ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

ELECTRICAL-ELECTRONICS ENGINEERING

DEPARTMENT

EE400 SUMMER PRACTICE REPORT

Student Name: Defne ODABAŞI

Student ID: 2443604

SP Company Name: Eindhoven University of Technology

University Department: Biomedical Engineering

SP Beginning-End Dates: 10.08.2023-15.09.2023

Supervisor Name: Carlijn Buck

Supervisor's Contact Information:
c.m.a.buck@tue.nl

Contents

1 Introduction

I have performed my summer practice in the Biomedical Department of Eindhoven Technical University in Netherlands, one of the leading universities in research and education. My internship lasted for 10 weeks and 50 working days. The reason for me to intern in a research laboratory is my desire to experience such a research environment. In my research group we were with 4 people, a professor, a PhD student of the professor and medical engineer who is a researcher at the Catharina hospital.

The beginning point for the research is that abnormalities in cardiac depolarization and repolarization as measured from 12-lead ECG, have been closely linked to adverse clinical outcomes. Traditional methods such as examining the duration of QRS and QT interval, have long been employed to analyze these abnormalities. However, the effectiveness of these methods has been questioned due to various methodological limitations and the perceived inefficiency of predicting capabilities for ventricular tachycardia (VT).

In the light of these challenges, in my summer practice, I aimed to explore new descriptors that can enhance the understanding of these cardiac abnormalities. I particularly focused on assessing the development of VT in post-myocardial (MI) patients using beat-to-beat Vectorcardiographic parameters. Specifically, this research seeks to calculate 'Total Cosine R to T (TCRT)' and 'T-wave Morphology Dispersion (TMD)' in a beat-to-beat manner.

This research builds upon the foundation laid by Margot Reijnen in her thesis, 'Analysis of electrocardiogram- and vectorcardiogram-derived parameters in post-myocardial infarction patients.'. In this report theoretical background, the methodologies applied in signal processing, parameter calculations as well as the presentation of the results and discussions are included. At the end, the code written for this research is provided.

2 Theoretical Background

2.1 The Vectorcardiogram (VCG)

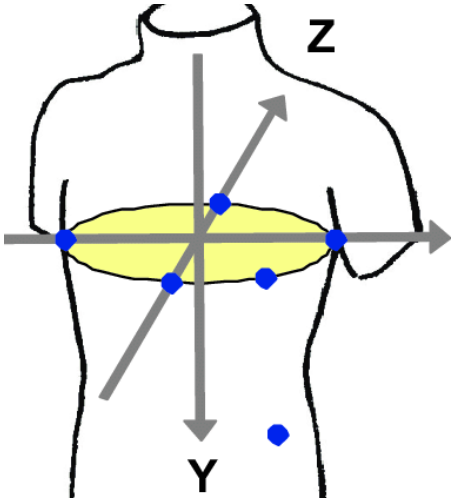
VCG aims at a spatial representation that reflects the electrical activity of the heart in the three orthogonal directions X, Y and Z [1]. While the standard 12-lead ECG represents magnitude information of the heart, VCG representation includes magnitude and spatial information of the heart signal.

2.1.1 VCG from Frank Lead

In Frank VCG lead system, seven electrodes are placed on the patient's body. Even though the placements of the electrodes are considered non-traditional, this configuration provides a three-dimensional representation of the heart's electrical activity.

The Frank VCG lead system measures electrical signals in three orthogonal directions mainly, from right to left (X axis), from head to feet (Y-axis) and from front to back (Z-axis) creating a spatial understanding

of the electrical activity on heart. The lead placements for the 3 orthogonal components are illustrated on a human body model as in Figure 1.



(Thygesen, 2023)

Figure 1. Lead placements show the 3 orthogonal components, X, Y and Z are derived [8].

In today's world, the Frank lead system is no longer used in clinical practice. The main reason for this is considered as the non-standard lead configuration of the practice. As an additional note, this practice is even more difficult for patients who are bed laid and it is very common for the patients to be bed laid in hospitals.

2.1.2 VCG from 12-lead ECG

In the derivation of VCG from standard 12-lead ECG, Kors Regression Transformation matrix is used to obtain VCG leads. Matrix multiplication is performed with eight-lead ECG (I, II, V1, V2, V3, V4, V5, and V6) and the coefficients of the Kors Regression Transformation Matrix. The result of the matrix multiplication is orthogonal Frank leads (X, Y and Z). The transformation coefficients of the Kors Regression Matrix are shown in Table 1.

Table 1 Kors Regression Transformation Matrix Coefficients.

Leads	I	II	V1	V2	V3	V4	V5	V6
X	0.38	-0.07	-0.13	0.05	-0.01	0.14	0.06	0.54
Y	-0.07	0.93	0.06	-0.02	-0.05	0.06	-0.17	0.13
Z	0.11	-0.23	-0.43	-0.06	-0.14	-0.20	-0.11	0.31

The frontal, sagittal, and transverse planes of the human body can be visualized using the three orthogonal leads of the Frank VCG. This gives a 3D graphical representation of the phases of cardiac cycle in the form of several loops. In addition to the Kors matrix, this study also utilizes singular value decomposition (SVD) following the approach by Acar et al. (1999). SVD is employed to reduce the number of dimensions in the input matrix while preserving as much variability as possible. In Figure 2, VCG of a control group patient is illustrated.

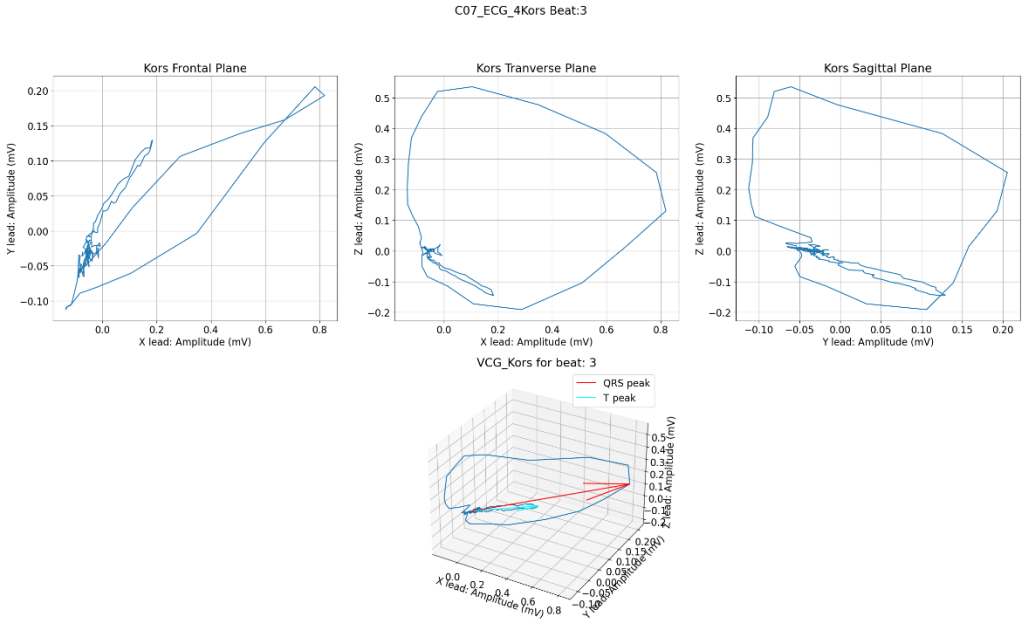


Figure 2 VCG Kors plot is illustrated. 3 figures on top shows the projected VCG plots. On the bottom 3d VCG is represented.

2.2 Beat-to-Beat Parameter Definitions

In this project, two VCG-based parameters are utilized to differentiate between control and VT group patients based on their electrocardiogram (ECG) data. These chosen data sets for the patients turned out to be significantly different in the C and VT group based on the analyzed parameters for the median signal. The first parameter is selected as Total Cosine R to T (TCRT). This parameter assesses the alignment between depolarization (QRS complex) and repolarization (T-wave) vectors. The computation is performed by taking the mean of all cosines of the angle between the vector carrying the maximum T-wave energy and vectors in the QRS complex within an interval where the ECG energy is higher than 70% of the maximum energy within a beat. The second parameter is T-wave morphology dispersion (TMD). This parameter represents the differences in T-wave morphology in different ECG leads during ventricular repolarization. After calculations are conducted for every beat within a rhythm ECG, the findings are kept in a DataFrame ,a data structure in Python, for further analysis.

3 Methods

The aim of this research is risk assessment of the development of ventricular tachycardia (VT) in post-MI patients using beat-to-beat Vectorcardiographic parameters. This section focuses on computational methods to derive beat-to-beat VCG-based parameters such as TCRT and TMD. This section starts with the dataset used in this research and followed by how automatic beat-to-beat analysis is conducted in this research.

3.1 Dataset

The data employed in this research is a subset from the IMPACT (IMPlementation of Artificial intelligence in Cardiology, a Test study) study, which serves as an initial investigation for the COMBAT-VT project [3]. All data used has been retropspectivly retrieved from and consecutively anonymized by Catharina Hospital Eindhoven. The study group comprises 46 post-myocardial infarction (MI) patients, among whom 36 patients did not experience ventricular tachycardia (VT), while 10 patients did. Inclusion criteria encompassed patients aged 18 or older who provided informed consent. VT patients were included based on specific criteria, including a history of ST-elevation Myocardial Infarction, treatment by ablation for ischemic VT at Catharina Hospital Eindhoven in 2019 or 2020, and the availability of various medical data. Similar criteria were applied to patients who did not develop VT (referred to as control patients), with the exception of ablation treatment for ischemic VT. Control patients who passed away prior to 2019 were also excluded. The pilot study encompassed 1253 electrocardiograms (ECGs), each with a duration of 10 seconds.

3.1.1 Data Selection

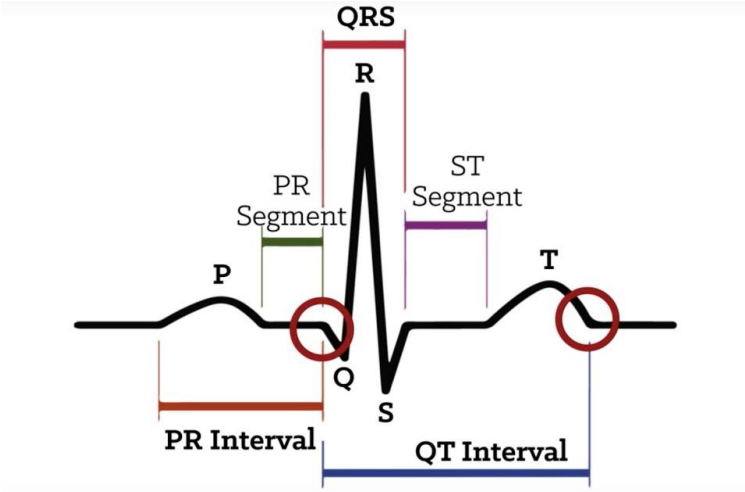


Figure 3 P-wave, QRS-wave and T-wave representation for a single beat

Criteria were defined to determine which data files from the pilot study align with the study's objectives. Firstly, since the study revolves around post-myocardial infarction (MI) patients, data acquired outside the MI-VT interval was excluded. This entailed removing all files obtained before or on the day of the MI event and after or on the day of the first VT event occurrence. For patients in the control group, the date of data reception was considered as the maximal date, so their ECG acquisition dates were required to fall within the MI to data receipt interval. This step led to the elimination of 580 ECG recording files. Secondly, the presence of specific fiducial points, namely QRS onset, QRS offset, and T offset as

represented in Figure 3 are calculated by The Marquette™ 12SL™ ECG Analysis Program (GE Healthcare, Milwaukee, WI), was crucial for subsequent parameter calculations. Data files lacking these fiducial points were excluded, resulting in the removal of 85 files from the analysis. Following this, all remaining files meeting the earlier criteria underwent visual inspection and individual review. An additional 216 files were excluded from the analysis due to issues like disturbances, incomplete leads, fewer than six remaining ECG recordings for the patient of interest, or the presence of pacing during ECG recording. Given the patient population consists of post-MI patients, the presence of pacing was not uncommon. However, ECG recordings with pacing artifacts were unsuitable for analysis and parameter calculation in the vectorcardiogram (VCG). To ensure multiple ECG measurements over time were available, a minimum of six ECGs were required. The ultimate selection of data files per patient over time is depicted in Figure 18 in Section 5. In the end, this process resulted in the inclusion of 45 ECG recordings from four patients who developed VT and 327 ECG recordings from 24 patients who did not develop VT, meeting the established criteria. This subset comprises three men and one woman in the VT group, while the non-VT group consists of 20 men and four women, with ages ranging from 55 to 86 for control patients and 69 to 87 for VT patients at the time of infarction. The mean age for the control and VT groups was 70.8 ± 9.40 years and 71.8 ± 6.65 years, respectively, indicating comparability. Additionally, the time elapsed between the MI event and VT development ranged from 4 to 24 years for VT patients.

3.1.2 Preprocessing IMPACT Study Dataset

ECG recordings are included in the data from IMPACT research and are given as files with .xml extensions. A python (3.8.13) function was created to translate the Base64 bit encoded data into numerical data in the previous study. This was to guarantee the proper data format. The ECG recordings consist of 8 separate lead recordings including Lead I, Lead II, as well as Leads V1 through V6. Each ECG recording has two distinct ECG signals: the median ECG and rhythm ECG. The rhythm ECG is recorded for 10 seconds while the median ECG contains only one beat. The development of the median ECG is explained under section 3.2.1. Additionally, Marquette™ 12SL™ ECG Analysis Program (GE Healthcare, Milwaukee, WI) is used to filter the ECG data. The median ECG, sample frequency, date and time of acquisition, gender, QT interval, and fiducial points were all extracted from the ECG recordings.

3.2 Automatic Beat-to-Beat Analysis

The automatic beat-to-beat analysis is the core of this study. In that manner, complexes within each rhythm ECG recording meticulously investigated to be used in the parameter calculations. When the analysis for the parameters is performed using a single beat (median beat), a static point of view for the repolarization abnormalities is established. On the other hand, observing and performing parameter calculations for each beat in the rhythm ECG enable us to understand the dynamical changes within the ECG.

3.2.1 Median ECG

The median ECG beat is used to derive parameters such as T wave offset for the rhythm ECG. This median ECG per recording was generated by Marquette™ 12SL™ ECG Analysis Program (GE Healthcare, Milwaukee, WI) and it is considered as the most informative beat for further analysis [4]. In the process, the program selects a specific beat type as the most informative beat type with a minimum of 3 complexes in the rhythm ECG. This is considered as the 'primary beat type'. Then, utilizing sample templates produced by the QRS detector (GE Healthcare Wilwaukee, WI), all beats with this kind of morphology are time aligned [4]. The median beat is determined by calculating the center value of all beats designated as the primary beat at each sample time after superimposing these beats [4].

3.2.2 T-wave Offset

To accurately determine the T-wave offset points for each beat in a rhythm ECG is a real challenge in this project. Nevertheless, it is crucial to pinpoint T-wave offset points for each beat since TCRT parameter depends on this point which will be explained in further details in the following sections.

In Figure 4, the process of finding the T-wave offset points for the rhythm ECG is illustrated with a flowchart. In the flowchart it can be inspected that T-wave Offset values for the beats in a rhythm ECG is estimated through the information gathered from the median ECG. In this project T-wave offset point for median ECG as well as R-wave top points for the rhythm ECG are extracted from the .xml files which are obtained by MUSE (GE Healtcare).

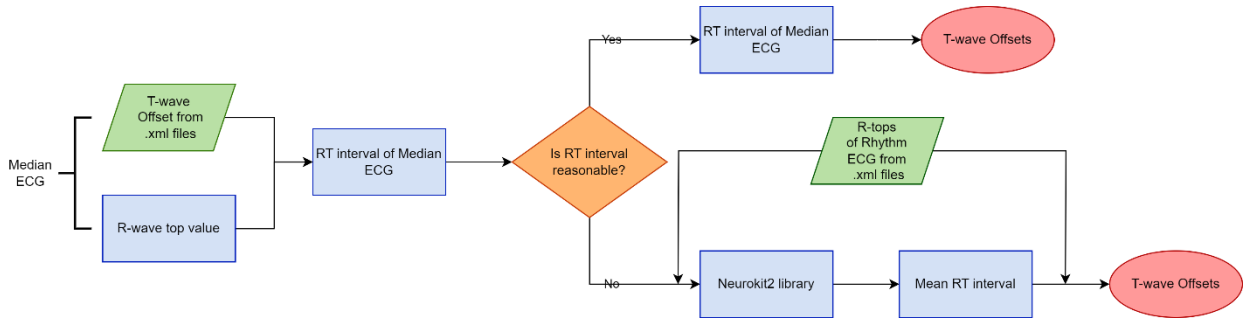


Figure 4 Flowchart of T-wave Offset Detection for beats in the Rhythm ECG

Firstly, the R peak of the Median ECG signal is detected automatically by identifying the maximum signal value of the median signal. The difference between R-wave top and T-wave offset value is taken for the median ECG. From R-wave top to T-wave offset interval abbreviated as RT is assumed to be more or less the same throughout all the beats in the same rhythm ECG. For each beat in the rhythm ECG, the RT value obtained from the median ECG is added to the corresponding R-wave top point. The resultant sum represents the estimated T-wave offset point for each beat in the rhythm ECG. It is useful to note that

the same RT interval is used for every beat within the same ECG, yet different ECG recordings may have different RT intervals.

Upon further analysis among the patients, it is found that using median ECG for determining T-wave offset is consistent in most cases. However, for a small number of ECG recordings the T-wave offset value appeared to be either earlier or later than expected. This resulted in errors or wrong fiducial point calculations followed by inconsistent TCRT and TMD results. For instance, in Figure 5, it can be seen that T-wave offset point is estimated to be earlier than the observation of the T-wave in every lead. The red dot in the figure is the expected T-wave offset point for the particular beat. This wrong calculation is followed by an error in TCRT calculation.

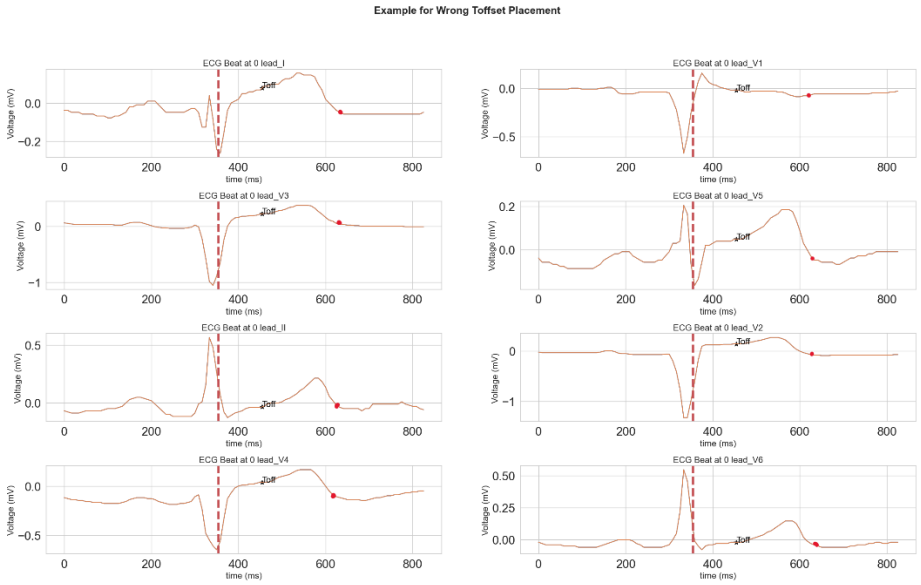


Figure 5 Incorrect T-wave Offset Placement shown with the '' and the expected T-wave offset placement is illustrated with a '.'.*

Recognizing the potential errors or inconsistencies in the further analysis, neurokit2 library which is a reliable tool for ECG signal processing is included in the analysis. This library takes the rhythm ECG signal and attempts to detect onset and offset points for the QRS complexes. Since the R-top values are known for the rhythm ECG, RT differences are collected among these beats and the average is taken. If this average RT difference significantly deviated from the corresponding value derived from the median ECG, the average RT difference found from neurokit2 library is used as the reference RT interval for the rhythm ECG. In Figure 6, the corrected T-wave offset for the same beat mentioned above can be observed.

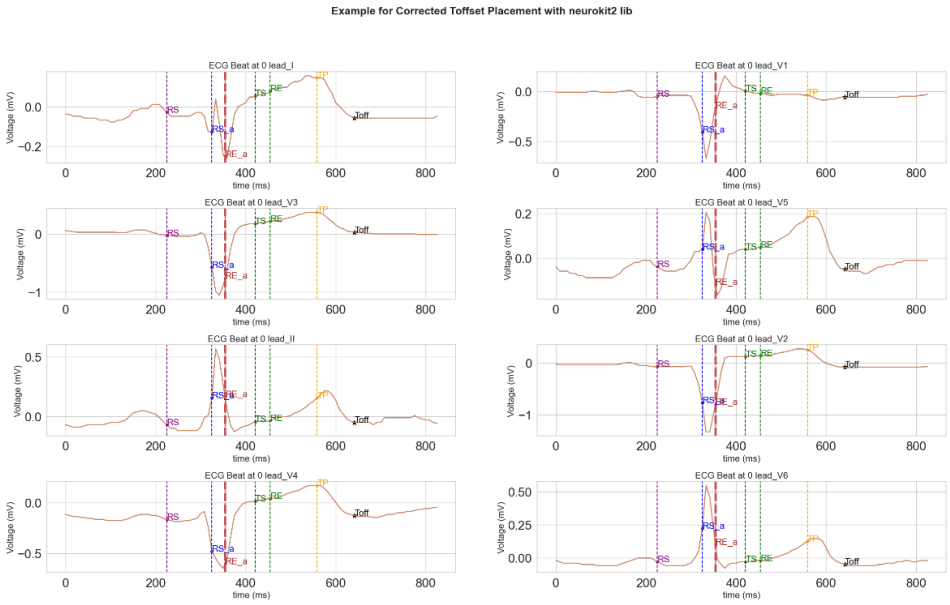


Figure 6 Corrected T-wave Offset Placement with neurokit2 library

Finally, it is important to note that neurokit2 library may fail to find the onset and offset points for QRS complexes. Therefore, it is only used as a second source to validate and hopefully correct the RT difference in the calculations.

3.2.3 Pan-Tompkins QRS Detection

In this project, TCRT and TMD parameters are automatically calculated in a beat-to-beat manner; thus, it is essential to segment the beats within the rhythm ECG by detecting each QRS complex effectively. The first approach explored for this task involved implementing Pan-Tompkins Algorithm.

The Pan-Tompkins Algorithm is a useful method for the detection of QRS complexes, it involves a serie of steps designated to smooth the signal and then amplify the QRS complexes. Since the ECGs used in this project are relatively clean and mostly noise-free, bandpass filtering is not applied to the signal as the initial step.

Firstly, the ECG rhythm signal is differentiated in order to emphasize the sharper peaks such as R peaks in the signal. Then the resulting signal is squared to further enhance sharper edges throughout the signal. As a result, smooth edges are attenuated while sharp edges are highlighted with the help of differentiating and squaring the signal.

Secondly, the moving average window method is applied. This step is necessary since the output signal of the previous signal still contains sharp edges which makes the QRS complex detection difficult. The logic behind this process is that by the moving average window, the nearby peaks merge together and cancel out the sharp points. In the implementation, a rectangular window is convolved with the signal.

Mathematically the operation can be shown as,

$$(s_{ECG} * W)[n] = \sum_{m=0}^{N+M+1} s_{ECG}[m]W[n-m]$$

Here, W is the moving window with length M and N is the input signal length and n are the sample points.

The window size for the operation is chosen as 0.15 times the Sample Frequency of the signal. This window size ratio is found to be appropriate for the signals in this project. The resulting signal from the moving average method possesses a slight phase shift.

Finally, having completed the steps regarding signal enhancement, peak detection is required to find the location of QRS complexes. The threshold for peak detection is set to mean of the signal, meaning that mean value of the signal can be considered as a minimum reference point for identifying the significant peak. In addition, distances between the peaks are determined to be at least 200 ms away for each detected QRS complex. This constraint is set since the minimum time between two consecutive QRS peaks are usually not closer than 200 msec physiologically. In Figure 7, step by step demonstration of how this algorithm does work is demonstrated.

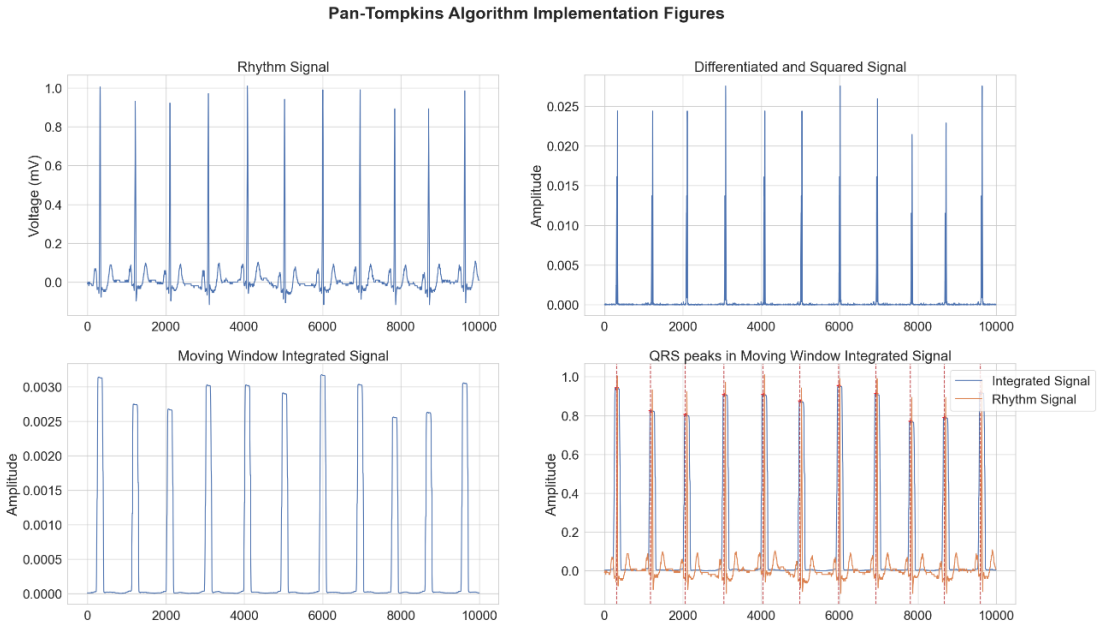


Figure 7 Step by step Pan-Tompkins Algorithm Implementation

Even though using Pan-Tompkins algorithm in the detection of QRS complexes gives promising results for some patients, it is detected that non-standard ECG morphologies in some patients' ECG result in inaccurate QRS complex detection. Figure 8 , provides a visual comparison highlighting two significant issues. The first issue is that the algorithm occasionally introduces false peaks which is not desired. The

second issue is that algorithm may exhibit phase differences relative to true peak locations. Therefore, using Pan-Tompkins method for beat segmentation is abandoned in the earlier phase of the project.

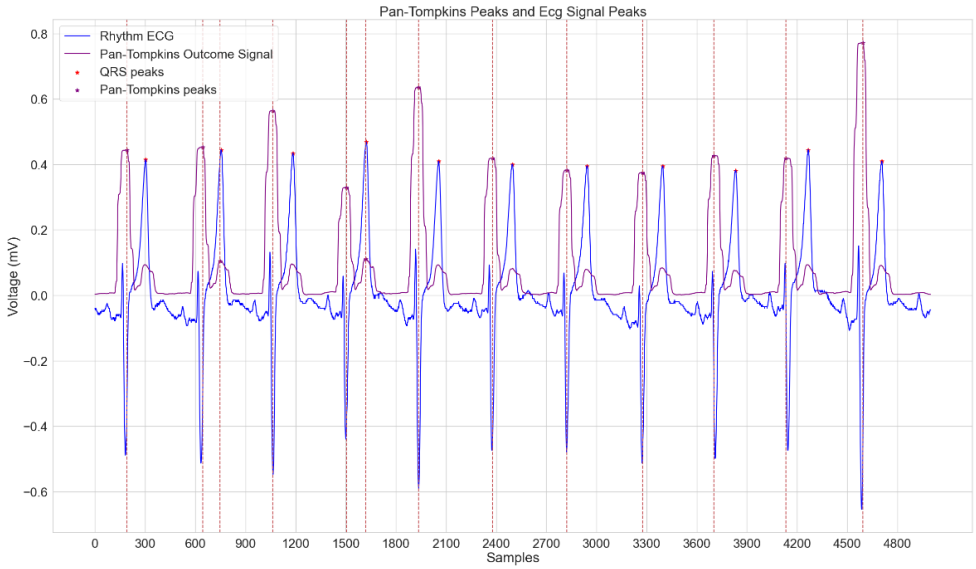


Figure 8 Pan Tompkins Algorithm Results for peaks and Rhythm ECG peaks

3.2.4 Beat Segmentation

Beat segmentation plays a very important role in this implementation, primarily because a different approach is employed compared to the previous study [6]. In the previous study, the calculations for TCRT (Total Cosine R to T) and TMD (T-wave Morphology Dispersion) relied on analyzing the median beat which is determined by calculating the center value of all primary beats also explained under the subsection 'Median ECG'.

However, the objective of this research is to extend the analysis by including all the beats within a rhythm ECG. This expansion introduces a critical requirement for the study, which is accurate beat segmentation. Identifying the boundaries of QRS complex is important for determining the fiducial points within the complex. Accurate fiducial point calculation within the complex forms the foundation of calculating TCRT and TMD parameters in an effective way. Therefore, this study puts an emphasis on correct beat segmentation as well as ensuring the boundaries within the complex are reasonable.

The segmentation process is demonstrated by a flow chart as in Figure 9. Firstly, the segmentation starts with the extraction of R-top values of the rhythm ECG from .xml files. Then, the distances between consecutive beats are computed by calculating the intervals between successive R-top values, referred to as RR intervals. The RR intervals lay the foundation of the beat segmentation.

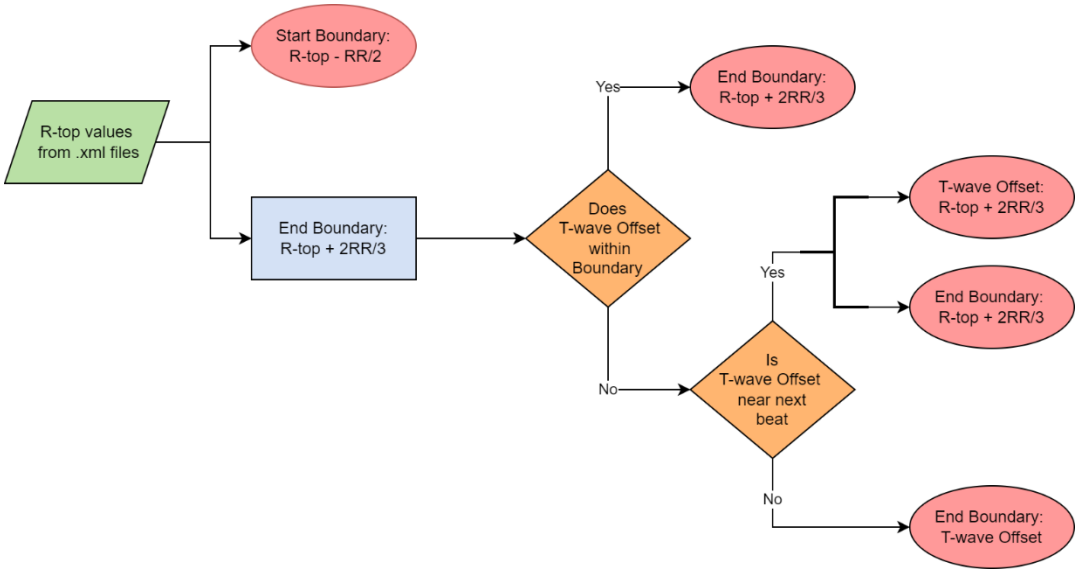


Figure 9 Beat-Segmentation flow chart

The starting boundary for each beat is selected to be in a distance RR/2 prior to the analyzed beat. This choice is made to ensure that the P-wave is included within the boundaries of the analyzed beat. Concurrently, the ending boundary is strategically selected at a distance 2RR/3 away from the beat's R-wave top value. The end point decision, different from setting the end boundary as RR/2, is observed to increase the likelihood of capturing T-wave within the boundaries of the beat. In Figure 10, a visual illustration for RR intervals can be observed for 3 consecutive beats. It should be noted that R-top values extracted from .xml files are not exact but yet very close to the real R-top instances.

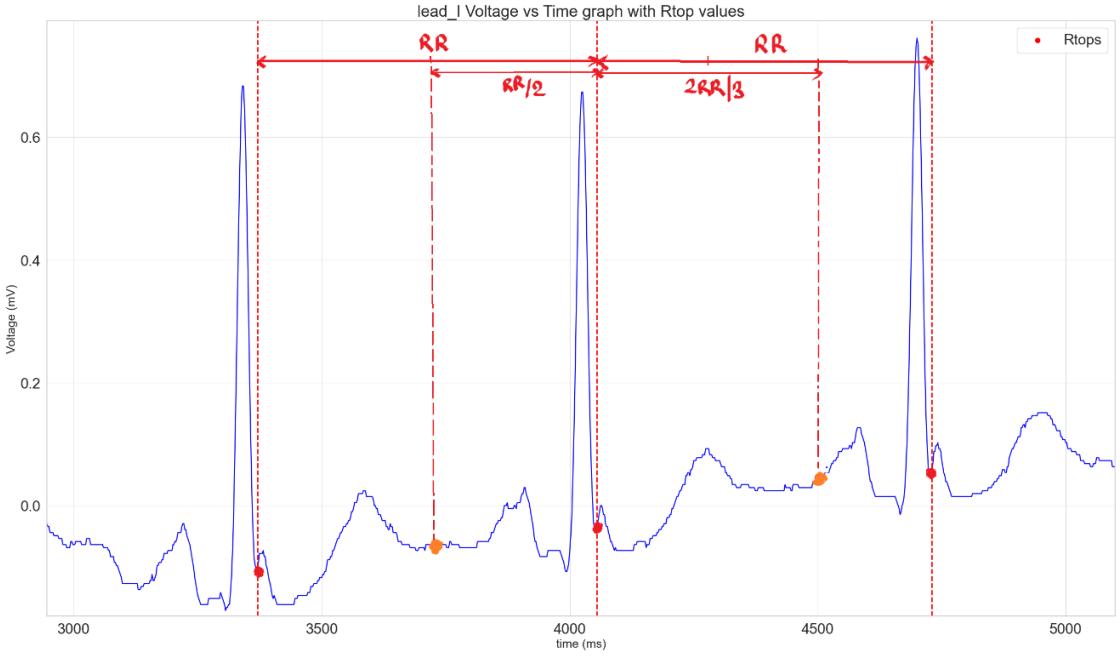


Figure 10 RR interval representation

After determining the intervals for the boundaries, a pivotal assessment regarding whether the T-wave offset point resides within the confined boundaries of the beat or not is conducted. This is considered as an important step since the presence of the T-wave within the boundaries can be controlled through checking for the T-wave offset within this boundary. When the result from the conducted assessment is that the T-wave offset instance does reside within the boundary, then the ending boundary for the beat can be safely set to 2RR/3 distance.

However, it is observed that in some cases T-wave offset appears in close proximity to the subsequent beat's R-top point as demonstrated in Figure 11 . This raises the question of 'Does T-wave offset set to be too close to the next beat?'. Even though T-wave offset placement for the beats is conducted elaborately as mentioned under 'T-wave Offset' section, misplacements of the T-wave offset point is possible for some recordings. The reasons for these misplacements may vary for recordings but it is essential to understand that these recordings may belong to patients with abnormalities in their ECGs. Going back to the question; if the answer to the question is a positive, then it means that T-wave offset point is not set correctly and it is changed to 2RR/3 value as the end point. On the other hand, if the answer is negative, then this means that T-wave offset point is reasonable and the ending boundary is set to be the T-wave offset point. An example for such a case is illustrated in Figure 12, it can be inspected from the figure that the first beat is abnormally close to its consecutive beat. If the end boundary is taken as 2RR/3, then this distance from R-wave top would not include the T-wave of the observed beat. Thus, the ending boundary for this beat is set as the T-wave offset point.

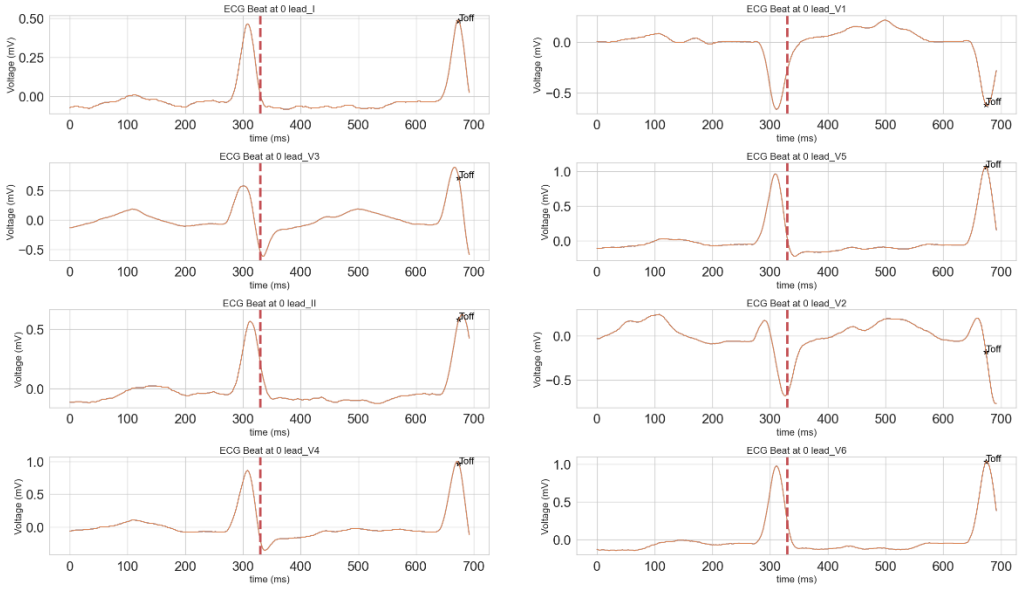


Figure 11 Example of T-wave Offset value close to consecutive beat's R-top

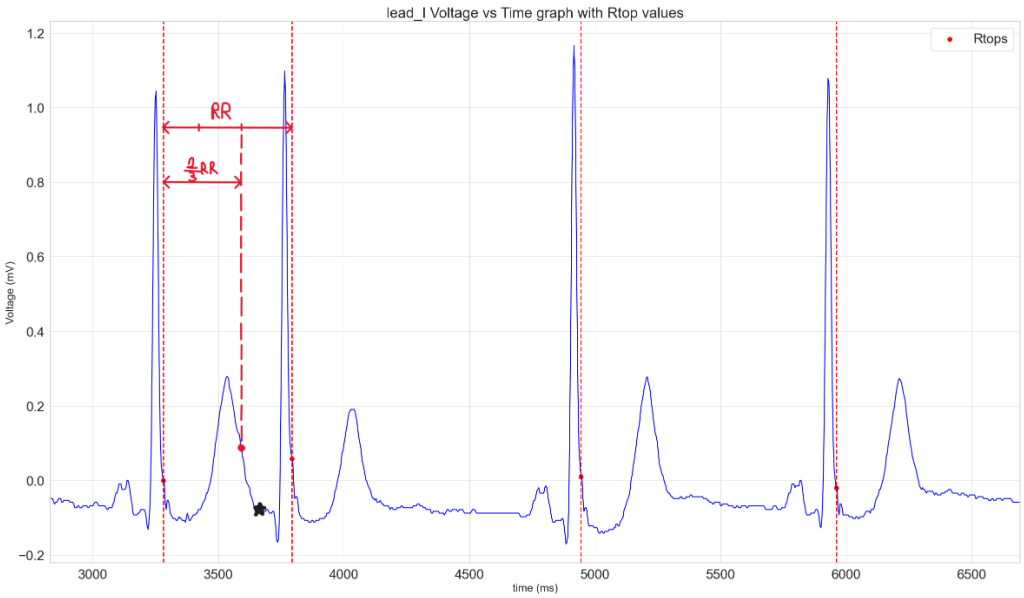


Figure 12 Example of setting ending boundary to T-wave offset instance to include T-wave in the analysis for he first beat. T-wave point is pinpointed with '' and the 2RR/3 point line is shown.*

Lastly, it is worth to note that the entire beat segmentation process is conducted with conditional statements within the algorithm. This methodology and logic are considered to guarantee the inclusion of P, Q, R, S and T waves while taking into account the various quality and morphologies in ECG recordings. As a result, precise TCRT and TMD parameter calculations are possible in a beat-to-beat manner.

3.2.5 Beat Exclusion

Regarding beat-to-beat calculations, it is important to ensure that each beat under analysis contains all the essential components for accurate parameter calculations. In some cases, particularly concerning the last beat of a rhythm ECG, the signal may be truncated before the full morphology of the beat is observed. An example for such a case is demonstrated in Figure 13. It can be seen that the incomplete beat is lack in T-wave morphology.

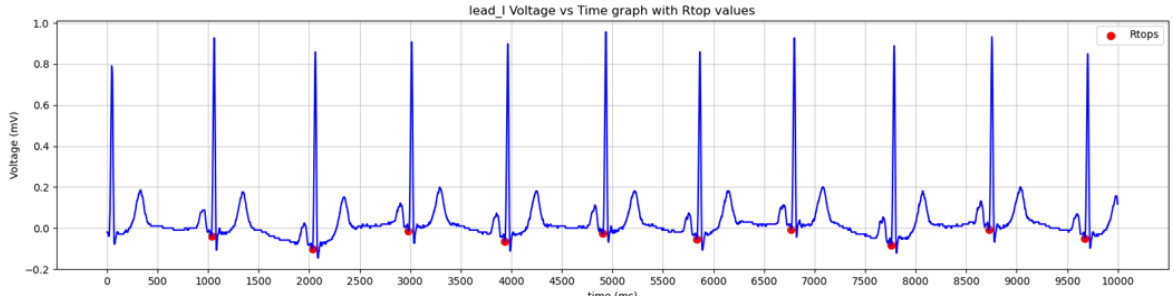


Figure 13 Incomplete beat Morphology Example in a full Rhythm ECG

In order to address this issue as well as maintaining the integrity of the calculations, beat exclusion is conducted for the last beat. In other words, if the last beat in the rhythm ECG does not fully include the T-wave, then this beat is discarded from the parameter calculations automatically.

Another example for beat exclusion is conducted due to abnormalities in the ECG signal within a specific lead. One scenario that is observed in a few recordings is that when an electrode attached to the patient's body is displaced, the potential difference recorded by this electrode may change suddenly. For instance, in Figure 14 it is shown that signal in lead V2 shows an abnormal potential rise. Thus, two beats from this rhythm ECG are excluded manually.

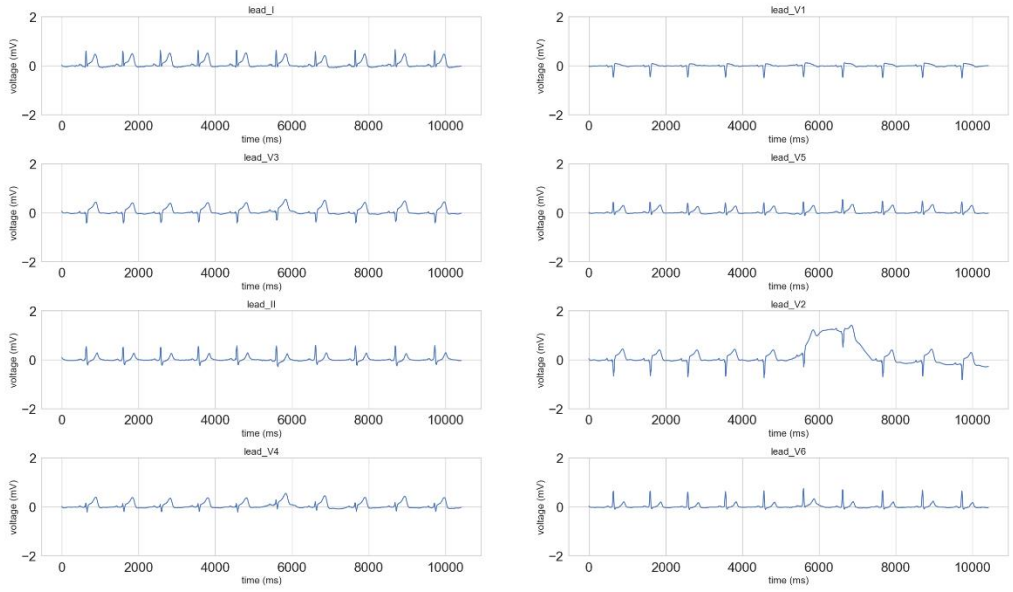


Figure 14 Example for a potential change in the lead ECG

3.3 Singular Value Decomposition

In the study 'Analysis of electrocardiogram- and vectorcardiogram- derived parameters in post-myocardial infarction patients' by M. Reijnen [6], singular value decomposition (SVD) is implemented to extend the parameter set. Therefore, the function of the previous study is used.

SVD is used in order to minimize the subspace dimension without sacrificing significant ECG energy. If a M^{mxn} matrix in R space exists, it can be expressed as the product to three distinct matrices: U^{mxn}, Σ^{rxr} and V^{nxr}. In this manner, m represents the number of ECG leads (I, II, V1, V2, V3, V4, V5, V6), n denotes the number of sample points and r represents the rank of the matrix which is set to eight in this cases. The Σ matrix is a diagonal matrix, with its values $_{\text{arranged}}$ as $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_8 \geq 0$. In these matrices, the columns of U are referred to as the left singular vectors, while those of V are the right singular vectors. Σ

contains the singular values, which signify the amount of energy associated with each individual vector u. This mathematical process is known as singular value decomposition, as illustrated below.

$$M = U\Sigma V^T$$

In this provided equation, the matrices are orthogonal. It can be noted that in the previous study it is found that 84% of the ECG energy can be effectively represented using only the first three dimensions, creating a compact three-dimensional subspace [2]. From this observation the effective rank of M is 3. Consequently, M is projected onto a reduced subspace denoted $\widetilde{U}^{mxn}$, yielding the matrix S^{3xn} which is equivalent to $S = \widetilde{U}M$. In Figure 15, it is shown ECG can be represented in a three-dimensional minimum subspace with s1, s2 and s3.

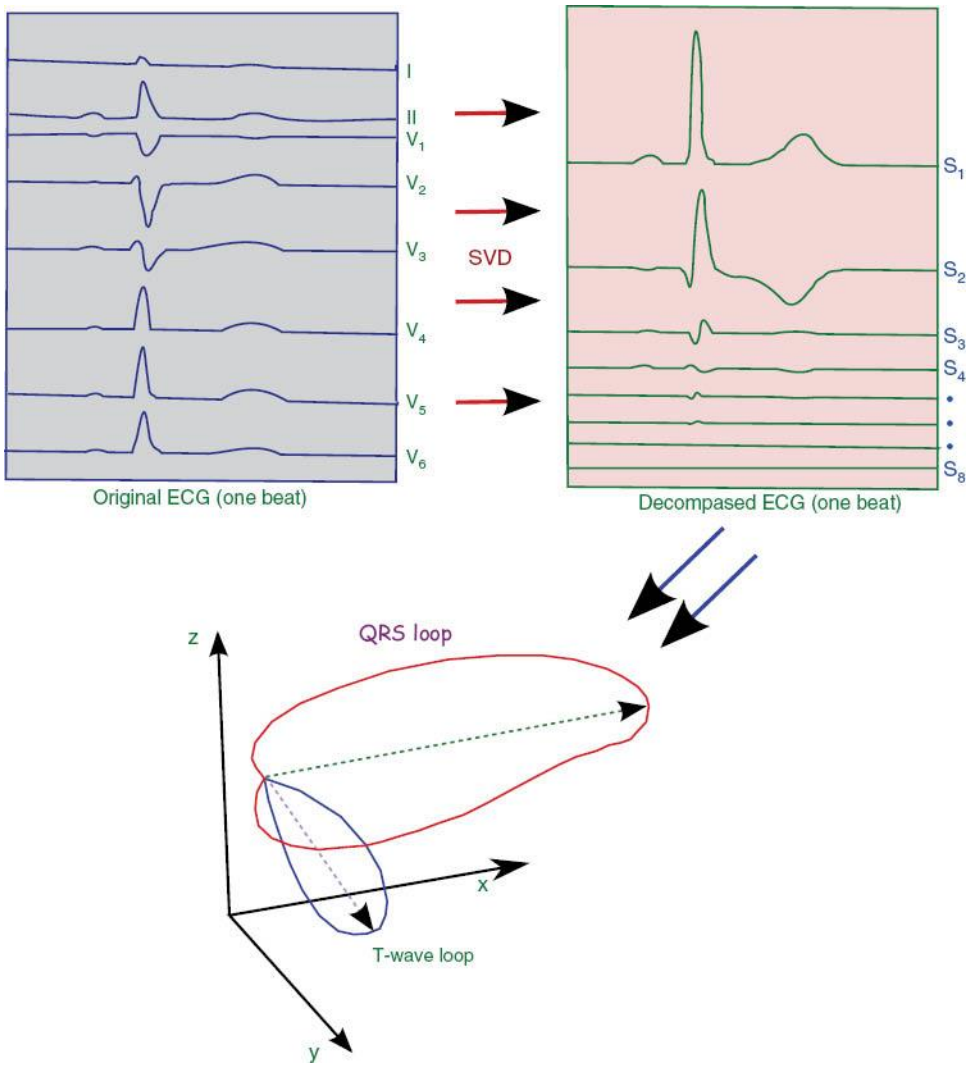


Figure 15 Original ECG and single beat ECG after SVD transformation. Below, The QRS and T wave loop as well as vectors of QRS and T wave are shown with axis x, y, z are s1, s2 and s3 in the implementation.

3.3.1 QRS- and T-wave Detection

In the QRS- and T-wave detection process, several key points are utilized to precisely locate important points in the ECG signal.

Firstly, the R-wave points are extracted from the .xml file. However, R wave detection can also be conducted based on the SVD. The time instance $t_{RE_a}$is identified, representing the time point where the ECG energy drops below 70% of the maximum ECG energy which is described as the maximum energy value of the three most information containing dimensions. The 70% point is established in the prior research and used in this research. Additionally, the symmetric point of this denoted as t_{RS_a} is determined corresponding to the instance with 70% of the maximum ECG energy. Then the beginning of the QRS complex is found using these two time points. To do this, 48 ms is subtracted from t_{RS_a} to obtain t_{RS}, and 48 ms is added to t_{RE_a} to obtain t_{RE} [2].

Secondly, the peak of the T-wave is pinpointed. This is achieved by identifying the point with the highest energy in the three-dimensional subspace after the endpoint of the QRS complex, which is earlier referred to as t_{RE}. The search interval to locate t_{TP}, the point corresponding the peak of the T-wave, extends from the end of the QRS complex (t_{RE}) to the T-wave offset point, a value determined in the section 3.2.2. The start of the T-wave is calculated by a simple formula: $t_{TS} = t_{RE\_a} + \frac{1}{3} * \left(t_{TP} + t_{RE_a}\right)$ [2].

In Figure 16, a clearer visual representation of all the key points in all the leads is illustrated.

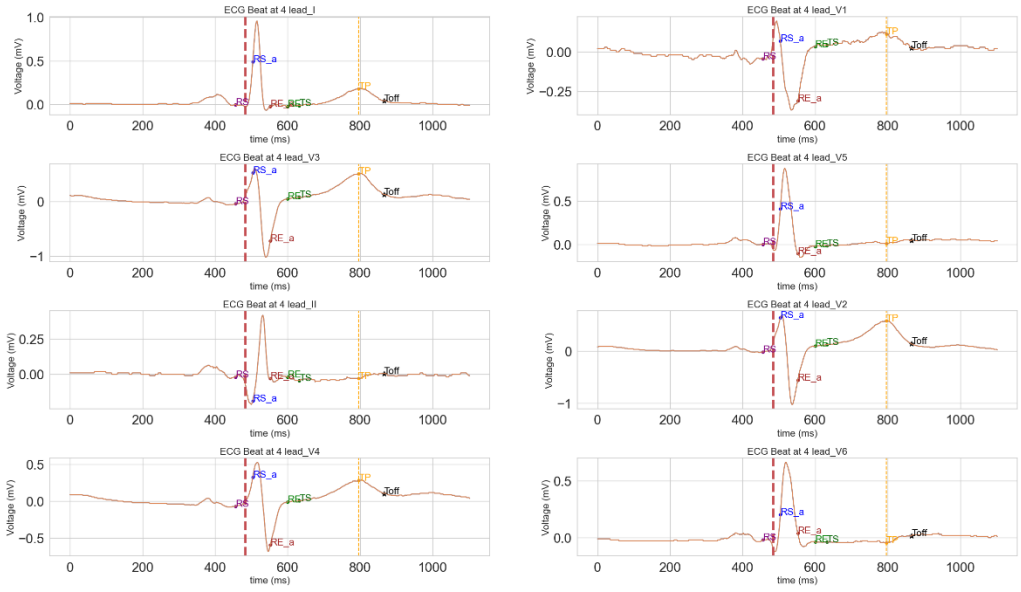


Figure 16 Fiducial points from beat number 4 is illustrated

3.3.2 Signal Normalization, QRS- and T-wave Extraction, and DC compensation

After the Singular Value Decomposition, energy of the subspace is subsequently calculated by measuring the Euclidean norm of the components of s_{3D}. Then the signal is subject to normalization by dividing vector s_{3D} by the maximum energy. Following this, normalized signals denoted by S_{QRS} and S_T are generated. While S_{QRS} corresponds to the normalized signal of the time interval between the start of the QRS complex (t_{RS}) to end of QRS complex (t_{RE}), S_T corresponds to the normalized signal of the interval from start (t_{TS}) and the end $t_{Toffset}$ of the T-wave.

After these steps, DC compensation is implemented by subtracting the average of s_{3D} vectors in times t_{RS}, t_{RE}, t_{TS} and $t_{Toffset}$ from S_{QRS} and S_T. S_{QRS} matrix is used for parameter extractions in the following sections. However, S_T matrix is used in the reconstruction of the T-wave back into its original ECG domain. This step is followed by singular value decomposition of the resulting matrix.

In summary, SVD in the calculations enables the usage of three signals with the highest energy and the determination of the QRS and T wave boundaries. The DC compensation improves the ECG signal in terms of its reliability. In order to perform T-wave morphology dispersion (TMD), reconstruction back into the ECG domain is performed. Finally, reconstructed ECG signal is decomposed by SVD to enable parameter calculations [6].

3.3.3 T-wave Morphology Dispersion (TMD)

T-wave Morphology Dispersion (TMD) is a parameter that quantifies the variations in the shape of T-waves across different leads in an ECG during the process of ventricular repolarization. To calculate TMD, the angles between all potential pairs of the reconstruction vectors are computed. These reconstruction vectors essentially describe the morphology of T-waves in different ECG leads except lead V1. The average of the sum of all the angles in pairs of two for leads from {I, II, V2, V3, V4, V5, V6} is calculated. The equation for TMD calculation is as follows:

$$TMD = \frac{1}{21} \sum_{i,j\ in\ leads} \emptyset_{ij}$$

Here, $\emptyset$ is the angle between individual reconstruction coefficient of the T wave in the ECG leads. When the angle $\emptyset_{ij}$ is smaller, it can be concluded that the shapes of T waves in i-th and j-th ECG leads are more similar. As a result, small values of TMD parameter can be correlated to similar T-wave morphologies in different ECG leads.

3.3.4 Total Cosine R to T (TCRT)

The total cosine R to T is defined as the mean value of all cosines of the angles between the unit vector carrying the maximum T-wave energy ($e_{T,1}$) and the normalized, DC-compensated ($S_{QRS}(i)$) signal matrix. The interval is where the maximum energy is higher than the 70% of the maximum energy of the beat. Below the equation for the TCRT equation is shown.

$$\emptyset_{TCRT} = \frac{1}{t_{RE}{}' - t'_{RS}} \int_{i=t'_{RS}}^{t_{RE}{}'} \cos\left(L\left(e_{T,1}, S_{QRS}(i)\right)\right)$$

As mentioned before, the TCRT measures the difference between the depolarization and repolarization vector. The result always falls between -1 and 1. While -1 represents an opposite direction between these two vectors, +1 shows that the two vectors are pointing in the same direction. In beat-to-beat calculations, TCRT calculation is performed after the beats are segmented.

3.3.5 Statistical Analysis

In this study Python 3.8.13 is used for the overall algorithm as well as the statistical analysis. In the statistical analysis, firstly, Shapiro-Wilk test is performed in order to evaluate whether the data sets from the Control and VT patients are normally distributed or not. From this evaluation, two conclusions can be made for the individual data sets. If the result of the test has a large p-value ($p \geq 0.05$), this indicates that the data set might be normally distributed. On the other hand, if the p-value is low ($p < 0.05$), then this indicates that the corresponding data set is not normally distributed [5]. In case the data sets are identified to be not normally distributed, then Mann-Whitney U test is conducted. The Mann-Whitney test is used to test whether two samples are likely to have the same shape or not [7]. However, if there exists evidence that the two data sets might be normally distributed, then Levene test is performed. This test characterizes the equal or unequal variances for the data sets. Finally, if there exists evidence for equal variances, then a t-test with equal variances is executed. On the other hand, if the evidence shows unequal variances for the data sets, then t-test with unequal variances is executed. For each test a significant level of 0.05 is used.

4 Results

4.1 Automatic Beat-to-Beat Analysis

In the context of our study involving 28 patients, we observed 4 patients and 24 recorded ECGs from patients developed Ventricular Tachycardia (VT). From the 24 patients who did not developed VT, we observed a total of 327 ECG recordings for this control group. At the end of the project, automatic beat segmentation successfully segmented 4162 beats. Some beats are excluded from the parameter calculation based on inspection. Last beats which do not fully have a T-wave are automatically excluded from the parameter calculation. Additionally, 4 beats from the control group ECGs are manually excluded from the parameter calculation due to having abnormal peaks in one of their leads. Finally, 3724 beats which belong to the control group patients are successfully segmented. On the other hand, 473 beats are successfully segmented for the VT group patients.

4.2 TCRT

TCRT parameter measures the difference between depolarization and repolarization complexes and the result is always in between -1 and 1. In Figure 17, the value range and the median value for TCRT is illustrated per Control and VT patient groups. The results for the two groups are found to be statistically different from the Mann-Whitney U test with a p value $1.621 * 10^{-79}$. It can be observed that the TCRT values for the Control group is closer to +1 compared to VT patients which has TCRT parameter values closer to -1.

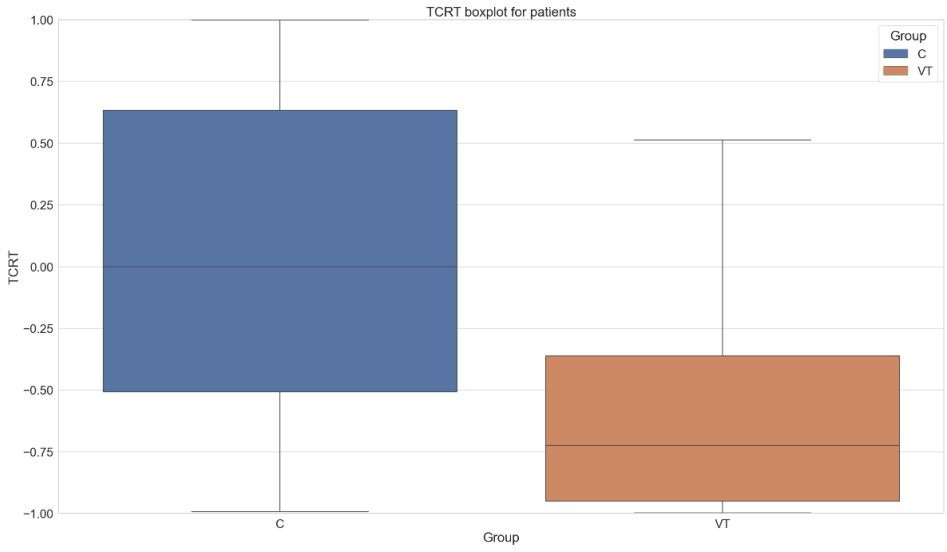


Figure 17 Boxplot showing the TCRT parameter result per patient group

In Figure 18, the results of TCRT calculation per patients is illustrated. It can be inspected that some patients' TCRT results per beats are distributed in a large scale. For example, for patients C18, C19 and C25 the large range for the TCRT results is observed. This indicates a high intra-patient variability. On the other hand, for some patients, TCRT values are seen to be in a more consistent range. For example, VT04, C06 have TCRT results closer to -1.

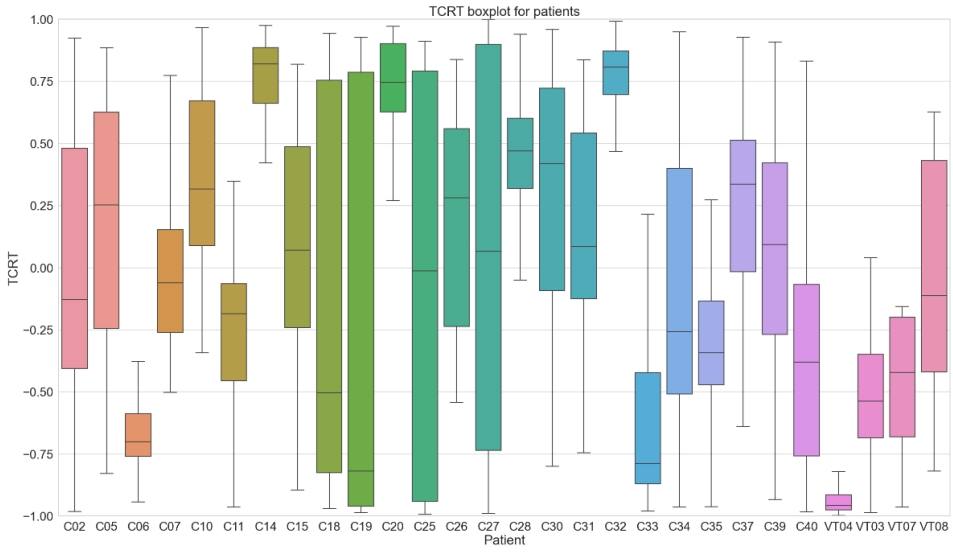


Figure 18 Boxplots showing TCRT per patient

In the swarmplots the x-axis represent a date ratio, a measure derived from ECG recording dates. For patients with VT, this ratio is computed as follows: The dates of MI and VT onset are identified from the particular patient file. The MI date is designated as 0 while the VT date is marked as 1. Each ECG recording is also associated with a date which is also extracted from the corresponding patient file. To calculate the rate ratio for each recording, the number of days passed between MI date and the specific ECG recording is calculated, and this duration is divided by the total number of days between the MI and VT dates. This process is applied consistently to all the ECG recordings and used as an axis for the swarmplots. This process allows a meaningful representation of temporal relation within data.

In Figure 19, the swarmplot for VT04 patient is illustrated. It can be seen that for TCRT parameter is mainly distributed consistently with some minor inconsistencies for beats. For further research, the inconsistent beats can be investigated and some other parameter test can be performed on these beats.

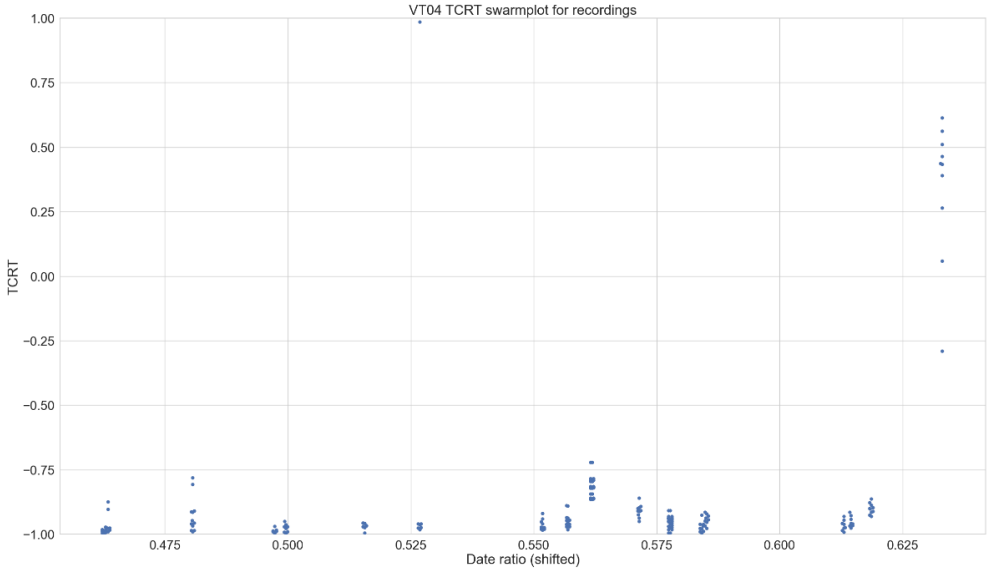


Figure 19 Swarmplot for VT04 patient

The date ratio for Control group patients is calculated in a similar fashion except since these patients did not develop VT, the date for the VT date is taken as a future value. It can be observed that at closer to 0.00 ratio point in the x-axis, there are many beats showing that most of the recordings are collected in this time period. It can also be said that the beats are again have consist values between each other.

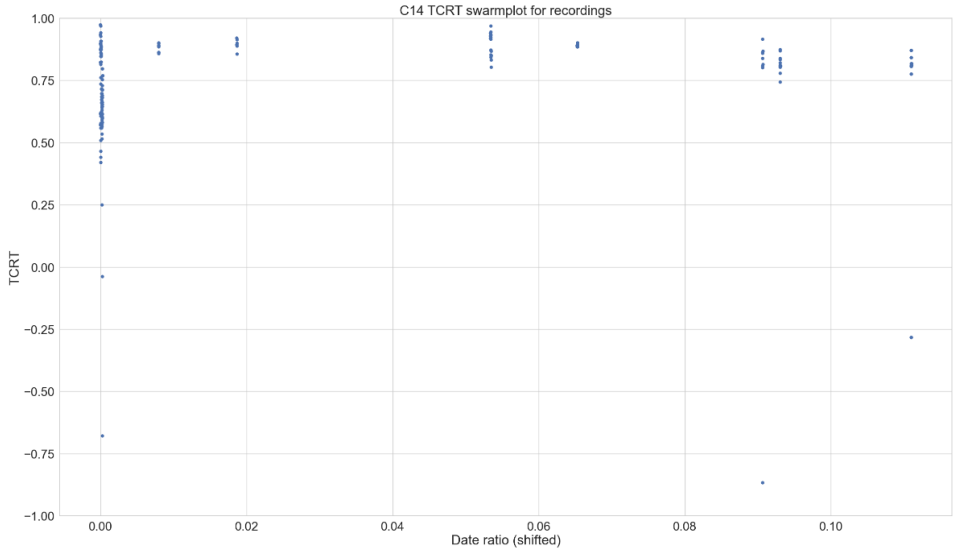


Figure 20 Swarmplot for C14 patient

In Figure 21, there exists lots of recordings for this patient. However it can be seen that the TCRT values for beats are not close to one value specifically but within a range of values.

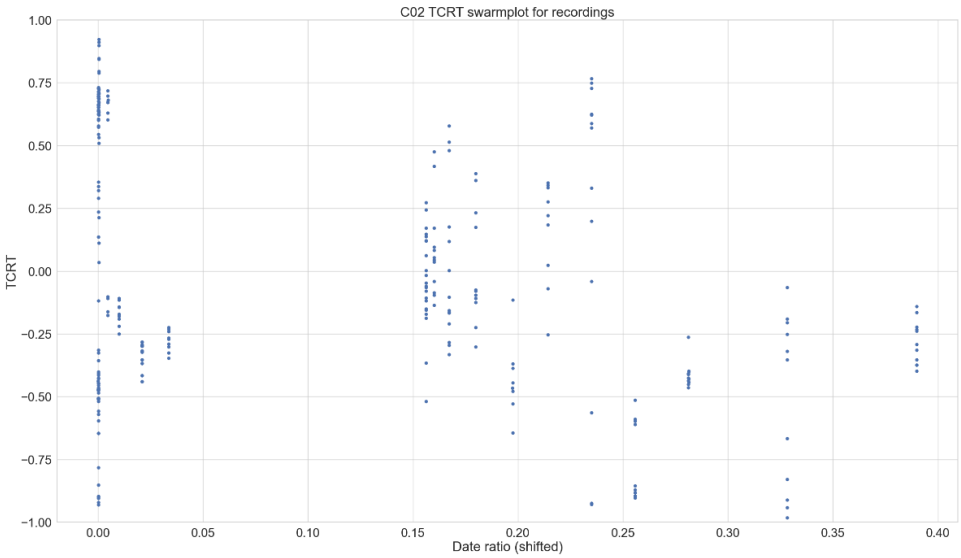


Figure 21 Swarmplot for patient C02

At the end, it is observed that while some patients show a downward trend over the time, some show a upward trend. Therefore, the trends are not found to be consistent and no trend is identified.

4.3 TMD

TMD refers to T-wave morphology dispersion which describes the spatial variation in the wave. In Figure 22, the results per patient group is shown. It is observed that VT patients show higher TMD values compared to Control patients. The differences in TMD values are found to be statistically different based on the Mann-Whitney U test with p-value $1.96 * 10^{-89}$.

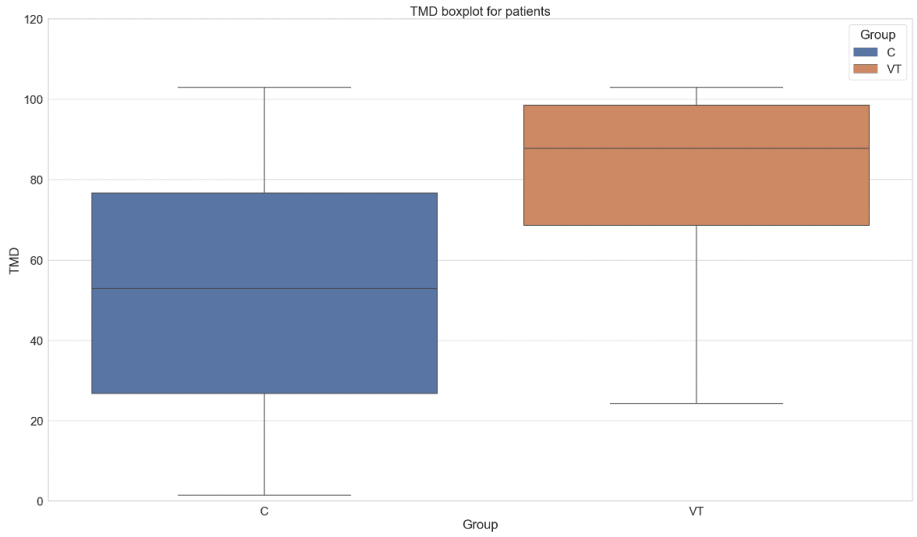


Figure 22 Boxplot of TMD per group

In Figure 23, TMD values per patients are represented with a boxplot. Similarly, the parameter values for some patients are within a small range, and some are distributed in a large range. For example, C06 and VT07 are relatively have a large range. While patients such as VT04 and C15 have values within a small confined range.

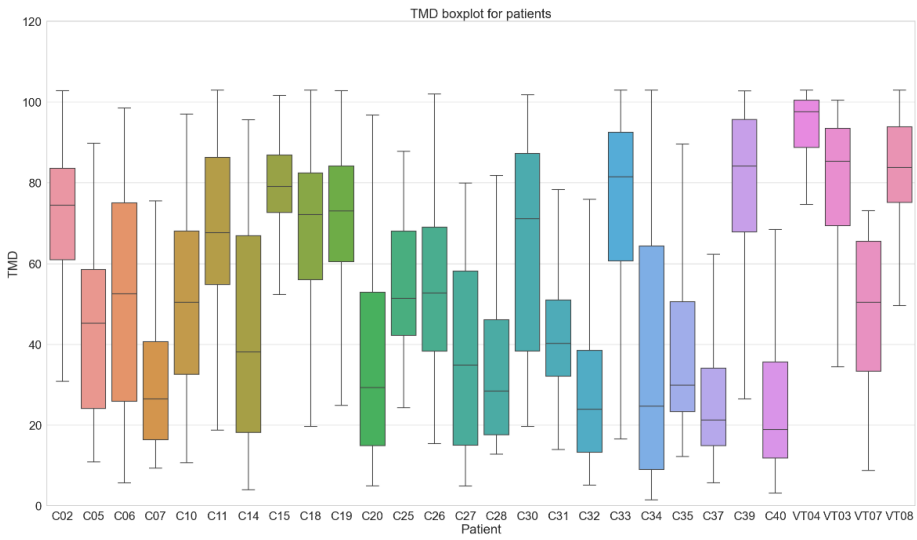


Figure 23 TMD boxplot per patients

In Figure 24, the swarmplot of VT04 patient is illustrated with each dot represents a TMD value for the corresponding beat. It can be observed that the values are consistent and TMD results are closer to 100.

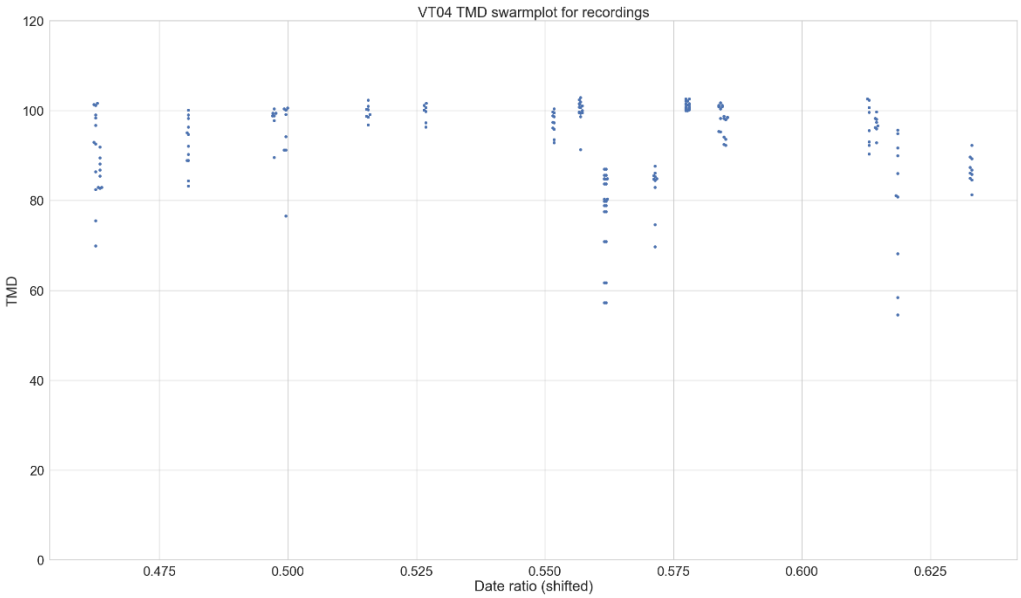


Figure 24 TMD swarmplot for VT04

In Figure 25, the swarmplot of C14 is illustrated. It can be observed from the plot that the TMD values for beats are distributed in a large range. Only for some beats within the plot show consistent results. The beats that have similar results for the same recording can be compared with the ones where the TMD results are within a large range.

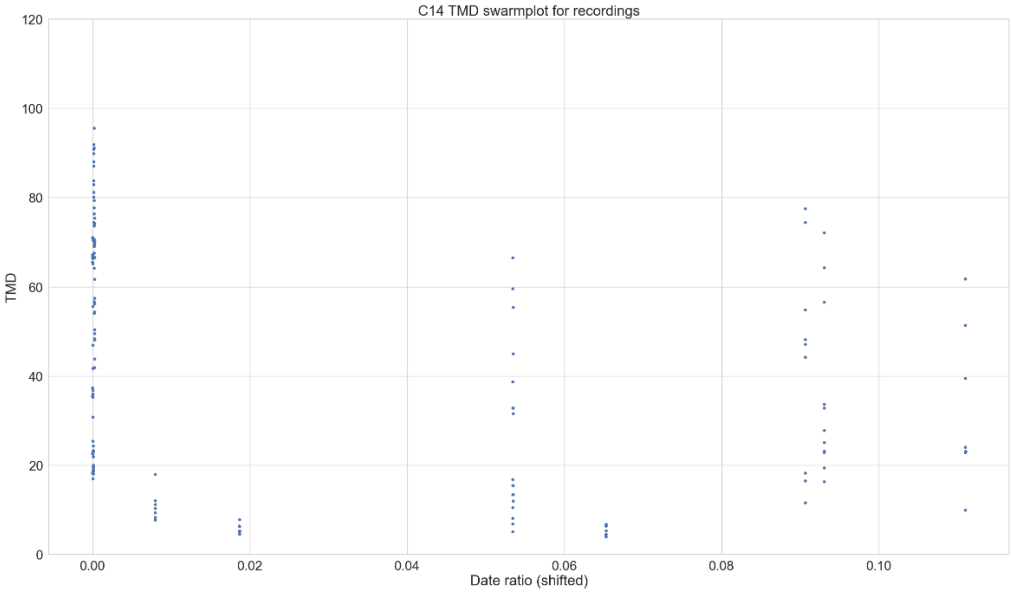


Figure 25 Swarmplot for patient C14

In Figure 26, swarmplot for patient C02 is illustrated. This patient file consists of many recordings therefore there are relatively more number of beat values within the plot. The beats are within a large range. Most of the recordings are closer to the time period where MI has happened.

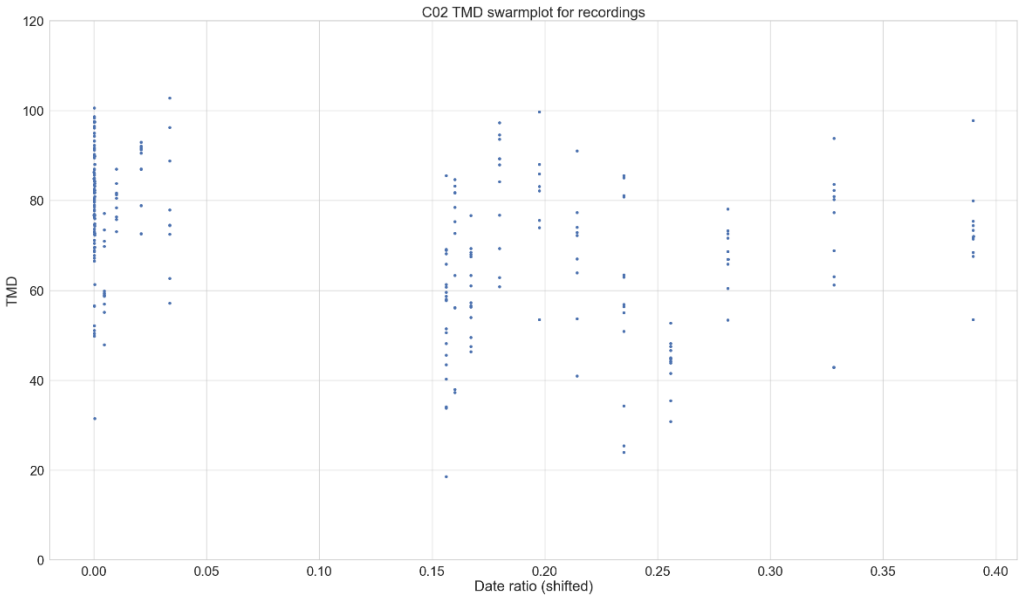


Figure 26 TMD swarmplot for patient C02

5 Brief Discussion

Before continuing with the conclusion, a brief analysis of the results can be discussed under this section. It can be inspected from the boxplots for both TCRT and TMD that there is a clear difference between the two groups of patients. This provides an overview of the distributions for each group and allows a broader assessment of the dat. This claim is also supported by the statistical test result. Since there exists lack of consistencies when the data is examined across patients, one patient having low TCRT parameter result from an ECG recording may not strongly indicate a possible VT development. However, the results are promising in the sense that some precautions can be taken based on the results obtained from TCRT and TMD.

Additionally, from the VT04 swarmplot, it can be inspected that the consistent results for the parameter calculation for each beat is possible and it can be further investigated for beats that do not show such consistencies for the same recording. One possible explanation for such kind of inconsistencies regarding the parameter results for some beats may be that the fiducial points within these beats may have been miscalculated and the miscalculation resulted in incorrect parameter calculation. Another aspect of this situation may be that these inconsistencies within the beats should not be ignored but

closely investigated for the analysis. Maybe, the beats which appear to be inconsistent from the other beats are the ones with showing a real trend in the development of VT.

Finally, I believe beat-to-beat analysis is significant in terms of understanding the dynamic characteristics within the beat. When the median ECG is selected for the analysis, only one aspect from the rhythm ECG can be shown. However, when all the beats are considered, these beats may end up showing different parameter results. Inspecting from the swarm plots, variability and trends across multiple beats can be considered. Moreover, irregularities within the cardiac activity can also be investigated. On the other hand, median ECG may be unable to indicate the dynamic changes or irregularities within the rhythm ECG.

6 Conclusion

In conclusion, I completed my internship in Eindhoven University of Technology in Biomedical Department. This internship was a great experience for me. All the team members of the research group were very welcoming and experienced.

During my internship the critical realm of risk assessment for developing ventricular tachycardia in post-myocardial infraction (MI) patients is explored. The study employed the derivation of beat-to-beat Vectorcardiogram (VCG) parameters as TCRT and TMD, to explore the potential for distinguishing between the Control and VT groups.

Meticulous attention and effort are dedicated to the beat segmentation process. At the end nearly all the beats are segmented with precision and accuracy, with only 4 beats showing irregularities in the segmentation.

Throughout this research, I gained numerous insights including understanding how research is conducted. This provided valuable lessons in conducting scientific research and reading academic sources. Additionally, this study allowed for practical skill development in coding using Python. I learned more about vital libraries such as Numpy, Pandas, Matplotlib, Seaborn and specialized statistics and signal processing libraries to process and analyze data effectively.

The result of the study is promising, it is found that TCRT and TMD parameter results are statistically different in control and VT patient groups. The results can be further investigated, and other beat-to-beat parameters can be calculated through further research. Parameter results for each beat can be analyzed as well as optimization to beat segmentation and parameter calculations can be made.

References

[1] Abbott, M. A. (2016). A review of beat-to-beat vectorcardiographic (VCG) parameters for analyzing repolarization variability in ECG signals. *Biomedical Engineering / Biomedizinische Technik*, 3-17.

[2] B. Acar, G. Y. (1999). Spatial, temporal and wavefront direction characteristics of 12-lead t-wave morphology. *Medical & biological engineering & computing*, 574-584.

[3] COMBAT-VT project. (2022). Retrieved from https://combatvt.nl/

[4] Healthcare, G. (2015). Marquette™ 12SL™ ECG Analysis Program: Physician's Guide. *General Electric Company*.

[5] Malato, G. (2023, May 02). *An Introduction to the Shapiro-Wilk Test for Normality*. Retrieved from built in: https://builtin.com/data-science/shapiro-wilk-test

[6] Reijnen, M. (2023). *Analysis of electrocardiogram- and vectorcardiogram- derived parameters in post-myocardial infarction patients.* Eindhoven.

[7] Sullivan, L. (2017, May 4). *Mann Whitney U Test (Wilcoxon Rank Sum Test)*. Retrieved from Boston University: https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_nonparametric/index.html

[8] Thygesen, J. N. (2023). ST-elevation myocardial infarction Studies of outcome in relation to fibrinolysis and ischemia monitoring with on-line vectorcardiography.