Use of Approximation Entropy for Stratification of Risk in Patients With Chagas Disease

Conference Paper - December 2018
DOI: 10.22489/CinC.2018.234

2 authors, including:

Miguel Vizcardo
National University of St Agustin
4 PUBLICATIONS 0 CITATIONS

Some of the authors of this publication are also working on these related projects:

Project Analysis of dysautonomia in patients with Chagas Cardiomyopathy View project
Use of Approximation Entropy for Stratification of Risk in Patients With Chagas Disease

M Vizcardo†, A Ravelo‡
†Departamento de Física, Universidad Nacional San Agustín, Arequipa, Perú
‡Instituto for Technological Development and Innovation in Communications, Universidad de Las Palmas de Gran Canaria, Spain

Abstract

According to the World Health Organization, the number of people infected with Trypanosoma Cruzi is estimated between 6 and 7 million, the causative agent of Chagas disease, and in 550,000 people exposed to the risk of affection. The approximate entropy was used to quantify the regularity of the tachograms of patients with Chagas disease. The study population consisted of three groups of volunteers: 92 controls (C), 102 patients with positive serology without cardiac involvement diagnosed by conventional non-invasive methods (CH1) and 107 patients with positive serology and mild to moderate incipient heart failure (CH2). We analyzed RR segments of 5 minutes, 288 segments, corresponding to 24 hours per patient. We found significant differences between the Control and CH2 groups, which is used to stratify risk in the CH1 group.

1. Introduction

Chagas disease is caused by a flagellated parasite: Trypanosoma cruzi, transmitted by an insect of the genus Triatoma and also by blood transfusions. In Latin America the number of people infected is approximately 6 million, with a population exposed to the risk of infection of 568,000[1]. In 40% of the population infected with Trypanosoma cruzi (T. cruzi) there is cardiac involvement [2, 3, 11]. These estimates explain why this disease is a serious public health problem in the countries where it is endemic. In the evolution of Chagas disease we can distinguish an initial acute phase of infection and a prolonged intermediate chronic phase, in which the disease is often clinically silent, and the usual diagnostic techniques do not provide a robust criterion to predict whether a seropositive asymptomatic patient will suffer cardiac involvement. It is our interest to develop a non-invasive methodology low cost, that allows to see the dysautononia or dysfunction in the course of the 24 hours and with this it could be used to detect early any cardiac alterations produced by the T. cruzi.

2. Database

We have used the Electrocardiogram (ECG) database of the Institute of Tropical Meditation (IMT) of the Central University of Venezuela. Tachograms were obtained by processing the electrocardiogram (ECG) and obtaining the RR interval. The following test is carried out: clinical evaluation, serological Machado-Gerreiro test, chest x-rays, echocardiogram, electrocardiogram and ambulatory Holter registration (24 hours). The volunteers are classified into three groups: 92 healthy people called group control C; 102 patients infected with only positive serology (clinical evaluation, chest x-rays, echocardiogram, electrocardiogram and Holter were normal) called CH1 group; and 107 seropositive patients with incipient cardiac involvement first-degree atrioventricular block (BAV), sinus bradycardia (BS) or right bundle branch block of the bundle of His (BRDHH), that were not being treated with medications, called the CH2 group. All were outpatients, and informed consent was obtained from all of them. The ECG signals were recorded at 500 Hz with 12 bits of resolution, a set of 288 frames of 5 minutes was obtained.

3. Method

3.1. Preprocessed

For the detection of the QRS complex the program based on the Pan-Tompkins algorithm[9] was used, then the 288 tachograms of the RRIs of 5 minutes were generated. A post processing based on an adaptive filtering procedure and details in [10], was also carried out.

3.2. Feature extraction

The approximate entropy (ApEn) was developed by Steve Pincus[4, 5] is based on the entropy of Kolmogorov-Sinai KS, proposed by Grassberger and Procaccia[6] and modified by Takens[7]. A modification of ApEn was made by Richman[8] which is the sample entropy (Sampen), both of which measures the irregularity and complexity of
a time series of data.

ApEn and SampEn are based on the comparisons of the component-to-component embedding vectors \((m = 2)\) and with a threshold of 20\% of the standard deviation of the intervals \(RR\) \((r = 0.20)\). The difference between the entropies is that ApEn does not take into account the comparisons of the embedment vectors with itself and the way to calculate the logarithm. To calculate ApEn:

\[
ApEn(m, r, N) = -\frac{1}{N - m} \sum_{i=1}^{N-m} \log \left( \frac{A_i}{B_i} \right) \quad (1)
\]

where \(B_i\) are the vectors of embedding dimension \(m\) and \(A_i\) are the vectors of dimension of embedding \(m + 1\), \(r\) is the threshold that is typically 20\% of the value of the standard deviation and \(N\) is the number of elements.

SampEn is calculated as:

\[
SampEn(m, r, N) = -\log \left( \frac{\sum_{i=1}^{N-m} A_i}{\sum_{i=1}^{N-m} B_i} \right) = -\log \left( \frac{A}{B} \right) \quad (2)
\]

similar a ApEn \(m\), \(r\) and \(N\).

### 3.3. Testing samples

We will use the Kruskal Wallis test that matches the 288 frames corresponding to the Control-CH1, Control-CH2 and CH1-CH2 groups to test differences between groups. Also a logistic regression is used to evaluate circadian profiles of the average values of ApEn, SampEn. Finally boxplot representations are used to analyze a frame in special.

### 4. Results

#### 4.1. Approximate Entropy (ApEn)

ApEn was calculated at 288 tachograms of 5 minutes and 24-hour circadian profiles were obtained (Fig 1a).

- Between 00:00 and 06:00 the control group C has slightly higher values than CH1, CH1 ApEn values are slightly greater than CH2.
- Between 06:30 and 11:20 there is no greater differences between the three groups.
- We noticed that between 11:20 and 21:30 there is a sustained difference of group C.
- At 14:30 it is the biggest difference between groups C and CH2.

We have used the logistic regression to discriminate groups C and CH2. For this task 50\% of frames are selected randomly (144) for training and 50\% specificity and 80\% of sensitivity between 14:00 h. and 15:40 h.

We applied the Kruskal-Wallis test where we can see in figure 1b that around frame 175 (14:30 hours) \(p\)-value < 0.05, for the groups: Control-CH1 and Control-CH2. This not the case for CH1-CH2 as shown in the figure 1.

#### 4.2. Sample Entropy (SampEn)

We calculated the SampEn to the 288 segments of 5 minutes to the three groups and we obtain the average val-

![Figure 1](image-url1)

![Figure 2](image-url2)

![Figure 3](image-url3)
ues of the RR intervals. This can be seen in Fig 4a,

### Figure 4. a) SampEn values of the 24 hours (288 frames)
b) p-value of the groups: Control-CH1, Control-CH2, CH1-CH2

- Between 00:00 hours and 08:30 the value of SampEn there is no greater difference between the three groups. The control group C has slightly higher values than CH1, CH1 is slightly greater than CH2.
- Between 08:30 and 10:00 the groups CH1 and CH2 have a higher value than group C.
- However, from 12:30 to 16:15 we observe that SampEn values are higher in group C than groups CH1 and CH2.
- At 14:10 it is the biggest difference between groups C and CH2.

Now we will use the logistic regression classifier for groups C and CH2, with this it is possible to reach 80% of specificity and 70% of sensitivity between 14:00 h. at 15:40 h. We applied the Kruskal-Wallis test to the values of SampEn to the groups: Control-CH1, Control-CH2, where we noticed that around frame 175 p-values lower than 0.05.

A similar result is obtained with ApEn.

### Figure 5.

Now we will use the logistic classifier for groups C and CH2, for this we will use 50% of frames at random (144) for training and 50% for validation, with this it is possible to classify up to 80% of specificity of group C and 70% of sensitivity in the CH2 group between 14:00 h. at 15:40 h. We applied the Kruskal-Wallis test to the values of SampEn to the groups: Control-CH1, Control-CH2, where we noticed that around frame 175 the p-values lower than 0.05.

The average ApEn values between 11:20 and 21:30 show (Fig. 1a) that group Control presents higher values than the values corresponding to group CH1 and CH2. This could be due to a decrease or enervation in its barometric response, this is also observed in SampEn (Fig 4a) at shorter time intervals from 12:30 to 16:15.

The approximation entropy (ApEn) and the sampling entropy (SampEn) measure the irregularity and the complexity, when we apply these features to the circadian profiles of the Control, CH1 and CH2 groups, we observe that in the afternoon hours the values of the Control group has greater irregularity and complexity than the values of the CH1 and CH2 groups that show decrease or enervation of these groups, this could be used as a measure of the variability of the heart rate.

The logistic classifier and Kruskal-Wallis test allows us to differentiate the both groups (control and CH2) both in ApEn and SampEn in the afternoon hours (18 frames), indicating that it can be used in the stratification of risk of group CH1, therefore it is possible to use to classifier to can be used in the stratify group CH1.

The logistic regression and the Kruskal-Wallis test allow
Acknowledgements

Virrectorado de Investigación de la Universidad Nacional San Agustín de Arequipa, contrato de subvención 16-2018-UNSA.

Conflicts of Interest

The authors declare no conflict of interest.

References


Address for correspondence:
Miguel Vizcardo Cornejo
Calle Santa Catalina N 117 Arequipa, Peru
mvizcardoc@unsa.edu.pe