Clinical Cancer Investigation Journal-ISSN: 2278-0513

Evaluating RQA Characteristics of ECG Signal to Enhance Sudden Cardiac Death Prediction Time

Abstract

The study presented a method to enhance SCD prediction using one of the genetic algorithms and to evaluate the characteristics of Recurrence quantification analysis (RQA). This study used data before and after the occurrence of previous attacks. The signal must be isolated a few minutes before the heart attack (SCD) to distinguish between unhealthy and normal people. In the next stage, the features related to recursive mapping were extracted using RQA. Then in the feature selection stage, the best features are selected using a genetic algorithm so that one can use them to distinguish between two groups of healthy people and those on the verge of heart death with great accuracy and predict SCD by revealing the increased risk. In the final stage, the MLP classifier was used with the help of a neural network to show the difference between the ECG signal of a healthy person and a person with SCD. Ultimately, the evaluation of feature changes according to signal RQA to improve SCD prediction needs to be considered. Performance evaluation for the DET method presents an acceptable predictive response. L method presents a good performance in prediction too. TT method has a lower performance than the three methods stated in this algorithm and is associated with estimation errors. Lmax, T1, T2, and Trans methods work almost the same, and all showed high errors in estimation in one case; if more accurate performance is required, DET and L-based evaluations are the best answer for implementing this algorithm.

Keywords: *Genetic algorithm, neural networks, SCD, RQA.*

Mohsen Abdi*, Mehdi Azarnoosh

Department of Medical Engineering, Faculty of Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran.

Corresponding authors :Mohsenabdi95@gmail.com

Introduction

SCD accounts for half of all deaths from coronary heart disease (CHD), which also accounts for half of all deaths worldwide. Most cases of SCD are thought to be the result of an acute event or a triggering operation in the context of a history of previous illnesses. Many medical conditions can result in SCD, yet the most common in men and women is CHD. Other common conditions are valvular heart disease, dilated cardiomyopathy, and hypertrophic cardiomyopathy. One group of rare but well-known underlying conditions is Long QT Syndrome (LQTS). CHD is more common in men than in women. A similar study indicates that the second most common underlying cause of SCD is dilated cardiomyopathy. Other less common forms of SCD that alternate between men and women are heart valve disease and long QT syndrome.

Iwaniec and. Iwaniec. (2017) presented a machine learning technique for classifying low-risk and high-risk individuals in patients with heart problems using HRV analysis. In this report, a classification accuracy of 5% was obtained using a support vector machine (SVMs) classifier, and RBF kernel and an accuracy of 85% was obtained using random forest classification (1). In a study of several young Italian athletes and screening with ECG, Maron et al. (2016) reported that 81 athletes had abnormal ECGs, of which 6% had cardiomyopathy (2). Shen et al. (2007) proposed a home and personal protection system for detecting SCD prediction of ECG signals. In their report, they reported an accuracy of 44% (minimum mean squares), 14% (decision-oriented neural network) and 55% (post-publication neural network) in SCD prediction (3). Rashed, U. Mirza. (2008) used time domain and HRV frequency parameters to predict SCD. They revealed that short-term low-frequency power during controlled breathing is a good tool for predicting SCD in patients with chronic heart disease (4). Goldstein et al. (2014) showed the nature of the EGC signal by short-term fractal scaling coefficient, revealing that this coefficient is an independent and suitable element for predicting SCD in 446 heart attack survivors with an assessment of the left ventricular function of less than 35% (5). Chat Clean (2007) et al. indicated that linear methods of time and frequency could not predict SCD (6).

Ebrahimzadeh (7) predicted SCD from HRV and ECG signals with time-frequency processing. The results indicated features in the HRV signal of people at risk of SCD, making them completely different from healthy people. Moradi et al. (8) examined the effect of the circadian cycle on the ST segment response of athlete heart electrocardiograms to a sub-maximal test. The findings indicated that the maximal test significantly reduced the response of the ST segment of the electrocardiogram of the athlete's heart in both morning and evening to rest. Nevertheless, no significant differences were observed between the mean of the morning and evening measurements.

Estimating and evaluating the risk of SCD in the global population is significant. About half of all SCDs happen in people without a history of heart disease, and the percentage and rate of successful resuscitation are very low at about 8%. An electrocardiogram (ECG) is used to predict SCD now, yet it has not reached a high capability in predicting SCD to date. Nonetheless, studying and combining electrocardiogram markers used so far and the potential for new cases could help predict SCD in the future.

The study uses a dataset under 24-hour ECG (Holter Database) signal records prior to and following previous attacks. These patients ' ECG signals before heart attack (SCD) should be categorized at 1-minute intervals. Then in the feature selection stage, the best features are selected using a genetic algorithm so that one can use them to distinguish between two groups of healthy people and those on the verge of heart death with great accuracy and predict SCD by revealing the increased risk. In the final stage, the MLP classifier was used with the help of a neural network to show the difference between the ECG signal of a healthy person and a person with SCD. The MLP network is trained using an error back propagation algorithm with a variable learning rate. Finally, it has to be considered to evaluate the feature changes based on the signal RQA to enhance the prediction of heart attack (SCD).

Methods

The study was applied in terms of purpose and deductive-inductive in terms of method, where the theoretical foundations and research background were collected through libraries, papers and the Internet. Inductive reasoning using mathematical methods and inductive reasoning was used to generalize the results.

1. Database: Research data, including ECG signals received from a set of cardiac signals of heart patients and ordinary people, extracted from valid Internet databases such as MIT-BIH, SCD Holter and Sinus Rhythm Database Normal .

2. Data processing: data preprocessing was done using a dataset under 24-hour ECG (Holter Database) signal records prior to and following previous attacks. ECG signals before these patients' heart attack (SCD) should be categorized at 1-minute intervals. To distinguish between unhealthy and normal people, we need to isolate the signal a few minutes before the heart attack (SCD). Then we determine the HRV signal by removing the noise from the ECG signal.

3. Recursive quantification analysis (RQA)

The method was first implemented in 1998 by Ackman et al. to display non-static EEG signal components at various times. Table (1) is some features of recursive mapping.

Table 1: The features extracted from recursive mapping analysis

Feature	Explanation
RR	Return rate
Det	Determinism
L	Average lengths of the diagonal lines
Lmax	The length of the longest diagonal line
ENTR	Entropy
LAM	Laminarity
TT	Trap time
Vmax	The length of the longest vertical line
T_1	Type 1 recurrence
T_2	Type 2 recurrence
RTE	Recurrence time entropy
Trans	Transferability

4. Genetic algorithm

Genetic algorithm is of the intelligent methods that can be applied to a wide range of problems. This algorithm is based on Darwin's theory of evolution. In this method, a population of candidate subsets is generated. The features extracted in the previous step are considered as the initial population. Each time the algorithm is iterated, new elements are generated using the mutation and recombination operators on the elements of the previous population. The competency of the elements of the current population is determined using an evaluation function, and the better elements are selected as the population of the new generation. The features selected in this step are used to classify the neural network.

5. MLP neural network

Multilayer perceptron has many applications in classifying patterns and approximating functions. The study used a nonlinear sigmoid family excitation function (hyperbolic tangent) in the hidden layer and a linear stimulus function in the network output layer. Moreover, the number of hidden layer neurons is 5.

Results

1. Data extraction

Sixteen samples of one-hour ECG signals were extracted from the database at archieve.physionet.org/cgi-bin/atm/ATM. The time format of each of these signals is sample, and they have a sampling frequency of 250Hz. 900,000 (3600 × 250) records were recorded for each signal.

2. HRV signal extraction

ECG signals were divided into one-minute intervals, and the signal level at each interval was calculated. Then ECG signal was sampled from each sample for 5 minutes before the lowest ECG level (6 minutes of sampling for each signal), and the ECG signal peak points were calculated using the ECG_TO_HRV function and the HRV of each signal (for 6 minutes) was calculated.

3. RQA feature extraction

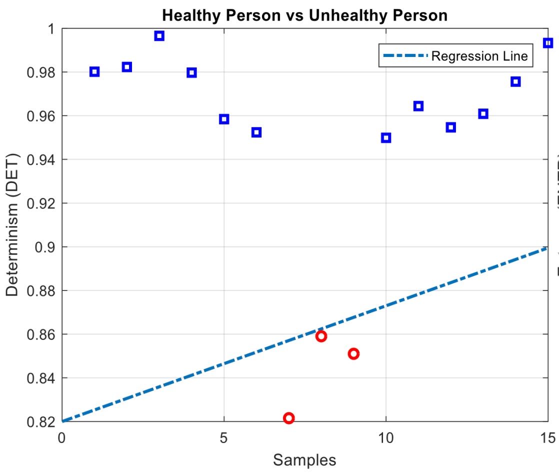
At this step, the best signal features are calculated. The similarity threshold is $10^{-7,}$ and the signal is divided into 2000 intervals to construct the R-matrix. The results are stored in the RQA matrix.(Table 2)

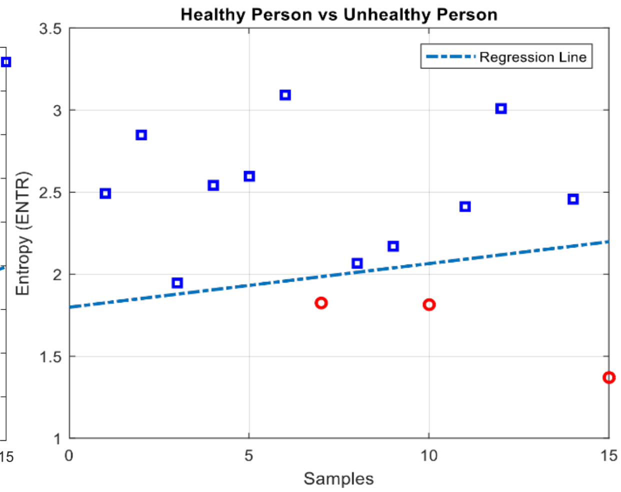
Table 2: The values obtained for RQA features

Lmax	T1	T2	Trans	RR	DET	L	ENTR	LAM	TT
0.9600	6.3000	0.6417	0.0044	0.4297	0.9802	10.6000	3.6260	0.9987	9.8182
1.4040	6.0028	0.5584	0.0037	0.5319	0.9986	10.7949	5.0156	0.9995	10.7500
2.0720	5.0598	0.4131	0.0206	0.5042	0.9928	12.6976	4.9356	0.9996	12.2477
1.4280	5.9804	0.4538	0.0043	0.4820	0.9949	13.2484	3.9914	0.9958	13.1790
1.0800	7.5741	0.6296	0.0032	0.4598	0.9946	12.0888	3.4247	0.9956	12.0294
2.5200	4.0778	0.3524	0.0181	0.4432	0.9864	12.0711	4.5750	0.9950	11.5721
1.4880	6.6747	0.5484	0.0023	0.5180	0.9892	12.6410	5.7717	0.9964	12.1716
3.6800	8.3391	0.5870	0.0017	0.4882	0.9906	15.0277	6.1056	0.9995	14.2074
2.4320	7.2862	0.5263	0.0046	0.4497	0.9919	14.4671	6.4052	0.9991	13.8438
2.4480	10.055	0.6830	0.0049	0.4793	0.9945	15.0985	6.6351	0.9984	14.7225
2.7200	5.8853	0.4529	0.0036	0.4556	0.9871	13.5959	6.0653	0.9950	12.9935
2.8560	6.5658	0.5350	0.0022	0.4497	0.9876	13.0847	5.7284	0.9996	12.2723
2.2000	4.5927	0.3382	0.0025	0.4970	0.9905	14.2386	6.5581	0.9984	13.5806
4.6000	6.0678	0.3722	0.0042	0.5414	0.9957	16.7874	6.3194	0.9997	16.3037
2.2880	8.5594	0.5734	0.0020	0.4941	0.9963	15.1677	5.6602	0.9988	12.9268
2.0000	7.9640	0.5400	0.0235	0.4231	0.9975	14.9398	5.5586	0.9995	12.7481

4. Dimension reduction using genetic algorithm

After extracting the RQA features, the best signal features should be selected to enhance the resolution. After obtaining the value of 10 attributes for all samples in this step, the task is to predict the two methods of genetic algorithm and neural network. Besides the neural network, this section uses a genetic algorithm to separate patients from healthy people and create a grammatical pattern.(figure 1)


Healthy Person vs Unhealthy Person
Regression Line
Determinism (DET)
Samples

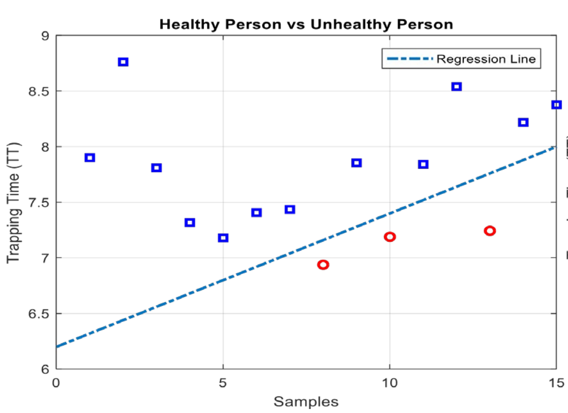

Healthy Person vs Unhealthy Person
Regression Line
Entropy (ENTR)
Samples

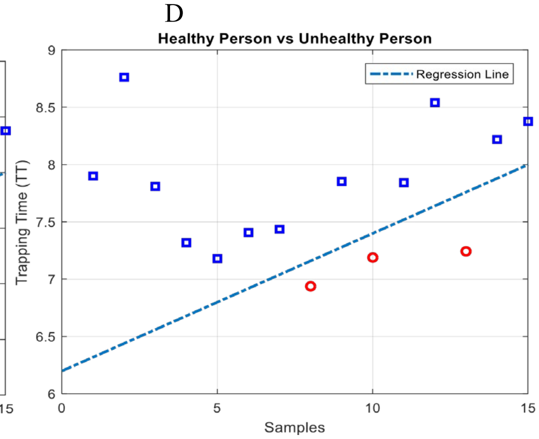
A B


Healthy Person vs Unhealthy Person
Regression Line
Laminarity (LAM)
Samples

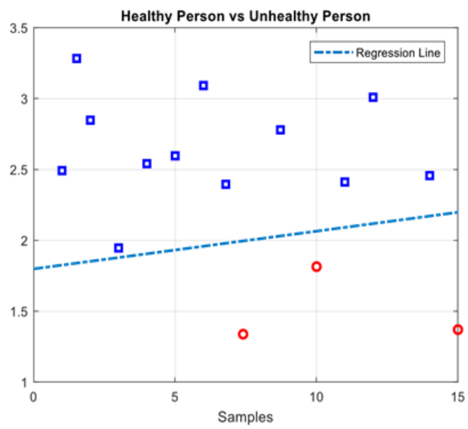

Healthy Person vs Unhealthy Person
Regression Line
Reccurence Rate (RR)
Samples

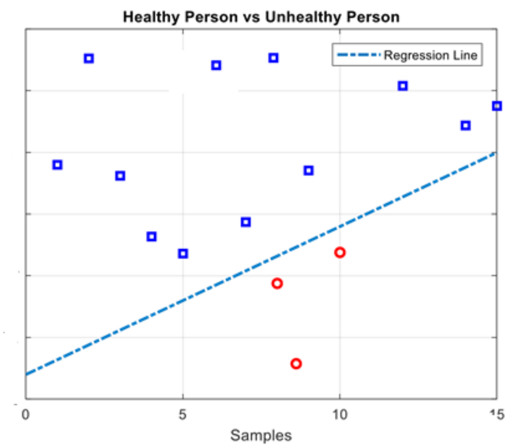
C D


Healthy Person vs Unhealthy Person
Regression Line
Trapping Time (TT)
Samples


Healthy Person vs Unhealthy Person
Regression Line
Trapping Time (TT)
Samples

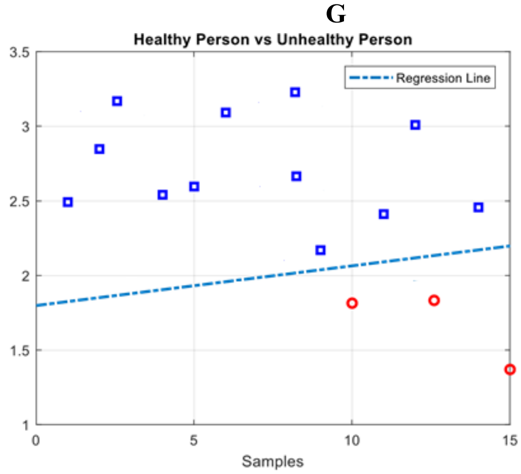
E **F**

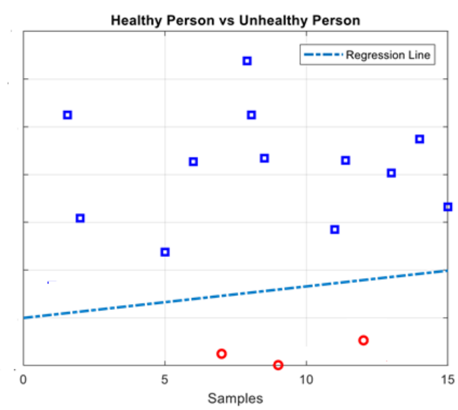




G

H





I

J

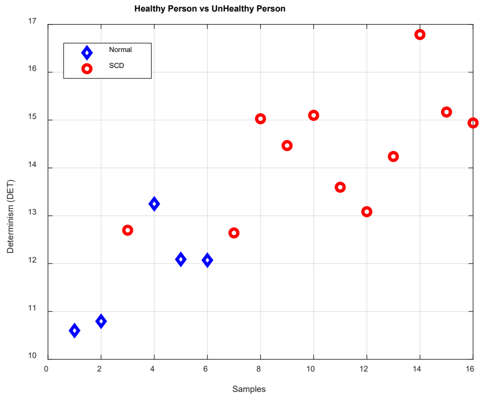
Figure 1. A)Calculation of recurrence rate with a separator line
B) Calculation of certainty along the separator line
C) Calculation of entropy with separator line
D) Calculation of laminarity with a separator line
E) Calculation of trapping time with separator line
F) Calculating the average length of diagonal lines
G) Calculation of the recurrence time of type 2 with the separator line
H) Calculation of the recurrence time of type 1 with the separator line
I) Calculation of the maximum length of diagonal lines along the separator line
J) Calculation of transmittance with a separator line

5. Classification using MLP neural network

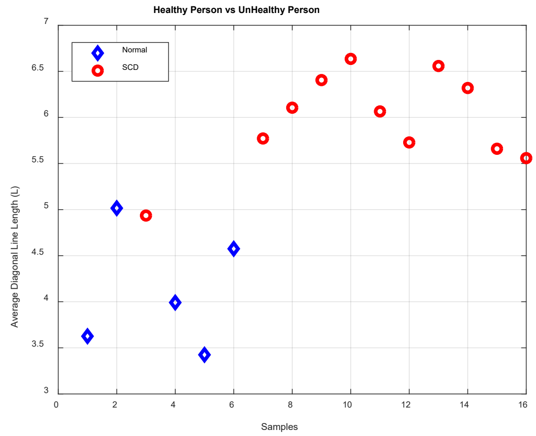
At this stage, a neural network with a hidden layer (with 5 neurons) is used to classify the samples into healthy and unhealthy people. By changing the neurons of the middle layer, it is tried to optimize the network. The result is repeated for many iterations of random values of hidden layer coefficients.

6. Evaluation

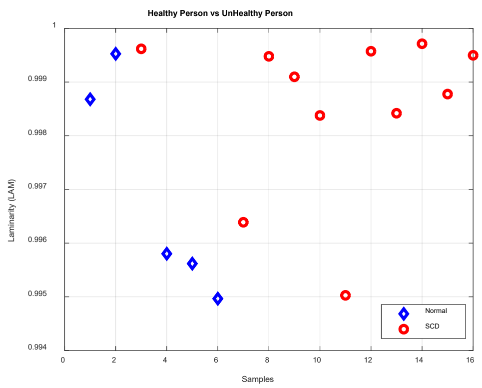
As was seen, RQA features can well indicate the difference between the EKG signal of healthy and unhealthy people.
Figures 2 indicate the method evaluation for RQA features by the neural network.


Healthy Person vs UnHealthy Person
Normal
SCD
Determinism (DET)
Samples

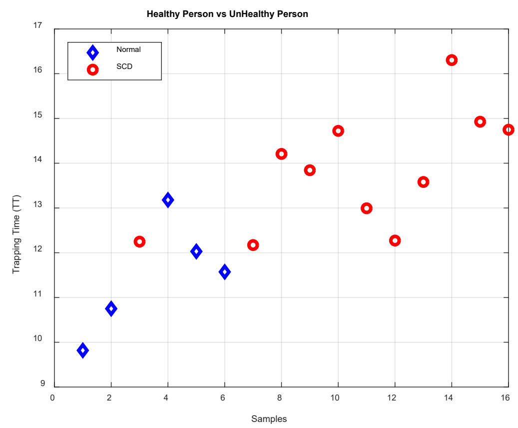
A


Healthy Person vs UnHealthy Person
Normal
SCD
Average Diagonal Line Length (L)
Samples

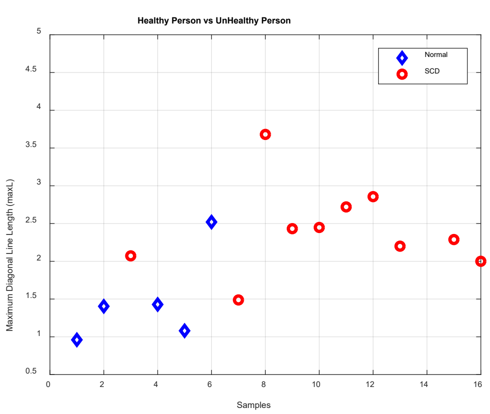
B


Healthy Person vs UnHealthy Person
Normal
SCD
Laminarity (LAM)
Samples

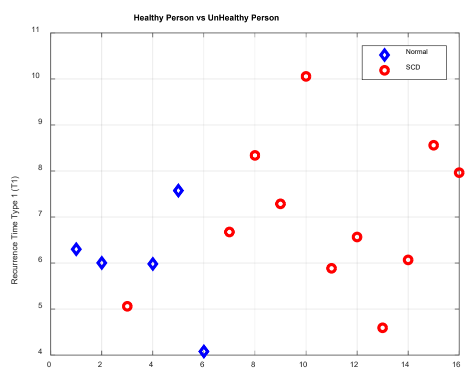
C


Healthy Person vs UnHealthy Person
Normal
SCD
Trapping Time (TT)
Samples

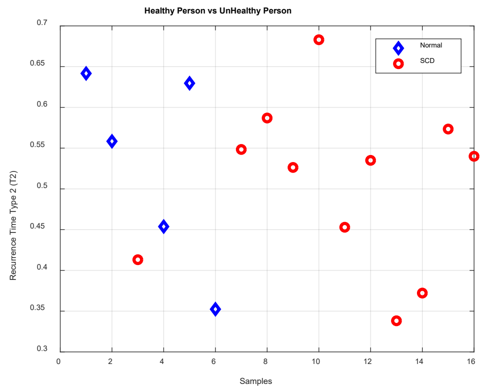
D


Healthy Person vs UnHealthy Person
Normal
SCD
Maximum Diagonal Line Length (maxL)
Samples

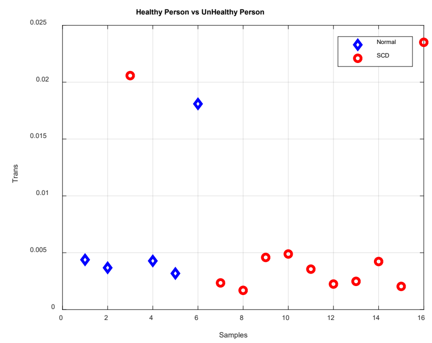
E


Healthy Person vs UnHealthy Person
Normal
SCD
Recurrence Time Type 1 (T1)
Samples

F



G



H

Figure 2. A)Method evaluation for DET feature
B. Method evaluation for L feature
C. Method evaluation for LAM property
D. Method evaluation for TT feature
E. Method evaluation for Lmax feature
F. Method evaluation for T_1 feature
G. Method evaluation for T_2 feature
H. Method evaluation for Trans feature

Discussion

The findings indicate that in the HRV signal, those on the verge of SCD have features significantly different from healthy people. The study method has a remarkable ability to detect this difference. The findings indicate a possibility of danger detection four minutes before the event's occurrence - as we get closer to the scene, the probability of detection increases. A fourth-minute examination before SCD shows that there is still a high risk of SCD. After the end of the fifth minute, the accuracy decreased by 90%. As various linear methods have been used for HRV signal analysis so far, it has recently been proven that nonlinear processing methods can provide more information than conventional linear methods (time domain analysis) and be a good complement to them.

Nonlinear analysis of the HRV signal has received much attention for two main reasons: the nonlinear nature of the signal observed from the heart, which acts as a nonlinear dynamic oscillator, and the need to know enough about this real phenomenon. Some scholars, such as Woss et al., Larver et al., Shen et al., George et al., and Elias et al. reported results for periods less than four minutes before SCD. Using nonlinear and time-frequency analyses of heart rate oscillation signals, Elias et al. predicted the occurrence of SCD up to four minutes before it happens. The mean energy was extracted as time-frequency properties, Poincaré plot, and Detrended oscillation analysis as nonlinear heart rate fluctuation signal features. A multi-layered k-neighbor search and perceptron (MLP) search classification method. An accuracy of 83.96% was observed on their findings for the fourth minute before SCD. Rajendra et al. proposed a method using the discrete wavelet transform and nonlinear analysis of electrocardiogram signals. They divided the electrocardiogram signal to four minutes prior to SCD. For each time interval, nonlinear features were extracted from the combined signal using discrete wavelet transform and then KNN and SVM methods. Classification methods were used to separate cases of SCD from normal cases. They reported 92.11% accuracy (using the SVM method) for four minutes before SCD. Although some SCD predictions have been carried out accurately in some studies, the proportionality of recursive quantification analysis with the nature of turbulent systems provides the necessary interpretations. It analyzes from a chaotic perspective besides using it in automated systems.

This study's performance evaluation for the DET method provides an acceptable predictive response. Based on the realization of each sample in the simulation, the rate of return points based on the prediction set for each sample is an acceptable criterion that compares the total recursion points. Only one case of individuals estimated to be abnormal by this method is within the range of normal individuals. The boundary between the two groups is well distinguishable according to this approach.

L method provides good performance in prediction too. Since this method is the average time when two trajectories are close to each other, it can be interpreted as the average time of forecasting time, which is the average time of predicting in the implementation of this program, which has left only one case. The estimated one should not be in its proper range, and the estimation of normal people from others should be applied with high accuracy and a suitable border.

Although the LAM feature shows lower accuracy than the DET and L methods, it shows a good evaluation in estimation. That is, the ratio between the return points of vertical structures to the total return points in this method is calculated. TT method has a lower performance than the three methods stated in this algorithm and is associated with estimation errors. The Lmax, T_1, T_2, and Trans methods work almost identically, showing high error in one case and performing well in others. Nonetheless, DET, L and LAM methods are good for estimation, and this algorithm and the evaluation can be done based on the DET, L and LAM methods. The DET and L evaluation is the best answer to implement this algorithm if we want to be more precise.

Conclusion

The findings reveal that evaluating the RQA features of the ECG signal can have a significant effect on enhancing the prediction of SCD. In the meantime, using neural networks can help facilitate reaching this. The suggestions for expanding this work in the future are suggested.

Acknowledgments:Non
Conflict of Interests: Non
Ethical Considerations: Non
Financial Disclosure: Non
Funding/Support: Non

References

1. J. Iwaniec and M. Iwaniec, "Heart Work Analysis Means of Reccurence-Based Method" Diagnostyka. 2017; 18(4):89-96.
2. Maron BJ, Bonow RO, Cannon RO. Hypertrophic cardiomyopathy. Interrelations of clinical manifestations, pathophysiology, and therapy (2). N Engl J Med. Apr 2 2016;316(14):844-52
3. Shen, T. W., Lin, C. C., Shen, H. P., Ou, Y. L., and Lin, C. H., "A personal SCD (SCD) detector based on ECG biometric technology". World Congress on Medical Physics and Biomedical Engineering, Vol.14, , pp 1249-1252.2007**.**
4. A.Rashed, U. Mirza, M.J.,"Identification of SCD using spectral domain analysis of Electrocardiogram(ECG)"Emerging Technologies, 2008. ICET 2008.
5. Goldstein S, Landis JR, Leighton R. Characteristics of the resuscitated out-of-hospital cardiac arrest victim with coronary heart disease. Circulation. Nov 2014;64(5):977-84
6. N. Chattipakoin, T. Incharoen, N. Kanlop , S. Chattipakoin, "Heart rate variability in myocardial infarction and heart failure" ,International Journal of Cardiology 120 (2007),EISEVIER
7. Ebrahimzadeh, E., Pooyan, M."Early detection of SCD by using classical linear techniques and timefrequency methods on electrocardiogram signals".Biomedical Science and Engineering, Vol. 4,No.11, pp . 699-706. 2011.
8. Moradi, F., Tartibian, B. (2014). The effect of circadian cycle on the response of ST segment electrocardiogram of athletes's heart to PWC195 maximal test. Research in Sports Science, 23, 49-60