasets are considered for this basic statistical analysis: a) the unaveraged dataset, consisting of 3 sets of features for each patient, therefore of 3x37 samples and b) the averaged dataset, consisting of 37 samples, each one having the average of the three samples responding to the same patient.

t-test

When t-test is applied to the averaged data, no parameter is selected, while on the unaveraged data, the parameters selected are displayed in Table 2

Table 2. Features selected by the t-test

	X	Y	Z	V
Width	**3**	4	**5**	6
Mean1	7	**13**	19	**25**
Max1	8	**14**	20	**26**
Mean2	9	**15**	21	27
Max2	10	**16**	22	**28**
Mean3	11	**17**	23	**29**
Max3	12	**18**	24	30

Wilcoxon test 1

The features selected by Wilcoxon signed rank test, that returns the significance probability that the medians of two matched samples are equal, are less than in the previous test. It seems that P- durations on X and Z-axis as well as X and V Wavelet parameters are significant, as displayed in Table 3.

Table 3. Features selected by the Wilcoxon signed rank test

	X	Y	Z	V
Width	**3**	4	**5**	6
Mean1	**7**	13	19	**25**
Max1	**8**	14	20	**26**
Mean2	9	15	21	27
Max2	10	16	22	**28**
Mean3	11	17	23	**29**
Max3	12	18	24	30

When the test is applied to the unaveraged data (Table 4), basically X and V parameters are selected again.

Table 4 Features selected by the Wilcoxon signed rank test, applied on the unaveraged dataset.

	X	Y	Z	V
Width	**3**	4	**5**	6
Mean1	**7**	13	19	**25**
Max1	**8**	14	20	**26**
Mean2	**9**	15	21	**27**
Max2	**10**	16	22	**28**
Mean3	**11**	17	23	**29**
Max3	12	**18**	24	**30**

Wilcoxon test 2

The Wilcoxon rank sum test, investigating whether two populations are identical, is also applied. In this case, again the same P-wave duration on X and Z-axis is selected along with some V parameters (Table 5).

Table 5. Features selected by the Wilcoxon rank sum test

	X	Y	Z	V
Width	**3**	4	**5**	6
Mean1	7	13	19	**25**
Max1	8	14	20	26
Mean2	9	15	21	27
Max2	10	16	22	**28**
Mean3	11	17	23	**29**
Max3	12	18	24	30

3.2. Classification

The initial statistical analysis shows that there is a difference between using averaged and unaveraged parameters, and specifically, when using the unaveraged set, more parameters are selected. In order to have a moderate and robust approach, the averaged set is used for training, while the unaveraged one, having more noise, is used to test the reproducibility of the classification, since there is no other test group available. The case where parameters are normalized to P-wave duration had been initially investigated and rejected, because it did not produce satisfactory results. The cases described below are:

- Case A. Start with the features depicted by sign rank test for medians (Table 3), which are X and Z durations, mean1 and max1 on X-axis, mean1, max1 max2, mean3 on vector magnitude (features 3, 5, 7, 8, 25, 26, 28, and 29)
- Case B. Starting with the same features, the iterative procedure described in 2.3, is applied and parameters 9, 27 and 19 are also added to the feature set. The parameters added are also from X, Z and V axes.

Table 6. Classification results for Cases A and B, as described above. AF stands for Atrial Fibrillation. Group and NAF stands for No Atrial Fibrillation Group

	Features	Error AF (training)	Error NAF (training)	Error AF (test)	Error NAF (test)
A	3, 5, 7, 8, 25, 26, 28, 29	16.6%	30.7%	23.61	25.64
B	3, 5, 7, 8, **9**, **19**, 25, 26, **27**, 28, 29	8.3	15.38	19.44	25.64

Therefore there is some benefit from the application of the iterative procedure, since the classification results are improved both in the training set and in the test set by the addition of three more features, as displayed in Table 6. Besides, the clusters became more compact and the distance between them larger. Eleven parameters are finally selected: P wave duration on X and Z axis, (mean1, max1, mean2) on X axis, (mean1) on Z axis and (mean1, max1, mean2, max2, mean3) on vector magnitude V. Using these features, the two groups can be classified with satisfactory sensitivity and specificity, as displayed in Table 8, taking into account the complication of the problem.

Table 7. Characteristics of the clustering of the two groups. “I-Mah” is the intra-mahalanobis distance, “Eukl” is the Euclidean distance between the groups and “Mah” is the Mahalanobis distance between the groups.

	Features	I-Mah AF	I-Mah NAF	Eukl / Mah
A	3, 5, 7, 8, 25, 26, 28, 29	25.96	33.28	7.08 64.30
B	3, 5, 7, 8, **9**, **19**, 25, 26, **27**, 28, 29	23.81	17.05	8.73 61.61

Table 8. Classification measures for training and test set.

	Training set	Test set
Sensitivity	91.6667	80.5556
Specificity	84.6154	74.359
Positive prediction index	91.6667	85.2941
Negative prediction index	84.6154	67.4419

4. Conclusion

Wavelet analysis and classification of the P wave are presented. This type of analysis proves to have high sensitivity and specificity for the prediction of post-CABG Atrial Fibrillation.

It is interesting that the parameters selected by the procedure are duration, also documented elsewhere and also X, Z and V-axis features. The reason for specific differences in these axes may be the changes in intra-atrial conduction path and velocity [13]. Sustained Atrial Fibrillation causes "remodeling" of the atrium, and therefore conduction changes including fibrosis and a decrease of the atrial refractory period. However, in order to reach more reliable results about the conduction differences, and also conclude to a classification scheme, it is crucial use a larger dataset for evaluation.

Acknowledgements

This work was supported in part by a CEC funded project IST-1999-13352 with acronym CHS.

References

[1] Amar D. ‘Postoperative atrial fibrillation’, *Heart Dis* 2002;4(2):117-23.

[2] Guo Y, Hu S, Wu Q, Xu J, Song Y, Zhu X, ‘Predictors of atrial fibrillation after coronary artery bypass graft surgery.’, *Chin Med J (Engl)* 2002;115(2):232-4.

[3] Inamura N, Matsushita T, Fukunami M, Sano T, Takeuchi M, Kurotobi S, Miura T, Matsuda H, Okada S, ‘P wave signal averaged electrocardiography in patients undergoing the Fontan operation.’, *J Cardiol* 2002;39(2):101-7.

[4] Passman R, Beshai J, Pavri B, Kimmel S, ‘Predicting post-coronary bypass surgery atrial arrhythmias from the preoperative electrocardiogram.’, *Am Heart J*, 2001;142(5):806-10.

[5] Zaman AG, Archbold RA, Helft G, Paul EA, Curzen NP, Mills PG, ‘Atrial fibrillation after coronary artery bypass surgery: a model for preoperative risk stratification’, *Circulation*, 2001;103(8):E48.

[6] Guidera SA, Steinberg JS, ‘The signal-averaged P wave duration: a rapid and noninvasive marker of risk of atrial fibrillation.’, *J Am Coll Cardiol*,. 1994;23(2):549-50.

[7] Chang CM, Lee SH, Lu MJ, Lin CH, Chao HH, Cheng JJ, Kuan P, Hung CR, ‘The role of P wave in prediction of atrial fibrillation after coronary artery surgery.’, *Int J Cardiol,* 1999;68(3):303-8.

[8] Duverney D, Gaspoz JM, Pichot V, Roche F, Brion R, Antoniadis A, Barthelemy JC, ‘High accuracy of automatic detection of atrial fibrillation using wavelet transform of heart rate intervals.’, *Pacing Clin Electrophysiol,* 2002;25(4 Pt 1):457-62.

[9] Khadra L, al-Fahoum AS, al-Nashash H, ‘Detection of life-threatening cardiac arrhythmias using the wavelet transformation.’, *Med Biol Eng Comput,* 1997;35(6):626-32.

[10] Boufidou A, Maglaveras N, Chouvarda I, Mohlas S, Kontopoulos A, Louridas G, ‘Morlet wavelet analysis of Holter recordings in acute anterior myocardial infarction patients may be used for early prediction of patent infarct related artery after thrombolysis.’, *Eur. Heart J,* 1999;20 (abs suppl):338-9.

[12] Jain A.K., Duin P.W and Mao J, ‘Statistical Pattern Recognition: A Review.’, *IEEE PAMI,* 2000;22(1):4-37.

[11] Zoubir A.M. and B. Boashash, ‘The Bootstrap and its Application in Signal Processing.’, *IEEE Sig Proc Mag,* 1998;15(1):56-75.

[13] Franz MR, Karasik PL, Li C, Moubarak J, Chavez M, ‘Electrical remodelling of the human atrium: similar effects in patients with chronic atrial fibrillation and atrial flutter.’, *J Am Coll Cardiol,* 1997;30:1785-92.

Address for correspondence

Nicos Maglaveras. PhD
Associate Professor
Aristotle University – The Medical School
Lab. of Medical Informatics – Box 323
54124 Thessaloniki, GREECE
EMAIL : nicmag@med.auth.gr