FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Minna Moilanen

AUTOMATIC DETECTION OF EARLY REPOLARIZATION IN ECG SIGNAL

Master's Thesis
Degree Programme in Biomedical Engineering
November 2020

Moilanen M. (2020) Automatic detection of early repolarization in ECG signal. University of Oulu, Degree Programme in Biomedical Engineering. Master's Thesis, 44 p.

ABSTRACT

The early repolarization is one form of heart's electrical disorder. The scope of this thesis is to develop an algorithm, which detects the marks of the early repolarization from the electrocardiography data. The definition of the early repolarization was fine-tuned in 2015 and the updated definition is used in this thesis. The implementation of the algorithm is done with Matlab.

The theory part of this work includes description of the heart's structure, physiology and review of the most common heart diseases. The heart's electrical functionality is explained in more detail and the principle of the electrocardiography is viewed, including the main precepts of the analysis of electrocardiography data. The definition of the early repolarization is presented in detail and the significance of this phenomenon is evaluated based on the research data. This work also includes short survey of some of the existing methods for detecting the early repolarization from the electrocardiography data. Thesis includes also the description of the algorithm developed in this work and the analysis of the results.

The performance of the algorithm is evaluated with manually classified ECG test set. The sensitivity of the algorithm is 94.0% and the specificity is 92.2%. The correlation to mortality was also studied for few different versions of the algorithm with the Health 2000 data. The correlation to mortality is found with two algorithm versions. The algorithm version with slightly relaxed early repolarization definition shows increased risk for all-cause-mortality in inferior leads, when the slur detection is deactivated. The algorithm version with precise thresholds of the early repolarization definition shows increased risk for all-cause-mortality and for cardiac death in inferior leads, when the slur detection is deactivated.

Keywords: ER, ERP, J-wave syndrome, cardiac death, sudden cardiac death

Moilanen M. (2020) Automatic detection of early repolarization in ECG signal. Oulun yliopisto, lääketieteen tekniikan tutkinto-ohjelma. Diplomityö, 44 s.

TIIVISTELMÄ

Tämä diplomityö käsittelee sydämen sähköisen toiminnan häiriötilaa, jota kutsutaan aikaiseksi repolarisaatioksi. Työn tavoitteena on kehittää algoritmi havaitsemaan aikaisen repolarisaation merkit sydämen elektrokardiografia mittausdatasta. Tämä työ perustuu vuonna 2015 tarkennettuun aikaisen repolarisaation määritelmään. Työssä kehitetty algoritmi on toteutettu Matlabilla.

Diplomityön teoriaosuudessa käydään läpi sydämen rakennetta, fysiologiaa ja yleisimpiä sydänsairauksia. Työssä tutustutaan tarkemmin sydämen sähköiseen toimintaan, elektrokardiografian tuottamaan dataan ja siihen, miten tätä dataa voidaan tulkita. Aikainen repolarisaatio käsitellään omana osionaan, jossa käydään läpi sen tarkka määritelmä, arvioidaan tutkimuksiin pohjautuen ilmiön merkitsevyyttä sekä esitellään muutamia olemassa olevia menetelmiä aikaisen repolarisaation havaitsemiseen elektrokardiografia datasta. Työ sisältää myös kehitetyn algoritmin esittelyn ja tulosten analysointia.

Algoritmin suorituskyky todettiin testisetillä, joka sisältää manuaalisesti luokiteltuja elektrokardiografia signaaleita. Algoritmin sensitiivisyys on 94,0% ja spesifisyys 92,2%. Tämän lisäksi ajettiin testejä kuolleisuus korrelaation selvittämiseksi muutamalla algoritmin variaatiolla Terveys 2000 datalle. Korrelaatio kuolleisuuteen löytyi kahdella algoritmivariaatiolla. Algoritmiversio hieman väljennetyillä aikaisen repolarisaation kynnysarvoilla ennustaa kohonnutta riskiä kokonaiskuolleisuuteen inferiorisissa signaaleissa, slur-tunnistuksen ollessa pois käytöstä. Algoritmiversio aikaisen repolarisaation määritelmän mukaisilla tarkoilla kynnysarvoilla ennustaa kohonnutta riskiä sekä kokonaiskuolleisuuteen että sydänperäiseen kuolemaan inferiorisissa signaaleissa, slur-tunnistuksen ollessa pois käytöstä.

Avainsanat: ER, ERP, J-aalto syndrooma, sydänperäinen kuolema, sydänperäinen äkkikuolema

TABLE OF CONTENTS

FOREWORD

Writing this chapter means that my studies in Biomedical Engineering at the University of Oulu are ending. The time spent with my studies and with this thesis has been interesting and educational in many ways.

This master's thesis describes a journey which target is to create an algorithm for detecting the early repolarization from the ECG data. This work consists of exploring the research in the field, studying heart's anatomy and physiology, learning how to read ECG signal, creating the algorithm and its implementation with Matlab and reporting the results.

As this project is ending, I would like to express my gratitude for Professor Tapio Seppänen, PhD Tuomas Kenttä and Professor Heikki Huikuri for proposing this topic and contributing to defining the scope for this thesis. I would like to thank Professor Seppänen for supervising, PhD Kenttä for technical supervising and second examiner D.Sc. (Tech.) Juha Partala for his review. Special thanks to Ari, my family and friends for their encouragement and support during this project.

Oulu, 30th October 2020

Minna Moilanen

ABBREVIATIONS

ACS	Acute coronary syndrome
ACM	All-cause-mortality
AF	Atrial fibrillation
CD	Cardiac death
CI	Confidence interval
CAD	Coronary artery disease
DCM	Dilated cardiomyopathy
ECG	Electrocardiography
ER	Early repolarization
ERP	Early repolarization pattern
ERS	Early repolarization syndrome
Exp(B)	Hazard ratio
HCM	Hypertrophic cardiomyopathy
HF	Heart failure
IVF	Idiopathic ventricular fibrillation
INF	Inferior
LAT	Lateral
MI	Myocardial infarction
RCM	Restrictive cardiomyopathy
SCD	Sudden cardiac death
Sig.	Significance, p value
SVT	Supraventricular tachycardia
VF	Ventricular fibrillation
VT	Ventricular tachycardia

1. INTRODUCTION

Inside our chest, behind the sternum, beats a vital organ called heart. The heart is central component of the blood circulation system. The beating of the heart is enabled by its complex structure. The heart is a combination of many different tissue types, signaling networks and biomechanical components. Malfunctions in this organ can have severe effect to our body. The disorders of this organ can be related to heart's own metabolism, vascular problems of the body, structural defect in the heart or disorder in the heart's electrical system. The anatomy and the physiology of the heart are described in Chapter 2.

In this thesis, one form of the heart's electrical dysfunction is studied. The early repolarization (ER) is a condition where electrical cycle of the heart is disturbed. The ER is described in more detail in Chapter 2. Issues in the heart's electrical conductivity can be detected by the electrocardiogram (ECG) measurement. The ECG is the most commonly used tool when studying cardiac problems and its basic principle is described in Chapter 2. The main goal of this work is to design and implement an algorithm, which enables automatic detection of ER from the ECG data. The algorithm is implemented in Matlab environment.

Over the years, studies regarding the ER have presented contradictory results about the severity of the condition. Early studies by Klatsky et al. (2003) concluded the ER to be benign phenomenon [1]. A few years later Haïssaguerre et al. (2008) found the connection between ER and increased risk of cardiac death [2]. After the study by Haïssaguerre et al. the maleficent nature of ER has been manifested also by many other research groups for example Nam et al. [3], Rosso et al. [4], Tikkanen et al. [5] and Antzelevitch and Yan [6]. In 2016 Roten et al. proposed ER to be divided in to harmless and maleficent forms [7].

The research regarding the ER long suffered from the vague ER definition. The research groups around the world have applied slightly different principles in their studies what comes to this phenomenon. In 2015 and 2016 the group of leading cardiologists created a consensus papers about the precise definition of the ER [8, 9]. In the past there have been several algorithms presented for finding the ER from the ECG signal. Some of these methods are presented in Chapter 2. In the algorithm development of this work, the ER definition from the consensus is used.

This thesis consists of six chapters. In the Chapter 2, the structure of the heart as well as its basic functionality, relation to the nervous system and electrical activity are described. In addition, the second chapter includes a brief review of the most common heart diseases and more detailed information about the ER condition. In the Chapter 3, the design process of the algorithm is described. The third chapter includes also description of the algorithm and information about the ECG data used in this thesis. The results of the thesis are presented in Chapter 4 and discussed in the Chapter 5. Thesis is summarized in Chapter 6.

2. THE ANATOMY AND PHYSIOLOGY OF THE HEART

The blood circulates through the body providing it with the vital oxygen and nutrients. The blood also collects waste products from the organs and carries them to waste disposal system. The blood is kept moving by the heart, which acts as a pump. The heart has complicated structure consisting of several blood inputs and outputs, four chambers, multiple valves, different tissue types and electrical network. The heart also has interfaces to other functional entities of the body such as nervous system.

2.1. The structure of the heart

The heart is an organ that is a little bit larger than its owner's fist. The weight of a male heart is approximately 280 g to 340 g and female heart is approximately 230 g to 280 g. Anatomically the heart can be divided in to two parts. The upper part of the heart is called basis and the lower part is called apex. Apex points towards the left hip and basis points towards the right shoulder. In the surface of the heart, there are grooves where the arteries and veins travel. [13 p.13]

The heart has four chambers; right atrium, left atrium, right ventricle and left ventricle (Figure 1). The right atrium receives blood from systemic circulation and from the heart's own coronary veins, (the systemic- and pulmonary circulations are explained in more detail in Chapter 2.2). The right ventricle forms the majority of the front side of the heart. The right atrium and the right ventricle are separated by the atrioventricular valve called tricuspid valve. The tricuspid valve enables the blood flowing only from atrium to ventricle direction. The right ventricle and the pulmonary artery are separated by the semilunar valve called pulmonic valve. The pulmonic valve enables the blood flow only from ventricle to artery direction. The left atrium forms major part of the back surface of the heart. The left atrium receives blood from the pulmonary circulation via three to five pulmonary veins. The left ventricle forms the lower part of the heart. The left atrium and the left ventricle are separated by the atrioventricular valve called mitral valve. The mitral valve enables blood flowing only from atrium to ventricle direction. The left ventricle and the aorta are separated by the semilunar valve called aortic valve. Muscular walls around the ventricles are thicker than walls around the atriums. The walls of the left atrium and the left ventricle are thicker than the walls at the right side. This is due to the left side of the heart deals with larger pressure. [13 p.13-14]

The tree-part wall separates the right and the left side of the heart. The uppermost wall is separating the atriums, bottom wall is separating the ventricles and in-between is the atrioventricular wall. The wall between the atriums is developed in two phases during the development of the fetus. In the first phase, the blood flow between the atriums is enabled via hole in the connecting wall. This hole is built twice in different locations during this development phase. At the second phase there is valve structure replacing the hole. Valve enables the blood flow from the right atrium to the left atrium. For most people, after the birth, this channel is closed and small dip remains in the tissue. The wall between the ventricles is mainly muscle tissue. The upper part of this wall changes into connective tissue and continues as atrioventricular area wall. The atrioventricular valves are connected into this wall. The triangle of Koch is located in this area. The triangle of Koch is formed by the connective tissue in the attachment location of the atrioventricular valve and the valve of the inferior vena cava. One of

the main conduction system parts of the heart, the atrioventricular node, is located in the area of the triangle of Koch. [13 p.14-15]

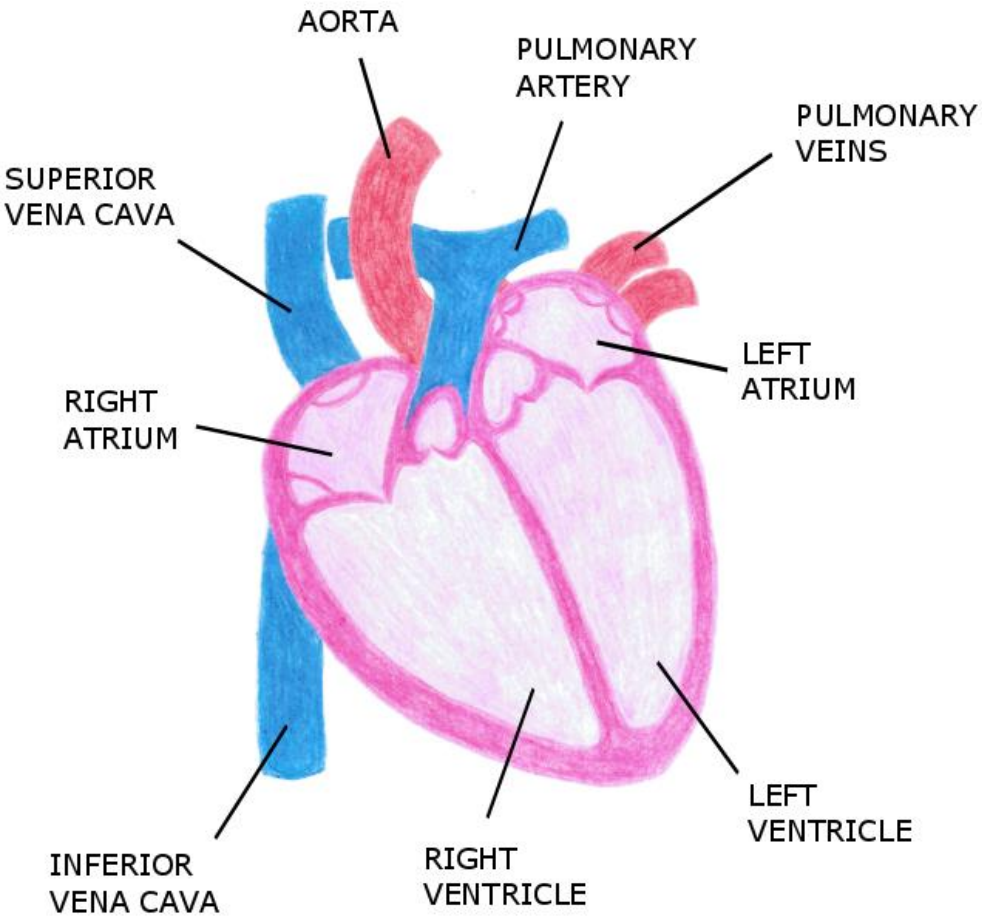

Figure 1. The simplified image of the structure of the heart.

The heart is covered by the pericardium. The purpose of the pericardium is to protect the heart muscle and to prevent too large and sudden movements. The pericardium protects also the beginnings of the aorta and the pulmonary artery. The pericardium consists of two layers; parietal leaf and visceral leaf. Between these layers, there is an intermediate layer containing lubricant. Under the pericardium is the heart wall consisting of the epicardium, the myocardium and the endocardium layers. The epicardium forms from the visceral leaf. The epicardium consists of mesothelial cells, connective tissue and fat tissue. The heart's own blood vessels and nerves coming to the heart are located at the epicardium. The myocardium layer consists of heart muscle. The endocardium creates the inner surface of the heart's chambers. It consists of the endothelial layer, that is the surface of the chambers and also the surface of the connective tissue layer. The connective tissue layer is below the endothelial layer and it is connected to the myocardium connective cells. The connective tissue layer includes main parts of the conduction system of the heart. [13 p.12]

The heart's own metabolism is secured by several coronary arteries and coronary veins. The coronary arteries branch from the beginning of the aorta and in resting state, if healthy, they transfer 4 - 5% of the circulating blood. There are usually two main coronary artery branches in the heart. Left main branch of the coronary artery is called arteria coronaria sinistra and it is divided into front and downward going branch and rotating branch. These branches are divided again into several smaller branches. Right main branch of the coronary artery is called arteria coronaria dextra and its smaller branches rotate to the backside of the heart. Four main coronary veins transfer the blood mainly to the right atrium. [13 p.15-16]

2.2. The basic functionality of the heart

The blood that has released the oxygen and the nutrients to the body and received waste products from the body is returned to the lungs. The heart pumps the blood between two circulations; the systemic circulation and the pulmonary circulation. In Figure 2 the systemic- and the pulmonary circulations are visualized. The blood from the systemic circulation arrives at the right atrium via inferior and superior vena cava. The chamber contraction is controlled by the electrical system of the heart (electrical system is described in more detail in Chapter 2.4). When the electric signal is initiated in the heart, the atriums depolarize and contract. The contraction makes the blood flow from the atriums to the ventricles. After the contraction of the atriums, electrical signal travels to the ventricles, which are usually filled at this point, and they depolarize and contract. As the ventricles depolarize the pressure in the chambers closes the atrioventricular valves. Increasing pressure in the ventricles opens the semilunar valves and the blood flows into the arteries. In repolarization, chamber pressure drops below the artery pressure and the semilunar valves are closed. From the right ventricle, the blood is pumped towards the pulmonary circulation via the pulmonary artery. In the pulmonary circulation, the blood is cleaned and oxygenized. The blood returning from the pulmonary circulation travels via pulmonary veins into the left atrium and from there to the left ventricle. The blood is pumped into the systemic circulation via aorta. [13 p.16-17, 28-31]

The cardiac cycle is divided into the systole and the diastole phases. The closing of the atrioventricular valves starts the systole phase of the cardiac cycle. When listening the heart with for example a stethoscope, the sound coming from the closing atrioventricular valves can be heard as the first heart sound. During the systole phase, the semilunar valves open and the blood flows to the arteries. The closing of the semilunar valves ends the systole phase and starts the diastole phase of the cardiac cycle. The closing of the aortic valve can be heard as the second heart sound [14 p.340]. Soon after this the atrioventricular valves opens, and the blood starts to flow from the atriums to the ventricles. [13 p.28-31]

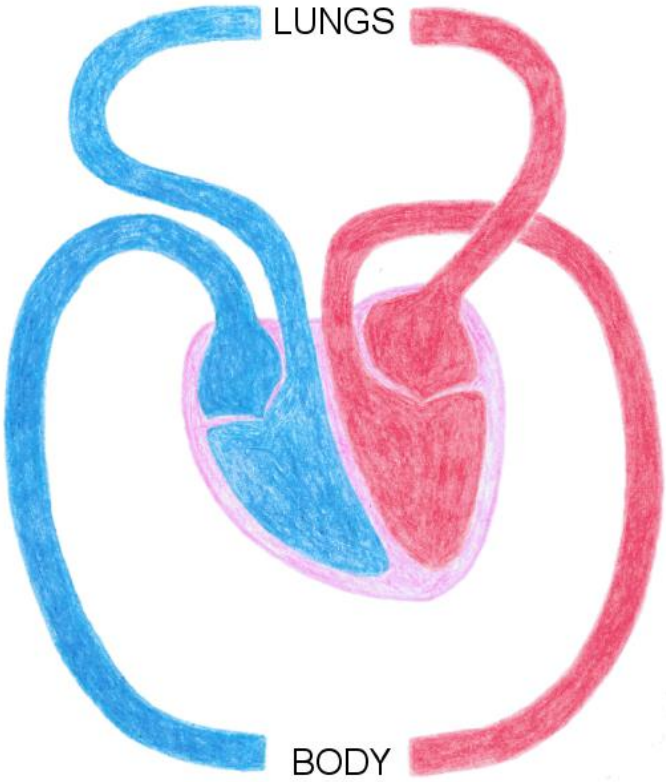

Figure 2. The simplified image of the systemic- and the pulmonary circulation.

2.3. The nervous system and the heart

The autonomous nervous system affects the functionality of the heart. It influences the heart rate, the contractibility of the heart muscle and the tone of the peripheral arteries and veins. The autonomous nervous system influences the heart also in humoral manner, for example via catecholamine and renin. The catecholamine is a neurotransmitter and hormone from the adrenal gland. The renin is an enzyme emitted by kidney. The autonomous nervous system is divided in to the sympathetic and the parasympathetic nervous systems. [13 p.58]

The activation of the sympathetic nervous system accelerates the heart functions. The sympathetic nervous system is activated when the body is experiencing stress. In stress the heart rate and the contractibility of the heart muscle is increased. Sympathetic nervous system influences the heart functionality mainly via beta_1-receptors in the sinoatrial node area, the atrioventricular node area and in the heart muscle surrounding the chambers. [13 p.58]

The parasympathetic nervous system acts as counterforce for the sympathetic nervous system. It controls the bodily functions such as the digestion and the diuresis during hibernation. The parasympathetic nervous system pacifies the heart functions. When this system is activated, heart rate slows down and the electrical conductivity decreases. The parasympathetic nervous system activation comes to the heart via

muscarinic receptors located mainly in the sinoatrial node area and the atrioventricular node area. [13 p.58]

The baroreceptors are sensors that sense the stretching in the artery wall. The baroreceptors are located in the carotid arteries and at the arch of aorta. The baroreceptors have neural connection to the central nervous system. The signals from the baroreceptors indicate the status of the arteries and affect for example to the blood pressure. In addition to the baroreceptors, there are also chemical receptors in the same area. The carotid body is a main group of chemoreceptors and it is located in the carotid artery. The chemoreceptors supervise gases and acidity of the blood and sends signals to the nervous system. [13 p.59-60]

2.4. The electrical activity of the heart

The majority of the heart cells are contractile myocardial cells. In addition to the myocardial cells, the heart muscle includes cells that are optimized for electrical activity. Through these electrical conducting cells, electrical impulse can travel in heart muscle and cause contractions in the right location at the right time. Tissue around the heart's four chambers acts as an electrical isolation. [13 p.16-18]

The sinoatrial node, which is located at the upper part of the back wall of the right atrium, generates the electrical impulses. In Figure 3 the electrical network of the heart is visualized. From the sinoatrial node, the signal spreads to the walls of the right and the left atriums. This causes the contraction of the atriums and the blood flow into the ventricles. From the atriums, the impulse travels via the atrioventricular node and the bundle of His towards the ventricles. The bundle of His is an electrical conductor between the ventricle and the atria. The bundle of His spreads into the right and the left branches. The right and the left branches are again spread into the thin Purkinje fibers surrounding the ventricles. [13 p.17, 48]

The flow of ions inside the cells and in the cell membranes enables the electrical activity of the heart. The sodium, the potassium and the calcium are the main ions that are involved in this electro-chemical process. The electrical conducting cells are electrically polarized having negative charge inside the cell in resting state. The potential difference over the cell membrane in resting state is usually -90 mV. The pumps in the cell membrane maintain this polarity. During the depolarization caused by the electrical impulse originated by the sinoatrial node, the electrical conducting cells charge changes momentarily. After the depolarization has travelled through the electrical conducting cells, the repolarization occurs, and cells return to their resting state. [13 p.48-49]

The depolarization and the repolarization are processes, which can be measured noninvasively with electrodes from the skin surface. The ECG is a commonly used method for measuring the electrical activity of the heart. [15 p.11]

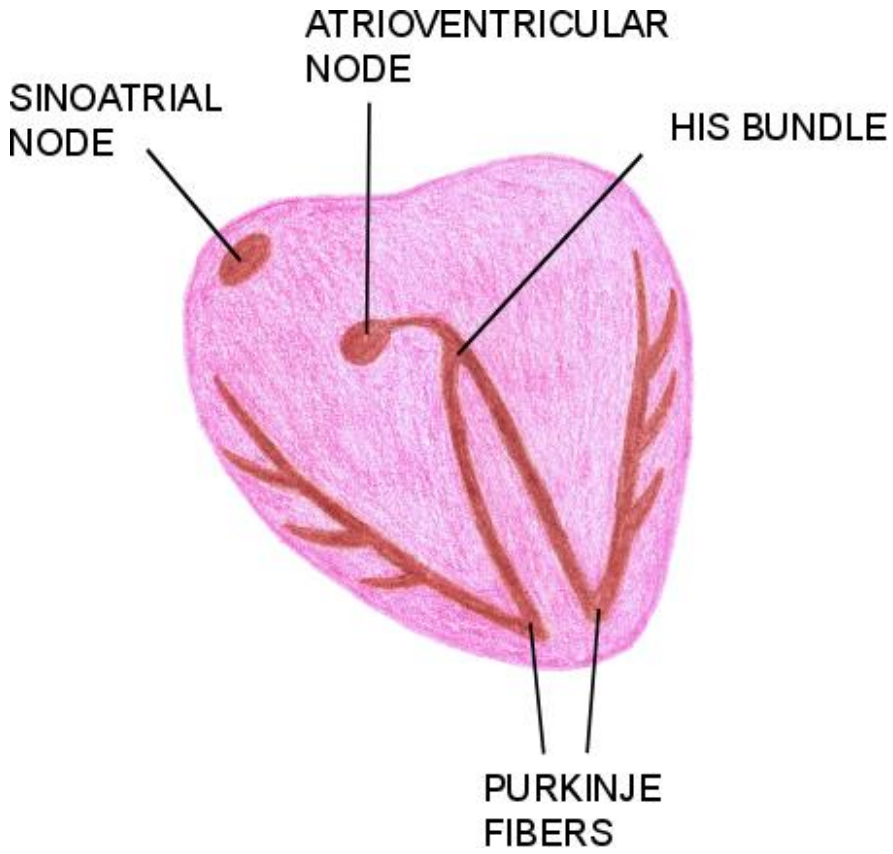

Figure 3. The simplified image of the electrical system of the heart.

2.4.1. ECG measurement

The ECG is the most common tool in investigating the functionality of the heart. The ECG system measures non-invasively changes in the voltage produced by the myocardial cells of the heart. In the ECG measurement, 10 electrodes are connected to the body and with these electrodes the 12-lead ECG is created [15 p.38]. The peripheral leads I, II, III, aVR, aVL and aVF measure potential differences between the limbs and show heart's electrical changes in frontal plane. The chest leads, V1-V6, are positioned to the left side of the chest area and measure heart's electrical activity in horizontal and sagittal plane. When studying some specific syndromes, some of the chest leads can be attached to alternative positions. For example, when studying the right ventricular infarction, the additional chest leads are positioned on the right side of the chest. The positioning of the chest leads too high or low will affect to ECG waveform. In addition, the position of the patient, lying or sitting, has an effect to ECG. Figure 4 shows the example of the signals of 12-lead ECG system. [16 p.31]

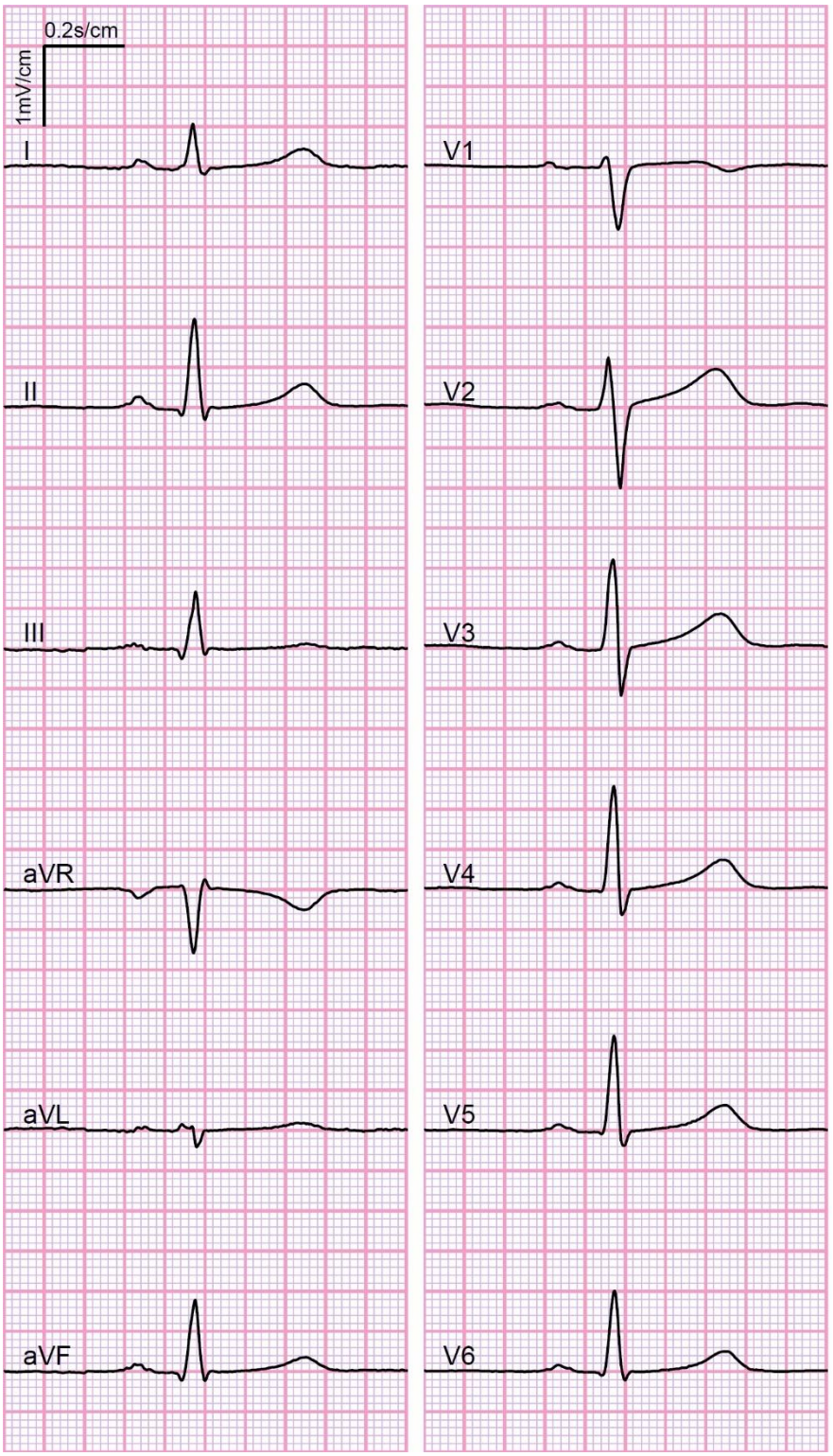

Figure 4. Example of the 12-lead ECG signals, permission to use this image acquired from the owner.

The basic structure of the ECG signal is visualized in Figure 5. When the electrical impulse from the sinoatrial node arrives to the atriums, they contract. This contraction can be seen in the ECG as the P-wave. The atrial cells' depolarizing characteristics in both atriums define the duration and the shape of the wave. During the P-Q interval of

the ECG signal, electrical impulse is traveling via the atrioventricular node and the bundle of His. This path is electrically narrow and the ventricles have enough time to fill with the blood. The depolarization of the ventricles forms the QRS complex in the ECG signal. The repolarization of the ventricles can be seen in the ECG signal as the S-T interval and the T-wave [17]. The origin of the occasionally seen U-wave is controversial; in different studies it has been associated to both repolarization processes of the heart and electromechanical functionality [17]. [16 p.31-32].

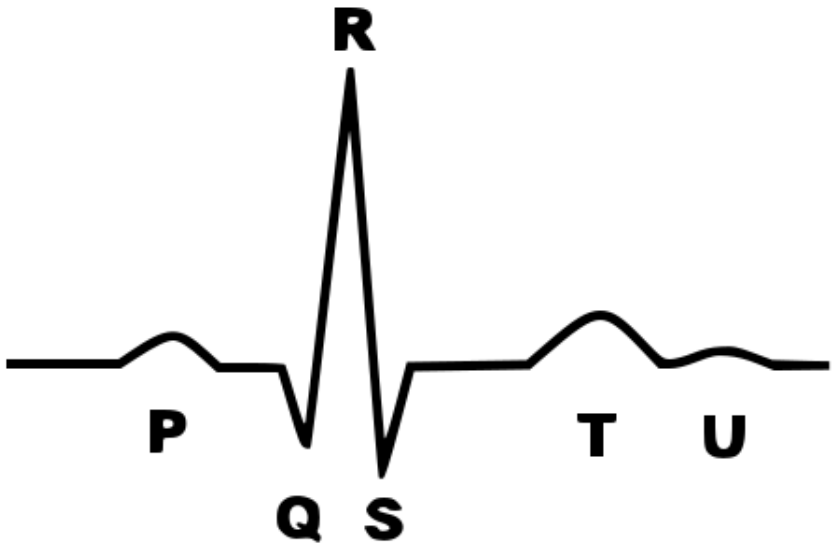

Figure 5. Simplified image of the ECG signal.

The ECG signal is used for analyzing several different cardiac conditions. The form of the different parts of the ECG signal depends on the structure of the heart muscle and can expose hypertrophy, heart failure, ischemia, myocarditis and certain system level diseases. [13 p.126]

The ECG measurement is prone to disturbances. Signal quality can be compromised for example by the muscle tension, the movement of the patient, the incorrectly connected electrodes or the alternating current. This type of disturbances can be seen in the ECG signal for instance as the fluctuation of the baseline, the high frequency noise and the inversion in the signal. [13 p.126-130]

2.4.2. Medical analysis of the ECG signal

The medical analysis of the ECG signal starts with the evaluation of the signal quality, finding possible disturbances and taking an overview of the signal. The next step is to study the speed and the regularity of the ventricular rhythm, the heart rate. The heart rate is defined from the sequential R peaks. If there is irregularity in RR-intervals,

several consecutive R peaks are used for calculating the average. The resting heart rate is usually between 60-100 beats per minute. [13 p.130-132]

After the heartrate is defined, the location, the form and the duration of electrical activation of the atriums, the P-wave, is examined. Due to the structure of the heart and the ECG measurement locations the P-wave is positive in leads I, II and aVF. The P-wave morphology can reveal arrhythmia and abnormalities in the atriums such as the atrial hypertrophy or the interatrial and the intra-atrial conduction issues. [13 p.132]

After the P-wave analysis the PQ-time, the conductance between the atrium and the ventricular, is measured. The duration of the PQ-time can be used for discovering the atrioventricular block or electrical shortcut. The long PQ-time usually refers to conduction issues in the atrioventricular node. This type of conduction issues can be caused by the drop in the sympathetic tone or the rise in vagal tone [16 p.32]. Cardiac drugs like digitalis or beta-adrenergic blocking agents cause similar affect as well as certain diseases that affect to atrioventricular junction [16 p.32]. The shortened PQ-time can be caused by the impulses bypassing the atrioventricular node [16 p.32]. [13 p.133-134]

Next in the ECG signal comes the QRS complex. The QRS complex reflects the ventricle depolarization phase of the heart. The left ventricle is larger in size than the right ventricle. The depolarization of the left ventricle dominates in the QRS complex. From the QRS complex the duration and the amplitude are analyzed. The high amplitude might be a sign of the ventricular hypertrophy [13 p.134]. The fascicular blocks can be detected as changes in the early part of the QRS complex and as changes in the electrical axis in the frontal plane. The bundle branch blocks can cause wide QRS complex. The bundle branch blocks can be caused for example by the calcification or the fibrosis. [16 p.32-33]

After the QRS complex, the T-wave is studied. The T-wave is caused by the ventricular repolarization. From the T-wave, the starting point and the peak are analyzed [13 p.137]. The T-wave can have one or two peaks [13 p.137]. The amplitude of the peak is measured. The T-wave can be symmetric, asymmetric or trench-like. It can be neutral, positive or negative. The T-wave direction is the same as the QRS complex direction. The T-wave is usually positive in leads I, V_5 and V_6 and negative in leads aVR and V_1. The purpose of the T-wave analysis is to detect for example hypertrophy, myocardial ischemia or the stress in the right or the left side of the heart. [16 p.34-35]

The T-wave is sometimes followed by the smaller U-wave and it can be difficult to separate the two. The source of the U-wave is disputable [17]. It can be detected easiest in leads V_2, V_3 and V_4. The high U-wave amplitude can refer to the hypokalemia, but it can also correlate to usage of some cardiac drugs. [16 p.35]

Next, the direction of the signal between the S and the T is interpreted. The ST area can reveal diseases like ischemia of the heart muscle, the acute pericarditis, the stress of the left chamber and the ER. The rising ST can be a sign of the ER, but it can be also found from healthy individuals. [13 p.137]

As the final step of the ECG analysis the QT-time is measured. The QT-time includes both the depolarization and the repolarization phases. It is measured from the onset of the Q-wave to the end of the T-wave. Based on the literature the QT-time is longer with females than with males. Along with some heart diseases, the duration of the QT-time depends also on the heart rate. With fast heart rate the QT-time is shorter and with slow heart rate it gets longer. Due to this, there are correction factors used in

the QT-time analysis. The long QT-time can refer to hereditary ion channel disease, drug effect, ischemia or electrolyte imbalance. [16 p.35]

2.5. Common heart conditions

Due to the complexity of the heart, there are many structures that can be defected. In this chapter, some common heart diseases are shortly presented.

The atherosclerotic coronary artery disease (CAD) is a group of coronary heart diseases. The atherosclerotic CAD is getting more common as obesity and type 2 diabetes mellitus increases in population. The CAD group consists of the chronic stable angina, the acute coronary syndromes, the congestive heart failure, the sudden cardiac death (SCD) and the cardiogenic shock. [16 p.97].

The myocardial diseases and the cardiomyopathy are conditions related to the heart muscle. There are three types of the cardiomyopathies; dilated, hypertrophic and restrictive. The dilated cardiomyopathy (DCM) is a functional problem of the left or both ventricles. In the DCM, the ventricles are dilated and have impaired contraction capability. The hypertrophic cardiomyopathy (HCM) is a heart disease where the left ventricle suffers from the hypertrophy. The HCM does not always cause symptoms and first sign of the disease can be the sudden cardiac arrest. The rarer form of the cardiomyopathy is the restrictive cardiomyopathy (RCM) which causes stiffness in the myocardium. The RCM prevents the heart tissue to stretch and fill with the blood normally [13 p.796]. Another heart muscle related disorder is he myocarditis. The myocarditis is inflammatory process. In the ECG, the myocarditis is usually seen as a nonspecific ST-segment and the T-wave changes, the atrial and the ventricular arrhythmias, the atrioventricular blocks and the wide QRS complex caused by the intra ventricular conduction issues. The heart failure (HF) is one form of the myocardial diseases. The HF is a syndrome where the left ventricle's capability to fill or eject blood is decreased. The HF can be caused by underlying heart disease, for example the dysfunction of the myocardial muscle with simultaneous dilation or hypertrophy of the left ventricle. [16 p.145, 155, 161, 181, 184]

The heart can suffer from the cardiac rhythm abnormalities. In the bradyarrhythmia the heart beats less than 60 beats per minute. The reason for the slow heartbeat in the bradyarrhythmia is a delay or a block in the conduction system of the heart. The problem can be in the electrical conduction network of the heart, drugs that change the functionality of the conductive cardiomyocytes or for example, diseases that affect the blood supply or the electrical conductivity of the heart. The supraventricular tachycardia (SVT) is the disease where the heartbeat is accelerated. The reasons behind the SVT are related to the disorder in the electrical system of the heart, locating above the His bundle. The SVT can be caused by the reentrant electrical circuit or by the atrial impulse outside the sinus node. The atrial fibrillation (AF) is the supraventricular tachyarrhythmia, which is originating from the random atrial activation. In the ECG the AF can be noticed as the P-wave abnormality. In the AF, the P-wave is replaced by the rapid oscillation. The ventricular tachycardia (VT) is the disease where the heart beats over 100 beats per minute. The cause of the VT is located in the conduction system distal to the His bundle or in the ventricular myocardium. In most of the VT cases, the ECG finding is a wide QRS complex. Patients with the VT usually have other heart diseases in the background. The syncope and the sudden cardiac death are also associated to the VT. [16 p.215, 223, 233, 241]

There are several different types of the valvular heart diseases. The blood flow from the left ventricle to the aorta can be limited by the aortic stenosis. The valvural aortic stenosis is disturbing the normal movement of the valve. Limited opening of the valve increases pressure in the left ventricle. This condition usually leads to the left ventricle hypertrophy and can cause abnormality of the left atrial. Another valve related condition is the aortic valve regurgitation. During the diastole, some of the blood from the aorta flows back in to the left ventricle. If the regurgitation is severe, it can cause the ventricle hypertrophy and lead to the systolic malfunction and the heart failure. In the mitral valve disease, the functionality of the mitral valve is decreased. This can be due to the mitral valve stenosis where the valve opening during diastole is not complete or the mitral regurgitation where the closing of the valve is not complete during the systole. This condition causes hypertrophy of the left atrium and the left ventricle and predisposes to the atrial fibrillation. [13 p.818-823, 835, 846-853]

The two-layered membrane called the pericardium surrounds the heart. The acute pericarditis is the inflammation of the pericardium usually caused by virus or bacteria. The acute pericarditis shows in the ECG as various changes in the T wave morphology in different stages of the disease. [13 p.912-915]

High blood pressure, the hypertension, is a contributor in many coronary events. The hypertension is caused by the increased resistance of the vessels, which forces the heart to use excess energy for pumping the blood. In time, the hypertension can lead to the left ventricle hypertrophy. [13 p.945]

2.6. The early repolarization

The electrical impulse is initiated in the sinoatrial node and it travels through the heart tissue causing contraction in the chambers as explained in Chapter 2.4. The ventricular depolarization shows as the QRS complex in the ECG signal [16 p.32]. The ER is premature regression of the polarization and is seen as a certain type of fragmentation at the end of the QRS complex [8]. The terminology related to the ER is two folded. The ER syndrome (ERS) is the condition where the VF or the VT is present together with the ER and patient does not have organic heart disease [9]. Otherwise, the ER condition can be referred to as the ER pattern (ERP) [9].

2.6.1. Significance of the ER

The prevalence of the ER varies between 1 - 24% of the population depending on the study and most likely also on the applied ER definition [1, 5, 18, 19]. The ER has originally been considered harmless [1], but the study of Haïssaguerre et al. in 2008 showed that there is a connection between the ER and the increased risk of cardiac death [2]. In the study by Haïssaguerre et al. the ECG signals of the group of patients with the idiopathic ventricular fibrillation (IVF) were analyzed. The finding was that the 31% patients having the IVF had the ER compared to the control group where the prevalence of the ER was 5%. The connection between the IVF and the ER was found also in the studies by Nam et al. in 2008 [3] and by Rosso et al. in 2008 [4].

There have been many publications made concerning the ER during the last decade. Tikkanen et al. concluded in 2009 that for middle-aged people the ER in the inferior

leads of the ECG can be connected to the increased risk of the cardiac death [5]. In 2010, the ER was connected with the ventricular fibrillation storms by Antzelevitch and Yan [6]. The ER notching in the inferior leads has been associated with the increased risk of the CAD patients' life-threatening ventricular arrhythmias in research by Patel et al. in 2010 [20]. Based on the research by Tikkanen et al. in 2012, the ER is one component causing the fatal arrhythmia during the acute myocardial ischemia [21]. Kawata et al. reported 2013 that the ER in the inferolateral leads has aggravating effect on the Brugada syndrome in patients with the underlying VF [22]. Coexistence of the long QT time and the ER with high J-point elevation has been associated with cardiac symptoms in study of Laksman et al. in 2014 [23]. Stumpf at al. studied group of long-term endurance sport athletes and reported in 2016 that the ER is suspected to be one factor inducing the AF in athletes [24]. Research by Roten et al. 2016 proposes dividing the ER into the malignant and the benign forms based on the T-wave analysis [7]. The harmfulness of the ER can depend on the race based on Walsh et al. 2019 [25]. In 2019, Haïssaguerre et al. reported that also late depolarization can appear as the J-wave in the ECG signal [26]. Holkeri et al. studied connection between the ER and the SCD [27]. Their result shows that in the case of Caucasian adults of age under 50 years, the ER can be associated with the SCD.

The definition of the ER was imprecise until 2015 when group of lead scientists in the area of cardiology published the first consensus paper to clarify the characteristics of the ER [8]. The work continued with the consensus conference report in 2016 [9]. The purpose of the consensus work was to provide guidelines for the future research. The content of the consensus is described in the Chapter 2.6.2.

2.6.2. Definition of the ER

The ER can be seen in the ECG signal as the notch (Figure 6) or the slur (Figure 7) in the downslope of the R peak. MacFarlane et al. have presented the definition for the ER pattern in the consensus paper [8]. In the consensus paper, the recommended terminology is following: the *J onset* means the onset of the notch, the *J peak* means the peak of the notch or the onset of the slur and the *J termination* means the end of the notch or the slur. [8]

In the consensus paper, it is stated that the ER notch or slur is visible on the final 50% of the downslope of the R peak. The notch and the onset of the slur are located above the baseline. The amplitude of the J peak must be at least 0.1 mV. Based on the consensus the notch or the slur is present in at least two contiguous inferior and/or lateral signals, excluding signals V_1-V_3. The consensus team's proposal is that the QRS duration is measured from the signal, which does not have notch or slur. In the case of the ER, the QRS duration should be smaller than 120 ms. If the ER includes the ST segment rising towards the T-wave it is called the "early repolarization with an ascending ST segment". Again, if ER includes the horizontal or downhill ST segment it is called the "early repolarization with a horizontal/descending ST segment". [8]

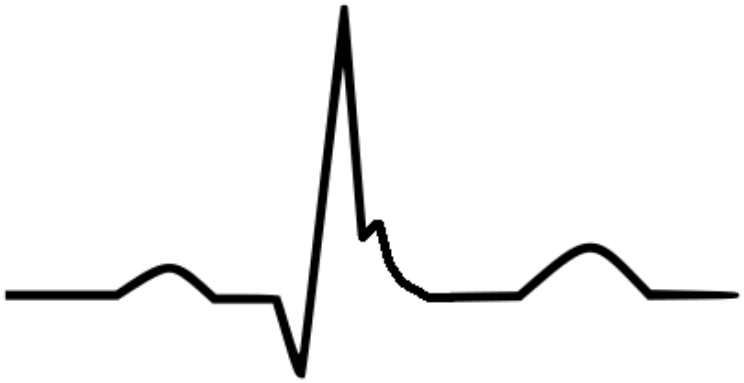

Figure 6. The simplified image of the ECG with the ER notch morphology.

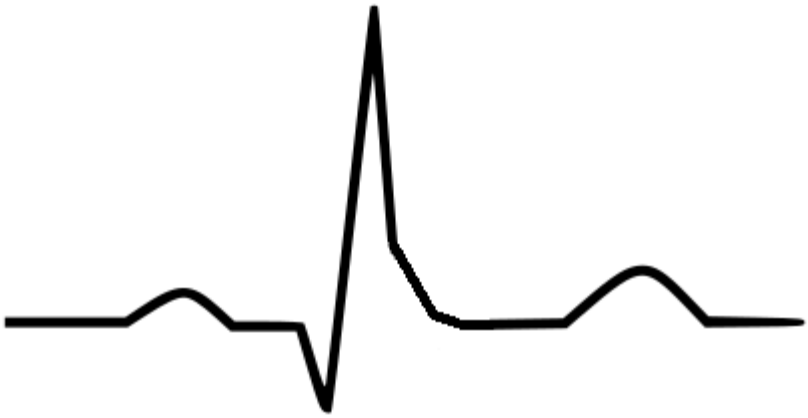

Figure 7. The simplified image of the ECG with the ER slur morphology.

The proposal in the consensus paper is to take following measurements when studying possible notch

- the amplitude of the J onset
- the amplitude the J peak
- the amplitude the J termination of the notch
- the duration from the J onset to the J peak of the notch
- the duration from the J onset to the J termination of the notch

The proposal in the consensus paper is to take following measurements when studying possible slur

- the amplitude of the J peak
- the amplitude of the J termination of the slur
- the duration from the J peak to the J termination of the slur

In defining the ST-segment slope, the consensus proposal is to measure the amplitude difference between the J termination of the notch or the slur and the ST-segment level at 100 ms after the J termination. If the ST-segment value is the same or less than the J termination value, the ST-segment is considered to be horizontal or downward sloping. If the ST-segment value is greater than the J termination value, the ST-segment is classified as upward sloping. [8]

2.6.3. Methods for detecting the ER

There are several different algorithms developed for the automatic detection of the ER over the years. Some of the existing methods are presented in this chapter. Maheshwari et al. developed 2013 a method for the automatic detection of the general QRS fragmentations [10]. In this method, the ECG is first preprocessed by correcting the baseline wandering and removing the noise. Then the QRS complex is extracted from the ECG by utilizing the information of the QRS morphology in time-domain and using gradient-based feature extraction algorithm. The QRS fragmentation is detected with the discrete wavelet transform using Haar wavelet. The morphology of the detected fragmentation is studied. The performance of this method was tested the with data from 31 patients whose data was evaluated by two cardiologists. The ECG data used in the study is from PhysioNet's PTB data set. Out of the 372 leads, method correctly analyzed 334 leads. The sensitivity of this approach is 89.7% and the specificity is 89.9%. [10]

Clark et al. introduced 2014 an automated algorithm for detecting the ER [11]. They utilized previously developed method for finding the notch by tracking changes in the ECG signal. The slur onset is detected by using two lines collinear to the R peak slope; one locked at the top of the R peak and another locked at the end of the slope. These two lines are settled over the R slope the way that the combined area between the signal and the lines is minimized. In this study, 100 ECGs from healthy young men (circa 22 – 28 years) from West of Scotland were used and they were randomly split into the equal sized training and testing sets. The result is measured lead-wise and for the test set the sensitivity is 90.5 % and the specificity is 96.5 %. [11]

Kenttä et al. developed an automated ER detection method 2015. As the preprocessing the studied signal is resampled to 500 Hz, the 50 Hz disturbance is removed, and signal is smoothed by using Savitzky-Golay filter. Next, the baseline of the signal is corrected, and the signal's morphology type is analyzed. For detecting the notch, local peaks are searched from the R peak's slope. If local peaks are found those are analyzed to determine if the notch criteria are met. If the local peaks are not detected the slur option is studied. The slur onset is searched by studying the changes in the angle of the R slope. If the angle change is larger than the defined threshold, the slur is detected. The ECG data used in this study is from Health 2000 data set. The sensitivity of the method is 96.2 % and specificity 90.1 %. [12]

Tobón-Cardona et al. developed 2018 waveform prototype-based feature learning method for detecting the ER. In this approach, feature vectors are generated from the ER pattern location from the ECG signals. In the research, three classifiers were compared: the linear discriminant analysis, the k-nearest neighbor algorithm and the support vector machine. The support vector machine produced best results; the sensitivity of the method was 91.80% and specificity was 92.73%. [28]

3. AUTOMATIC DETECTION OF EARLY REPOLARIZATION

The objective of this work is to create a method for the automatic detection of the ER from the ECG data. The ER definition used in this work is based on the consensus paper [8]. The algorithm implementation is done in the Matlab environment. The ECG data utilized in this work is pre-classified into the ER and the non-ER groups. The pre-classified data is divided into the training set and the test set. Signals from the training set are used for the algorithm development. The signals from the test set are used only for the testing purpose. The description of the data is presented in Chapter 3.2.

The algorithm design work started with the literature review concerning the ER phenomenon. The basic idea of the algorithm started to form by visually studying the ECG signals of the training set, analyzing the different shapes of the QRS configurations of the ER and the non-ER signals and reflecting these signals in to the ER definition from the consensus paper [8]. The ECG data consists of 12 signals: I, II, III, aVR, aVL, aVF, V_1, V_2, V_3, V_4, V_5 and V_6. The consensus paper defines eight ECG signals to be used in the ER detection: the five lateral signals I, aVL, V_4, V_5, V_6 and the three inferior signals II, III, aVF. [8]

3.1. Description of the algorithm

The structure of the automatic ER detection algorithm developed in this master's thesis is presented in this chapter. The different functional entities of the implementation, such as the R peak location estimation, the baseline determination, the notch detection and the slur detection, are separated as sub-functions inside the main function. During the development of the implementation, the training set of the ECGs is used for development phase testing and improving the functionality.

At the early stage of the processing there are quality checks for estimating if the signal is good enough for the automatic ER detection. If the ECG signal has quality issues, for example one or more of the eight signals are empty; the automatic ER detection cannot be done reliably. In case of the quality issues, the signal is not automatically processed but the signal ID is saved to text file for manual evaluation.

3.1.1. Preprocessing of the ECG

The first step is to read in the patient's 12 lead ECG. Next, the signals are arranged into the matrix so that first eight signals contain the INF and the LAT leads. After this the content of the first eight signals are checked for validity. If any of these signals are containing only zeros, the ECG is not further processed. In this case, the signal is added to the list of signals to be manually examined. After this, the signals' sampling frequency is checked. If the sampling frequency is under 500 Hz, then signals are resampled to achieve 500 Hz sampling frequency.

Some ECG signals used in the development suffer from the 50 Hz disturbance. Based on the trials during the algorithm development it became evident that the removal of this disturbance is not straightforward. The signal morphology is slightly affected by the 50 Hz notch filtering. The need for the filtering is decided based on the power spectrum of the signal. If there is clear peak at the 50 Hz then the filtering can be performed. Since the signal morphology is affected by the 50 Hz notch filtering,

filtering was left as the optional step and it was not utilized for test set results. Example of the 50 Hz notch filtering in Figure 8.

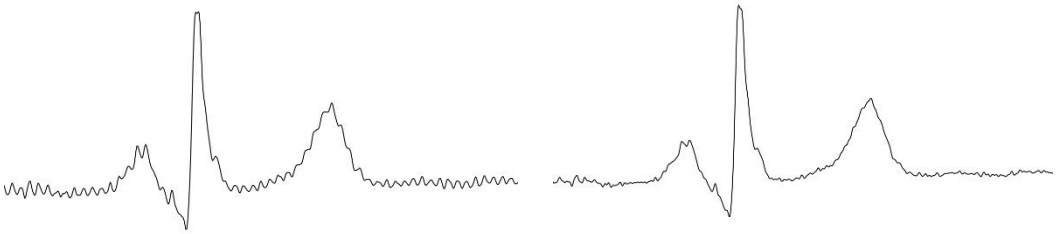

Figure 8. Left signal with interference, right the result of 50 Hz notch filtering.

3.1.2. Locating the R peak

The estimate about the R peak location is useful information for the later processing steps. Most commonly, the normal ECG signal has one high peak, which is the R peak. The R peak location has generally the same timing in all the 12 ECG signals. However, there are ECG signals, which might have fragmentation in top area of the R peak. Fragmentation can appear for example as there would be two or more lower peaks instead of one higher peak. In some signals, also the T wave top can be higher than the R peak. Due to these reasons, locating the R peak is more complicated than locating the highest peak of the ECG signal. In this implementation, the R peak location estimating is performed by utilizing all eight signals. First, the highest point and its location is searched for each of the eight INF and LAT signals. After this, the mean value of the locations is calculated. This is done for increasing the reliability of the general estimate of the R peak location. The local maximum value is then searched for each signal of the ECG, using the mean value of the previously calculated R location estimations as the center value and the experimentally selected range around it. After the R peak location is estimated, the information about the signals' possible abnormality is evaluated. This is done by smoothing the signals, for the purpose of this processing step, and finding all the peaks. Then the largest peaks are compared to the R peak estimate achieved earlier. If the signal obtains more than two almost as high peaks as the R peak, the signal is defined as abnormal and added to list of signals to be manually examined. Estimations of the R peak location of the ECG are utilized in baseline defining step. Example of the R peak location estimation in Figure 9.

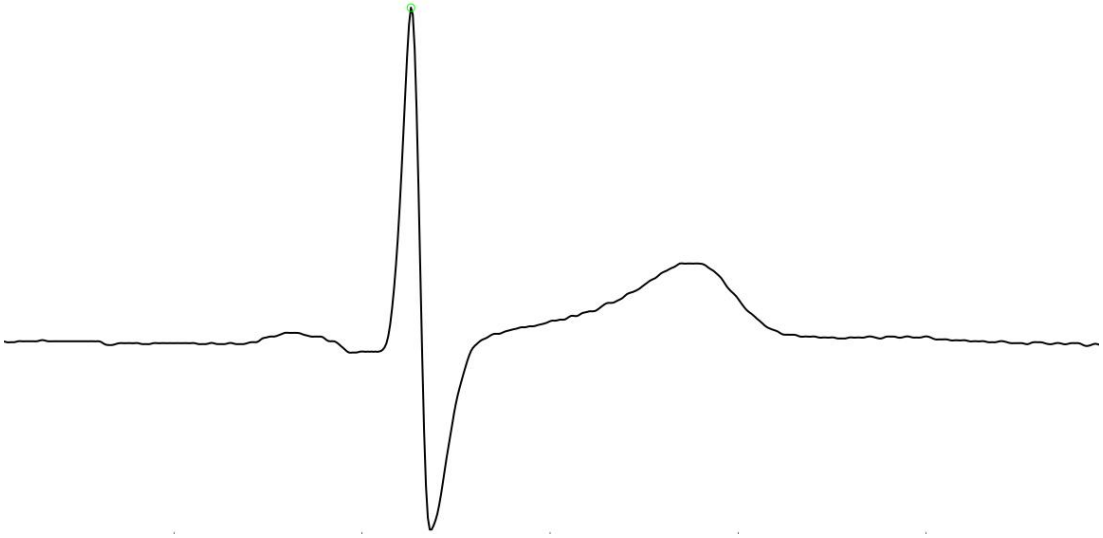

Figure 9. Example of the R peak location estimation, R peak location is the green circle in the image.

3.1.3. The baseline level

Defining the baseline level is important for achieving the as accurate as possible the slur and the notch related amplitude measurements. In this work, the baseline (zero level) of the ECG is decided to be the even area between the end of the P wave and the start of the QRS wave. The baseline calculation is performed by using the histogram data of the area between the P-end and the QRS start. Before calculating the histogram, the signal is smoothed from the studied area for the purpose of this step. The idea is that the most common value in the histogram data is considered to point out the true baseline. If there are more than one as common values found in the histogram these are studied further. In case these values are close to each other, their average is selected to be the baseline. If these values have large difference, the one that is closest to the original zero level is selected to be the baseline. In most cases, in the training set signals, there is only one most common value in the histogram data. In the case of the one common value, its validity is further evaluated. If the most common value differs very much from the original zero level, it is not trusted. In this case, the next most common values from the histogram are evaluated and the one closest to original zero level is selected to be the baseline. After the baseline is found, the zero level of the ECG signals is corrected for further processing steps. Example of the baseline level defining in Figure 10.

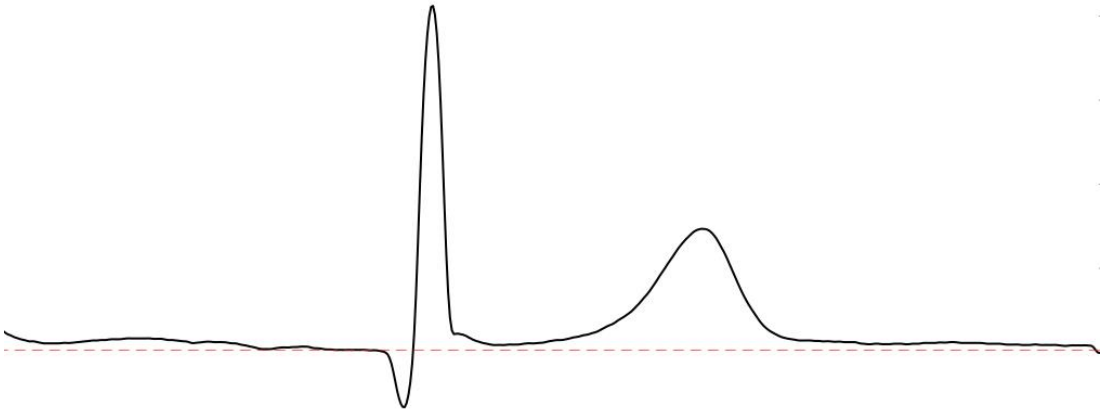

Figure 10. The example of the baseline level defining, the new baseline is dotted red line in the image.

3.1.4. Duration of the QRS complex

Next step is the QRS complex duration estimation. This starts by finding the Q starting position. By utilizing the R peak location estimation, the R peak top is defined. After this, using the R peak's amplitude, the root of the R peak's rising edge is located and the area from the R peak's root towards the P wave is selected for the evaluation. The adjacent value differences are defined, and signal's sign changes are analyzed for estimating the Q bottom location. After the estimation of the Q bottom is available, same principle is used for locating the Q start. Finding the QRS complex end position is performed with the same principles as the locating of the Q bottom and the Q start. From some ECG signals, it is difficult to locate both the Q start and the S end position reliably. In addition to this, the QRS complex duration should be measured from signals free from the ER. In this implementation, the QRS duration calculation is performed to all eight ECG signals. If from one of the eight ECG signals the Q start can be found reliably and the S end can be detected from another signal, this information is in this implementation combined to get the estimate of the QRS complex duration. In this work, the QRS complex calculation is left as an optional step and is not utilized in official test set results. This is due to the ECG signals in training and testing sets have been selected so that all signals have the QRS complex duration less than 120 ms. Example of the Q-start and S-end estimation in Figure 11.

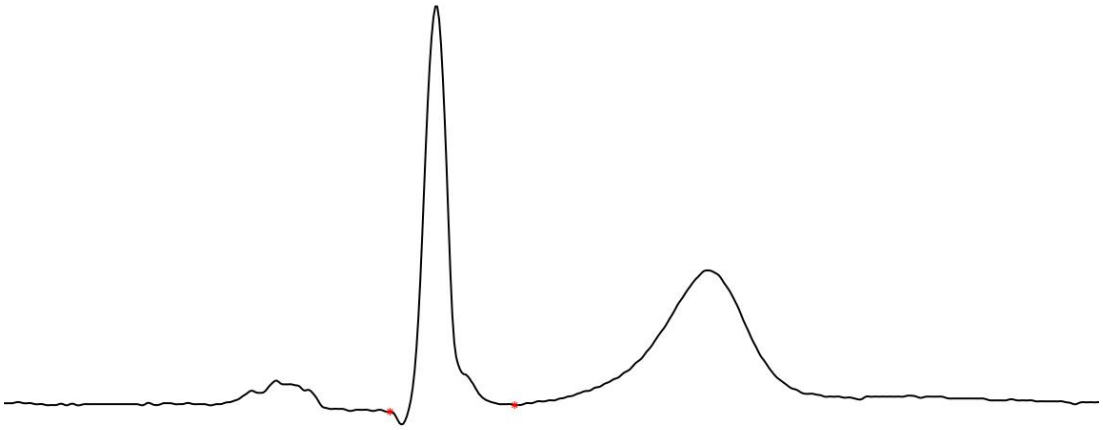

Figure 11. Example of Q-start and S-end estimation, red asterisk marks these positions.

3.1.5. Detecting the ER notch

Next step is to detect the possible notches. All the eight ECG signals are processed consecutively. Detecting the ER notch starts with determining the R peak amplitude. Next, the 50% value of the R peak amplitude is calculated for later use. After this, area from the R peak towards the T wave is processed. For this area of the ECG signal, the adjacent value differences are defined and from that data sign changes are evaluated. Sign change locations are further studied to discover the possible notch morphologies. If the notch morphologies are found, the J peak (notch peak) value is checked. The notch, which is under 50% of the R peak amplitude and over the set ER notch amplitude threshold is marked as the notch candidate. If the signal has too high or too low notches, this information is saved for later processing steps. In case there is the notch candidate found processing continues with ensuring that the J onset (onset of the notch) and the J termination (notch termination) are above the baseline. If these rules are met, the notch related measurements are defined. The measurements include the amplitude of the J onset, the amplitude the J peak, the duration from the J onset to the J peak of the notch “D1”, the duration from the J onset to the J termination of the notch “D2”and the ST type “M”. Example of the J peak and J termination presented in Figure 12.

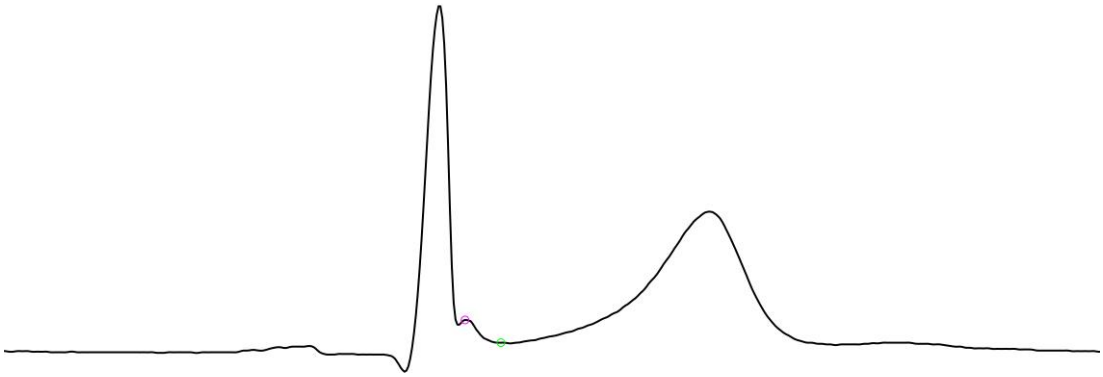

Figure 12. Example of notch detection, J peak is the red circle and J termination is the green circle.

3.1.6. Detecting the ER slur

After the possible notches are found the processing proceeds to the slur detection. All the eight ECG signals are processed consecutively, excluding the signals which have valid notch, too high or too low notch and the signals which S bottom is significantly under the baseline. The slur detection starts with defining the signal's maximum steepness from the R peak top onwards. After the maximum steepness is defined, signal's steepness changes are monitored from 52% location of the R peak's amplitude towards the T wave. Sudden change in the steepness in this area is treated as the possible J peak of the ER slur. Next, the amplitude of the J peak is defined, and the slur angle measured. The slur morphology can be either the shoulder shaped or the gentle slope. If signal has the shoulder, the slur angle is measured between the J peak to the shoulder end. If the signal has the gentle slope, the measurement is performed from the J peak to the J termination. The slur angle is compared to the angle measured from the area of the maximum steepness. If the difference of these angles is larger than the ER slur definition limit, the J peak's amplitude is fulfilling the ER slur amplitude definition limit and the J peak's amplitude is less than 50% of the R peak's amplitude, the signal is judged to have the ER slur. In case the signal is judged to have the ER slur, rest of the measurements are performed. The duration from the J peak to the J termination of the slur and the ST type "M" are measured. For all the signals which are processed, and which have the slur morphology, even if the ER slur definition is not met, the J peak amplitude and the angles are saved for the later use. An example of the slur detection measuring points in Figure 13. Reference steepness is the slope between blue and yellow marks and slur steepness is measured between red and green marks.

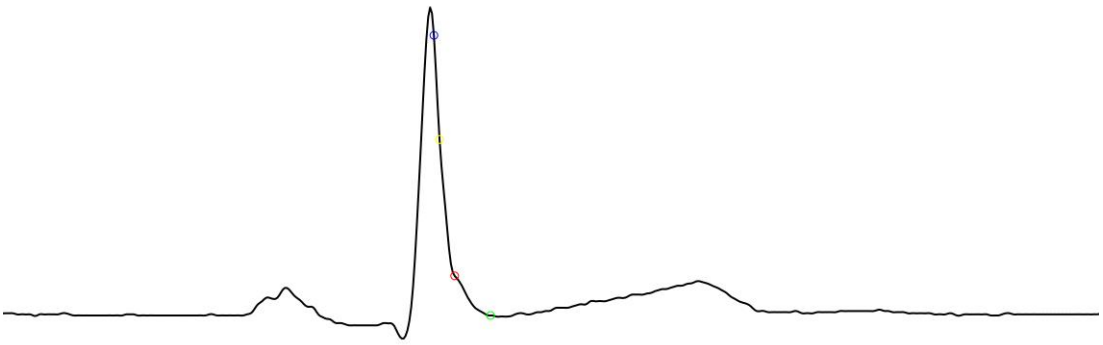

Figure 13. The example of the slur measurement locations.

3.1.7. The ER classification result

At this point, the ECG data is studied, and the results can be combined. Merging the notch and the slur information from the processed eight signals reveals how many ER positive signals are in the INF leads and the LAT leads. The consensus defines that the ER exists if two or more contiguous leads in the INF leads or in the LAT leads fulfills the ER requirement [8]. It is not made truly clear if there is actual importance of the lead order within the INF and the LAT leads and the information of the neighboring signals is not clarified in the consensus [8]. In the basic form of this implementation, the ER definition is fulfilled if either the INF leads have two or more ER positive signals or the LAT leads have two or more ER positive signals. In case of one ER positive signal is detected within the eight signals, situation is re-evaluated. This is due to the J peak amplitude measurement is heavily depending for example on the correctly defined baseline level and there is room for misjudgment. Depending on which leads, the INF or the LAT, the ER positive signal has been found all other signals inside that entity are re-reviewed. In the previously performed notch and slur searching steps, data from all the processed signals has been saved for this purpose. This extra check is performed with the slightly relaxed ER limits. This is done for catching the signals which are just under the ER limit. If there are new ER positive signals found, by using slightly relaxed ER limits, the ECG is defined as ER positive.

3.1.8. Reporting the results

As the ER conclusion is now achieved and all the required measurements are performed result can be printed. In case a single ECG is processed, the results such as ER or not ER and the ST type can be directly printed on the Matlab's command window. If larger amount of the ECGs is processed at once, it is more convenient to get the results collected in the text file. This implementation provides two alternatives for the results saving to the text files. The first approach is the simplified content, which includes the signal ID and the ER judgement for the INF and the LAT areas separately. The second option is to print more detailed information including the signal ID, the ER judgement for the ECG, the ER result in the INF, the ER result in the LAT, the QRS duration in ms, the ST analysis, the R peak location for the each eight leads,

the notch J onset amplitude as uV, the notch J peak amplitude as uV, the duration from the notch J onset to the notch J peak as ms, the duration from the notch J onset to the notch J termination as ms, the slur J peak amplitude as uV, the slur J termination amplitude as uV, the duration from the slur J peak to the slur J termination as ms.

3.2. The ECG data used in this work

The ECG signals used in this thesis are from the Health 2000 study. The Health 2000 study was coordinated by the National Public Health Institute, KTL, and was carried out in Finland in 2000-2001. The population of the study was 30-year-old and older people living in mainland Finland. The study consisted of the health interview and the health examination, which included a 12-lead resting ECG measurement. The health examination data was collected from 6354 persons. In this data set each lead of the 12-lead ECG is averaged from 10 second recording.

In this work, the 480 ECGs from the Health 2000 are used for the algorithm training and testing. These ECG signals were manually classified according to the 2015 ER consensus [8]. The ECGs with low quality, existing pre-exitation syndrome, prolonged QRS complex (>120 ms) or existing nonsinus rhythm were excluded from the set. For this work, the ECG signals were divided into the training and the test sets, both sets including the 150 ER positive and the 90 ER negative signals. The full Health 2000 data set is also utilized in this work for studying the algorithm predictability for the sudden cardiac death (SCD), the cardiac death (CD) and the all-cause-mortality (ACM).

4. RESULTS

The ER detection algorithm is a binary classifier which has four alternative results; true positive, true negative, false positive and false negative which are explained in the Table 1. The algorithm's ER classification result is compared to the manual ER classification of the test data set.

Table 1. Explanations for true positive, true negative, false positive and false negative

True positive	The algorithm implementation has detected the ER in the ECG and the ECG has been classified as the ER positive in the manual inspection.
True negative	The algorithm implementation has not detected the ER in the ECG and the ECG has been classified as the ER negative in the manual inspection.
False positive	The algorithm implementation has detected the ER in the ECG, but ECG has been classified as the ER negative in the manual inspection.
False negative	The algorithm implementation has not detected the ER in the ECG, but ECG has been classified as the ER positive in the manual inspection.

The result of the algorithm performance is presented as the sensitivity and the specificity values. The sensitivity describes how reliably the algorithm detects the existing ER and the specificity describes how reliably the algorithm detects the absence of ER. The sensitivity is the ratio between the true positive and the sum of the true positive and the false negative results:

Sensitivity = true positives / (true positives + false negatives).

The specificity is the ratio between the true negatives and the sum of the true negatives and the false positives:

Specificity = true negatives / (true negatives + false positives). [29 p.273]

4.1. The performance of the algorithm

The algorithm developed in this thesis is tested with the 240 manually classified ECG signals. The result of the algorithm is the sensitivity 94.0% and the specificity 92.2% (Table 2). Although the manual classification of the ECG signals was done based on the consensus definition, the precise consensus threshold was not used in this algorithm version. This is due to the manual classification was evaluated not be exact and therefore the algorithm thresholds are slightly relaxed, a compromise between the consensus and the manual classification. The differences between the manual classification and the algorithm following the precise consensus thresholds can be caused for example by the divergence in evaluation of the baseline level.

Table 2. Algorithm results compared to manual ER classification

Result type	Algorithm's classification	Manual classification
True positive	141	150
True negative	83	90
False positive	7	-
False negative	9	-

In the Table 3 the results of the algorithm developed in this thesis and the results of the few selected ER detection algorithms are compared. The descriptions of the compared algorithms are presented in Chapter 2.6.3. The approach of Maheshwari et al. is a general QRS fragmentation detection tool [10]. The algorithm by Clark et al. is specified for ER detection and presents enhanced method for previous slur detection approaches [11]. Kenttä et al. analyses the ECG signal morphology for ER detection [12]. Tobón-Cardona et al. compares supervised classification methods (waveform prototype-based feature vector) for ER detection [28]. The algorithm developed in this work has similar performance as these other methods when comparing the sensitivity and the specificity results. The direct comparison is however complicated due to many differences between the approaches, for example the amount of data used in training and testing, the quality of the pre-classification of the ECGs, the prevalence of the ER in the data set and different ER definition used.

Table 3. Comparison of the results of different approaches

Research group	Year	Sensitivity	Specificity
Maheshwari et al.	2013	89.7%	89.9%
Clark et al.	2014	90.5%	96.5%
Kenttä et al.	2015	96.2 %	90.1 %
Tobón-Cardona et al.	2018	91.80%	92.73%
Algorithm developed in this thesis	2020	94.0%	92.2%

4.2. The predictability of mortality

The algorithm results are compared to the Health 2000 ECG data to see if the algorithm predicts the ACM, the CD or the SCD. The results for the six variations of the algorithm are presented. The motivation for testing several algorithm versions is to get an idea of the significance of the different parts of the consensus-based ER definition and to reflect the results to earlier studies in the area.

The version 1 is the algorithm with the relaxed consensus thresholds. This is the same version as used for results in the Table 2. The version 2 is the algorithm with the relaxed consensus thresholds having the slur detection deactivated. At the early phase of the algorithm development, when the slur detection was not yet functioning properly, mortality correlation test was performed with Health 2000 data. At that point, the results showed statistically significant correlation to mortality. Therefore, couple of versions of the final algorithm which have the slur detection deactivated are included in this correlation study. The version 3 is the algorithm with the precise

consensus thresholds. This version is included to get information about the mortality correlation of the algorithm considering the ER consensus definition. The version 4 is the algorithm with the precise consensus thresholds having the slur detection deactivated. The version 5 is the algorithm with the otherwise precise consensus thresholds, but the J peak threshold set to 0.2 mV and the slur detection deactivated. This version is included in this test to compare the results with the findings by Tikkanen et al. [5]. The version 6 is the algorithm with the precise consensus thresholds with lead neighbors considered. In the consensus paper it is not very unambiguous whether the ER positive leads (within INF or LAT) should be adjacent to fulfill the ER definition. Therefore, the mortality correlation is tested also with the algorithm version considering only the adjacent ER positive leads. The short descriptions of the different versions are also shown in the Table 4.

Table 4. The description of the algorithm versions used in mortality correlation study

Version 1	The algorithm with the relaxed consensus thresholds (same version as used for results in the Table 2)
Version 2	The algorithm with the relaxed consensus thresholds and the slur detection deactivated
Version 3	The algorithm with the precise consensus thresholds
Version 4	The algorithm with the precise consensus thresholds and the slur detection deactivated
Version 5	The algorithm with the otherwise precise consensus thresholds but the J peak threshold set to 0.2 mV and the slur detection deactivated
Version 6	The algorithm with the precise consensus thresholds and lead neighbors considered (neighboring order used: INF: II/aVF and aVF/III, LAT: I/aVL, V4/V5 and V5/V6)

The statistical calculations with the SPSS for this chapter are provided by PhD Tuomas Kenttä from Medical Research Center, Institute of Clinical Medicine, University of Oulu. The Health 2000 ECG signals having the QRS length > 120 ms are excluded and depending on the case (ACM, CD or SCD) around 5557 ECGs were included. The results with the significance < 0.05 proceeded into the multivariate analysis. The multivariate analysis considers gender, age, regular smoking habit, body mass index, systolic blood pressure, cholesterol information, heart rate, left ventricular hypertrophy (combined Cornell and Sokolow-Lyon criteria), high blood pressure, diabetes, CAD, MI and the ER result.

The results indicate that two variations of the algorithm predict mortality. The results for the ACM are presented in the Table 5. The algorithm with the relaxed consensus thresholds with the slur detection deactivated (Table 5, Version 2) has correlation to the ACM in inferior leads. The results are Exp(B) 2.194, 95.0% CI 1.345 - 3.580 with p = 0.002 and after the multivariate adjustment Exp(B) 2.225, 95.0% CI 1.353 - 3.656 with p = 0.002. The algorithm with the precise consensus thresholds with the slur detection deactivated has correlation to the ACM in inferior leads (Table 5, Version 4). The results are Exp(B) 2.278, 95.0% CI 1.355 - 3.828 with p = 0.002 and after the multivariate adjustment Exp(B) 2.742, 95.0% CI 1.619 - 4.645 with p = 0.000. The algorithm version 2 indicates 2.2-fold increased risk for ACM and algorithm version 4 indicates 2.7-fold increased risk for ACM.

Table 5. The correlation to the ACM in the Health 2000 data set

Algorithm	ER/INF/LAT	Exp(B)	95.0% CI for Exp(B)	Sig.
Version 1	ER	0.774	0.574 - 1.044	0.094
	INF	0.752	0.496 - 1.141	0.181
	LAT	0.792	0.543 - 1.155	0.226
Version 2	ER	1.359	0.844 - 2.189	0.207
	INF	2.194	1.345 - 3.580	0.002
	Multivariate adjusted INF	2.225	1.353 - 3.656	0.002
	LAT	0.170	0.024 - 1.208	0.076
Version 3	ER	0.753	0.552 - 1.028	0.074
	INF	0.759	0.511 - 1.127	0.171
	LAT	0.723	0.464 - 1.126	0.151
Version 4	ER	1.461	0.883 - 2.418	0.140
	INF	2.278	1.355 - 3.828	0.002
	Multivariate adjusted INF	2.742	1.619 - 4.645	0.000
	LAT	0.210	0.029 - 1.496	0.119
Version 5	ER	1.726	0.430 - 6.935	0.442
	INF	3.381	0.841 - 13.581	0.086
	LAT	0.050	0.000 - 5795.215	0.614
Version 6	ER	0.900	0.635 - 1.275	0.553
	INF	0.979	0.655 - 1.463	0.916
	LAT	0.718	0.393 - 1.311	0.281

The results for the CD are presented in the Table 6. The algorithm with the precise consensus thresholds with the slur detection deactivated has correlation to the CD in inferior leads (Table 6, Version 4). The results for the CD correlation are Exp(B) 2.552, 95.0% CI 1.035 - 6.292 with $p = 0.042$ and after the multivariate adjustment Exp(B) 3.002, 95.0% CI 1.189 - 7.582 with $p = 0.02$. This result means that algorithm version 4 indicates 3-fold increased risk for CD. This algorithm version could be used as a tool for detecting patients with higher risk for CD.

Table 6. The correlation to the CD in the Health 2000 data set

Algorithm	ER/INF/LAT	Exp(B)	95.0% CI for Exp(B)	Sig.
Version 1	ER	0.575	0.313 - 1.056	0.074
	INF	0.513	0.208 - 1.264	0.147
	LAT	0.609	0.281 - 1.317	0.208
Version 2	ER	1.261	0.512 - 3.109	0.614
	INF	2.157	0.875 - 5.318	0.095
	LAT	0.048	0.000 - 40.693	0.378
Version 3	ER	0.571	0.304 - 1.073	0.082
	INF	0.655	0.303 - 1.418	0.283
	LAT	0.450	0.165 - 1.227	0.119
Version 4	ER	1.534	0.622 - 3.781	0.353
	INF	2.552	1.035 - 6.292	0.042
	Multivariate adjusted INF	3.002	1.189 - 7.582	0.020
	LAT	0.049	0.000 - 85.375	0.428
Version 5	ER	0.049	0.000 - 140159.341	0.692
	INF	0.050	0.000 - 43015099.393	0.775
	LAT	0.050	0.000 - 101857011.545	0.784
Version 6	ER	0.651	0.315 - 1.345	0.246
	INF	0.879	0.406 - 1.903	0.744
	LAT	0.214	0.030 - 1.536	0.125

The results for the SCD are presented in the Table 7. The algorithm with the precise consensus thresholds with the slur detection deactivated has correlation to both the ACM and the CD, it is interesting to evaluate the result also for the SCD for inferior leads. The result for the SCD case for this algorithm version in inferior leads is Exp(B) 2.642, 95.0% CI 0.963 - 7.252 with p = 0.059 (Table 7, Version 4). Since the significance value concerning the SCD is above 0.05, the result is considered non-significant and the multivariate analysis is not performed.

Table 7. The correlation to the SCD in the Health 2000 data set

Algorithm	ER/INF/LAT	Exp(B)	95.0% CI for Exp(B)	Sig.
Version 1	ER	0.625	0.320 - 1.221	0.169
	INF	0.530	0.193 - 1.456	0.218
	LAT	0.679	0.294 - 1.568	0.364
Version 2	ER	1.304	0.475 - 3.579	0.607
	INF	2.233	0.814 - 6.130	0.119
	LAT	0.048	0.000 - 100.831	0.437
Version 3	ER	0.607	0.301 - 1.223	0.162
	INF	0.730	0.316 - 1.687	0.462
	LAT	0.435	0.137 - 1.383	0.158
Version 4	ER	1.585	0.577 - 4.351	0.371
	INF	2.642	0.963 - 7.252	0.059
	LAT	0.049	0.000 – 232.522	0.484
Version 5	ER	0.049	0.000 - 1084562.551	0.727
	INF	0.050	0.000 - 797082500.445	0.802
	LAT	0.050	0.000 - 1775442872.474	0.809
Version 6	ER	0.743	0.340 - 1.623	0.456
	INF	0.980	0.424 - 2.264	0.962
	LAT	0.277	0.038 - 1.993	0.202

The results in this chapter indicates that the two variation of the implementation has correlation to the mortality within population participated in the Health 2000 study. The algorithm version with the relaxed ER consensus thresholds correlates with the ACM in inferior leads when the slur detection is deactivated. The algorithm version with the precise ER consensus thresholds correlates with the ACM and the CD in inferior leads when the slur detection is deactivated.

Algorithm version 1 with relaxed consensus thresholds and version 3 with precise consensus thresholds do not correlate to the ACM, the CD or the SCD in this study. The ER detection results in Chapter 4.1 indicate that this algorithm version can detect manually classified ER quite reliably. Although the amount of test data used for the results in Chapter 4.1 is relatively small, it is expected to include also signals with slur morphology. Considering this it seems based on this study, that the ER slur does not have connection to mortality.

Algorithm version 5 represents precise consensus threshold with the exceptions of J peak threshold raised to 0.2 mV and slur detection deactivated. This version does not correlate to the ACM, the CD or the SCD in this study. Same conclusion could be drawn from the results of the algorithm version 6 which has precise consensus thresholds with neighboring order considered. The selected signal order within INF and LAT used in this work is represented in Table 4. In this version the slur detection is included so that will probably have some impact on the results.

The mortality correlation statistics were generated as a very last step in this project with preselected algorithm versions. Although six different versions of the algorithm were included, while analyzing these results, it becomes evident that even more versions could have been tested to get more clear understanding of the limits of this implementation.

5. DISCUSSION

The purpose of this thesis was to create the automatic detection of the ER from the ECG signal. One main point in this work was to use the 2015 updated definition of the ER [8]. In the beginning of the project, the algorithm performance was agreed to be officially evaluated utilizing the manually classified ECG test set. Although the goal for this work was relatively clear, the journey turned out to be extensive learning experience.

One of the limitations of this work is the small data amount used in the training and the testing. The original idea was to use almost the full set of ECGs of the Health 2000 for official testing, but the existing ER classification of the Health 2000 required re-examination and due to time constraints, a smaller data set was agreed to be re-classified to comply the consensus definition [8].

The consensus paper [8] was used as the basic specification for the implementation. In addition to that, manually classified training set of the ECG signals was utilized for the development. During the implementation it became clear that the manual evaluation is not, maybe ever cannot be, as systematic in measurements as the Matlab code. Due to this dualism, the implementation is a compromise between the ER thresholds defined in the consensus paper [8] and the manually classified training set. Even though the official results were run with the slightly modified consensus thresholds, the strict consensus thresholds were included in set of the algorithm variations, which were tested for correlation to mortality with the Health 2000 data set.

The results for the automatic ER detection developed in this work, are presented in the Chapter 4.1. The algorithm performance compared to test set of the manual ER classification is the sensitivity 94.0% and the specificity 92.2%. In the Table 3, this result is compared to the few other published ER detection implementation results. There are some aspects to be considered when comparing the results between different implementations. Maheshwari et al. are detecting fragmentations in QRS complex and have used data from 31 patients (372 leads) for testing the lead-based performance [10]. The classification of the data is done by two cardiologists. The result of this approach is the sensitivity of 89.7% and the specificity of 89.9%. Clark et al. are detecting ER introducing enhanced method for slur detection [11]. The ECG data is measured from 100 young male and data set is divided to equal sized training and test sets. The lead-based result, with ER definition prior to the consensus, is the sensitivity 90.5 % and the specificity 96.5 %. Kenttä et al. are analyzing the signal morphology for detecting the ER [12]. The ER definition used in this study is prior to the consensus. The data set used for testing is the Health 2000 (over 6000 ECGs) and is ER classified by two experienced reader. The result of this approach is the sensitivity of 96.2 % and the specificity of 90.1 %. Tobón-Cardona et al. compared three machine learning solutions for detecting the ER [28]. The data used in the study is Health 2000 data set. The data is preprocessed to remove unsuitable signals for the machine learning methods. The ER definition in the study is based on the consensus. The best performer of the machine learning options is found to be the support vector machine with the sensitivity of 91.80% and the specificity of 92.73%. The direct comparison of the results of these different implementations is not straightforward. There are differences in ER definition used, amount of data used in training and testing, probable quality differences in the manual classification, the prevalence of the ER in the used data, different rules applied for excluding part of the data as unsuitable for the method and the results are reported either lead level or ECG level. The algorithm developed in this

work is classifying the ECG for ER or non-ER group at INF and LAT level. Compared to that the results of the lead-based approaches can be considered stricter.

During the development of the algorithm of this work, it was noticed that quite small adjustments in the thresholds affect to the classification results of this implementation. In addition, the effect of the absence of the slur detection was found to be interesting already at the early point of the development. Due to these observations, Chapter 4.2 includes results of correlation to the ACM, the CD and the SCD within the Health 2000 data set for the six variations of the implementation. The contents of these six different variations are explained in Chapter 4.2 and listed in the Table 4. The results in the Table 5 and the Table 6 shows that with this algorithm implementation there are correlations found with the mortality if the slur detection has been deactivated. The algorithm version 2, with the relaxed consensus thresholds with the slur detection deactivated, has correlation to the ACM in inferior leads. The result after the multivariate adjustment is Exp(B) 2.225, 95.0% CI 1.353 - 3.656 with p = 0.002. This result indicates 2.2-fold increased risk for the ACM. The algorithm version 4, with the precise consensus thresholds with the slur detection deactivated, shows correlation to the ACM and the CD in inferior leads. The result after the multivariate adjustment for the ACM is Exp(B) 2.742, 95.0% CI 1.619 - 4.645 with p = 0.000. The result suggests 2.7-fold increased risk for the ACM. The result after the multivariate adjustment for the CD is Exp(B) 3.002, 95.0% CI 1.189 - 7.582 with p = 0.02. This result refers to the 3-fold increased risk for the CD. The correlation results for the SCD are presented in the Table 7. There are no significant correlations found. With the SCD, the algorithm version 4 in inferior leads is closest to significance level (Exp(B) 2.642, 95.0% CI 0.963 - 7.252 with p = 0.059), but since the p value is > 0.05 the multivariate analysis is not performed. The correlation to mortality when the slur detection is deactivated, could indicate that either the slur detection implementation is not working accurately enough, the consensus definition of the slur is not optimal or that the existence of the ER slur does not predict mortality at all or within the Health 2000 population. Considering the relatively high sensitivity and specificity in the results in the Chapter 4.1. the slur detection could be expected to work tolerably. The presented results concerning the slur are conflicting with findings by Tikkanen et al. [5], but similar with findings from Rollin et al. [30].

Another observation from the results with this implementation is that the ER in the lateral leads seems not to correlate with the mortality. The importance of the ER in the inferior leads has been reported also by Tikkanen et al. [5] and Sinner et al. [18]. In the research of Tikkanen et al. [5], the connection between cardiac based death and the inferior ER (definition prior consensus) with threshold 0.2 mV has been detected within middle age subjects. Due to this, the algorithm version 5 in this work has the 0.2 mV J peak threshold along with the slur detection deactivation. With this algorithm within the full Health 2000 population, there was no significant correlation found with the higher J peak threshold and the mortality. This could be studied further with different age groups. The result for algorithm version having the 0.2 mV threshold with the slur detection activated is not available. Rollin et al. have studied the effect of higher J peak amplitude and they did not find significant difference in mortality rates between J peak under or above 0.2 mV [30]. Algorithm version 6 is included for investigating the effect of the significance of the lead neighboring order. The results of the algorithm version 6 in the Table 5, the Table 6 and the Table 7 are non-

significant. In the future, neighboring effect could be tried with the slur detection deactivated.

It would be interesting to see what kind of results the algorithm developed in this work would give for different age groups and different data sets. Tikkanen et al. [5] and Holkeri et al. [27] have indicated higher risk for the SCD in certain age groups with the ER. Sinner et al. [18] have also connected certain age groups having the ER with the increased risk of the cardiac death. There are also many possibilities for the further developing this implementation. The automatic ER detection algorithm can be optimized in many ways to support the ECG data varieties and signal qualities. During the development, the most challenging phases were the slur detection, the notch and the slur termination detection and defining the correct baseline. All these functions have room for further optimization. In the future, it would also be interesting to refine the QRS length evaluation to be less sensitive to signal quality issues and to develop an improved method for alternating current interference removal.

The ER detection implementation developed in this thesis could be considered as prescreening tool for detecting the consensus [8] based ER from the ECG signal. For improving the prescreening performance, the sensitivity could be increased by adjusting the thresholds of the main variables of the implementation. This would lead to having more false positives, but as a prescreening tool, it would catch true positives more comprehensively. Holkeri et al. have developed the ECG risk score for predicting the increased risk for the SCD [31]. The developed score considers several different heart related conditions and is able to identify patients with increased risk for the SCD. The variations of the implementation developed in this thesis could be considered to be used an additional tool for evaluating the increased risk for the ACM or the CD, independently or as a part of some combined tool such as the ECG risk score.

One very intriguing part of this work has been the possibility to study the heart and its electrical functionality. The complexity of this organ is astonishing. Studying the heart's electrical conduction system, the ER definition and going through large amount of ECG signals raised some questions. Since the R peak's amplitude can vary from patient to patient and also lead to lead, is the fixed notch J peak amplitude threshold optimal approach in the ER definition or should it be in proportion to the R peak's amplitude? Another obscurity is related to preprocessing of the ECG signals. The implementation created in this work takes the 12 ECG signals as an input. Like ECGs used in this work, some ECG measuring systems create 12 lead ECG signals by combining the signal from several consecutive waves from each lead. In this case, each final signal is an average of these consecutive pulses. Even though this is one way of reducing the noise of the signal, it also has disadvantages. In case the patient would have for example arrhythmia, which would appear during the measurement, the resulting signal morphology would be distorted. Same issue would appear if one of the consecutive signals would be deformed due to momentary heart event. When trying to detect automatically features or abnormalities from the ECG signal it would be beneficial to have the raw data for optimal results.

The time spent with this project has taught many things about the heart, the ECG and the signal processing. The most importantly this thesis has taught how much there is still to learn. For developing high quality biomedical signal processing algorithms, the medical knowledge of the measured system is as important as is the understanding of the measurement system. Cardiology is a fascinating field, and it has been a privilege to get a glance of this complex area during this project. The current research

around predicting the increased risk for the CD and the SCD is intense. The constantly increasing knowledge about the significant markers of the CD and the SCD are enabling better analysis-tools for preventing the premature deaths.

6. SUMMARY

In this thesis, one type of the heart's electrical malfunction, the ER, is studied. The primary purpose of this thesis is to develop and implement an algorithm for automatic detection of the ER from the ECG data, considering the 2015 refined ER definition. The implementation of the algorithm is done in Matlab. Since the ER is visible and defined in time domain in the ECG signal, the time domain approach is also selected to be used in the automatic detection algorithm developed in this thesis.

This thesis includes basic information about the heart; its structure, functionality and common heart related diseases. In more detail the electrical activity of the heart, fundamentals of the ECG measurement, the ECG signal interpretation and the ER phenomenon are described. Some of the current methods for the automated ER detection are reviewed and the algorithm developed in this thesis is described. The presentation of the results and the analysis of the results conclude the thesis.

The performance of the ER detection algorithm developed in this thesis is the sensitivity 94.0% and the specificity 92.2%. The algorithm results' correlation to mortality is also studied with the Health 2000 data. The two variations of the algorithm show correlation to mortality within the Health 2000 population. The algorithm version with the slightly relaxed consensus thresholds indicates increased risk for all-cause-mortality in inferior leads, when the slur detection is deactivated. The algorithm version with the precise consensus thresholds indicates increased risk for all-cause-mortality and cardiac death in inferior leads when the slur detection is deactivated.

The algorithm developed in this work can be considered as a prescreening tool for detecting consensus-based ER. The algorithm versions which have correlation to mortality could be used for evaluating the increased risk for the ER based ACM or CD.

7. REFERENCES

[1] Klatsky A.L., Oehm R., Cooper R.A., Udalstova N., Armstrong M.A. (2003) The early repolarization normal variant electrocardiogram: correlates and consequences. Am J Med 115:171-7.

[2] Haïssaguerre M., Derval N., Sacher F., Jesel L., Deisenhofer I., de Roy L., Pasquié J.-L., Nogami A., Babuty D., Yli-Mäyry S., De Chillou C., Scanu P., Mabo P., Matsuo S., Probst V., Le Scouarnec S., Defaye P., Schlaepfer J., Rostock T., Lacroix D., Lamaison D., Lavergne T., Aizawa Y., Englund A., Anselme F., O'neill M., Hocini M., Lim K.T., Knecht S., Veenhuyzen G.D., Bordachar P., Chauvin M., Jais P., Coureau G., Chene G., Klein G.J., Clémenty J. (2008) Sudden cardiac arrest associated with early repolarization. The New England Journal of Medicine 358:2016-23.

[3] Nam GB., Kim YH., Antzelevitch C. (2008) Augmentation of J Waves and Electrical Storms in Patients with Early Repolarization. The New England Journal of Medicine 358:2078–2079.

[4] Rosso R., Kogan E., Belhassen B., Rozovski U., Scheinman M.M., Zeltser D., Halkin A., Steinvil A., Heller K., Glikson M., Katz A., Viskin S. (2008) J-Point Elevation in Survivors of Primary Ventricular Fibrillation and Matched Control Subjects. Journal of the American College of Cardiology, 52(15):1231-1238.

[5] Tikkanen J.T., Anttonen O., Junttila J.M., Aro A.L., Kerola T., Rissanen H.A., Reunanen A., Huikuri H.V. (2009) Long-Term Outcome Associated with Early Repolarization on Electrocardiography. The New England Journal of Medicine 361:2529-37.

[6] Antzelevitch C, Yan GX. (2010) J wave syndromes. Heart Rhythm 7:549–558.

[7] Roten L., Derval N., Maury P., Mahida S., Pascale P., Leenhardt A., Jesel L., Deisenhofer I., Kautzner J., Probst V., Rollin A., Ruidavets JB., Ferrières J., Sacher F., Heg D., Scherr D., Komatsu Y., Daly M., Denis A., Shaha A., Hocini M., Jaïs P., Haïssaguerre M. (2016) Benign vs malignant inferolateral early repolarization: Focus on the T wave, Heart Rhythm 13(4):894–902.

[8] MacFarlane P.W., Antzelevitch C., Haissaguerre M., Huikuri, H.V., Potse M., Rosso R., Sacher F., Tikkanen J.T., Wellens H., Yan GX. (2015) The early repolarization pattern: A consensus paper (Review). Journal of the American College of Cardiology 66(4):470-477.

[9] Antzelevitch C., Yan G.-X., Ackerman M.J., Borggrefe M., Corrado D., Guo J., Gussak I., Hasdemir C., Horie M., Huikuri H., Ma C., Morita H., Nam G.-B., Sacher F., Shimizu W., Viskin S., Wilde A.A.M. (2016) J-Wave syndromes expert consensus conference report: Emerging concepts and gaps in knowledge. Journal of Arrhythmia 32:315–339.

[10] Maheshwari S., Acharyya A., Puddy P.E., Mazomenos E.B., Leekha G., Maharatna K., Schiariti M. (2013) An automated algorithm for online detection of fragmented QRS and identification of its various morphologies. Journal of the Royal Society Interface 10:20130761.

[11] Clark E.N., Katibi I., MacFarlane P.W. (2014) Automatic detection of end QRS notching or slurring. Journal of Electrocardiology 47(2):151-154.

[12] Kenttä T., Porthan K., Tikkanen J.T., Väänänen H., Oikarinen L., Viitasalo M., Karanko H., Laaksonen M., Huikuri H.V. (2015) Sensitivity and Specificity of Automated Detection of Early Repolarization in Standard 12-lead Electrocardiography. Annals of Noninvasive Electrocardiology 20(4):355-361.

[13] Airaksinen J., Aalto-Setälä K., Hartikainen J, Huikuri H., Laine M., Lommi J., Raatikainen P., Saraste A. (2016) Kardiologia. Helsinki, Kustannus Oy Duodecim, 3rd ed., 1296 pages.

[14] Katz, A.M. (2011) Physiology of the Heart. Wolters Kluwer Health/Lippincott Williams & Wilkins cop,. 5th ed, 576 pages.

[15] Malcolm S., Thaler M.D. (2003) The only EKG book you'll ever need. Wolters Kluwer Health/Lippincott Williams & Wilkins cop., 4th ed., 318 pages.

[16] Runge, M.S., Patterson, C., Stouffer, G.A., Netter, F.H. (2010) Netter's cardiology. Elsevier, 2nd ed., 650 pages.

[17] Rainer S., Charles A., Dariush H., Carla G., Alfredo P., Fiorenzo G., Christian V., Christian W., Martin B. (2008) Electromechanical coupling in patients with the short QT syndrome: Further insights into the mechanoelectrical hypothesis of the U wave. Heart Rhythm 5(2):241–245.

[18] Sinner M.F., Reinhard W., Müller M., Beckmann B.-M., Martens E., Perz S., Pfeufer A., Winogradow J., Stark K., Meisinger C., Wichmann H.-E., Peters A., Riegger G.A.J, Steinbeck G., Hengstenberg C., Kääb S. (2010) Association of early repolarization pattern on ecg with risk of cardiac and all-cause mortality: a population-based prospective cohort study (MONICA/KORA). PLoS Med. 7:e1000314.

[19] Haruta D., Matsuo K., Tsuneto A., Ichimaru S., Hida A., Sera N., Imaizumi M., Nakashima E., Maemura K., Akahoshi M. (2011) Incidence and prognostic value of early repolarization pattern in the 12-lead electrocardiogram. Circulation 123:2931–2937.

[20] Patel R.B., Ng J., Reddy V., Chokshi M., Parikh K., Subacius H., Alsheikh-Ali A.A., Nguyen T., Link M.S., Goldberger J.J., Ilkhanoff L., Kadish A.H., (2010) Early Repolarization Associated With Ventricular Arrhythmias in Patients With Chronic Coronary Artery Disease, Circulation: Arrhythmia and Electrophysiology 3:489-495.

[21] Tikkanen J.T., Wichmann V., Junttila M.J., Rainio M., Hookana E., Lappi OP., Kortelainen ML., Anttonen O., Huikuri H.V. (2012) Association of Early Repolarization and Sudden Cardiac Death During an Acute Coronary Event, Circulation: Arrhythmia and Electrophysiology 5:714-718.

[22] Kawata H., Morita H., Yamada Y., Noda T., Satomi K., Aiba T., Isobe M., Nagase S., Nakamura K., Kusano K.F., Ito H., Kamakura S., Shimizu W. (2013) Prognostic significance of early repolarization in inferolateral leads in Brugada patients with documented ventricular fibrillation: A novel risk factor for Brugada syndrome with ventricular fibrillation, Heart Rhythm, 10(8):1161–1168.

[23] Laksman Z.W.M., Gula L.J., Saklani P., Cassagneau R., Steinberg C., Conacher S., Yee R., Skanes A., Leong-Sit P., Manlucu J., Klein G.J., Krahn A.D. (2014) Early repolarization is associated with symptoms in patients with type 1 and type 2 long QT syndrome. Heart Rhythm Society 11:1632-1638.

[24] Stumpf C., Simon M., Wilhelm M., Zimmermann S., Rost C., Achenbach S., Brem M.H. (2016) Left Atrial remodeling, early repolarization pattern, and inflammatory cytokines in professional soccer players. Journal of Cardiology 68:64-70.

[25] Walsh B., Macfarlane P.W., Prutkin J.M., Smith S.W. (2019) Distinctive ECG patterns in healthy black adults, Journal of Electrocardiology 56:15–23.

[26] Haïssaguerre, M., Nademanee, K., Hocini, M., Cheniti G., Duchateau J., Frontera A., Sacher F., Derval N., Denis A., Pambrun T., Dubois R., Jaïs P., Benoist D., Walton R.D., Nogami A., Coronel R., Potse, M., Bernus, O. (2019) Depolarization versus repolarization abnormality underlying inferolateral J-wave syndromes: New concepts in sudden cardiac death with apparently normal hearts. Heart Rhythm 16(5):781-790.

[27] Holkeri A., Eranti A., Haukilahti M.A.E., Kerola T., Kentta T.V., Tikkanen J.T., Rissanen H., Helio-Vaara M., Knekt P., Junttila M.J., Aro A.L., Huikuri H.V. (2020) Impact of age and sex on the long-term prognosis associated with early repolarization in the general population. Heart Rhythm 17(4):621-628.

[28] Tobón-Cardona M., Kenttä T., Porthan K., Tikkanen J.T., Oikarinen L., Viitasalo M., Salomaa V., Huikuri H.V., Junttila J.M., Seppänen T. (2018) Waveform prototype-based feature learning for automatic detection of the early repolarization pattern in ECG signals. Physiological Measurement, 39(11):115010.

[29] Bland M. (1995) An introduction to medical statistics. New York, Oxford University Press Inc., 2nd ed, 396 pages.

[30] Rollin A., Maury P., Bongard V., Sacher F., Delay M., Duparc A., Mondoly P., Carrié D., Ferriéres J., Ruidavets J.-B. (2012) Prevalence, Prognosis, and Identification of the Malignant Form of Early Repolarization Pattern in a Population-Based Study. American Journal of Cardiology 110(9):1302-1308.

[31] Holkeri A., Eranti A., Haukilahti M.A.E., Kerola T., Kentta T.V., Tikkanen J.T., Anttonen O., Noponen K., Seppänen T., Rissanen H., HeliöVaara M., Knekt P., Junttila M.J., Huikuri H.V., Aro A.L. (2020) Predicting sudden cardiac death in a general population using an electrocardiographic risk score. Heart 106(6):427-433.