

Cardiac Risk Assessment Based on QTc Speculation and Trending from Past References

Thomas Ho Chee Tat, Chen Xiang

Institute for Infocomm Research, Singapore, Singapore

Abstract

A method to construct a predictive time series index based on QTc-intervals is proposed in this paper. Monitored electrocardiography (ECG) data is converted into a root mean square of successive difference (RMSSD) trend-line by first finding the QT intervals [1] and then using [2]. The trend-line is then used as a priori in extrapolating the predictive trend. The next unknown RMSSD is extrapolated using a modified implementation of the Pearson's Correlation. The last $(m-1)^{\text{th}}$ amount of RMSSD in the priori is the 1st of the 2 time-series being compared in Pearson's Correlation; the second time-series is a running window of m^{th} number of RMSSD from the 1st RMSSD to the $(n-m-1)^{\text{th}}$ RMSSD. The extrapolated RMSSD is substituted into the priori to compute the next RMSSD. The procedures are reapplied recursively until the required time series has been found. Using our method, the average Mahalanobis distance for all the records in the MIT-Arrhythmia Database is 0.1899.

1. Introduction

In this paper, the extension of the previous publication [2] to make it an even more purposeful method for long term cardiac monitoring is proposed. In this method, an approach to create a predictive index based on QTc-intervals and root mean square of successive differences is proposed. The motivation to create a predictive index was based on our observation that the current technologies for cardiac monitoring is already quite advanced and that now there is more than ever a need to create pre-emptive, if not preventive, usable lifestyle solutions to educate people about their cardiac health.

In a short span of time, industrialization and urbanization has made Singapore from a relatively unknown fishing village into a global and modernized city. It is not surprising that related cardiac health problems accompanied and plague our nation. Over the years, there has been a steady increase of sudden cardiac deaths. From people dropping dead for no apparently reasons to healthy and physically trained people dying in the midst of a marathon or a triathlon [3,4], the statistics

are rather astounding. The Singapore General Hospital oversees the post mortem cases that happened in Singapore over the last years. The statistics were as follows for the year 2003 [5].

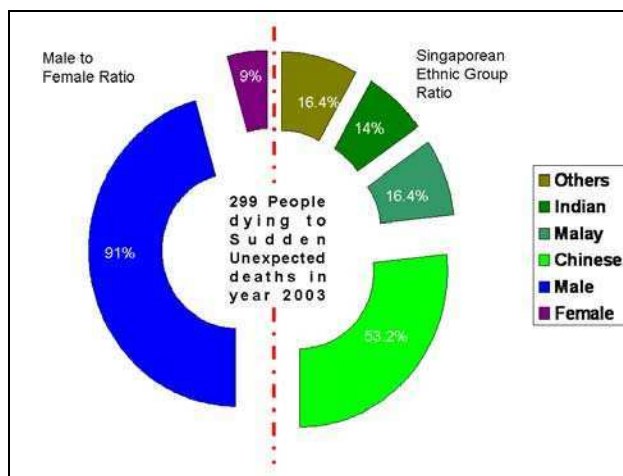


Figure 1. Statistics for the number of sudden unexpected deaths in Singapore in 2003.

In the statistics, the mean age for males is 46 and the mean age for females is 49. And also it was important to note that in the post mortem analysis, 81% of these cases were caused by coronary artery diseases. This implies that all these people who died probably had silent precursors of their conditions that inevitably led to their deaths without priori detection. Hence, this motivated the need to create an application that would be able to take in a limited period of ECG, extrapolate the ECG and surmise with a certain degree of accuracy to determine if one is in danger of suffering from coronary heart disease in the future. The prediction is of course given that the person taking the test is maintaining his or her current lifestyle and habits. And also, if one is found to be in danger of coronary heart diseases, one can be advised to alter his or her lifestyle habits even though it may not seemed to be apparently in the current situation.

In the next sections, the methodology of achieving this index is discussed. After which, the experiments of running the algorithm against the MIT-Arrhythmia

database and its analysis is deliberated upon in section 3. Finally, the paper would be concluded in section 4 together with a future works in section 5.

2. Methodology

2.1. Overview

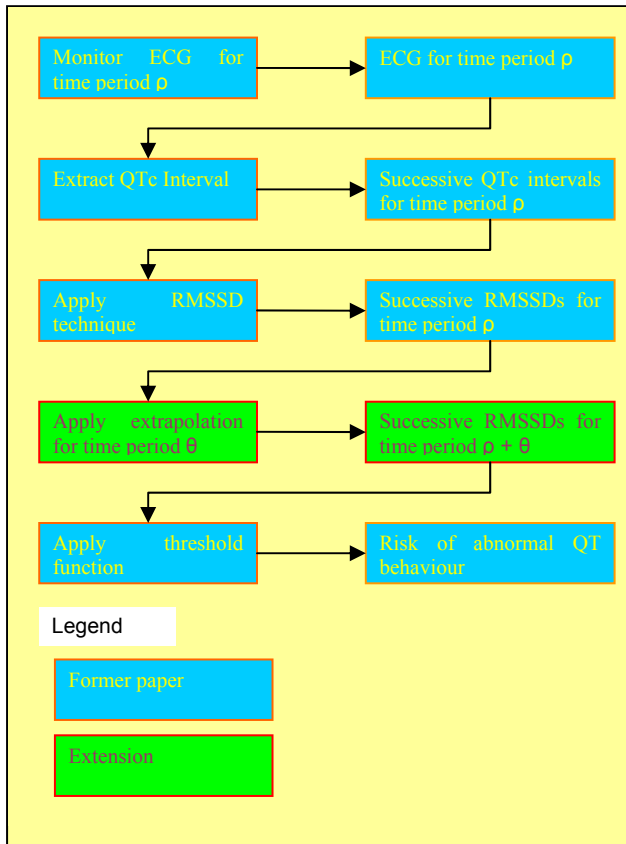


Figure 2. General process of detecting future abnormal risks.

In the previous paper [2], an index that is indicative of LQTS or abnormal QT syndrome during long term monitoring was discussed. The previous paper's shortcoming was that it only focuses on available data. It does offer somewhat an insight to the cardiac well-being of a person but usually the conclusion may have been made at a stage that might already be too late.

Therefore, there is a need to derive unknown data that could perhaps fulfil up this shortfall. It must be re-iterated again that in this paper, we are not looking for an extrapolation of ECG data but rather an extrapolation of the indicative index that was previously discussed. And after the extrapolation, the previous analytical algorithm is reapplied on the new data to determine the cardiac well-being of the person.

The whole process of determining one's cardiac well-being will start with taking the person's ECG for a period

of time. It will then follow by the calculation of the QTc-intervals. Next, the RMSSD method on the QTc-intervals is applied. After the time series of RMSSD are obtained, the previous publication [2] would straight away apply the threshold function to determine the risk of abnormal behaviour. However, in this paper, the extension would now be applied prior to the threshold function.

2.2. Extension

In the extension, there was a need to reason under uncertainty for the creation of the predictive index. This predictive index is based on the previous RMSSD time series. We do want to draw relationships between the current series of RMSSD and the next RMSSD; therefore we would do the prediction one index at the time to maintain the relationship. Hence, given nth number of RMSSD in a sequential time series, the intent is to predict the (n+1)th RMSSD as the extension.

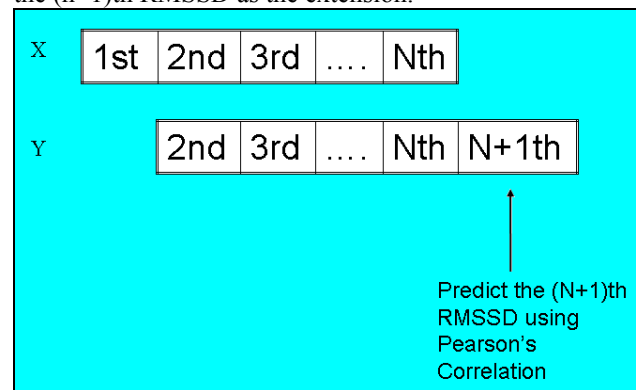


Figure 3. Extension contributed in this paper

By repeating the process over and over again predicting 1 RMSSD at a time, the approach extrapolates a time series of RMSSD based on the current ones.

```

// To predict the next series RMSSD based on prior knowledge aka monitored ECG
Take list of calculated RMSSDs of the QTc Interval and using a window size of 50:
While (Require Prediction)
{
    Take last 49 RMSSDs; //This is the query
    Scan from 1st RMSSD to (N-49)th RMSSD // This is the search window
    {
        Take 50 consecutive RMSSD;
        Take the 50th RMSSD value and concatenate to the last of the query
        Perform Pearson's Correlation;
        Maximize the correlation during the scan;
    }
    Take the last value on the query that produced the highest Pearson's
    Correlation and append it to the search window;
    If length of required prediction is reached, stop while loop else continue
}
  
```

Figure 4. Algorithm to generate the predictive index

The threshold function can then be applied upon the predicted RMSSD to give a risk assessment based on the

QT behavior. A specified time period into the future can be given to the algorithm depending how long into the future it needs to be observed. However, the caveat is that as the distance between the predicted indices and the current time series grow, the probability of the accuracy decreases. This is because each prediction forward removes away 1 current RMSSD from the basis for the prediction. And if this continues, it will reach a state that the time series is predicted based on already predicted values. Hence, it may be presumptuous to assume that no new trends were to surface since the current time series did not capture any trends that might surface in the future due to other health factors and/or commitments.

2.3. Modified pearson’s correlation

The actual use of Pearson’s correlation is to measure of the correlation of linear dependence between two variables X and Y , giving a value between +1 and -1 inclusive. However, in our case, we sought to determine the variable Y such that it’s computed co-efficient together with variable X would exist between +1 and -1 inclusive. The formula needs to be modified and it is not that simple to re-balance the equation and let Y be the unknown to be determined.

$$\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - (E(X))^2} \sqrt{E(Y^2) - (E(Y))^2}}$$

Figure 5. Pearson’s product moment correlation

However, since the allowable RMSSD indices are already pre-fixed in the previous publication [2] and that the computation can be conducted offline; this approach will use substitution and make an exhaustive search on all allowable RMSSD.

Then the next problem would be to determine the coefficient that would be best to use. And in our experiments we have tried 4 methods to determine the best co-efficient that gives the best Euclidean distance.

3. Experiments and results

3.1. Experiments

The prediction algorithm was tested against all the records in the MIT-Arrhythmia database. 4 methods of determining the best co-efficient was tested concurrently to determine the best. The 4 methods were maximizing the coefficient, using the mean co-efficient, using the median co-efficient and using the mode of the co-efficient.

This was possible because during each generation of the next RMSSD, a list of co-efficient is also generated.

Hence, with this set of co-efficient, the approach is then able to select the next RMSSD based on the former mentioned criteria.

Maximizing the co-efficient involves selecting the RMSSD that has its respective Pearson’s coefficient closest to 1. Using the mean involves averaging all the generated Pearson’s co-efficient and selecting the RMSSD that has its respective Pearson’s co-efficient closest to the mean. Using the median involves selecting the RMSSD that has it respective Pearson’s co-efficient closest to the median of the generated set and also similarly in fashion for using the mode.

After each method’s reconstruction, all are matched accordingly to the RMSSD trend generated from the MIT-Arrhythmia database. The results are then collated. Table 1. Euclidean distance measure of the reconstruction

	Maximized	Mean	Median	Mode
Normalized Total Euclidean Distance (Highest Possible score is 48)	9.11	10.64	10.69	10.58

3.2. Results

The reconstruction results of using the 4 different co-efficient to reconstruct the data are shown below.

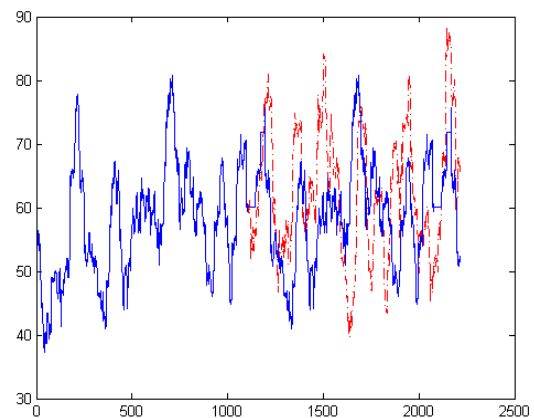


Figure 6. Reconstruction using maximised coefficient

The red-dotted line is the ground truth while the blue line is the reconstructed line. It is observed that by using the maximized coefficient that the best reconstruction of the signal is obtained. Although the reconstruction may seem to be still quite some distant away but it yielded the best reconstruction. It proved that the approach to reconstruct the trend-line under uncertainty is feasible. It

also proved that this modified Pearson's product moment correlation is feasible in this approach.

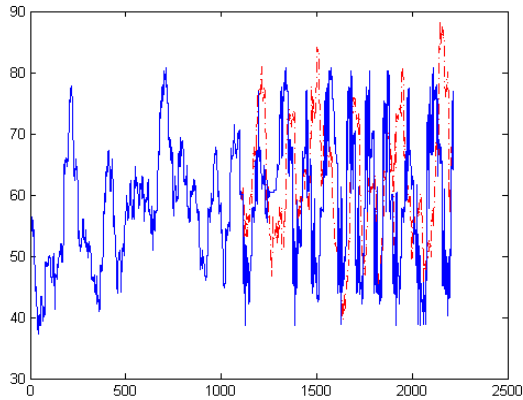


Figure 7. Reconstruction using mean coefficient

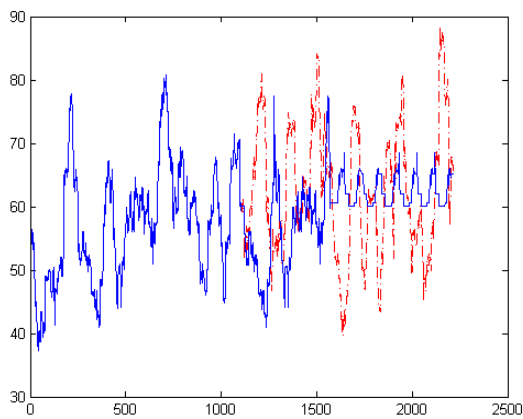


Figure 8. Reconstruction using median coefficient

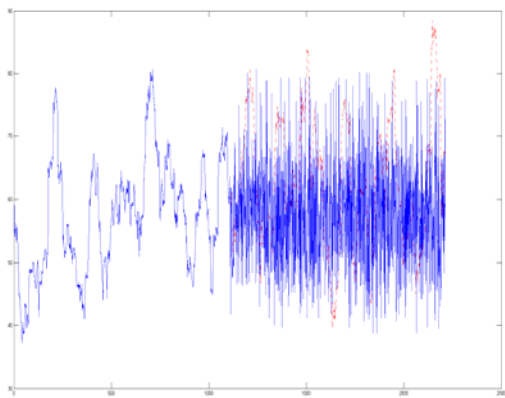


Figure 9. Reconstruction using mode coefficient

4. Conclusion

Our approach of creating a predictive index is able to predict the trends with certain amount of accuracy. The best average Mahalanobis distance for the MIT-

Arrhythmia database that is 0.1899. The best reconstruction is obtained by maximizing the coefficient factor in the Pearson's correlation equation as shown below from a record collected from a random co-worker.

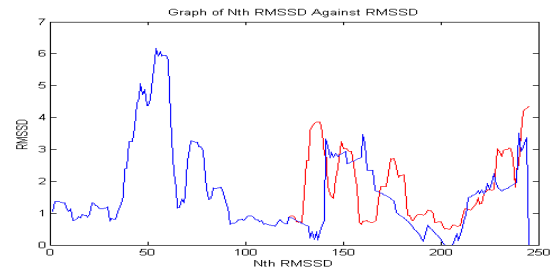


Figure 10. Reconstruction of real life QTc RMSSD trend line using maximized Pearson's co-efficient

5. Future works

This approach for the creation of a predictive indexed based on actual cardiac signals is still very preliminary. As with all applications dealing with health and general well-being, medically or non-medically approved, any marginal of error is unacceptable. Therefore, moving forward there will be more research and evaluation being conducted to close the gap of 0.1899 of Mahalanobis distance with the MIT-Arrhythmia database.

Next, this approach would be useless if not developed into a usable application. As such using our current expertise in micro devices, we would embed this approach into our current mobile applications and promote it for feasibility studies on the masses

References

- [1] Lim E.T, Chen X, Tin Z.K, Meena S, "Smart Phone-based Automatic QT Interval Measurement" 34th Computer in Cardiology, Sept 30-October 3, 2007. Durham, North Carolina, USA
- [2] Ho C.T.T, Chen X, "An Augmentative and Portable QTc-Observer(QTO-Q2) to Facilitate More Purposeful Outpatient Monitoring" 37th Computing in Cardiology, Sept 26-29, 2010, Belfast, Northern Ireland, United Kingdom
- [3] Army regular collapses and dies after half-marathon, <http://www.asiaone.com/Health/News/Story/A1Story20070826-23153.html>
- [4] Know your limits - and stay alive, http://www.straitstimes.com/print/Free/Story/STIStory_134689.html
- [5] Sudden Cardiac Death - Singapore's Growing Problem, <http://www.nhcs.com.sg/aboutnhc/Newsroom/publications/year2007/Documents/NHCNewsMar07.pdf>

Address for correspondence.

Thomas Chee Tat Ho
 1 Fusionopolis Way, #21-01 Connexis. Singapore 138632
 E-mail: ctho@i2r.a-star.edu.sg