# ORIGINAL ARTICLE

# Abnormality Detection in ECG Signal applying Poincare and Entropy-based Approaches

**Shabdik Chakraborty[1], Shreya Saha[1], Sibeswar Prasad Singha[1], Sweta Sarkar[1], Kingshuk Chatterjee[1*], Deboleena Sadhukhan[2, 3], Alok Mukherjee[1], Tanmay Sarkar[4]**

## Abstract

Detection of abnormality in heart is of major importance for early and appropriate clinical medication. In this work, we have proposed two models for detection of abnormality in ECG signals. The normal ECG signals are closely repetitive in nature to a large extent, whereas ECG signals with abnormalities tend to differ from cycle to cycle. Hence, repetitive plot like the Poincare is efficient to detect such non-repetitiveness of the signal; thereby, indicating abnormalities. Hence, we have used Poincare plot to develop the two proposed models. One of the models uses direct analysis of the binary image of the plot to detect the difference in retracing, between the healthy and unhealthy samples. The other model uses entropy of the Poincare plot to detect the difference in randomness of plots between the two classes. Most importantly, only lead II ECG signal is used here for analysis. This ensures ease of computation as it uses signal of only a single lead instead of the 12 leads of the complete ECG signal. We have validated the proposed models using ECG signals from the 'ptb database'. We have observed that the entropy analysis of the Poincare plots gives the best results with 90% accuracy of abnormality detection. This high accuracy of classification, combined with less computational burden enables its practical implementation for the development of a real-life abnormality detection scheme.

**Key Words:** *Electrocardiogram (ECG); Myocardial infarction (MI); Poincare plot; Entropy analysis*

[1]*Government College of Engineering and Ceramic Technology, Kolkata, India*

[2]*Applied Physics, University of Calcutta, India*

[3]*Institut Langevin, Université de Paris, Paris, France*

[4]*Malda Polytechnic, West Bengal State Council of Technical Education, Government of West Bengal, Malda, India*

*[*]Corresponding author: Kingshuk Chatterjee, Assistant Professor, Computer Science & Engineering, Government College of Engineering and Ceramic Technology, Kolkata, India, Tel: +91-9883098632; E-mail: kingshukchaterjee@gmail.com*

# Introduction

Myocardial infarction (MI) is of the most fatal diseases that occur in the human heart. Hence, several researchers have used different techniques for the detection and analysis of ECG signals for easy and authentic detection of myocardial infarction (MI) in the human heart [1,2]. The standard 12-lead electrocardiogram (ECG) is one of the most commonly used medical studies in the assessment of cardiovascular disease. It is the most important test for interpretation of the cardiac rhythm, detection of myocardial ischemia and infarction, conduction system abnormalities, pre-excitation, long QT syndromes, atrial abnormalities, ventricular hypertrophy, pericarditis and other conditions. A standardized system has been developed for electrode placement for a routine ECG. Ten electrodes are needed to produce 12 electrical views of the heart. An electrode lead, or patch, is placed on each arm and leg and six are placed across the chest wall. The signals received from each electrode are recorded. The printed view of these recordings is the electrocardiogram. By comparison, a heart monitor requires only three electrode leads – one each on the right arm left arm and left chest. It only measures the rate and rhythm of the heartbeat. This kind of monitoring does not constitute a complete ECG. But can determine whether a person needs the supervision of the doctor for future heart diseases. Most of the available automated MI detection techniques rely on the use of advanced signal processing tools, a large number of ECG features, and complex classifiers, which increase their computational load; thus, making them unsuitable for implementation on portable healthcare devices.

Most of the available automated MI detection techniques rely on the use of advanced signal processing tools, a large number of ECG features, and complex classifiers, which increase their computational load; thus, making them unsuitable for implementation on portable healthcare devices. In this work, we have developed two schemes of identifying the maloperation of the human heart by inspecting only the lead II data of the ECG signals. Both the schemes employ the Poincare plot as the basis of feature extraction. Poincare plot is a kind of repetitive plot, which is used to identify the non-repetitiveness of any signal. This method has been used in abundance in the field of ECG based heart abnormality detection [3-7], as it can easily detect any abruptly different and non-repetitive nature of the heart cycle of the ECG plot of a diseased patient. Besides, the use of a single lead greatly reduces the complexity and cost of hardware implementation. These models of analyzing only the lead II data makes the schemes simple in computation, apart from ease of acquisition of the ECG signals, since, one of our major objectives is to achieve portability so that it can become a common household item for someone with a history of cardiac disorders. This also makes the cost-effective implementation of the schemes [8].

Many such heart disease detection techniques using ECG signals include different supervised learning algorithms such as artificial neural network (ANN) or methods incorporating wavelet transform, Fourier transform fuzzy inference system (FIS) or others which include intense mathematical analysis. Most of these methods have a high computational burden; hence, a high associated time of computation. Artificial neural network (ANN), along with two of its most significant subsidiaries such as machine learning and deep learning, has provided excellent results of disease detection. The neural network is often hybridized with fuzzy inference system (FIS) to develop effective classifiers [9]. Many works have

preferred principal component analysis (PCA) to extract major features from the signal due to its effectiveness for reducing the dimension of data towards the most important directions of highest variance in the descending order of importance. Some other researchers have adopted the time-frequency domain approach such as wavelet transform and PCA to extract features from the signal and used numerous variants of neural networks to analyze the same features for disease detection [10,11]. Wavelet transform is also combined directly with the support vector machine (SVM) for successful MI detection [12]. These supervised learning models depend heavily on long and diverse training of the model; therefore, possess the requirement of the large volume of training data, as well as, appreciably high training time. Many of the research majorly focus on the QRS complex region, which is the most prominent feature part of the ECG signal [13-16], even employing supervised learning techniques.

Convolution neural network (CNN), which is an advanced analysis of the neural network, has become a popular tool of research due to its high efficiency; hence, used in ECG signal analysis as well [17-20]; as well as used as 16-layer CNN-LSTM deep learning model [1,21]. Radial neural network (RNN) has also been effective in detecting heart rate abnormalities [22]. The authors have used hybrid approach combining features of CNN and RNN simultaneously [23]. Fourier-Bessel expansion based on wavelet transform has been presented by the authors of [24], who have aided by deep learning approaches. Deep learning-based methods, although are high in accuracy, often suffer from the requirement of large training volume; hence, large training time. The authors of [25] have presented a comparative analysis of three major supervised learning methods such as Convolutional, Residual and Inceptional neural network to highlight the major features and outcomes of each model. Transfer learning-based approaches also have proved their effectiveness in disease identification [26,27], sometimes combined with image processing approaches.

## Methodology

### Source of data set

We have used Physikalisch-Technische Bundesanstalt (PTB) database, i.e., 'ptbdb' in the present work, which is compiled by the National Metrology Institute of Germany for the users of PhysioNet. This database contains ECG signals of the healthy and the diseased patients [28,29]. The proposed work is used to develop two algorithms to detect different myocardial infarction (MI) disorders. We have also presented a comparative analysis of the outcomes of the two proposed classifiers. The complete PTB database contains 148 ECG samples of patients having various MI disorders and 52 subjects with healthy control. We have considered a random set of 30 ECG signals of healthy people and 50 such samples of diseased patients for developing the proposed algorithms. Another random sample set of a similar number of ECG data has been used to validate the proposed schemes.

### Analysis of ECG signal

ECG is the primary non-invasive tool used by clinicians for evaluating patients with suspected Myocardial Infarction. In MI, the damage of myocardial tissues caused by the coronary artery blockage affects their electrical conduction ability and contractility, thus, altering the electrical conduction pathway through the heart. These cause morphological changes in the ECG wave pattern. Usually, changes are observed

in the T-wave, ST-segment, and eventually Q-wave, as different regions of the heart go through varying degrees of ischemia, injury, and infarction [30]. Figure 1 elaborates a prototype ECG signal. These changes are reflected in the particular ECG leads corresponding to the zone of infarction. Reciprocal changes are often observed in other leads. However, if timely medical intervention is provided, a patient may not undergo all the phases shown in the figure.
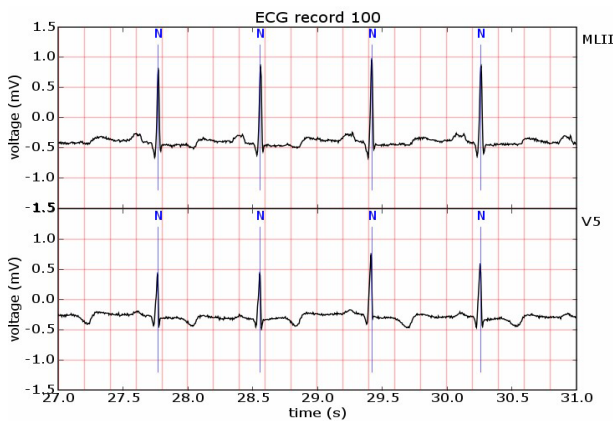


**Figure 1)** *A prototype ECG signal.*

The overall ECG changes that are induced due to the presence of acute MI can be summed up as follows:

- ST segment elevation over the infarcted zone
- ST segment depression in leads opposite to the zone of infarction
- Inverted T waves
- Reduced R waves
- Pathological Q waves

However, all these changes may not be prominent simultaneously

**Application of Poincare analysis on ECG signal**

In this work, only the lead II data of the ECG signal of both healthy and diseased patients is used, followed by Poincare analysis to split the signal into two distinct components. If the length of the signal is considered as $L$, we have segmented the signal $X$ into two consecutive time-split components $X_1$ and $X_2$ as follows:

$$X_1(L-1) = x_1, x_2, x_3, \dots, x_{L-1} \quad (1)$$

$$X_2(L-1) = x_2, x_3, x_4, \dots, x_L \quad (2)$$

The above expressions indicate that the first component $X_1$ contains the first *(L-1)* elements $x_1, x_2, x_3, \dots, x_{L-1}$ of the signal, where $L$ is the length of the original signal; and $X_1$ contains the final *(L-1)* elements $x_2, x_3, x_4, \dots, x_L$ of the signal. It indicates that $X_1$ and $X_2$ are time-shifted by a single element or sample point. In the present case, the array of ECG signal contains 30000 sample points. Hence, the signal is split for plotting the graph as above. The first set $X_1$ contains 1-29999 sample points and the second set $X_2$ contains 2-30000 sample points. $X_1$ is plotted along the *x*-axis of the graph and the *y*-axis contains the $X_2$ elements [4,14]. Both the healthy and the unhealthy signals are plotted in the following way and some of the plots are shown in Figure 2 and Figure 3 respectively as example cases.

The two-dimensional plots of Figure 2 and Figure 3 show a promising difference between healthy and unhealthy patients. Poincare plot of the second lead for a normal patient is more periodical than diseased patients. Therefore, the healthy patients exhibit more uniform, overlapping, and concise plots; whereas patients with certain cardiac abnormalities had an uneven and distributed plot with the data more spread out. This feature has been used in this work to explore the less repeatability of the abnormal ECG traces, compared to the healthy signals [18,31]. Hence, the following two models are proposed in the subsequent sections.
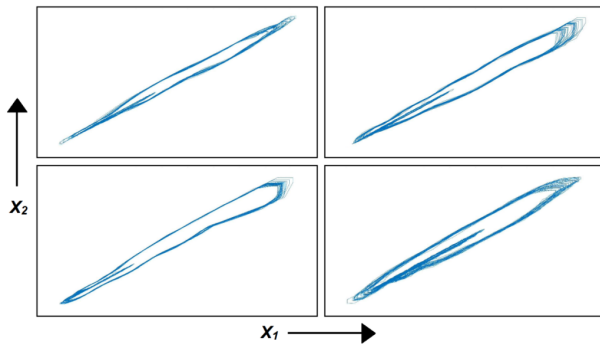
**Figure 2)** *Four random examples of Poincare plot of ECG single lead data of normal people.*
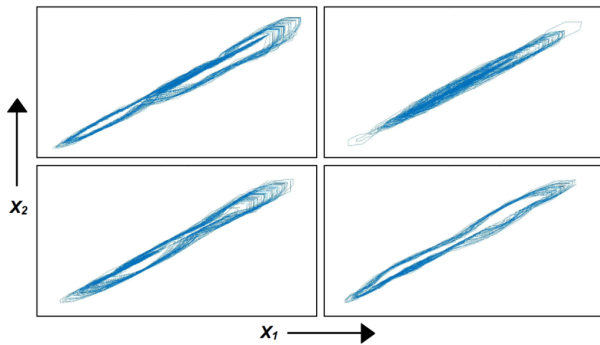


**Figure 3)** *Four random examples of Poincare plot of ECG single lead data of diseased patients.*

## Development of proposed schemes

### Binary image analysis model

An image contains a continuum of intensity values before it is quantized to obtain a digital image. ECG signals of a total number of 30 normal people and 50 diseased patients, chosen randomly, have been considered to develop the proposed scheme. The information in the image is in the gray values. Hence, we have converted the obtained graphs into grayscale which in turn was converted into a binary image using thresholding. Further, we have counted the number of black pixels in the image, which are the traces of the Poincare plot, and denoted this number by K-values. We have computed another parameter, denoted by B values, which are the K-values divided by the total resolution of the image. For the healthy patients, overlapping of the Poincare plot is more pronounced and hence the K-values are less compared to unhealthy patients. This is more pronouncedly observed

from the plots of Figure 2 and Figure 3. B-values, being directly proportional to K-value, show a similar trend. The B-values obtained with the healthy and unhealthy patients are shown using boxplots in Figure 4a. The single-variate analysis as shown in Figure 4a establishes a visible distinction between the data obtained from both sets of patients.
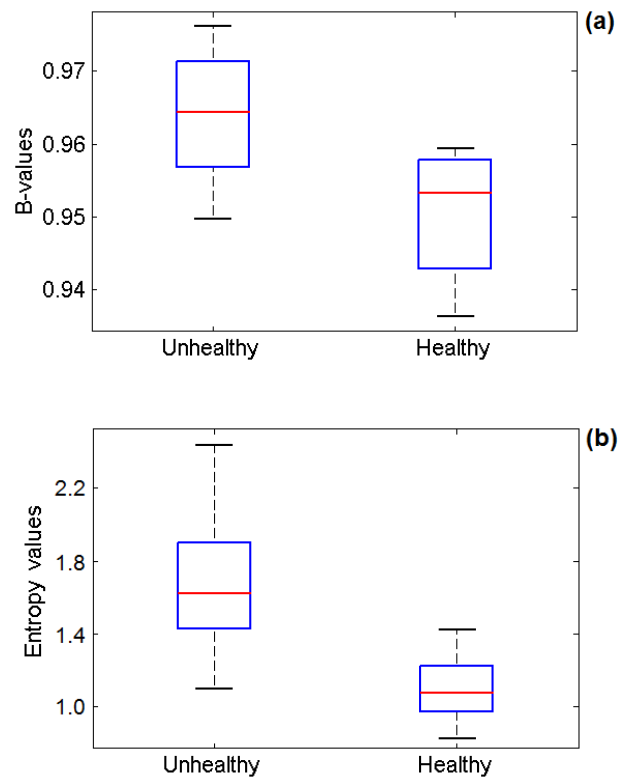


**Figure 4)** *Boxplot indicating (a) The B-values and (b) Entropy values of healthy and unhealthy patients.*

### Image entropy analysis model

We have analyzed the Poincare images using another major parameter such as entropy. Entropy is a statistical measure of randomness that characterizes the texture of an image [32]. It is defined by the following expression:

$$Entropy = sum(p*log_2(p)) \qquad (3)$$

where, $p$ contains the normalized histogram counts. The major steps associated with the analysis of the images are discussed here:

- The images obtained by applying Poincare Analysis of lead II ECG signal were cropped using a cropping algorithm to remove additional whitespace.

- The cropped images were further converted into grayscale.

- The entropy of the grayscale images was computed using the entropy function in MATLAB environment using expression (3).

The same set of ECG data as used in to develop the binary image analysis model, were used to develop the image entropy analysis model again. The ECG signals were selected randomly from the PTB database and were analyzed using the proposed entropy analysis based model. The entropy values obtained for both classes of patients are expressed using boxplot and are shown in Figure 4b, which easily distinguishes the two classes; although, a marginal overlap of the entropy values are observed.

## Result and Analysis

We have observed boxplots of the feature parameters such as the B-values and the entropy values, obtained using both the proposed methods, viz., Binary Image Analysis and Entropy Analysis respectively. We have proposed a threshold based approach and calculated a threshold value between the feature parameters of the healthy and the unhealthy patients, followed by classifying the patients. We have considered another separately chosen random set of 30 numbers of ECG signals of healthy people and 50 numbers of ECG signals of unhealthy patients to validate the proposed models. The results obtained from both the models are described in Table 1 and Table 2 respectively.

**TABLE 1**

**Accuracy of classification (%) obtained from Binary Image Analysis model**

| Class of patients | | Predicted class | | Total | Accuracy of classification (%) |
|---|---|---|---|---|---|
| | | Healthy | Diseased | | |
| True class | Healthy | 26 | 4 | 30 | 86.67 |
| | Diseased | 9 | 41 | 50 | 82 |

Overall Correct Classification: 67 out of 80
Overall accuracy of classification (%): 83.75

**TABLE 2**

**Accuracy of classification (%) obtained from Entropy Analysis model**

| Class of patients | | Predicted class | | Total | Accuracy of classification (%) |
|---|---|---|---|---|---|
| | | Healthy | Diseased | | |
| True class | Healthy | 26 | 4 | 30 | 86.67% |
| | Diseased | 4 | 46 | 50 | 92% |

Overall Correct Classification: 72 out of 80
Overall accuracy of classification: 90%

It is observed from the above results that the entropy analysis-based model is superior compared to the binary image analysis based direct detection scheme. The overall accuracy of detection is in the order of 83.75% and 90% respectively. Therefore, we can conclude that disease detection using entropy value is more precise. Hence, it could easily be inferred that the entropy analysis scheme applied over the point Poincare plot provides an appreciably high level of accuracy [33]. Besides, the proposed scheme is simple as it uses only a Poincare plot for processing the signal, which is a simple repetitive plot and does not include any complicated algorithm, thereby making the implementation of the scheme simple.

## Conclusion

In this work, we have developed abnormality detection algorithms from the ECG signals of the human heart. Only lead II data of the ECG signal has been used instead of the entire 12 channel ECG signal. This ensures ease of computation since only a single channel is analyzed. This is followed by application of Poincare analysis to develop the plots of the single-channel ECG signal which are further analyzed in two ways. This is a binary image analysis model, which is a direct detection method. We have also analyzed the Poincare plots using the entropy method. We have acquired sample ECG signals from 'ptb database'. We have analyzed the training results using boxplots and found a distinct difference between the healthy and unhealthy classes. We have further validated the models using a separate set of randomly chosen samples from the same database. Both the models proposed are simple as it uses only a repetitive plot followed by entropy analysis; as well as highly accurate in disease detection with a highest average accuracy of 90%. This simplicity of the schemes, combined with high accuracy of detection, widens the possibility of the proposed models being implemented in real-life devices.

## References

1. Lih OS, Jahmunah V, San TR, et al. Comprehensive electrocardiographic diagnosis based on deep learning. Artif Intell Med. 2020;103:101789.

2. Schneider M, Kraemmer MM, Weber B, et al. Life events are associated with elevated heart rate and reduced heart complexity to acute psychological stress. Biol Psychol. 2021;163:108116.

3. Mourot L, Bouhaddi M, Perrey S, et al. Quantitative Poincaré plot analysis of heart rate variability: effect of endurance training. Eur J Appl Physiol. 2004;91:79-87.

4. Park J, Lee S, Jeon M. Atrial fibrillation detection by heart rate variability in Poincare plot. Biomed Eng Online. 2009;8:1-2.

5. Guzik P, Piskorski J, Krauze T, et al. Heart rate asymmetry by Poincaré plots of RR intervals. Biomed Tech (Berl). 2006;51:272-5.

6. Strumillo P, Ruta J. Poincare mapping for detecting abnormal dynamics of cardiac repolarization. IEEE Eng Med Biol Mag. 2002;21:62-5.

7. Mukherjee A, Sadhukhan D, Chatterjee K, et al. Indoor cardiovascular health monitoring system under COVID 19 situations. 2022;12:3488-500.

8. Meek S, Morris F. ABC of clinical electrocardiography. Introduction. I-Leads, rate, rhythm, and cardiac axis. BMJ. 2002;324:415-8.

9. Lu HL, Ong K, Chia P. An automated ECG classification system based on a neurofuzzy system. Proceedings of the Computers in Cardiology. 2000. Vol.27 (Cat. 00CH37163), Cambridge, MA, USA. 2000.

10. Bhaskar NA. Performance analysis of support vector machine and neural networks in detection of myocardial infarction. Procedia Comput Sci. 2015;46:20-30.

11. Jafarian K, Vahdat V, Salehi S, et al. Automating detection and localization of myocardial infarction using shallow and end-to-end deep neural networks. Appl Soft Comput. 2020;93:106383.

12. Thatipelli T, Kora P. Classification of myocardial infarction using discrete wavelet transform and support vector machine. Int Res J Eng Technol. 2017;4:429-32.

13. Arif M, Malagore IA, Afsar FA. Automatic detection and localization of myocardial infarction using back propagation neural networks. Proceedings of the 2010 4th International Conference on Bioinformatics and Biomedical Engineering, Chengdu, China. 2010.

14. Safdarian N, Dabanloo N, Attarodi G. A new pattern recognition method for detection and localization of myocardial infarction using T-wave integral and total integral as extracted features from one cycle of ECG signal. J Biomed Sci Eng. 2014;7:818-24.

15. Liu W, Huang Q, Chang S, et al. Multiple-feature-branch convolutional neural network for myocardial infarction diagnosis using electrocardiogram. Biomed Signal Process Control. 2018;45:22-32.

16. Han C, Shi L. ML-ResNet: A novel network to detect and locate myocardial infarction using 12 leads ECG. Comput Methods Programs Biomed. 2020;185:105138.

17. Ansari S, Farzaneh N, Duda M, et al. A review of automated methods for detection of myocardial ischemia and infarction using electrocardiogram and electronic health records. IEEE Rev Biomed Eng. 2017;10:264-98.

18. Sharma LD, Sunkaria RK. Inferior myocardial infarction detection using stationary wavelet transform and machine learning approach. Signal, Image Video Process. 2018;12:199-206.

19. Muminov B, Nasimov R, Mirzahalilov S, et al. Localization and classification of myocardial infarction based on artificial neural network. Proceedings of the 2020 Information Communication Technologies Conference, Nanjing, China. 2020.

20. Fu L, Lu B, Nie B, et al. Hybrid network with attention mechanism for detection and location of myocardial infarction based on 12-lead electrocardiogram signals. Sensors (Basel). 2020;20:1020.

21. Darmawahyuni A, Nurmaini S, Sukemi. Deep learning with long short-term memory for enhancement myocardial infarction classification. Proceedings of the 2019 6th International Conference on Instrumentation, Control, and Automation, Bandung, Indonesia. 2019.

22. Prabhakararao E, Dandapat S. Myocardial infarction severity stages classification from ECG signals using attentional recurrent neural network. IEEE Sens J. 2020;20:8711-20.

23. Lui HW, Chow KL. Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ECG devices. Informatics Med Unlocked. 2018;13:26-33.

24. Tripathy RK, Bhattacharyya A, Pachori RB. A novel approach for detection of myocardial infarction from ECG signals of multiple electrodes. IEEE Sens J. 2019;19:4509-17.

25. Wang HM, Zhao W, Jia DY, et al. Myocardial infarction detection based on multi-lead ensemble neural network. 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Berlin, Germany. 2019.

26. Alghamdi A, Hammad M, Ugail H. et al. Detection of myocardial infarction based on novel deep transfer learning methods for urban healthcare in smart cities. Multimed Tools Appl. 2020.

27. Gupta A, Huerta E, Zhao Z, et al. Deep Learning for cardiologist-level myocardial infarction detection in electrocardiograms. European Medical and Biological Engineering Conference, Springer, Cham. 2020.

28. Bousseljot R, Kreiseler D, Schnabel A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. Biomed Tech (Berl). 1995;40:317-8.

29. Goldberger AL, Amaral LA, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation. 2000;101:E215-20.

30. Liu W, Wang F, Huang Q, et al. MFB-CBRNN: A hybrid network for MI detection using 12-lead ECGs. IEEE J Biomed Heal Informatics. 2020;24:503-14.

31. Padhy S, Dandapat S. Third-order tensor based analysis of multilead ECG for classification of myocardial infarction. Biomed Signal Process Control. 2017;31:71-8.

32. Mukherjee A, Chatterjee K, Sarkar T. Entropy-aided assessment of amla (emblica officinalis) quality using principal component analysis. Biointerface Res Appl Chem. 2022;12:2162-70.

33. Hoshi RA, Pastre CM, Vanderlei LCM, et al. Poincaré plot indexes of heart rate variability: relationships with other nonlinear variables. Auton Neurosci. 2013;177:271-4.