

Original Research Article

Prediction of Paroxysmal Atrial Fibrillation through Spectral and Statistical Techniques

ABSTRACT

Aims: The paper describes some methods based on statistical tools and time-frequency analysis for extracting features in frequency and temporal domain that may be used as predictors of Paroxysmal Atrial Fibrillation

Place and Duration of Study: Department of Computer Engineering, College of Computers and IT, Taif University, Saudi Arabia (Jan 2014-Feb 2015)

Methodology: The main focus is on the investigation of useful tools emerging from statistics or quadratic transformations that can predict the possible onset of PAF in order to provide a prompt remedy. Our study was based on a database of two-channel ECG recordings which has been created for use in the Computers in Cardiology Challenge 2001. Real ECG records from 10 normal subjects and 10 subjects having PAF are studied. The specific features from the surface electrocardiogram immediately prior to the onset of Paroxysmal Atrial Fibrillation (PAF), shows that the occurrence of PAF can be predicted to some extent by analysing the spectral components, statistical and time-frequency parameters. The presence of PAF in p-records makes the signal spectrum rich in energy. Alternatively, the variance and standard deviation of signals prior to the occurrence of PAF also give valuable information which indicates the possible occurrence of PAF in the subsequent signal part.

Results and Conclusion: The specific features from the surface electrocardiogram immediately prior to the onset of paroxysmal atrial fibrillation (PAF) shows that the occurrence of the latter can be predicted to some extent by analyzing the statistical and spectral parameters

Keywords: Biomedical Signal Processing, Electro Cardiogram (ECG), Paroxysmal Atrial Fibrillation (PAF), Power Spectral Density, Fourier Spectrum

1. INTRODUCTION

Atrial fibrillation results from irregular heart beat or arrhythmia which may cause no symptoms. But this is normally accompanied with palpitations, chest or congestive heart failure. Atrial fibrillation normally results in stroke which can be identified by taking heart pulse readings. It can last from minutes to days or can be permanently present in the patient. Atrial fibrillation (AF) mostly affects elder people. Cardiac arrhythmic disorders are known for more than hundred years and atrial fibrillation (AF) has now been recognized as the most common among all kinds of arrhythmias.

Many studies have been performed on the assessment of the risk for atrial fibrillation (AF) over the recent years. The prediction of AF has been investigated in different contexts: in patients without apparent heart disease[6],[10], in hypertensive ones[11],[16]), in patients with coronary artery disease or undergoing coronary artery bypass surgery[8] and in patients after cardiac surgery[17],[19]. Most of these studies used the 12 leads for measuring the electrocardiogram (ECG) signals, the signal-averaged ECG methods[13], [15], [17], [19], [14], [3], [4] or the Frank leads [10]. Different electrocardiographic markers have been proposed for the assessment of risk for AF: R-R intervals, maximum P wave duration, P index P wave dispersion, and morphological changes of the P waves [7]. It has been reported that patients developing AF after CABG (Coronary artery bypass grafting) had a significantly longer P-wave duration in standard ECG than patients who remained in sinus rhythm[5]. The presence of prolonged P-wave

duration ($>100\text{ms}$ in lead II) can be an independent predictor of AF with a 1.9-fold risk compared to a P-wave duration of less than 100 ms [9]. The P-wave duration in both signal-averaged ECG and surface ECG has been reported to be prolonged in patients who develop AF after CABG [12]. In that study, a significant correlation was found between the P-wave duration in standard ECG and signal-averaged ECG. In other studies, it has been shown that the P-wave duration in lead II and signs of left atrial enlargement were determined from standard ECG. They found that left atrial enlargement but not the P-wave duration was an independent predictor of AF after CABG[1]. Stafford and colleagues analyzed lead II P-wave duration, total P-wave duration, and P-terminal force in standard ECG. The total P wave duration was the time from the earliest onset of P-wave activity in any of leads I, II or III to the last P-wave activity in any of these leads. No significant differences were observed in any of these variables between patients who developed AF after CABG and those who did not[18]. Similarly, another study found no difference in the P-wave duration on standard ECG between patients with and those without AF after CABG [6].

The organization of this paper describes first about the ongoing and past works in this area as in section 2. The methodology adapted in this research paper is described in section 3. The discussion on the different results is detailed in section 4. The section 5 concludes the final result of the study and the possible future works in this area..

2. METHODOLOGY

The different methods are implemented in MATLAB using the datasets available from Physionet.org [2]. The database is divided into a learning set (records with names starting with n and p) and a test set (records with names starting with t). The learning set consists of 50 record sets [2]. Each record set contains two 30-minute records with consecutive record names (e.g., p11 and p12). The records with names beginning with p come from subjects who have paroxysmal atrial fibrillation (PAF). The even-numbered record in each pair of 30-minute records contains the ECG immediately preceding PAF. For example the PAF signal record p16 immediately precedes an episode of PAF [2]. The odd-numbered record of the set (for example, record p13) contains 30 minutes of the ECG during a period that is away from any episode of PAF (there is no PAF during the 45-minute period before the beginning or after the end of the 30-minute record). The records with names beginning with n come from subjects who do not have atrial fibrillation. The test set is similarly constructed of 50 record sets (from 50 different subjects). The test set records are named t01, t02 and so on [2].

The research is divided into two major steps; pre-processing and feature Extraction. During the analysis, time domain and statistical analysis are carried out. Then, the signals are filtered at different frequency bands and analyzed. The spectral analysis is also carried out for the same purpose of extracting specific features to support PAF prediction. The proposed methods are described in the Figure 1, Figure 2 and Figure 3.

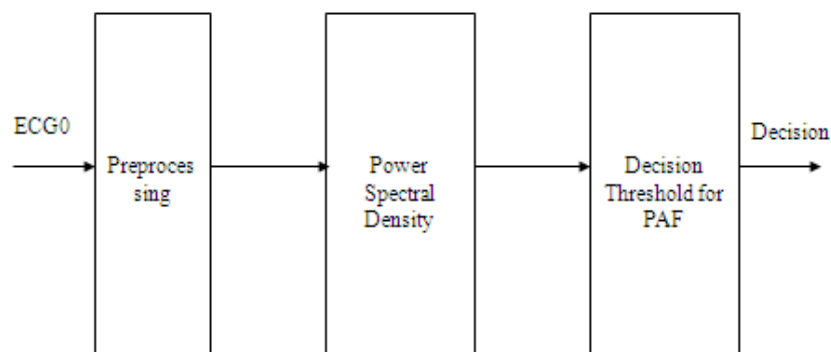


Figure 1: Analysis using Power Spectral Density

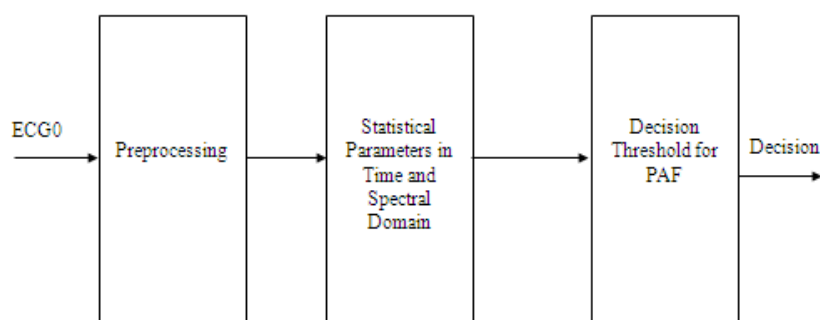


Figure 2: Analysis using Statistical Methods

2.1. Step 1: Pre-processing

The ECG signals collected using the electrodes usually are affected by different kinds of interference signals such as the 50 Hz power line interference and the baseline wandering. Low pass and band pass filters with appropriate cut-off frequencies are used to remove such artefacts.

2.2. Step 2: Feature Extraction

The feature extraction step includes applying suitable signal processing algorithms on the filtered signal in order to identify a parameter that describes the desired response. We use power spectral density (PSD) analysis and statistical parameter analysis for the same. Power Spectral Density (PSD) describes how the power of a signal is distributed over different frequencies.

The analysis is split up into multiple stages. In the first stage, the time domain analysis is carried out. In this stage, the signals are filtered at different frequency bands and are analysed. In the next stage the Spectral and PSD components are analysed. This is followed by statistical analysis in both time and frequency domain.

Time domain analysis can be carried out after splitting up the ECG signals into different frequency bands in order to see the different time domain contributions of PAF signals. As a first step, low pass filtering at 20Hz is carried out, and then band-pass filtering at 20-30Hz and 30-50Hz are carried out separately. The frequency range of 20-50Hz is more focused in this study. Spectral patterns and the concentration of energy are also analysed for all the records, in order to study if a particular frequency range contributes more to the PAF. Mean and standard deviation are also calculated for the time-series signals of p and n records, and the uniqueness in responses are noticed. Later the variance of spectral components is also studies by calculating the statistical parameters of the signals in the spectral domain.

3. RESULTS AND DISCUSSION

The above methods are implemented in MATLAB using the datasets available from Physionet.org [1]. 20 sets of signals are used in the analysis. Real ECG records from 10 normal subjects (n-records) and 10 subjects having PAF (p-records) are studied. The signals are analyzed after removing the 60Hz power line interference terms using a notch filter.

99 3.1. Time Domain Analysis

100 The p-records and n-records were analyzed in time domain after appropriate filtering at different
 101 frequencies. The Figure3 shows a view of a normal n-record n01, which is free of PAF. It shows signals
 102 from 2 electrodes ECG0 and ECG1. For convenience, ECG0 is mostly used and represented in this
 103 report. In the figure, the shape of PQRST wave of ECG can be noted.

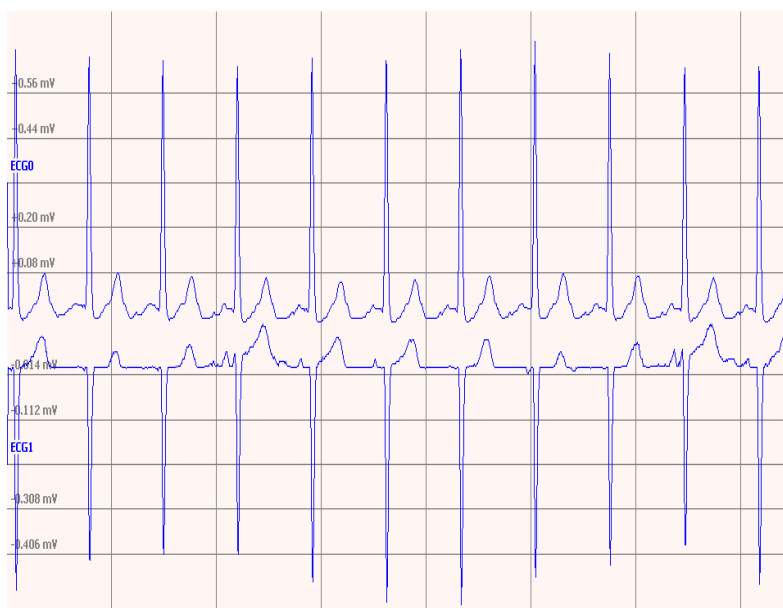


Figure 3: A Sample of Record n01 in Time Domain

106 A similar representation of a p-record p01 is shown in Figure4. The p-records are characterized by the
 107 presence of PAF immediately after the even numbered records. The Figure5 is an odd numbered p-
 108 record p01 that is at least 45 minutes before the even numbered counterpart. So it represents the nature
 109 of p signals very far from the possible occurrence of a PAF. Nevertheless, the presence of higher
 110 frequency components can be noticed in both ECG0 and ECG1 of p records in comparison with n-
 111 records.

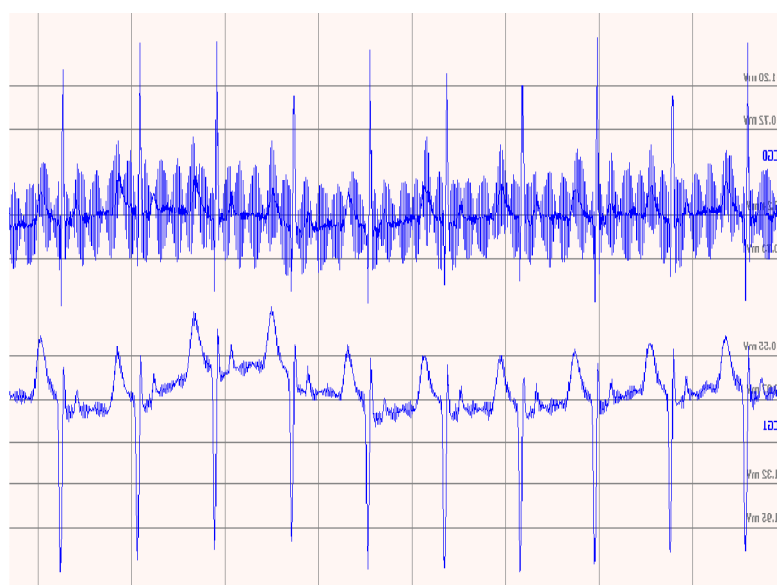


Figure 4: ECG0 and ECG1 Record p01 in Time Domain

Filtering is performed in order to analyze the frequency contents that are dominating in the even numbered p-records containing PAF, in comparison with the odd numbered records.

Low Pass Filtering at 20Hz

Low pass filtering was done for 10 subjects of both n and p-records. The Figure5 represents the record p01 after low pass filtering at a frequency 20Hz. We can see almost a normal shape of a PQRST waveform in it.

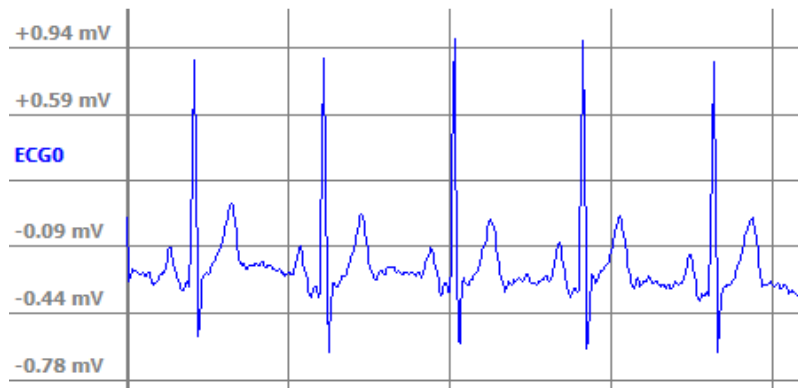


Figure 5: ECG0 of Record p01 after LPF at 20Hz

Figure6 shows the ECG0 of record p02 after low pass filtering at 20Hz. The shape is almost the same except that the peak values of signal amplitudes in p02 are slightly lesser than p01. There is also a higher negative DC value for p02 record in comparison with p01.

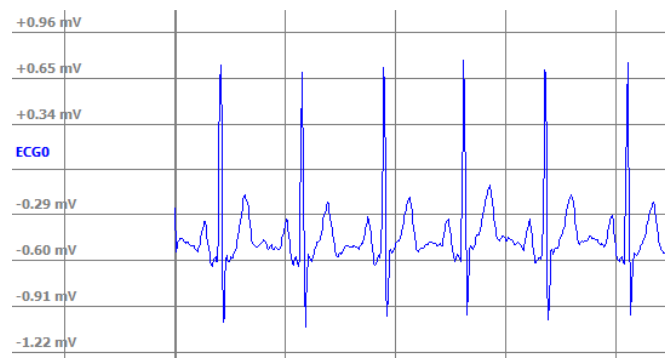


Figure 6: ECG0 of Record p02 after LPF at 20Hz

From the ECG0 of record n01 after low pass filtering at 20Hz as in Figure7, it can be noticed that both p01 and p02 signals show many low-amplitude oscillations in comparison with the n01 records. This is understood as an indication for the presence of PAF in p-records.

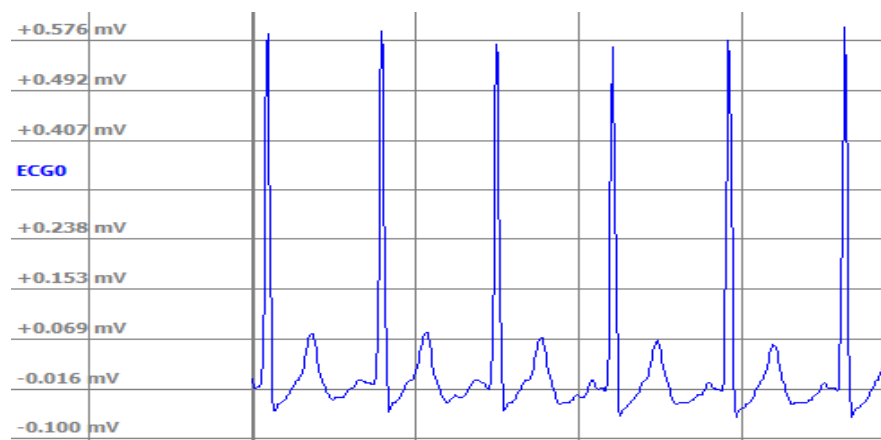


Figure 7: ECG0 of Record n01 after LPF at 20Hz

Band Pass Filtering in the Band 20-30Hz

The 10 subjects of n and p-records are undergone band pass filtering in the band of 20-30Hz. Some of the past studies in this area give an indication that, the PAF components usually contribute more to the signal in the range of 20Hz to 50Hz. So, in this work the contributions of PAF in the frequency band of 20-30Hz and 30-50Hz are studied independently. It is anticipated that, if the PAF contributes more to any of these bands, the corresponding signal amplitudes must show peak amplitude values and variations accordingly.

The results of a band pass filtering on p02 and n01 records in the frequency band 20-30Hz are shown in Figure8 and Figure9 respectively. It is noticed that there is a difference in peaks between BPF outputs for p01 and p02. The peak amplitudes in the n-records are noticed as around 0.1mV whereas in p-records they are in the range of 0.5mV; which is almost 5 times that of n-records.

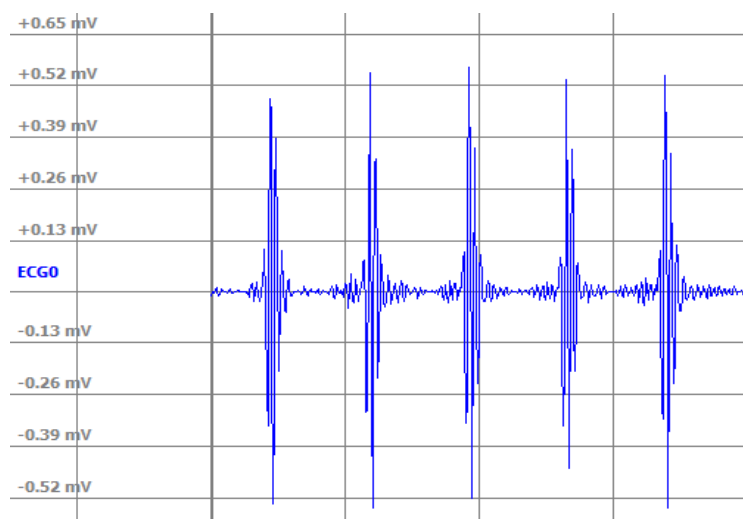


Figure 8: ECG0 of Record p02 after BPF at 20-30Hz

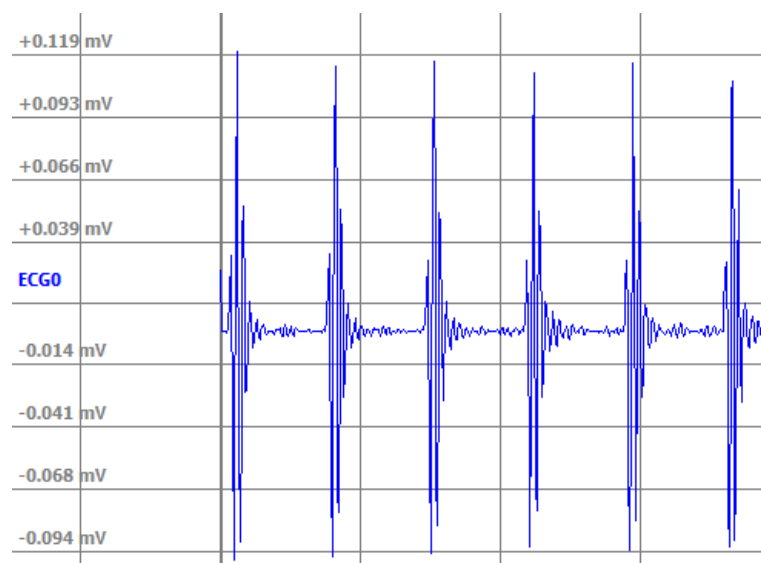


Figure 9: Record n01 after BPF between 20-30Hz

Band Pass Filtering in the Band30-50Hz

Band-pass filtering is done on the 10 subjects of n and p-records in the frequency range of 30-50Hz in order to see how much extra do the PAF components contribute here, in comparison with the band of 20-30Hz. It is planned that, if there is sufficient additional contributions in the band, further split-up studies will have to be conducted for different frequency sub-bands within the range of 20Hz to 50Hz. Figure10 and Figure11 show the outputs of band pass filtering at 30-50Hz on the records p02 and n01 respectively.

As in Figure10, it is noted that, there is a considerable peak amplitude values of p-records, in the band of 30-50Hz (peak amplitude is around 0.38mV) in comparison with that of n records as in Figure12 (peak amplitude around 0.04mV). The p-records also show oscillations in this range as before. These observations can be indicators for the occurrence of PAF. It is also noted that the n-records do not have sufficient signal components in the band of 30-50Hz (peak amplitude around 0.04mV) in comparison with that of 20-30Hz (peak amplitude around 0.11mV) as seen from Figure10 and Figure11.

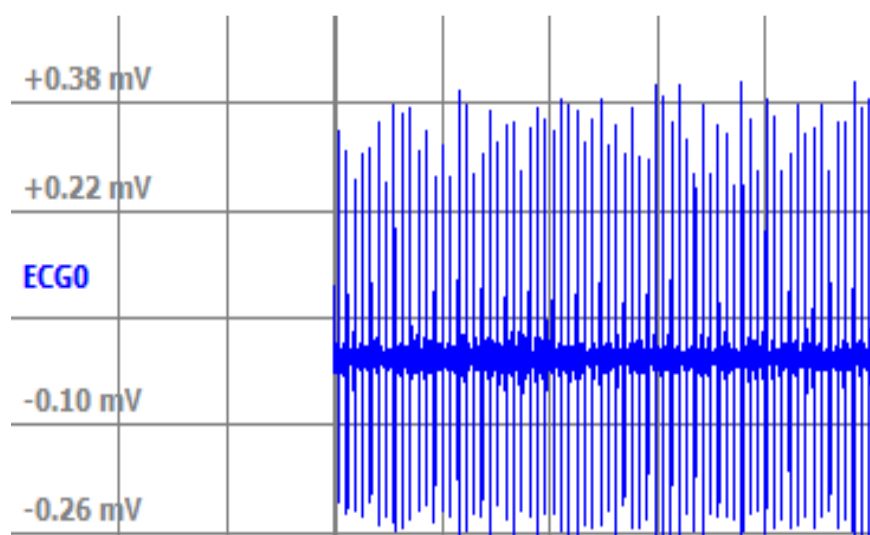


Figure 10: Record p02 after BPF at 30-50Hz

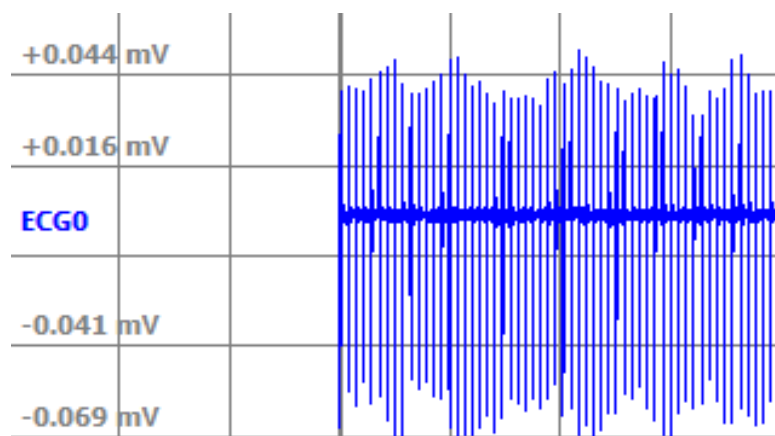


Figure 11: Record n01 after BPF at 30-50Hz

3.2. Spectral Analysis

The FFT of n and p-records of 10 subjects are studied and plotted. In order to concentrate the analysis in the band of 20-50Hz, only this range is plotted. Among most of the n and p-records, it is observed that the low frequency components are displaying higher peaks than the higher harmonics. Also the spectral peaks are found of considerably low in amplitude for n-records in comparison with that of p-records. This indicates that, the presence of PAF in p-records makes the signal spectrum rich in energy so that it is possible to identify or predict the possible occurrence PAF by analyzing the frequency spectrum records in the range of 20-50Hz.

The frequency spectra of p06 and p08 records are shown in Figure12 and Figure13 respectively. The n-records are also analyzed in a similar way as shown in Figure14 and Figure15. Since the n-records do not contain any PAF signal, it can be noticed that the spectral amplitudes are lower than that of p-records.

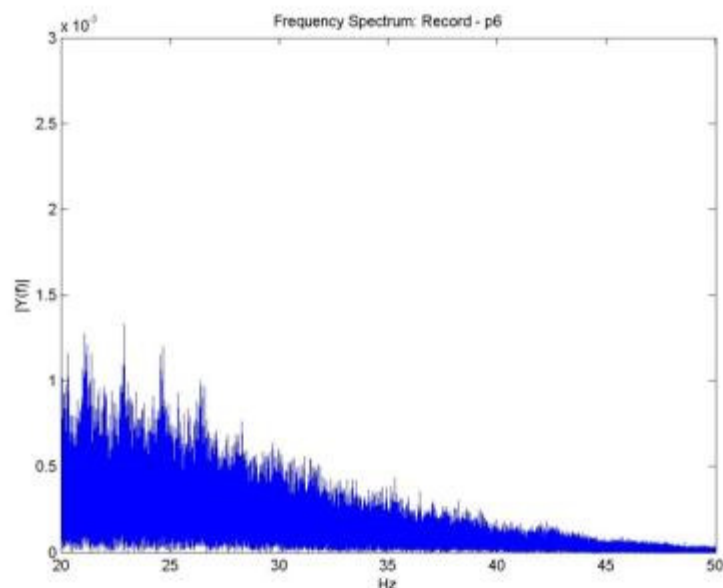


Figure 12: Spectrum(20-50Hz) of Record p06

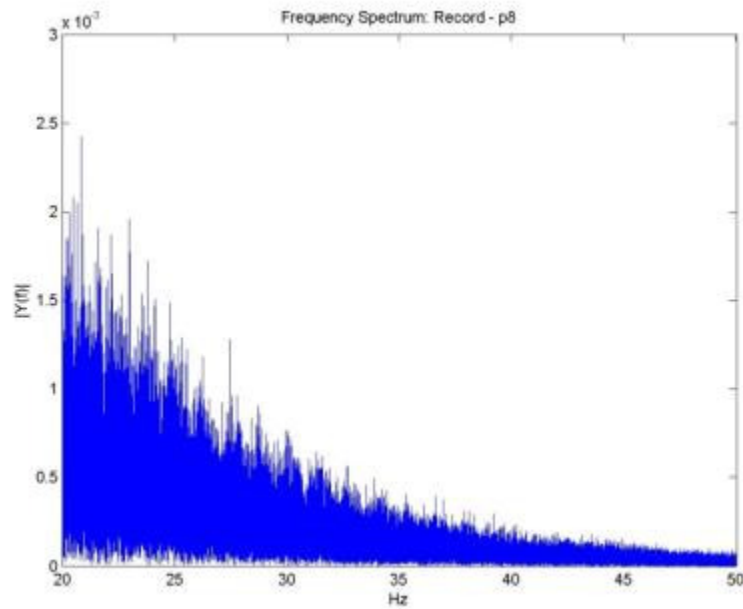


Figure 13: Spectrum(20-50Hz) of Record p08

In the frequency band of 20-35Hz, only 20% of the total n-records which are studied, show reasonable spectral amplitude in comparison with the spectral amplitudes of p-records. That means most of the n-records that we studied have very small spectral amplitude in this frequency range. In a similar analysis in the frequency range of 35-50Hz, almost 90% of the studied n-records show very small spectral amplitudes as in Figure14 and Figure15.

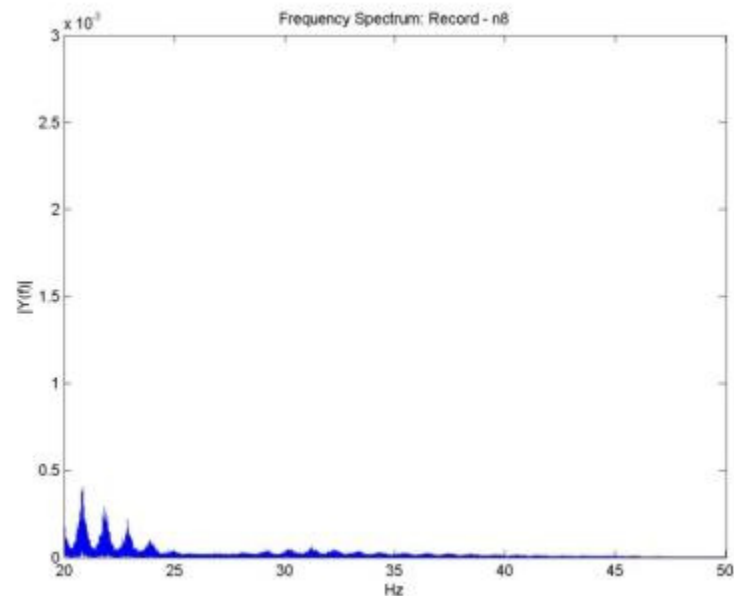


Figure 14: Spectrum(20-50Hz) of Record n08

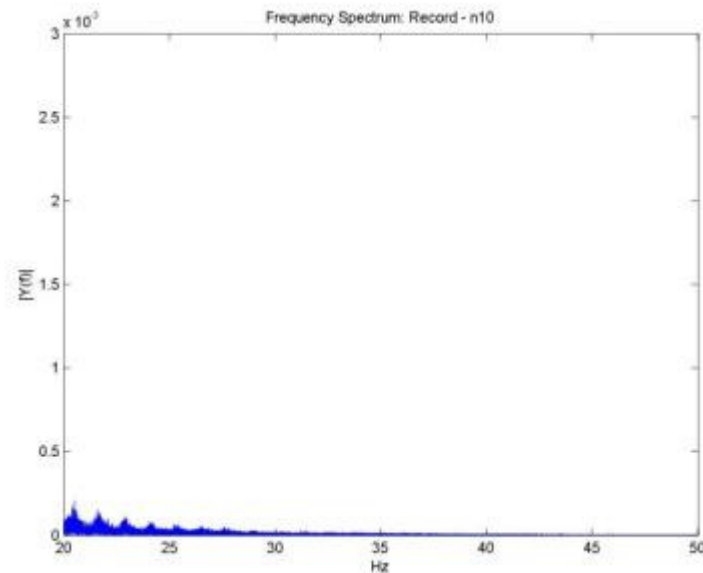


Figure 15: Spectrum(20-50Hz) of Record n10

3.3. Power Spectral Density

Power spectral density is calculated for 10 records (20 signals samples n01-n20 and p01-p20) after band pass filtering between 20-30Hz. Only the even numbered records are described in this paper since they are closer to the occurrence of PAF. Welch spectrum and 'hanning' window are used in the algorithms developed in MATLAB.

As in the Figure16 that shows the PSD plots of even numbered p-records (p02, p04, ..., p20), it is noticed that 100% of PSD peaks for the p-records are found above -50dB. In the case of n-records (n02, n04, ..., n20) as shown in Figure17, 60% of the PSD peaks are found below -50dB. The PSD peak value of -50dB can be taken as a threshold/decision criteria for predicting PAF.

In a similar way the PSD analysis of the n and p records after BPF between 20 - 40Hz shows that, 70% of the n-records are having the peak PSD values below the threshold, while considering -55dB as the threshold value.

After the band pass filtering of the n and p records between 20-50Hz, the decision based on PSD peak values is found to be improved. A higher stop-band ripple parameter (80 decibels) is used here. A threshold of around -65dB is taken. It is noticed that 80% of the peak PSD values are below the threshold for the n-records in comparison with the p-records. PSD analysis of the n and p records after low pass filtering at 20Hz is also carried out as shown in Figure18 and Figure19. A threshold of -30dB taken as the decision criteria result in 90% of prediction about the possible occurrence of PAF in the p records.

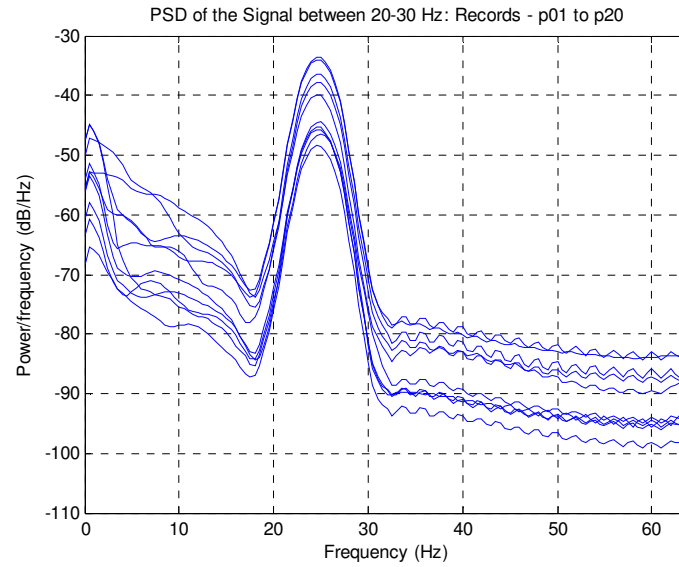


Figure 16: PSD of Even Records from p01-p20

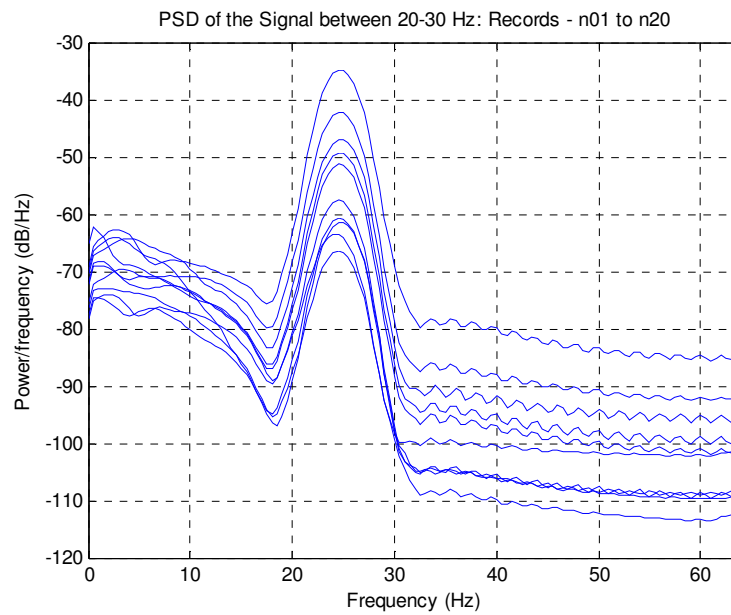


Figure 17: PSD of Even Records from n01-n20

The frequency band of 20-50Hz is found as a reasonably good band in order to analyze and detect the possibility of occurrence of PAF. A comparison of the above results is shown in Table I.

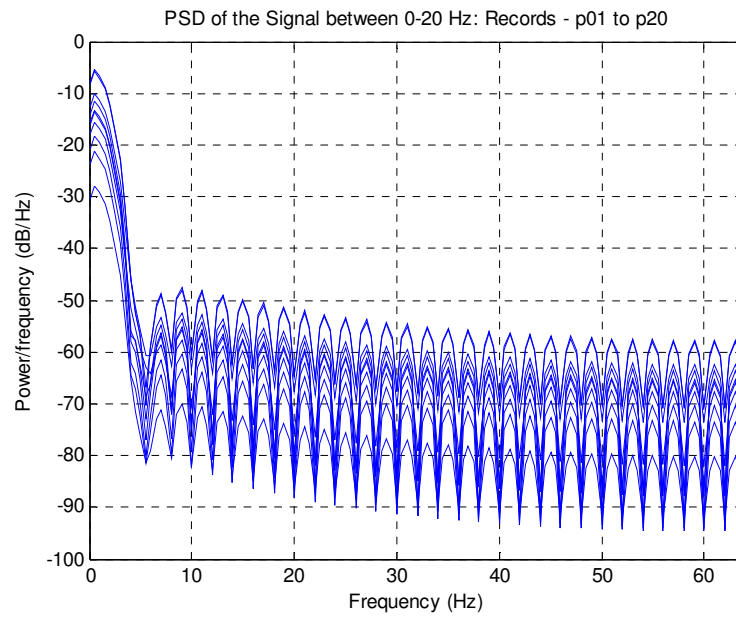


Figure 18: PSD of Even Records from n01-n20

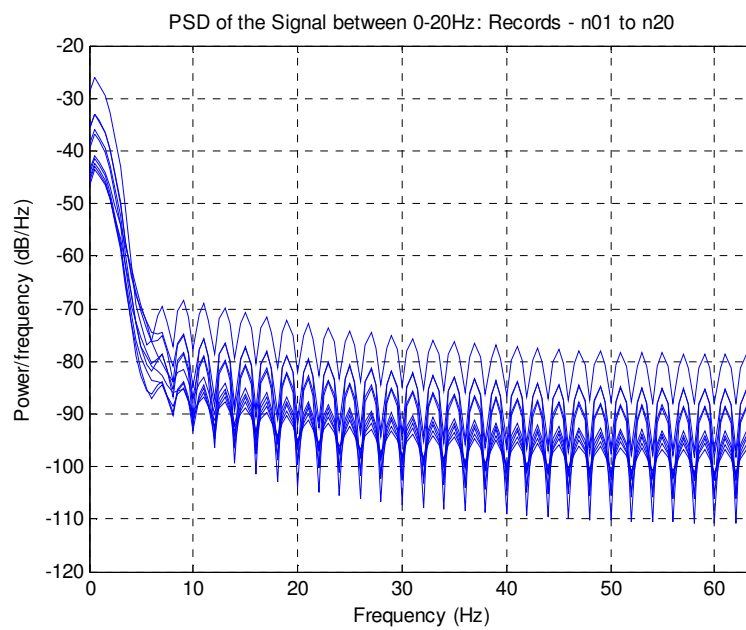


Figure 19: PSD of Even Records from n01-n20

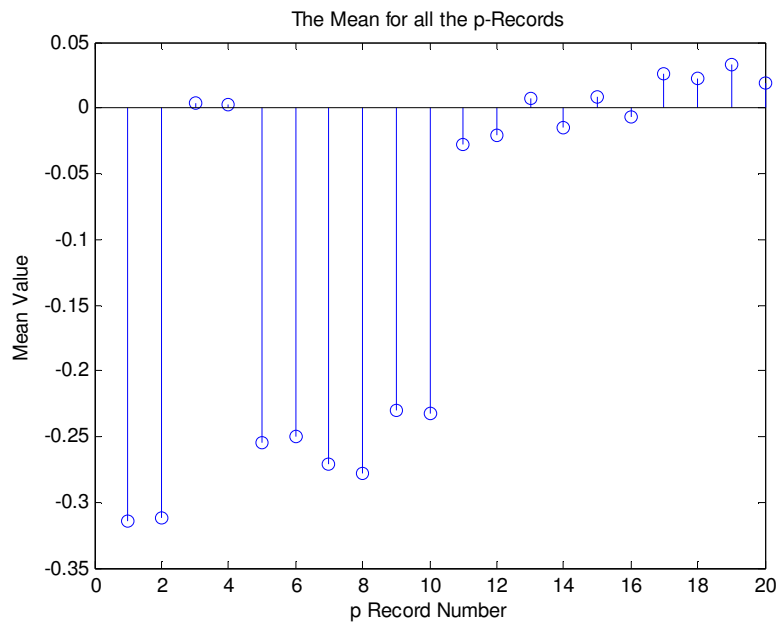
TABLE I: Comparison using different thresholds for different frequency bands

Frequency Band	Threshold PSD Peak	Accuracy of PAF Prediction
20-30Hz	-50dB	60%
20-40Hz	-55dB	70%
20-50Hz	-65dB	80%
0-20Hz	-30dB	90%

222
223

224 3.4. Statistical Methods

225 Statistical methods are also used in the analysis. When the mean signal amplitudes are analyzed for the
226 p-records (p01-p20), it is noticed that most of the p-records show negative values of mean as in Figure20.
227 That means that a negative dc value distinguishes the p-records from the n-records.



228
229

Figure 20: The Mean Values for Records p01 to p20

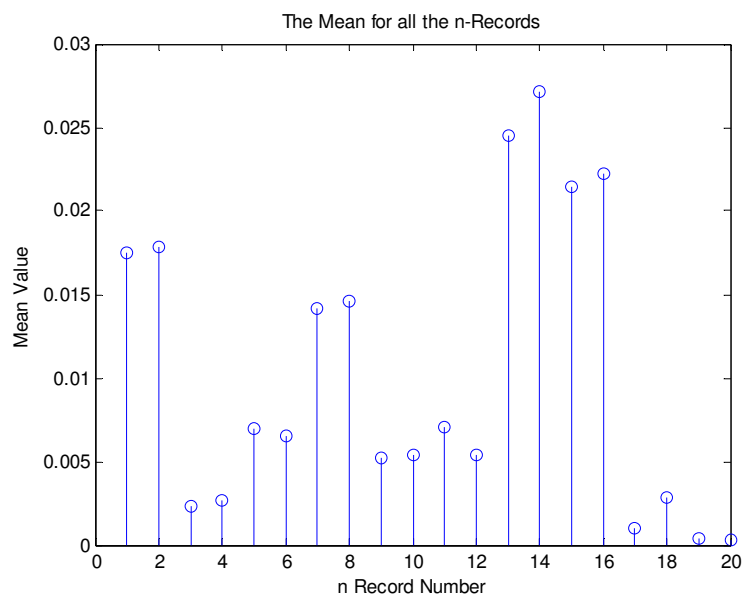
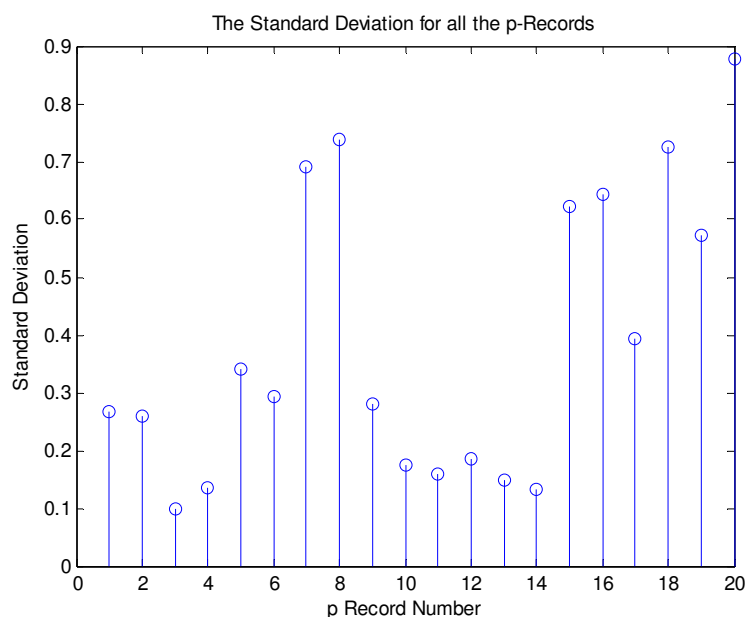


Figure 21: The Mean Values for Records n01 to n20

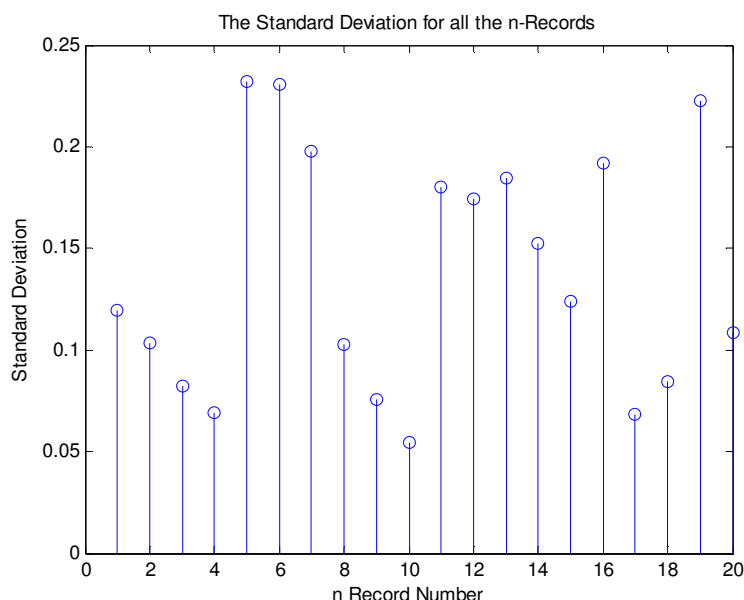
As in Figure21, all the n-records from n01 to n20 show a positive mean value whereas p records show mostly negative values. It is also noted that each pair of signals of a single subject (for example n01 and n02) as in the Figure21, have close mean values, since they are the part of the same subject.

In the next step, the standard deviation is measured and compared for 20 cases of n and p-records as in Figure22 and Figure23. In n-records, the standard deviation is found in the range of values less than 3.5, but for p records, nearly 70% of the cases analyzed show higher values than that of the n-records. This indicates a higher amount of signal variation in p-records in comparison with n-records, which predicts a possible PAF in the vicinity. Out of the 10 signals studied in both n-records and p-records; the average value of signal's standard deviation in n-record is found 0.1480, whereas in p-records it is 0.3280. This is more than 50% of that in n records.



243

Figure 22: Standard Deviation for Records p01 to p20



244

245

Figure 23: Standard Deviation of n-Records

246 A similar result was observed while studying the standard deviation of the p-records after band pass
 247 filtering between 20Hz and 30 Hz. In general it can be stated that, a higher standard deviation and
 248 negative mean values are noticed in p-records in comparison with that of n records. These features are
 249 identified as possible indicators of PAF.

250 *Standard Deviation of Spectral Values*

251 In the next step the standard deviation of spectral values is measured in order to study the variations of
 252 spectral amplitudes during PAF. In p-records the standard deviation of spectral amplitudes are found
 253 higher compared to n-records. It is noted that the average value of the standard deviation (a measure of
 254 variation) for the p-records for 10 subjects (i.e.; 20 records) is around 8.4349×10^{-4} . On the other hand,
 255 the average value of the standard deviation for the n-records for 10 subjects is 2.8874×10^{-4} . The lesser
 256 standard deviation in n-records in comparison with p records implies that the spectral amplitude variations
 257 are prominent in case of PAF. When a threshold of 5×10^{-4} is taken for classification of "PAF" and "non-
 258 PAF" based on the above concept, 70% of the p-samples are correctly classified as having PAF relevant
 259 contents. So, this implies that the variance parameter of spectral amplitudes is also a useful feature in
 260 predicting the possible occurrence of PAF.

261

262

263 **4. CONCLUSION**

264

265 From the above experiments and results, we conclude that, the occurrence of PAF can be predicted to
 266 some extent by analyzing the spectral components, statistical and spectral parameters. As a result of the
 267 time domain analysis and filtering, it is noticed that the band of frequencies of different p-records signals
 268 in the range 20-50Hz shows higher amplitudes and variations, due to the possibility of imminent
 269 occurrence of PAF. Spectral amplitude variations are also found higher in p records that contain PAF in
 270 comparison with the n-records. The presence of PAF in p-records makes the signal spectrum rich in
 271 energy so that it is possible to identify or predict the PAF in advance by analyzing the signal energy and
 272 power spectral density. The variance and standard deviation of the ECG signals prior to the occurrence of

PAF also give valuable information in the prediction process. Even the spectral amplitude variations are useful parameters in predicting the PAF. The average value of signal's standard deviation can be used as a parameter for PAF prediction. It is also noticed that the negative dc values of the spectrum (mean) distinguish the p-records from the n-records, which helps to indicate the possible occurrence of PAF.

REFERENCES

1. Aytemir K, Aksoyek S, Ozer N, Aslamaci S, Oto A. Atrial fibrillation after coronary artery bypass surgery: P wave signal averaged ECG, clinical and angiographic variables in risk assessment. *Int J Cardiol* 1999;69:49-56.
2. <http://www.physionet.org/physiobank/database/afpdb>
3. Budeus M, Felix O, Hennersdorf M, Wieneke H, Erbel R, Sack S. Prediction of Conversion from Paroxysmal to Permanent Atrial Fibrillation. *Pacing Clin Electrophysiol* 2007;30:243–252.
4. Budeus M, Hennersdorf M, Rohlen S et al. Prediction of atrial fibrillation after coronary artery bypass grafting: the role of chemoreflex-sensitivity and P wave signal averaged ECG. *Int J Cardiol* 2006;106:67-74.
5. Buxton AE, Josephson ME. The role of P wave duration as a predictor of postoperative atrial arrhythmias. *Chest* 1981;80:68-73.
6. Caravelli P, De Carlo M, Musumeci G et al. P-wave signal-averaged electrocardiogram predicts atrial fibrillation after coronary artery bypass grafting. *Ann Noninvasive Electrocardiol* 2002;7:198-203.
7. Cabasson, L Dang, JM Vesin, L Kappenberger, R Leber, R Abacherli, Susceptibility to Paroxysmal Atrial Fibrillation: A Study using Sinus Rhythm P Wave Parameters ; 2010, Author manuscript, published in "Computing in Cardiology, Belfast : Ireland (2010)"
8. Chandy J, Nakai T, Lee R, Bellows W, Dzankic S, Leung J. Increases in P-Wave Dispersion Predict Postoperative Atrial Fibrillation After Coronary Artery Bypass Graft Surgery. *Anesth Analg* 2004;98:303–310.
9. Chang CM, Lee SH, Lu MJ et al. The role of P wave in prediction of atrial fibrillation after coronary artery surgery. *Int J Cardiol* 1999;68:303-8.
10. Carlson J, Havmoller 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:S39–S48.
11. Ciaroni S, Cuenoud L, Bloch A. Clinical study to investigate the predictive parameters for the onset of atrial fibrillation in patients with essential hypertension. *American Heart Journal* 2000;139:814–819.
12. Dimmer C, Jordaens L, Gorgov N et al. Analysis of the P wave with signal averaging to assess the risk of atrial fibrillation after coronary artery bypass surgery. *Cardiology* 1998;89:19-24.
13. Holmqvist F, Platonov P, Carlson J, Zareba W, Moss A. Altered interatrial conduction detected in MADIT II patients bound to develop atrial fibrillation. *ANE* 2009;14:268–275.

- 313 14. Havmoller R, Carlson J, Holmqvist F, Olsson B, Platonov P. Evolution of P-Wave Morphology in
314 Healthy Individuals: A 3-Year Follow-Up Study. ANE 2009;14:226–233.
- 315 15. Holmqvist F, Platonov P, McNitt S, Polonsky S, Carlson J, Zareba W, Moss A. Abnormal P-wave
316 morphology is a predictor of atrial fibrillation development and cardiac death in MADIT II patients.
317 ANE 2010;15:63–72.
- 318 16. Magnani J, Mazzini M, Sullivan L, Williamson M, Ellinor P, Benjamin E. P-Wave Indices,
319 Distribution and Quality Control Assessment (from the Framingham Heart Study).ANE
320 2010;15:77–84.
- 321 17. Steinberg J, Zelenkofske S, Wong S, Gelernt M, Sciacca R, Menchavez E. Value of the P-Wave
322 Signal-Averaged ECG for Predicting Atrial Fibrillation After Cardiac
323 Surgery.Circulation1993;88:2618–2622
- 324 18. Stafford PJ, Kolvekar S, Cooper J et al. Signal averaged P wave compared with standard
325 electrocardiography or echocardiography for prediction of atrial fibrillation after coronary bypass
326 grafting. Heart 1997; 77:417-22.
- 327 19. Azeemsha Thacham Poyil, ShadiyaAlingalMeethal, "Cross-term Reduction Using Wigner Hough
328 Transform and Back Estimation", ICICEE2012: 2012 International Conference on Industrial
329 Control and Electronics Engineering", Aug. 2012
330
331
332