Classification of Electrocardiographic P-wave Morphology

Jonas Carlson[1], Rolf Johansson[1,2], S. Bertil Olsson[1]

[1]Department of Cardiology, Lund University, SE-221 85 LUND, Sweden;
Email: Jonas.Carlson@kard.lu.se, Bertil.Olsson@kard.lu.se
[2]Department of Automatic Control, Lund Institute of Technology, Lund University;
PO Box 118, SE-221 00 LUND, Sweden; Email: Rolf.Johansson@control.lth.se

Abstract

The atrial activity of the human heart is normally visible in the ECG as a P-wave. In patients with intermittent atrial fibrillation, a different P-wave morphology can sometimes be seen, indicating atrial conduction defects. The purpose of this study was to develop a method to discriminate between such P-waves and normal ones. 20 recordings of each type were used in a classification which, based on impulse response analysis of the P-wave and linear discrimination between various parameters, produced a correct classification in 37 of the 40 recordings (sensitivity 95%, specificity 90%).

1 Introduction

The normal cardiac rhythm is initiated from the sinus node, a structure with inborn electrical automaticity, situated close to the entrance of the superior caval vein into the right atrium. From there, the electrical impulse propagates along the entire atrial myocardium, thereby creating the electrocardiographic evidence of normal sinus rhythm, a P-wave. When the conduction within an atrium or between the two atria is different from the normal situation, the morphology of the P-wave changes. Recently, we observed that patients prone to attacks of atrial fibrillation, a global atrial re-excitation dysrhythmia, often had evidence of delayed impulse conduction between the right and the left atrium. This was evidenced during direct impulse conduction studies, using cardiac catheters and electrical stimulation technique [1] but also appeared as a distinct change of morphology of the P-wave during sinus rhythm, as compared with those recorded in patients without attacks of atrial fibrillation [2].

The standard ECG is an inexpensive and simple investigation that, with high-resolution acquisition in combination with signal averaging, can provide a lot of information. This was first applied to analysis of the ventricular activity, the QRS-complex, but is now also applied in investigations of the atrial activity, the P-wave. Whereas standard approaches deal with an analysis of the duration of the P-wave, there is an increasing interest for the P-wave morphology [2][3]. Differences in morphology are believed to reflect different activation patterns in the atria. Conduction defects could then be identified by differences in morphology between different patient recordings.

The purpose of the study was to investigate the possibility of using system identification as a tool to classify P-waves as being either normal or showing a different appearance indicating some kind of conduction defect within the atria.

2 Materials and methods

2.1 Acquisition of P-waves

ECG was recorded using modified Frank leads [4], a technique that produces three orthogonal leads, **X**, **Y** and **Z**, where **X** is positive from right to left, **Y** from up to down and **Z** from front to back. Sampling was performed at 1 kHz with 0.625 μV resolution (equipment from Siemens-Elema, Sweden).

The QRS-complexes in the ECG recordings were identified and a 400 ms long 'window' before the onset of the QRS was extracted, believed to hold the P-wave. In each recording, 200 P-waves were used in an averaging in order to reduce noise. A template matching with the criterion of correlation > 0.9 was used to include P-waves in the averaging. The onset and end of the P-wave were defined manually (Fig. 1).

2.2 ECG data

52 P-wave recordings were used in the analysis, all from different individuals. Each recording was inspected by a physician and defined as 'normal' (Type

1) or 'different' (Type 2) according to morphology differences observed in an earlier study [2]. Fig. 1 shows a normal P-wave with two leads (X and Y) being positive, with only one dominant peak, and one (Z) being negative. P-waves of Type 2 differed mostly in leads X and Z, with lead Z having two peaks, one being negative and the other positive. This P-wave morphology has been shown to be common in patients with paroxysmal atrial fibrillation [2]. An example of a P-wave of Type 2 is also seen in Fig. 1.

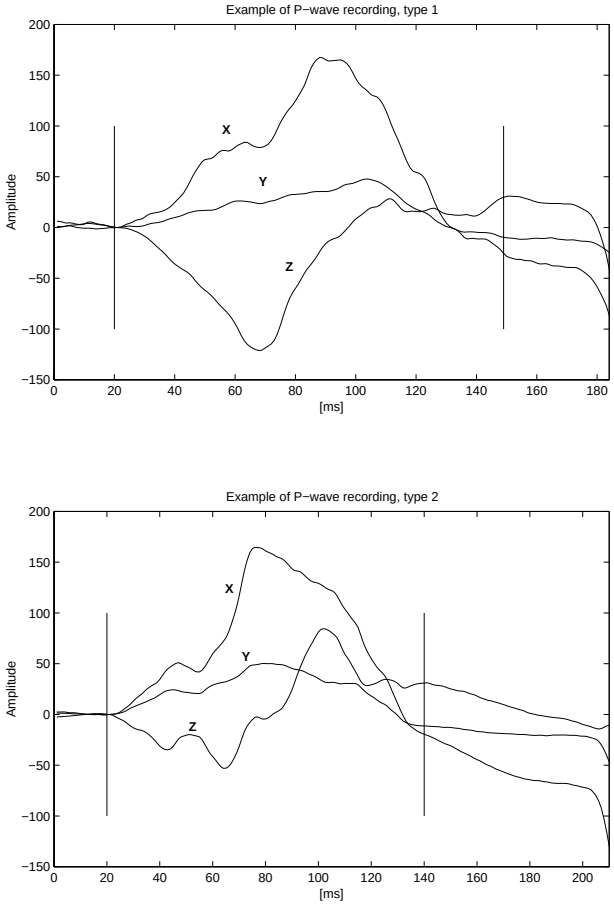

Figure 1: An example of the three orthogonal leads of P-wave recordings of Type 1 and Type 2 with onset and end marked. The deflection at the end is the beginning of the QRS-complex

27 of the recordings were considered to be of Type 1, and 25 of Type 2. The first 20 recordings of each type were used as a 'training set' to optimize the performance of the classifier, and the remaining 12 were used as a test set to evaluate the performance of the classifier.

2.3 System identification

MATLAB (The MathWorks, Inc. USA) for Linux, running on a standard PC, was used as a modeling and classification tool.

The system to be identified, the transfer of electrical impulses from the heart to the outside of the body measured in three orthogonal leads, is a system with three outputs but without an easily measurable input. It has recently been shown that atrial activity measured during sinus rhythm can be very well modeled when treated as an impulse response [5], therefore we chose a similar model for our study, the three leads of the P-wave recording being the impulse response.

Consider a discrete-time system of order n:

$$\begin{cases} x_{k+1} &= A_n x_k + B_n u_k \\ y_k &= C_n x_k + D u_k \end{cases} \quad (1)$$

where x denotes the states, u the impulse input and the output y the three orthogonal leads of a P-wave recording.

For the identification, the state-space model identification algorithm proposed by Ho and Kalman [7] and Juang and Pappa [6] was used to determine the matrices A_n, B_n, C_n and D in the state-space model of Eq. (1).

First, Hankel matrices were constructed as:

$$H_{r,s}^{(k)} = \begin{pmatrix} H_{k+1} & H_{k+2} & \cdots & H_{k+s} \\ H_{k+2} & H_{k+3} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ H_{k+r} & H_{k+r+1} & \cdots & H_{k+r+s-1} \end{pmatrix}, k = 0, 1$$

where r and s are chosen equal, $r+s$ is the length of the P-wave recording and

$$H_0 = \begin{bmatrix} y_{0(\mathbf{X})} \\ y_{0(\mathbf{Y})} \\ y_{0(\mathbf{Z})} \end{bmatrix}, H_1 = \begin{bmatrix} y_{1(\mathbf{X})} \\ y_{1(\mathbf{Y})} \\ y_{1(\mathbf{Z})} \end{bmatrix}, \cdots$$

i.e., the Markov parameters for the two Hankel matrices are the sample values of the P-wave data sequences. Of the 400 ms window, only the part containing the actual P-wave was used in the modeling. To assure stable model dynamics, the interval 20 ms after the defined end of the P-wave was included in the identification data. The baseline was defined as being 0 at the onset of the P-wave in all three leads.

Second, the model order, n, was determined by computing the singular value decomposition $H_{r,s}^{(0)} = U\Sigma V^T$ of the Hankel matrix $H_{r,s}^{(0)}$ and its singular values $\sigma_1, \sigma_2, \ldots, \sigma_s$ (Fig. 2).

The algorithm for an n-th order model with m inputs

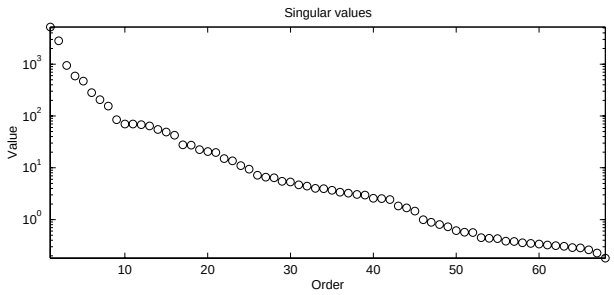

Figure 2: Example of singular values of the Hankel matrix $H_{r,s}^{(0)}$

and p outputs is:

$$\begin{aligned} A_n &= \Sigma_n^{-1/2} U_n^T H_{r,s}^{(1)} V_n \Sigma_n^{-1/2} \\ B_n &= \Sigma_n^{1/2} V_n^T E_u \\ C_n &= E_y^T U_n \Sigma_n^{1/2} \\ D &= H_0 \\ E_y^T &= [I_{p\times p} \;\; 0_{p\times(r-1)p}] \\ E_u^T &= [I_{m\times m} \;\; 0_{m\times(s-1)m}] \\ \Sigma_n &= \mathrm{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_n\} \\ U_n &= \text{matrix of first } n \text{ columns of } U \\ V_n &= \text{matrix of first } n \text{ columns of } V \end{aligned}$$

where, in this case, $m = 1$ and $p = 3$.

From the state-space model of Eq. (1), the discrete impulse responses for the three outputs were simulated and plotted together with the measured impulse responses for manual comparison (Fig. 3).

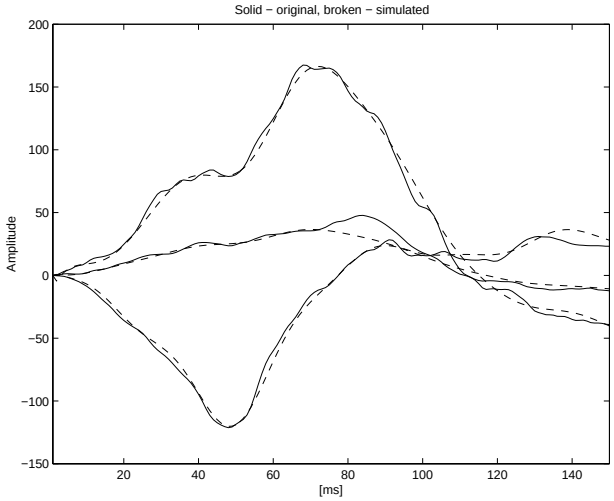

Figure 3: Example of the result of identification with model order 9. Solid lines are original data, dashed are the results of the simulation

The residuals between the modeled and the measured impulse responses were also plotted (Fig. 4) to verify that there were no visible definable components left (e.g. sinusoid-like behavior) indicating that not all system dynamics were accounted for in the model.

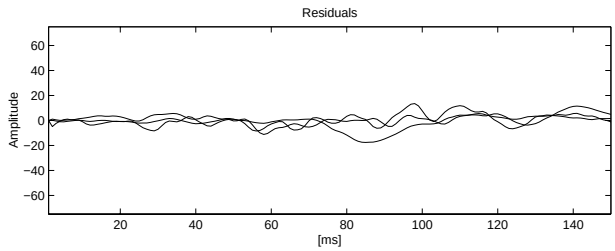

Figure 4: Residuals of the simulation in Fig. 3

2.4 Classification

Three independent methods to classify data were studied: the Fisher linear discriminant [8], the spectrum of the Discrete Fourier Transform (DFT) [8] and the duration of the P-wave.

To use the Fisher linear-discriminant method, a suitable parameter set, θ, must be chosen. In this study, three different choices of parameter sets were evaluated:

- The absolute value of the eigenvalues of the A-matrix
- The coefficients of the denominator polynomial of the transfer function evaluated from the state-space model
- The coefficients of the numerator polynomial of the **Z**-lead transfer function

The last choice was motivated from earlier manual classification of the same data material showing the biggest difference between the two different types of P-waves in the **Z**-lead [2].

The Fisher linear discriminant classifies a parameter estimate, $\hat{\theta}$, as belonging to one of two classes $\mathcal{A}$ or $\mathcal{B}$, in this case P-waves of Type 1 or Type 2. After computing the mean m and covariance R as

$$\begin{aligned} m_i &= \mathcal{E}\{\theta_i\} \\ R_i &= \mathcal{E}\{(\theta_i - m_i)(\theta_i - m_i)^T\} \end{aligned} \quad \text{, where } i = \mathcal{A} \text{ or } \mathcal{B}$$

and the average variance

$$R = \frac{1}{2}(R_{\mathcal{A}} + R_{\mathcal{B}})$$

the Fisher linear discriminant, λ, is calculated as

$$\lambda = R^{-1}(m_{\mathcal{B}} - m_{\mathcal{A}}) \qquad (2)$$

A parameter vector $\hat{\theta}$ can then be classified by calculating

$$\lambda^T \hat{\theta} = (m_{\mathcal{B}} - m_{\mathcal{A}}) R^{-1} \hat{\theta}$$

and performing the classification test

$$\hat{\theta} \in \begin{cases} \mathcal{B} & \text{, if } \lambda^T\hat{\theta} > \vartheta \\ \mathcal{A} & \text{, if } \lambda^T\hat{\theta} < \vartheta \end{cases}$$

for some threshold ϑ, in this study chosen as

$$\vartheta = \frac{(\sqrt{\lambda^T R_{\mathcal{A}}\lambda}\,\lambda^T m_{\mathcal{A}} + \sqrt{\lambda^T R_{\mathcal{B}}\lambda}\,\lambda^T m_{\mathcal{B}})}{\sqrt{\lambda^T R_{\mathcal{A}}\lambda} + \sqrt{\lambda^T R_{\mathcal{B}}\lambda}} \qquad (3)$$

under assumption of normal distribution of $\lambda^T\hat{\theta}$.

Classification was performed with parameters from models of either the best model order as indicated by the singular values or order 5. The latter was chosen in an attempt to reduce unnecessary information and only look at the major differences in appearance. For the same reason, decimation of the data series by a factor of 5 was also evaluated in combination with the higher model order. All three parameter sets were tested with the different model orders. To justify the use of the discriminant in Eq. (3), the parameter estimates $\hat{\theta}$ were investigated by means of a Kolmogorov-Smirnov test [9] to verify they had normal distribution.

As an alternative method, DFT was used to evaluate if the two types of P-waves had differences at any point in the frequency domain. To analyze the spectral data, three plots were made, one for each of the three leads, X, Y, and Z. In each plot, the spectra of all recordings were plotted together to evaluate visually if at any point (i.e., frequency), in any plot, the groups differed. A more formal analysis was performed as well, using the Fisher discriminant described above. The parameter estimates, $\hat{\theta}$, were chosen as a vector of all points of the DFT of the separate leads:

$$\hat{\theta} = \begin{pmatrix} \mathcal{F}_{\mathbf{X}}(1) \\ \vdots \\ \mathcal{F}_{\mathbf{X}}(k) \\ \mathcal{F}_{\mathbf{Y}}(1) \\ \vdots \\ \mathcal{F}_{\mathbf{Y}}(k) \\ \mathcal{F}_{\mathbf{Z}}(1) \\ \vdots \\ \mathcal{F}_{\mathbf{Z}}(k) \end{pmatrix}^T$$

A test of normal distribution of the parameter estimates, $\hat{\theta}$, was made by means of a Kolmogorov-Smirnov test.

Two different spectral resolutions were used: 32-point and 64-point. Since most P-waves have different amplitudes, an attempt to standardize the DFTs was made by normalizing the area under each individual curve to 1 before calculating $\hat{\theta}$.

To compare the results with the standard approach of P-wave discrimination, duration was measured in all recordings, defined manually as the time between the earliest onset and latest end in any of the three leads. Statistical analysis was made to evaluate if there was any difference in P-wave duration between the two groups.

For all analyses, sensitivity (the probability of correctly classifying a P-wave of Type 2) and specificity (the probability of correctly classifying a P-wave of Type 1) of the classification results were calculated.

3 Results

For model orders below $n = 9$, increasing model order showed increasing accuracy in the prediction of the three P-wave leads. Model orders above 9 did not give any further improvement which is in agreement with the information seen in the singular-value plot (Fig. 2).

The Kolmogorov-Smirnov test showed that the null hypothesis, i.e., variables being normally distributed, could not be rejected for any of the chosen parameter estimates at the significance level $p = 0.05$.

Table 1 shows the results of the various combinations of parameter sets and model orders. The best result was achieved with a 9th order model without decimation where 92.5% (37 of 40) of the recordings were classified correctly. Based on this material, the sensitivity and specificity of the method was 95% and 90% respectively.

Table 1: Result of Fisher discriminant classification with various choices of parameters

	Order 9, dec. 5	Order 9	Order 5
Eigenvalues	72.5%	75%	82.5%
Denominator	80%	72.5%	77.5%
Z numerator	87.5%	92.5%	87.5%

The DFT approach did not yield any useful results. No obvious patterns could be seen by visual evaluation of the three leads' spectra in any of the analyses. The Kolmogorov-Smirnov test showed that the null hypothesis, parameter estimates being normally distributed, could not be rejected at the significance level $p = 0.05$. The Fisher analysis showed that the DFT parameters failed to produce an acceptable classification, see Table 2.

Table 2: Result of Fisher discriminant classification with various choices of DFT parameters

	Sensitivity	Specificity
32-point DFT	0.45	0.55
64-point DFT	0.45	0.50
32-point DFT, area = 1	0.55	0.35
64-point DFT, area = 1	0.40	0.60

Also P-wave duration failed to discriminate the two groups. The duration was (mean±SD) 117±15 ms and 124±12 ms respectively, the difference not being significant ($p = 0.19$, Mann-Whitney U-test). No single value of P-wave duration produced a discrimination comparable to that of the Fisher linear discriminant. Using the mean of the duration of the two groups (120.5 ms) produced both sensitivity and specificity of 55%. A plot of sensitivity and specificity versus P-wave duration can be seen in Fig. 5.

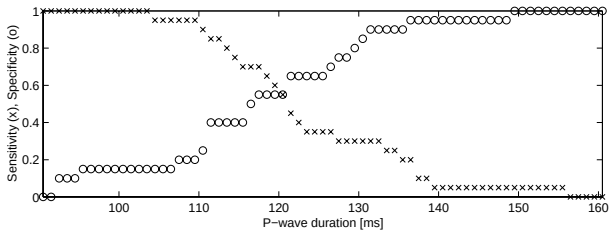

Figure 5: Sensitivity and specificity plotted as functions of the P-wave duration

4 Discussion

The idea of using the signal-averaged P-wave to discriminate between patients with different diseases is not new. For more than a decade, different groups have tried to find easily definable characteristics that can be used [10]. The one most often used is the duration of the P-wave, or its vector magnitude, $\sqrt{\mathbf{X}^2 + \mathbf{Y}^2 + \mathbf{Z}^2}$ [11]. It has, for example, been used in attempts to predict atrial fibrillation after cardiac surgery [12] or recurrence of atrial fibrillation after electrical cardioversion [13]. Another approach is to compute the P-wave duration in each lead, and calculate a "P-wave dispersion index", defined as 100×(SD of duration/mean duration). This has, in combination with standard P-wave duration measurements, been shown to produce better results than using the duration alone [14]. Another characteristic is the RMS amplitude of the last part of the bandpass-filtered (40-250 Hz) P-wave, a value denoted as RMSnn, where nn denotes the nn last milliseconds of the P-wave [13]. There has also been attempts to analyze the frequency content of the P-wave [15].

In our study, both P-wave duration and frequency content analysis failed to discriminate the two types of P-waves while the system identification was successful in 92.5% of the cases. This shows that a more advanced analysis can identify differences which are not easily defined in intuitive and visible characteristics such as amplitude or duration.

Due to the small material used, most recordings were used in the calculation of the Fisher linear discriminant and a prospective study is needed for purposes of validation. No results are shown on the analyses of the 7+5 remaining recordings since the result of one individual analysis changes the overall result too much. In general, however, the results were above 80% for the **Z**-lead denominator polynomial analyses.

The two types of P-waves had the most visible morphology differences in the **Z**-lead. This is also shown in the result as the best discrimination being found in parameters only incorporating information from this lead. These different forms of P-waves represent the presence or absence of an underlying pathophysiological condition in patients prone to attacks of atrial fibrillation. Interestingly, only few individuals without this arrhythmia have P-waves of Type 2, indicating that improper impulse conduction between the atria may be a basic prerequisite for the disease in some patients [2]. An interesting application of a tool capable to discriminate between the two types of P-waves would therefore be a prospective study to investigate the difference in occurrence of atrial conduction defects in patients with paroxysmal atrial fibrillation compared to normal subjects.

5 Acknowledgments

The authors wish to thank Dr Pyotr Platonov for providing the ECG-data used in the study.

The study was supported by grants provided by the Swedish Heart-Lung Foundation.

References

[1] P.G. Platonov, S. Yuan, E. Hertervig, O. Kongstad, L.V. Chireikin, and S.B. Olsson, "Presence of right atrial conduction disturbance in patients with lone atrial fibrillation," *Eur Heart J*, vol. 19 (Abstr suppl), 1998.

[2] P.G. Platonov, J. Carlson, M.P. Ingemansson, A. Roijer, A. Hansson, L.V. Chireikin, and S.B. Olsson, "Detection of inter-atrial conduction defects with unfiltered signal-averaged P-wave ECG in patients with

lone atrial fibrillation," *Europace*, vol. 2, pp. 32–41, January 2000.

[3] M.P. Ingemansson, J. Carlson, P.G. Platonov, and S.B. Olsson, "Effects of $MgSO_4$ and glucose, insulin and potassium (GIK) on atrial conduction during the first 12 hours after DC-conversion of chronic atrial fibrillation," in *Cellular electrophysiological modulation in chronic atrial fibrillation – Studies with magnesium and GIK solution*, chapter IV. Lund University Press, 1998.

[4] E. Frank, "An accurate, clinically practical system for spatial vectorcardiography," *Circulation*, vol. 13, pp. 737, 1956.

[5] R. Johansson, M. Holm, S.B. Olsson, and J. Brandt, "System identification of atrial activation during chronic atrial fibrillation in man," *IEEE Trans. Automat. Contr*, vol. 43, no. 6, pp. 790–799, June 1998.

[6] J.N. Juang and R.S. Pappa, "An eigensystem realization algorithm for model parameter identification and model reduction," *Journal of Guidance, Control and Dynamics*, vol. 8, pp. 620–627, 1985.

[7] B. L. Ho and R. E. Kalman, "Effective construction of linear state-variable models from input/output functions," *Regelungstechnik*, vol. 14, pp. 545–548, 1966.

[8] R. Johansson, *System modeling and identification*, Prentice Hall, Englewood Cliffs, NJ, 1993.

[9] Robert R. Sokal and F. James Rohlf, *Biometry – 3d ed.*, W. H. Freeman and Company, New York, 1995.

[10] S. Rosenheck, "Signal-averaged P wave in patients with paroxysmal atrial fibrillation," *Pacing Clin. Electrophysiol.*, vol. 20, pp. 2577–2586, 1997.

[11] M. Hofmann, L. Goedel-Meinen, A. Beckhoff, and A. Schomig, "Analysis of the P wave in the signal-averaged electrocardiogram: normal values and reproducibility," *Pacing Clin. Electrophysiol.*, vol. 19(11), pp. 1928–1932, November 1996.

[12] J.S. Steinberg, S. Zelenofske, S-C. Wong, M. Gelernt, R: Sciacca, and E. Menchavez, "Value of the P-wave signal-averaged ECG for predicting atrial fibrillation after cardiac surgery," *Circulation*, vol. 88(6), pp. 2618–2622, 1993.

[13] G. Opolski, P. Ścisło, J. Stanisławska, A. Górecki, R. Steckiewicz, and A. Torbicki, "Detection of patients at risk for recurrence of atrial fibrillation after successful electrical cardioversion by signal-averaged P-wave ECG," *Int J of Cardiol*, vol. 60, pp. 181–185, 1997.

[14] G.Q. Villani, M. Piepoli, A. Rosi, and A. Capucci, "P-wave dispersion index: a marker of patients with paroxysmal atrial fibrillation," *Int J of Cardiol*, vol. 55, pp. 169–175, 1996.

[15] T. Yamada, M. Fukunami, M. Ohmori, K. Kumagai, A. Sakai, N. Kondoh, T. Minamino, and N. Hoki, "Characteristics of frequency content of atrial signal-averaged electrocardiograms during sinus rhythm in patients with paroxysmal atrial fibrillation," *J Am Coll Cardiol*, vol. 19(3), pp. 559–563, 1992.