

A Principal Component Analysis Approach to Heart Rate Turbulence Assessment in Chagas Disease

Alex C. Alberto¹, Gabriel A. Limeira², Jurandir Nadal²

¹Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rio de Janeiro, Brazil

²Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Abstract

The analysis of heart rate turbulence (HRT) is a powerful method to estimate the baroreflex from the 24 h Holter ECG signals, by considering that an isolated premature ventricular contraction (PVC) causes an immediate cardiac acceleration followed by a deceleration in normal subjects. This study aims at developing a method for risk stratification of sudden death in chronic Chagas cardiomyopathy, by applying Principal Component Analysis (PCA) to the averaged tachogram segments extracted for HRT analysis. HRT analysis was applied to a database of high resolution ECG from Chagas disease patients, with 10 min signals in three leads, sampled with 16-bit resolution at 1000 Hz. From a set of 115 records that presented premature ventricular contractions (PVC), it was possible to extract at least one valid tachogram for HRT analysis in just 51 signals. The valid segments from each ECG record were taken to compute a coherent mean, used them for measuring the parameters turbulence onset (TO) and turbulence slope (TS). From this dataset, two groups of eight signals were extract, according to the estimated risk of sudden death: high risk ($TO \geq 0$ and $TS \leq 2,5$ ms/RR interval) and low risk ($TO > 0$ and $TS > 2,5$ ms/RR interval). PCA was thus applied to this 16 coherent means of 19 samples to reduce data representation to three principal components (PC), which represented 99.5% of the original variance. Applied to the respective PC scores, a logistic regression allowed the separation of groups with 94% accuracy, 88% sensibility and 100% specificity. As a conclusion, PCA has a potential for baroreflex assessment throughout HRT in Chagas disease, but this method should be validated with a larger sample with long duration ECG.

1. Introduction

Heart rate variability (HRV) has been widely investigated as a method to assess the heart control by the autonomic nervous system (ANS). The HRV analysis consists on the study of time series of consecutive inter-beat intervals (R-R). At rest, these series shows slow fluctuations that reflect the sympathetic and parasympathetic regulation of the sinus node [1].

HRV based indexes have been helpful in different clinical applications, such as in risk stratification of sudden death after myocardial infarction, in the prognosis of diabetes and in the monitoring of heart transplanted patients. However, few studies focused on Chagas disease (American Trypanosomiasis). This parasitic disease, caused by *Trypanosoma cruzi*, is endemic in Latin America with nearly 20 million infected subjects [2], being a serious and alarming public health problem. Due to immigration, it is estimated that more than 300,000 infection persons live in the United States [3]. The parasite infects the heart tissues and causes an immune response, which degenerates myocardial contractile cells and nerve endings. Consequently, it is associated to serious rhythm disturbances such as ventricular fibrillation and tachycardia, with increased risk of sudden death [4]. Particularly, the infection affects the beta autonomic and muscarinic receptors in different phases of the disease, contributing to increase this risk [5].

Schmidt et al. [6] described the biological phenomenon called heart rate turbulence (HRT). Unlike conventional HRV analysis, which removes the ectopic beats, the HRT analysis is focused on the changes of RR intervals due to isolated premature ventricular contractions (PVC). After a PVC there is an immediate cardiac acceleration followed by a progressive deceleration in subjects with low risk of sudden death. The HRT's mechanisms were not fully identified, and has been attributed to a baroreflex response to hemodynamic changes caused by the PVC. According to Watanabe et al. [7], the HRT trigger is a rapid blood pressure disturbance caused by PVC (Figure 1). The premature beat has a lower stroke volume and causes an instantaneous decrease in blood pressure, while the following beat occurring after a compensatory pause has a higher stroke volume, causing the opposite effect. If ANS is intact, the HRT corresponds to its immediate response in the sinus node. However, in impaired autonomic heart rate control the reaction is weak or does not occur.

The HRT analysis guidelines [8] proposes the use of two parameters: the turbulence onset (TO), that measures the immediate increase in heart rate; and the turbulence slope (TS), that measures the compensatory and progressive decrease in heart rate.

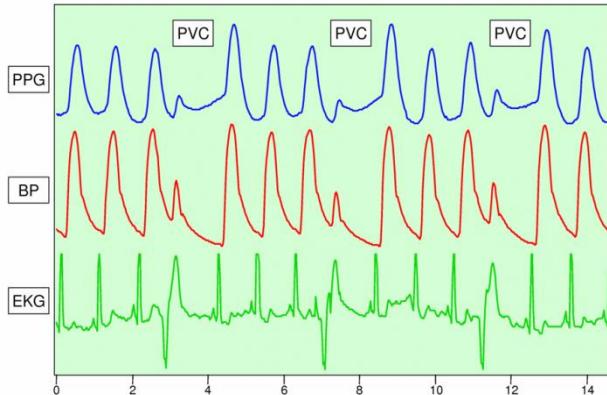


Figure 1 – Changes in blood pressure (BP) caused by PVCs (PPG = finger plethysmography), (Source: <https://en.wikipedia.org/wiki/Photoplethysmogram>)

TO and *TS* parameters are powerful indicators of sudden death risk, particularly when combined to other established parameters. According to some studies [9,10], *TS* shows to be an independent mortality predictor for subjects after myocardium infarction, however loses prognostic power in elderly [8]. Additionally, the occurrence of abnormal *TO* and *TS* is a more powerful risk stratifier than severe cardiac insufficiency (ejection fraction lower than 30%) [8].

Although proposed and validated for the assessment of patients after acute myocardial infarction, some studies are indicating that such parameters are also useful to risk stratification in other cardiac diseases [11,12] including Chagas cardiomyopathy [13].

This study aims at developing a method for risk stratification of sudden death in chronic Chagas cardiomyopathy, by applying Principal Component Analysis (PCA) to the tachogram segments extracted for HRT analysis. The long-term reasoning of this study is to produce an index that estimates risk even when just *TO* or *TS* are positive.

2. Material and methods

2.1. The data bank

It was used a CD-ROM database developed by the international cooperation “Proyecto SEARCH” (*Señales Eletrocardiográficas de Alta Resolución em Chagas*) between two Venezuelan Groups (Universidad Simón Bolívar and Universidad Central de Venezuela) and the University of Oklahoma (USA). This database contains 10 min records of high resolution ECG signals (three leads, 1000 Hz sample rate) of 19 control subjects and 164 patients with Chagas disease with and without cardiomyopathy [2]. Such signals are classified in a scale of disease evolution according to four different tests: serology, resting and Holter ECG, and echocardiography

(Table 1). The diagnosis of Chagas disease was based on Machado-Guerrero serology, in accordance with the guidelines of the World Health Organization [14].

Table 1. Rating groups and clinical data

Group	Serology	Resting ECG	Echocardiogram	Holter ECG
0 (Control)	-	Normal	Normal	Normal
1	+	Normal	Normal	Normal
2A	+	Abnormal	Normal	PVC
2B	+	Abnormal	Normal	Normal
3A	+	Abnormal	Normal	PVC
3B	+	Abnormal	Decreased ejection fraction	PVC
4A	+	Abnormal	Normal	Ventricular tachycardia
4B	+	Abnormal	Decreased ejection fraction	Ventricular tachycardia
unknown	+	Abnormal	Without review	Without review

2.2. Heart Rate Turbulence

The HRT analysis was performed with the original program developed by Schmidt et al. [6]. Each accepted tachogram (Figure 2) consisted of 19 RR intervals comprising a short one related to a PVC, preceded by two normal intervals, and the respective compensatory pause followed by 15 normal intervals. The several events from the same patient record were lined up with the PVC interval as the third one to produce coherent averaged segments that were used to measure *TO* and *TS* parameters.

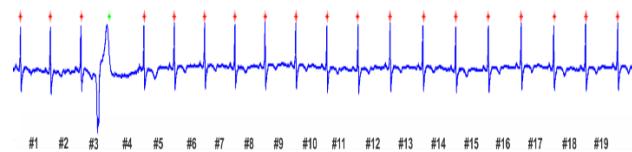


Figure 2 – ECG segment with the minimum requirement to produce an acceptable tachogram for HRT analysis.

TO was measured as the ratio between the two intervals just after and the two ones just before the PVC:

$$TO = \frac{(RR_1+RR_2)-(RR_{-1}+RR_{-2})}{((RR_{-1}+RR_{-2}))} \quad (1)$$

The *TS* parameter is retrieved from the maximum slope of the regression line obtained over any five consecutive normal RR intervals within the 15 intervals after the PVC [9].

HRT is considered as normal if both $TO < 0\%$ and $TS > 2.5 \text{ ms}/\text{beat}$.

2.3. PCA

The principal components analysis (PCA) is a multivariate statistical technique used to reduce the dimensionality of a data set where there are a large number of correlated variables. This reduction occurs by obtaining a new and reduced set of orthonormal variables, not correlated, called principal components (PC), which better preserves the original data variance.

From the 16 tachograms, 19 samples each, is obtained the covariance matrix C (19 x 19), where each element is given by

$$c_{i,j} = \frac{1}{N-1} \sum_{k=1}^N (x_{i,k} - \bar{x}_i)(x_{j,k} - \bar{x}_j) \quad \{i,j = 1, 2, \dots, M\} \quad (2)$$

where \bar{x} is the ensemble mean, $N = 16$ and $M = 19$.

The PC are obtained by the solutions of

$$Cw_p = k_p w_p \quad \{p = 1, 2, \dots, P\} \quad (3)$$

where $P \ll M$, w_p is the p -th eigenvector (or PC), and k_p is the corresponding eigenvalue. Each eigenvalue has a value proportional to the fraction of total variance represented by its CP. The eigenvectors concentrate as much as possible of the variance of the signals in the first CPs and form an orthonormal basis. The coefficient of a given tachogram is obtained by

$$a_{q,p} = \sum_{i=1}^M x_{q,i} w_{i,p} \quad \{q = 1, 2, \dots, Q; p = 1, 2, \dots, P\} \quad (4)$$

where Q is the total number of available tachograms and P is the number of CP, determined according the cumulative explained variance [15]. After the calculation of the coefficients, each tachogram can be reconstructed by the inverse transform

$$\hat{x}_{q,i} = \bar{x}_i + \sum_{p=1}^P a_{q,p} w_{p,i} \quad (5)$$

where $\hat{x}_{q,i}$ is an approximation of $x_{q,i}$.

After determining the number of CP, the separation between the groups was performed according to Muniz *et al.* [16]. Initially, it was determined the average values of all coefficients of the low-risk group, named centroid, and then calculated the standard distance (root of the Mahalanobis distance) between each case and this centroid. To this set of distance values it was applied the logistic regression, in order to separate the groups. Classification results were assessed in terms of total accuracy, sensitivity and specificity.

3. RESULTS

Not all the signals in the database were useful to the analysis due to the need of at least one isolated PVC

occurrence. This does not occur in the clinical classes 0, 1 and 2B. In addition to this fact, the indexes calculated by HRT were not obtained on all remaining signals. Considering the three channels from each subject in separate, it was obtained 51 valid tachograms, relating to 25 subjects. These segments were divided into low and high risk patients using the methodology described by Schmidt *et al.* [6]. After excluding the data with just one parameter (*TO* or *TS*) abnormal, it left just 35 HRT segments, with 27 in the low-risk group and eight in the high-risk group. For the present assessment, it was used just 8 of the 27 HRT segments from the first group, taken at random, so as to obtain two homogenous groups.

The cumulative explained variance according to eigenvalues is shown in Figure 3. Based on the scree criteria [15], only three PC would be enough for analysis. To be conservative, it was analyzed four PC whose eigenvectors are shown in Figure 4. These PC represented 99.5% of the total variance of the original data. The comparison of each PC coefficients between groups (Table 2) showed that only the second and forth PC have significant difference between groups ($p < 0.05$). The first eigenvector (Figure 4A) presented an almost uniform distribution of loading factors along the vector with little fluctuation around the PVC. This suggests that the large fraction of variance represented by this PC (97.3%) is due to differences in average heart rate of the subjects, and not the HRT caused by PVC. Thus, two distinct classifications were tested, each with three PCs: from the first to the third and from the second to the fourth. In both cases, classification results were exactly the same (Figure 5), with just one abnormal case (high risk) being classified as normal (low risk).

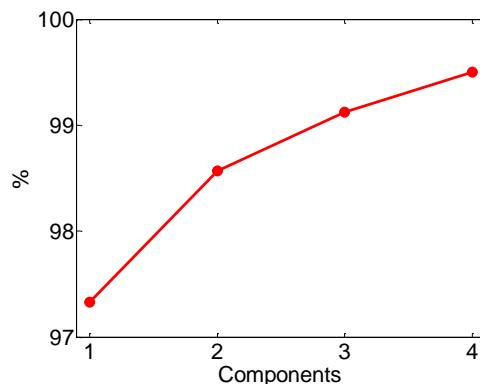


Figure 3. Cumulative variance explained by PCA.

Table 2 – Comparison of PC coefficients between groups

PC#	Low-Risk	High-Risk	P (t-test)
1	-151.20 ± 191.04	151.20 ± 670.14	0.2399
2	67.00 ± 84.70	-67.00 ± 85.79	0.0072
3	1.06 ± 88.82	-1.06 ± 50.19	0.9538
4	-28.28 ± 55.80	28.28 ± 22.19	0.0185

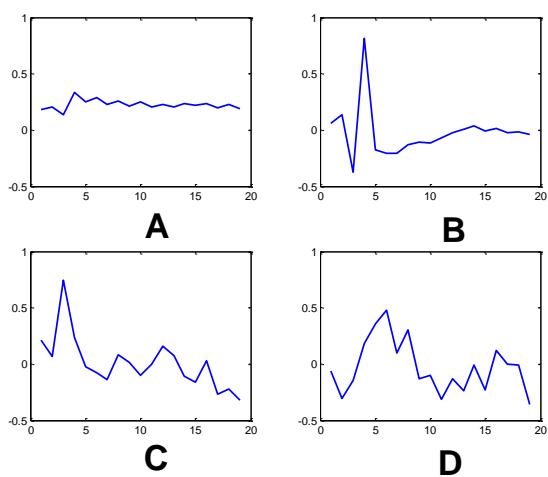


Figure 4 – Analyzed eigenvectors.

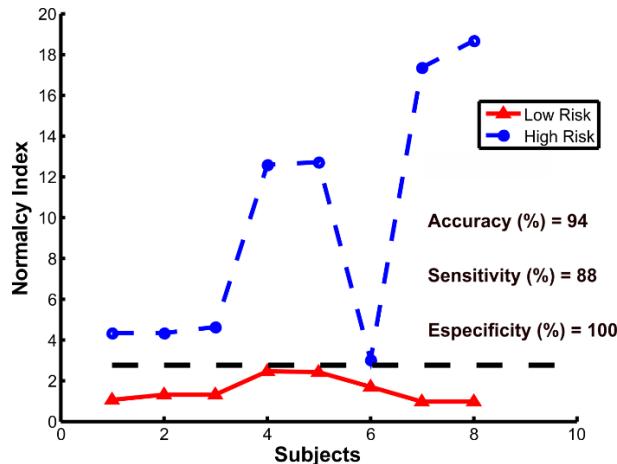


Figure 5. Classification results using the first three PC.

4. Discussion

The use of three PC allowed separating the cases of low and high risk, according to *TO* and *TS* in a reduced data set. Apparently, the second PC was the most sensitive to differences between groups. The eigenvector loading factors enhances the compensatory pause and shows a shape similar to TRC, with an immediate heart rate acceleration followed by a progressive deceleration [6]. The standard distance may also be used as a risk score or normalcy index, being applicable also to cases that have just one abnormal parameter (*TO* or *TS*). However, the absence of a gold standard as well as the reduced data set with short epochs of ECG may be viewed as study limitations, which not allows relating the occurrence of abnormal HRT with the clinical stage of the disease.

As a conclusion, the PCA appears to be suitable for enhance the HRT analysis, which has to be confirmed with an increased dataset with longer records.

Acknowledgements

This work received financial support from the Brazilian agencies CAPES, CNPq and FINEP.

References

- [1] Winchell RJ, Hoyt DB. Spectral analysis of heart rate variability in the ICU: A measure of autonomic function. *Journal of Surgical Research* 1996;63:11-16.
- [2] Grupo de Bioingeniería y Biofísica Aplicada – Universidad Simón Bolívar. Proyecto SEARCH – Base de Datos de Registros de Alta Resolución. CD-ROM, 1996:april.
- [3] Centers for Disease Control and Prevention. Chagas Disease. Available on line in CDC website accessed in 28/08/2015: <http://www.cdc.gov/parasites/chagas/epi.html>.
- [4] Hagar JM, Rahimtoola SH. Chagas' heart disease. *Curr. Probl. Cardiology* 1985;20:825-924.
- [5] Oliveira SF et al. Sera From Chronic Chagasic Patients With Complex Cardiac Arrhythmias Depress Electrogenesis and Conduction in Isolated Rabbit Hearts. *Circulation* 2007;96:2031-2037
- [6] Schmidt G, Malik M, et al. Heart-rate turbulence after ventricular premature beats as a predictor of mortality after acute myocardial infarction. *Lancet* 1999;353:1390-1396.
- [7] Watanabe, MA. et al. Effects of Ventricular Premature Stimulus Coupling Interval on Blood Pressure and Heart Rate Turbulence. *Circulation* 2002;106:325-330.
- [8] Bauer, A, Malik M, et al. Heart Rate Turbulence: Standards of Measurement, Physiological Interpretation, and Clinical Use. *Journal of the American College of Cardiology* 2008;52:1353-1365
- [9] Watanabe MA, Schmidt G. Heart rate turbulence: A 5-year review. *Heart Rhythm* 2004;1:732-738.
- [10] Klingenheben T, et al. Heart rate turbulence and other autonomic risk markers for arrhythmia risk stratification in dilated cardiomyopathy. *Journal of Electrocardiology* 2008;41:306-311.
- [11] Savelieva I, Wichterle D, et al. Heart Rate Turbulence After Atrial and Ventricular Premature Beats – Relation to Left Ventricular Function and Coupling Interval. *PACE* 2003;26:401-405.
- [12] Grimm W, Sharkova J, et al. Prognostic Significance of Heart Rate Turbulence Following Ventricular Premature Beats in Patients with Idiopathic Dilated Cardiomyopathy. *J Cardiovasc Electrophysiol* 2003;14:819-824.
- [13] Ribeiro ALP, et al. Heart Rate Turbulence in Chagas Disease, *PACE* 2003;26:406–410.
- [14] World Health Organization. Control of Chagas' Disease. Technical Report Series 811. Washington DC, 1991;1-95.
- [15] Jolliffe IT. Principal Component Analysis. Berlin:Springer. 2002
- [16] Muniz, AMS, Nadal J, Application of principal component analysis in vertical ground reaction force to discriminate normal and abnormal gait. *Gait & Posture*. 2008;29:31-35

Address for correspondence:

Jurandir Nadal
PO Box 68510; Zip code 21941-972; Rio de Janeiro, RJ, Brasil.
E-mail: jn@peb.ufrj.br