Computerized Analysis
of ST Recovery and ST Episodes

Henrik Johansson

TRITA-NA-E03090
Computerized Analysis of ST Recovery and ST Episodes

Henrik Johansson

Master’s Thesis in Biomedical Engineering (20 credits) at the School of Engineering Physics, Royal Institute of Technology year 2003
Supervisor at Nada was Anders Lansner
Examiner was Anders Lansner
Abstract

ECG is used in the diagnosis and surveillance of most heart patients and is often recorded continuously during several days. This generates an enormous amount of data for the doctor or nurse to interpret. To speed up this process computerized tools are used to compress the information. One such tool is MIDA™ that is used to monitor oxygen deficiency in the heart e.g. during a heart attack. MIDA™ extracts several parameters from the ECG signal and shows their development over time in trends. This master’s project deals with the interpretation of one of these trends, the ST-VM trend. We have developed two algorithms. The first algorithm calculates the time to 50% ST recovery and recovery at 90 minutes, two parameters that are used as indicators of whether clot dissolving medications used during a heart attack is effective. The results were tested against the Accent II plus study in which the patients were manually reviewed by two independent reviewers from Ischemia Corelab in Gothenburg, Sweden. The comparison showed that the automated analysis gave the same results in most cases and that when there was a difference it was often due to mistakes made by the reviewers in the study. The second algorithm identifies ST episodes that are used in the prognosis of patients with coronary artery disease, CAD. The algorithm uses a new method to define a baseline through the trend combining local features in the trend, such as slope and point density, with a global curve approximation method, smoothing cubic splines. The algorithm was tested against 406 patients from the Accent II study, another study from Ischemia Corelab where the patients are manually reviewed. The results show that both methods find approximately the same number of episodes, and that the performance of the algorithm is stable even when the data quality is bad.

Datroriserad analys av ST-återgång och ST-episoder

Sammanfattning

Preface

This paper is the result of my master’s project at the Department of Numerical Analysis and Computing Science (NADA) at the Royal Institute of Technology in Stockholm, Sweden. The project was commissioned by Ortivus AB and most of the work was done at their office in Täby. The project was supervised by Gunilla Lundahl at Ortivus and Professor Anders Lansner at NADA. Examiner was Professor Anders Lansner.

I would especially like to thank Gunilla Lundahl for her never-ending interest and for all the input and support she has given me during my work on the project.

Henrik Johansson
# Table of Contents

1 Introduction ..................................................................................................................................................1

2 The Heart - A Remarkable Workhorse ........................................................................................................3

## 2.1 HOW DOES IT WORK?
- The circulation systems ..........................................................................................................................3
- The myocardium ........................................................................................................................................3
- The conduction system ............................................................................................................................4
- The action potential ................................................................................................................................4

## 2.2 DISEASES OF THE HEART
- Coronary Artery Disease - CAD ..............................................................................................................5
- Myocardial ischemia ..................................................................................................................................6
- Myocardial Infarction - MI .......................................................................................................................6

3 ECG – Measuring the Heart's Electrical Activity ............................................................................................7

## 3.1 HOW DOES IT WORK?

## 3.2 INTERPRETATION OF THE ECG

## 3.3 LEADS

## 3.4 VECTORCARDIOGRAPHY - VCG

## 3.5 MIDA™

## 3.5.1 Interpretation of ST-VM trends

4 ST Recovery ...............................................................................................................................................11

## 4.1 COMPUTERIZED ANALYSIS

## 4.2 THE ALGORITHM

## 4.3 TESTING AGAINST THE ACCENT II PLUS STUDY

## 4.4 DISCUSSION

5 To Identify ST Episodes ..............................................................................................................................15

## 5.1 COMPUTERIZED DETECTION OF EPISODES

## 5.1.1 Detecting candidates

## 5.1.2 Draw a baseline

## 5.2 OUR ALGORITHM

## 5.2.1 Detecting candidates

## 5.2.2 Global approximation

## 5.2.3 Finding the episodes

## 5.3 EVALUATING THE RESULTS

## 5.3.1 Number and location of episodes

## 5.3.2 Evaluation of the baseline

## 5.3.3 Comparison between baselines

## 5.4 DISCUSSION

6 Final conclusions ........................................................................................................................................29

References .......................................................................................................................................................30

Appendix A: Deriving the Equations for Smoothing Cubic Splines ..........................................................A1

Appendix B: Finding the episodes ................................................................................................................B1
1 Introduction

The human heart pumps about five liters of blood each minute in order to supply the tissues of the body with oxygen and nutrients and to remove waste products [1]. The heart also supplies its own cells with blood through the coronary arteries. An obstruction in any of these vessels thus has serious implications. If the heart muscle doesn’t get enough oxygen it cannot work and the heart muscle tissue will soon start to die, a heart attack has begun.

Today ECG measurements are made on every patient who arrives at the hospital with chest pains. The ECG is then recorded continuously during several days, which generates an enormous amount of data for the doctor or nurse to interpret. To speed up this process computerized tools are used to compress the information and make it clearer so that a faster diagnosis can be made.

One such tool is MIDA™ (Myocardial Ischemia Dynamic Analysis) that is used to monitor myocardial ischemia, oxygen deficiency in the heart muscle. MIDA™ extracts several parameters that are used in the diagnosis and prognosis of patients suffering from myocardial ischemia. These parameters are plotted in trends that show the development of the oxygen deficiency over time. The interpretation of these trends requires some experience and is often quite time consuming. It would therefore be a good idea to automate the interpretation to make the information available faster.

One of the parameters extracted by MIDA™, ST-VM, provides valuable information about the effect of clot dissolving medications used in the treatment of myocardial infarction and about the prognosis for patients experiencing ischemic episodes, i.e. short periods of oxygen deficiency in the heart. In this master’s project I have developed two algorithms. The first extracts information about the effect of clot dissolving medications and the second finds the number and duration of ischemic episodes. The project was commissioned by Ortivus AB that has developed the MIDA™ system.

To be able to verify the correctness of the results given by the algorithm we have used the records from the Accent II and the Accent II plus studies both made by Ischemia Corelab, Gothenburg. These studies contain data on patients suffering from myocardial infarction treated at hospitals all over Sweden. The interpretations of the trends are made manually by two independent reviewers from the lab in Gothenburg.

This paper starts with a short review of the heart and how it works in chapter 2. We then introduce the ECG in chapter 3. This chapter also deals with vectorcardiography, VCG, and the MIDA™ system. Chapter 4 deals with the first algorithm, extracting information about the effect of clot dissolving medication whereas chapter 5 deals with the algorithm counting the ischemic episodes.
2 The Heart - A Remarkable Workhorse

The human heart is about the size of a closed fist and its mass averages 250 g in adult females and 300 g in adult males. Despite its moderate size the heart pumps more than 14,000 liters of blood through the body’s estimated 100,000 km of blood vessels each day, without ever taking a break [1]. This is possible only through nature’s extraordinary engineering capabilities.

2.1 HOW DOES IT WORK?

The heart rests on the diaphragm, near the midline of the thoracic cavity. About two-thirds of the mass of the heart is positioned to the left of the body’s midline. Surrounding the heart is the pericardium. The fibrous pericardium is a tough inelastic sack that confines the heart to its position in the chest and protects the heart muscle from over stretching. The serous pericardium is a thinner more delicate membrane that forms a double layer around the heart. Between these layers is a film of serous fluid that reduces friction between the membranes as the heart moves. Inside the heart are four chambers, two atria and two ventricles (see figure 1). The right atrium is connected to the right ventricle via the tricuspid valve, and the left atrium is connected to the left ventricle via the bicuspid valve. The right ventricle is connected to the pulmonary trunk via the pulmonary semilunar valve and the left ventricle is connected to aorta via the aortic semilunar valve. The valves prevent blood from flowing in the wrong direction and are thus essential for the function of the heart.

2.1.1 The circulation systems

With each beat the heart pumps blood in two closed circuits - the systemic circulation and the pulmonary circulation. The left side of the heart is the pump for the systemic circulation; it receives freshly oxygenated blood from the lungs and the left ventricle ejects the blood into the aorta (see figure 2). From the aorta, the blood divides into smaller and smaller streams, entering progressively smaller systemic arteries that carry it to all the organs of the body. As the blood divides into smaller and smaller vessels it finally reaches the capillaries. In the capillaries blood unloads O$_2$ (oxygen) and nutrients and picks up CO$_2$ (carbon dioxide). As the blood is on its way through the systemic circulation it enters progressively larger veins and ultimately flows back to the right atrium. The right side of the heart is the pump for the pulmonary circulation; it receives all the deoxygenated blood from the systemic circulation and blood ejected from the right ventricle flows into the pulmonary trunk. The pulmonary trunk branches into the right and left pulmonary arteries that carry blood into the right and left lungs. In the pulmonary capillaries the blood unloads CO$_2$ and picks up O$_2$ to and from the air sacs (alveoli) of the lungs. The freshly oxygenated blood then flows into pulmonary veins and returns to the left atrium.

2.1.2 The myocardium

The myocardium, which is cardiac muscle tissue, makes up the bulk of the heart and is responsible for its pumping action. Although it is striated like skeletal muscle, cardiac muscle is involuntary. Comp-
red to skeletal muscle fibers, cardiac muscle fibers are shorter in length, larger in diameter, and not as circular in transverse section. They also exhibit branching, which gives an individual fiber a Y-shaped appearance. Although cardiac muscle fibers branch and interconnect with each other, they form two separate functional networks. The muscular walls and partition of the atria compose one network, the muscular walls and partition of the ventricles the other. The ends of each fiber in a network connect to its neighbors via desmosomes, which hold the fibers together mechanically, and via gap junctions, which allow muscle action potentials to spread from one muscle fiber to another via electrical coupling. As a consequence, when a single fiber of either network is stimulated, all the other fibers in the network become stimulated as well. Thus, each network contracts as a functional unit.

2.1.3 The conduction system

A normal heartbeat is a very precisely synchronized muscle contraction. The key to this synchronized contraction is a network of specialized cardiac muscle fibers that are called autorythmic cells because they are self-excitable. These autorythmic cells repeatedly generate spontaneous action potentials that trigger heart contractions. Nerve impulses from the autonomic nervous system and blood-borne hormones (such as ephedrine) modify the heartbeat, but they do not establish the fundamental rhythm. Autorythmic cells have two important functions: they act as a pacemaker, setting the rhythm for the entire heart, and they form the conduction system, propagating the action potentials throughout the heart muscle. The conduction system (see figure 3) assures that cardiac chambers become stimulated in a coordinated manner, which makes the heart an efficient pump. Normally, cardiac excitation starts in the sinoatrial (SA) node propagating along the muscle fibers in the walls of the atria to the atrioventricular (AV) node. In the AV node the action potential slows considerably resulting in a 100-msec delay that give the atria time to fully contract. The action potential then enters the atrioventricular (AV) bundle (also known as the bundle of His), the only electrical connection between the two functional units of the heart. The AV bundle branches into the right and left bundle branches that course through the mid-wall of the heart down towards its apex. Finally, the large-diameter conduction myofibers (Purkinje fibers) rapidly conduct the action potential, first to the apex of the myocardium and then upward to the remainder of the ventricular myocardium. The myocardium is depolarized from the inside out, and repolarized in the opposite direction. About 200 msec after the atria contract, the ventricles contract. On their own the autorythmic fibers in the SA node initiate action potentials 90-100 times per minute. This rate is then decreased, under the influence of the parasympathetic nervous system, to a resting rate of 60-75 action potentials per minute.

2.1.4 The action potential

The action potential initiated by the SA node travels along the conduction system and spreads out to the “working” atrial and ventricular muscle fibers, which are called contractile fibers. An action potential occurs in a contractile fiber as follows (see figure 4):

1. **Depolarization.** Contractile fibers have a resting membrane potential close to -90 mV. When they are brought to threshold by excitation in neighboring fibers, voltage-gated fast Na⁺ channels open. When these channels open the permeability of the sarcolemma (plasma membrane) to sodium ions ($P_{Na^+}$) increases. The result is an inflow of Na⁺ along the electrochemical gradient that produces a rapid depolarization. Within a few milliseconds, the fast Na⁺ channels automatically inactivates and ($P_{Na^+}$) decreases.

2. **Plateau.** During the next phase voltage-gated slow Ca²⁺ channels in the sarcolemma and sarcoplasmic reticulum membrane open, increasing the permeability to calcium ions ($P_{Ca^{2+}}$). This leads
to an increased level of Ca$^{2+}$ in the cytosol. At the same time the membrane permeability to potassium ions ($P_K$) decreases due to the closing of $K^+$ channels. For about 250 ms the membrane potential stays close to 0 mV as a small outflow of $K^+$ just matches the inflow of Ca$^{2+}$.

3. **Repolarization.** After a delay voltage-gated $K^+$ channels open thereby increasing the membrane permeability to potassium ions. At the same time the calcium channels are closing. As more $K^+$ leave the fiber and fewer Ca$^{2+}$ enter, the resting potential of −90 mV is restored.

As Ca$^{2+}$ concentration rises inside a contractile fiber it binds to the regulator protein troponin, which allows the actin and myosin filaments to begin sliding past one another, and tension starts to develop.

![Figure 4](a) The membrane potential is measured over the membrane with the outside as a reference. (b) The membrane permeability for different ions vary during the action potential (values taken from [1]).

In muscle the refractory period is the time interval during which a second contraction cannot be triggered. The refractory period for a cardiac fiber lasts longer than the contraction of the fiber. As a result another contraction cannot begin until relaxation is well underway. For this reason tetanus (maintained contraction) cannot occur in cardiac muscle tissue.

### 2.2 Diseases of the Heart

During 1996 ischemic heart disease was the cause 24% of all deaths in Sweden [2]. It is the single largest cause of death in the industrial world. The coronary arteries, left and right, branch from the ascending aorta and supply oxygenated blood to the myocardium. An obstruction in one of these vessels or any of their branches could very quickly lead to a life-threatening situation. Without oxygen the contractile fibers cannot contract and thus do not contribute to the pumping motion. If the blood flow from the heart stops completely one becomes unconscious within seconds and within minutes many tissues suffer irreversible damage.

#### 2.2.1 Coronary Artery Disease - CAD

CAD is defined as the effects of the accumulation of atherosclerotic plaques in coronary arteries that lead to a reduction in blood flow to the myocardium [1]. Atherosclerotic plaques are initiated by one or more factors that cause damage to the endothelial lining of large and medium-sized arteries. Factors that contribute to this process include high levels of circulating cholesterol, cytomegalovirus (a common herpes virus), prolonged high blood pressure, smoking, and diabetes mellitus. The injury promotes the aggregation of platelets and attracts phagocytes, and cholesterol and triglycerides collect in the inner layer of the arterial wall. Contact with platelets, lipids, and other components of blood stimulates smooth muscle cells and collagen fibers in the arterial wall to proliferate abnormally causing an atherosclerotic plaque to develop. As the plaque enlarges it obstructs the blood flow.

In some atherosclerotic plaques the tissue covering the cholesterol and triglycerides is very thin and thus might burst if the strain on the artery wall is too big. When the cholesterol and triglycerides come into contact with the blood stream it initiates the formation of a blood clot. In our bodies there is a balance between the formation and dispersion of blood clots, if there were no blood clots we would
pretty soon bleed to death. If the blood clot is dispersed we get an ischemic episode. If it remains and causes a total obstruction of the blood flow it will cause a heart attack. Yet another danger is that part of the clot may dislodge to become an embolus and obstruct blood flow in other vessels.

2.2.2 Myocardial ischemia

Partial obstruction of the blood flow in the coronary arteries may cause myocardial ischemia. Ischemia often causes hypoxia (reduced oxygen supply), which may weaken cells without killing them. Myocardial ischemia may manifest itself as angina pectoris, typically described as a tightness or squeezing sensation, as though the chest were in a vise, accompanied by a severe pain referred to the neck, chin, or down the left arm to the elbow. Angina pectoris often occurs during exertion, when the heart demands more oxygen, and then disappears with rest. In some people, especially people with diabetes who suffer from neuropathy (disease of the peripheral or autonomous nervous system), ischemic episodes occur without producing pain. This is known as silent myocardial ischemia and is particularly dangerous, as the person has no forewarning of an impending heart attack.

2.2.3 Myocardial Infarction - MI

A complete obstruction of the blood flow in a vessel may cause a myocardial infarction, MI, commonly called a heart attack. Infarction means death of an area of tissue because of interrupted blood supply. The obstruction may be caused by a blood clot forming on an atherosclerotic plaque or an embolus that obstructs a smaller vessel. Because the heart tissue distal to the obstruction dies and is replaced by scar tissue the effects of a myocardial infarction are irreversible. The heart muscle loses some of its strength and furthermore, the infarction may disrupt the electrical conduction system of the heart and cause sudden death by triggering ventricular fibrillation. When the heart tissue dies the sarcolemma is disrupted and the contents of the cytosol are released into the intracellular space. Specific enzymes can later be found in the bloodstream and their presence there is used as a marker that the infarction has in fact occurred.
3 ECG – Measuring the Hearts Electrical Activity

3.1 HOW DOES IT WORK?

As the action potential is conducted through the myocardium the difference in membrane potential between the depolarized and non-depolarized fibers gives rise to an electric field. This electric field can be measured with electrodes on the body surface [3]. Let’s consider the following example: we mentioned earlier that the myocardium is depolarized from the inside out, this means that if an electrode is placed outside the myocardium it will register a rise in potential when the depolarization starts (see figure 5). The potential will reach its peak when about half of the cardiac muscle fibers are depolarized. When the whole myocardium has reached the plateau level the electrode potential will be back at zero. Since the myocardium is repolarized from the outside in the repolarization will be recorded as another peak. This second peak is smaller than the first since the repolarization is slower than the depolarization. In reality every electrode registers the electrical activities of the whole heart and the recording of a heartbeat thus looks a bit more complex.

3.2 INTERPRETATION OF THE ECG

There are normally three distinct features in the ECG signal that can be identified during every heartbeat (see figure 6). The P-wave represents the atrial depolarization that spreads from the SA node through the atrial wall. The second wave called the QRS complex begins as a downward deflection continues as a large upward triangular wave and ends as a downward wave. The QRS complex represents the depolarization of the ventricles. The third wave, an upward dome-shaped deflection called the T wave, represents the ventricular repolarization. In reading an ECG the size and timing of the waves can provide clues to abnormalities. Larger P wave, for example, indicate enlargement of an atrium and an enlarged R wave generally indicates enlarged ventricles. If the T wave is flatter than normal the heart muscle might not be receiving sufficient oxygen as, for example, in coronary artery disease.

The time-spans between the waves are called intervals or segments. For example, the P-Q interval is measured from the beginning of the P wave to the beginning of the QRS complex. It thus represents the conduction time from the beginning of atrial depolarization to the beginning of ventricular depolarization. In coronary artery disease and rheumatic fever, scar tissue forms in the heart. As the depolarization detours the scar tissue the P-Q interval lengthens. The Q-T interval begins at the onset of the QRS complex and ends at the end of the T wave. It represents the time from the beginning of the ventricular depolarization to the end of the ventricular repolarization. The Q-T interval is lengthened by myocardial damage, coronary ischemia, or conduction abnormalities. The S-T segment begins at the end of the QRS complex and ends at the beginning of the T wave and thus represents the time the ventricular heart muscle cells are at their plateau phase. The S-T segment is elevated (above the baseline) in acute myocardial infarction and depressed (below the baseline) when the heart muscle receives insufficient oxygen.
3.3 Leads

To register an ECG electrodes are placed on the body surface. These are used in different configurations called leads. The most common configuration, the 12 lead ECG, uses six unipolar chest leads (V1-V6), three unipolar extremity leads, and three bipolar extremity leads [4] (see figure 7). The three bipolar extremity leads (often referred to as standard leads) register the difference in electric potential between two electrodes. Lead I registers between the right and left arms (RA and LA in figure 7), where the left arm is the positive electrode, i.e. a depolarisation wave towards the left arm comes out as positive. Leads II and III register between the left leg (L leg) and the right and left arm respectively, the leg being the positive electrode. The three unipolar extremity leads use one of the extremity electrodes as the positive electrode and the two others combined as the negative electrode. The unipolar extremity lead using the left leg and left arm as negative electrode and the right arm as positive electrode is denoted aVR, R for right arm. The other two are denoted aVL and aVF, L for left arm and F for foot. The unipolar chest leads all use the same reference electrode. The most common is the Wilson electrode that connects the right and left arms and the left leg. Leads that use the Wilson electrode as reference are denoted with a V (V1, V2, etc). The 12 lead ECG can be interpreted as the heart seen from 12 different angles.

When a patient arrives at the hospital with heart problems a 12 lead ECG is normally connected for the examination and treatment. As the patient gets better but is under surveillance other lead system with fewer electrodes are often used.

3.4 Vectorcardiography - VCG

The conventional ECG depicts the potentials measured by the different leads in relation to time in as many curves as there are leads. In vectorcardiography the electrical activity is represented by a single dipole. The strength and spatial orientation of the dipole at each moment is depicted by a spatial vector [5]. The three characteristic features, the P wave, the QRS complex, and the T wave, mentioned above are now described as loops in a three dimensional space. The course of the vector is often shown as the projection on three orthogonal planes referred to as the sagittal, transverse, and frontal planes (see figure 8).

Figure 7 Three electrodes are placed on the left arm (LA), right arm (RA), and left leg (L leg) using one of the other or both the others combined as a reference electrode. Another six (V1-V6) are placed on the chest using the three extremity leads combined (the Wilson electrode) as a reference.

Figure 8 The projections of the vectorelectrocardiogram as seen in MIDA™. The three loops representing the P wave, the QRS complex, and the T wave are not clearly distinguishable here except for the QRS complex.
Four electrodes are the minimum number required theoretically in any system of vectorcardiography since three independent potential differences are needed to determine the heart vector [6]. The system proposed by Ernest Frank (see figure 9) uses seven electrodes that are interconnected to get six leads that are compensated for physiological differences. The six leads: right, left, front, back, head, and foot, in pairs make up the three orthogonal leads $V_x$, $V_y$, and $V_z$.

Figure 9 The Frank lead system uses seven electrodes connected in a resistance network to achieve an orthogonal system $V_x$, $V_y$, and $V_z$. Today no resistances are used, instead the compensation is done digitally.

3.5 MIDA™

The MIDA™ system was developed by Ortivus AB in Täby, Sweden. Today MIDA™ is a clinically approved monitoring system for patients with myocardial infarction and is used in hospitals all over the world. MIDA™ uses the Frank lead system to retrieve a vectorcardiogram describing the electrical activity of the heart. Each beat is examined with regard to morphology, artifacts, and noise. Accepted beats are then used to create an average beat in order to improve the signal-to-noise ratio [7-8]. Normally an averaging period of one minute is used in monitoring of patients with myocardial infarction. The average beats are used to retrieve a number of parameters, ST-vectormagnitude, ST-VM, being the one of interest in this paper. ST-VM is the geometric displacement of the ST segment,

$$ST-VM = \sqrt{ST_x^2 + ST_y^2 + ST_z^2},$$

and mirrors the oxygen levels in the myocardium. The notation ST-VM 60 indicates that ST-VM is measured 60 msec after the end of the QRS complex.

3.5.1 Interpretation of ST-VM trends

During the more than ten years that MIDA™ has been on the market several clinical studies have shown that ST-VM trends contain important information about the effects of trombolytic treatment and the prognosis for future complications. It has been shown that the time until ST-VM reaches 50% of the maximum after a myocardial infarction, time to 50% ST recovery, is an important parameter. The results from the ASSENT 2 trial show that patients alive at 30 days had a median time to 50% ST recovery of 65 minutes whereas patients dead at 30 days had a median time of 106 minutes [9].

ST episodes are another feature in the trends that is used to give a prognosis for the patient. In the TRIM vectorcardiographic sub-study it was shown that the presence of one or more ST episodes was associated with a 2.6 times higher risk for death or myocardial infarction within one year [10].
number and size of ST-episodes are thus used in the decision whether or not a patient should receive PTCA treatment or undergo bypass surgery. These parameters and how they are interpreted will be more extensively treated in the next two chapters.
4 ST Recovery

Treatment of a myocardial infarction often starts with injection of thrombolytic (clot-dissolving) agents such as streptokinase or t-PA and anticoagulants (normally heparin). The thrombolytic agents take some time to work and the result is thus hard to anticipate. The ST-VM level reflects the size of the ischemic area and is often used to follow the course of the myocardial infarction [4]. During an ongoing myocardial infarction the ST-VM level often reaches 200-600 µV. After some time the ST-VM level starts to diminish, either because the thrombolytic treatment is working or because the ischemic area has become necrotic (cells are dying). A commonly used criterion is that if the ST-VM level reaches under 50% of the maximum ST-VM level (50% ST recovery) within 90 minutes after treatment onset, the treatment is working [7]. If it doesn’t the patient should receive a second dose of clot dissolving medication or other options, such as PTCA or coronary artery bypass grafting should be considered.

4.1 COMPUTERIZED ANALYSIS

The task of finding 50% ST recovery is well suited for computerized analysis and is mainly a matter of defining strict criteria for the analysis. The criteria used in this paper were those of the Accent II study with some modifications. The reason we chose these conditions is that these criteria have been validated in a clinical study and we wanted to use the results from the study to confirm our results. The algorithm we have developed is meant to work in an online environment and thus reads the data vector value by value. The parameters analyzed are time to 50% recovery and percent recovery after 90 minutes.

ST-VM trends often contain gaps where the signal quality is too low. The gaps often occur when there is a fast change in the ST-VM level such as when ST-VM returns to its ischemia free level after an infarction. In the aftermath of an infarction the patient often experiences arrhythmia that also results in gaps. Gaps should be treated with caution. It is sometimes tempting to interpolate over a gap, but ST changes can sometimes be very rapid and there is no telling what happened during the gap. In the analysis of time to 50% recovery we have chosen to consequently use the value directly after the gap whereas in the analysis of recovery at 90 minutes we use the value in the interval 90 +/- 10 minutes closest to 90 minutes. If there is no such value, recovery after 90 minutes will not be analyzed.

The ST-VM trends also contain disturbances that sometimes deviate strongly from the surrounding values. To avoid the risk of setting time to 50% recovery at a temporary depression caused by a disturbance we have introduced two conditions, minimum number of values under the threshold and minimum time under the threshold. When both conditions are fulfilled the time to 50% recovery is set to the time where ST-VM first crossed the threshold. In the Accent II plus study no such conditions are used, the time to 50% recovery is set to the first time the ST-VM trend is equal to or under the threshold.

Patients with myocardial infarction often experience several periods of enhanced ST-VM level. To be able to handle this online we have had to introduce a condition for when to restart the analysis after 50% recovery has been reached. If the ST-VM level reaches the previous maximum value or 50 µV over the previous threshold for 50% recovery, a new ST-VM max is set and we start looking for a new 50% recovery. In the final analysis the recovery from the highest peak value is used.

4.2 THE ALGORITHM

Each new value, STVM.latest, is tested against the maximum value so far, STVM.max, and a level, the NewRecLevel that is 50 µV over the level where the latest 50% recovery occurred. If STVM.latest is higher than STVM.max or NewRecLevel, STVM.max is updated. Both these levels are initially set to zero. Each new value is also tested against 50% of STVM.max. If it is equal to or below this level a counter, Rcount is incremented, if it is over this level Rcount is set to zero. When Rcount reaches the minimum number of values below the limit, MINVALUESUNDER, we check to see for how long STVM has been under the level for 50% recovery. If this time is longer than MINTIMEUNDER 50% recovery is reached.
To find recovery at 90 minutes we check to see if time.latest is over 90 minutes and Rflag is zero. The flag is used to stop the analysis once recovery at 90 minutes is identified. If both conditions are true, we check the interval 80 to 100 minutes for the time closest to 90 minutes. If there are no values in this interval there will be no calculations of recovery at 90 minutes. If such a value exists percent recovery is calculated using STVM max in the interval 0 to 90 minutes. After calculating recovery at 90 minutes Rflag is set to 1. In pseudo code it looks like this,

```pseudocode
function findrecovery { STVM , time }
if STVM.latest > STVM.max or STVM.latest > NewRecLevel
    STVM.max ← STVM.latest
end
if STVM.latest <= STVM.max/2
    ++Rcount
else
    Rcount ← 0
end
if Rcount >= MINVALUNDER & time.latest - time(index.latest-flag) >= MINTIMEUNDER
    Recovery50percent( STVM( index.latest-Rcount ) , time( index.latest-Rcount ) )
    Rcount ← 0
    NewRecLevel ← STVM.max/2 + 50
end
if time.latest > 90 & Rflag == 0
    IndexScope ← find( time >= 80 & time <= 100 )
    Index90 ← min( IndexScope – 90 )
    Rec90min ← 100 * ( 1 - STVM( Index90 ) / max( STVM( 0:Index90 ) ) )
end
end
```

4.3 Testing against the Accent II Plus study

To make sure that the algorithm really is able to analyze real patient files we have compared the results with those in the Accent II plus study. The algorithm was implemented in MATLAB code and the online environment was simulated within the program. The program returned ST-VM max, time to 50% recovery, and recovery at 90 minutes for both ST-VM60 and ST-VM20. ST-VM20 max is not included in the study records but can be calculated.

![Graph showing ST-VM max, time to 50% recovery, and recovery at 90 minutes for both ST-VM60 and ST-VM20.].

Figure 10 The peak at about 1:15 hours is called the reperfusion peak and is quite common. For some reason the ischemia gets worse before it gets better. The second peak around 3:00 hours is a reocclusion. The vessel is cleared as the ST-VM level starts to fall after the first peak but is then obstructed again causing a second peak.
The first four hours of the files were analyzed. If more than one 50% recovery was found, the one with the highest maximum was used. The output from the program (see figure 11) contains both the trend itself and the data listed above. In the trend all maxima and recoveries are marked but only the data from the recovery from the highest peak is listed.

A total of 155 patients were compared. Of those 111 (72%) matched exactly the manual interpretation of the trend. In the remaining 44 (28%) there were deviations in one or more of the analyzed parameters. A more thorough analysis of these gave the following explanations for the deviations:

- In 9 patients (6%) the reviewer had concluded that the data were in such a condition that the analysis couldn’t be performed. The reason for this is either that there are too many disturbances or that there are too many gaps. This is something that our algorithm doesn’t take into account.

- In 22 patients (14%) we concluded that the reviewers had made mistakes in their analysis (not following their own criteria for the manual analysis). These files have been sent to Corelab in Gothenburg for a second review.

- In 12 patients (8%) the reviewer identified the maximum found by the algorithm as a disturbance and thus chose another maximum. The algorithm does not look for disturbances.

- One patient (0.6%) had three maxima with exactly the same amplitude. The reviewers in the Accent II plus study concluded that the first maximum was the only correct one. In our algorithm we consistently choose the last one to minimize the risk of giving false positive results.

The reason only part of the patients from the Accent II study (155 out of 200) were used is that we have had some problems with the transfer of files and records and thus only 155 patients were complete. That is, however, enough to receive reliable results. In order to be able to run the patient files in MATLAB they had to be converted to ASCII format. This was done with a program previously developed at Ortivus.

4.4 DISCUSSION

When a reviewer in the manual analysis of the files suspects that a value in the ST-VM trend might be a disturbance he/she looks at the mean ECG complex used to derive that ST-VM value. With a little practice most disturbances are quite easy to identify. In our automated analysis this is a lot harder to accomplish. Since there is no easy way to describe the disturbances in only a few parameters, trying to identify them could be a very tedious project. Another reason to leave this analysis out of our algorithm is that these disturbances have already slipped through the analysis in the MIDA™ system. It would thus be better to improve the algorithms in MIDA™ so that the disturbances were excluded from our input signal already from the beginning.

This kind of automated analysis must be used with caution. Our intention is not to create a program that tells the doctor what to do but to provide him/her with additional information, information that he/she otherwise would have had to find for himself. In this case, as our comparison to the Accent II plus study showed, computers are also less prone to making errors than, although skilled in this subject, are humans. As we mentioned earlier this algorithm is developed to run online and can thus be used to remind doctors when it is time to make a decision or warn that the patients ST-VM level isn’t going down fast enough.

In our algorithm we have introduced requirements on how long the ST-VM trend must be under the threshold and how many values must be under the threshold before we say that 50% recovery is accomplished. These requirements are meant to improve the certainty of the statement. The values for these two parameters are bound by two things, on one hand one wants to know that 50% recovery is accomplished as fast as possible. On the other hand one wants to be sure that it really is accomplished.
5 To Identify ST Episodes

Ischemic episodes are, as mentioned earlier, short periods of myocardial ischemia that end spontaneously (i.e. without treatment) before they lead to an infarction. They are characterized by a rapid increase in ST-VM of more than 50 µV [7-8]. Today most of the analysis of ST episodes is done manually. A baseline that is supposed to describe the patients normal ST level is drawn and all peaks that are more than 50 µV over this baseline for more than one minute are defined as ST episodes. In the study we have used to evaluate the results of our analysis, the Accent II study, the baseline is a horizontal line that is manually defined. It is, however, possible to define more than one such line for each patient.

5.1 COMPUTERIZED DETECTION OF EPISODES

There are mainly two ways to identify the kind of changes in ST-VM that are described above. Either you look for all changes in the ST-VM level that might be candidates for ST episodes and then try to figure out which ones actually are. Or you define a baseline that describes all changes in the signal level that are not episodes and subtract this baseline from the signal so that all that remains are the episodes.

5.1.1 Detecting candidates

There are several methods to detect episode candidates [11]. If the signal level is relatively stable it might be sufficient to look for changes in the signal level; if the signal level is higher than a certain threshold level you have a candidate. If the signal contains a lot of noise statistical methods can be of use to discover when a significant change has occurred. Another pretty straightforward method is to differentiate the signal. It is then easy to identify an episode as a zero crossing. This method tends to amplify high frequency noise so it might be a good idea to use a prefilter. The next step is to identify the episodes among the candidates. This can be done by looking at the height and duration of the candidate but this is not always enough. If you manage to describe the episodes you are looking for in just a few parameters you can train a neural network to recognize them [12]. This method is potentially very efficient and the built in ability of the neural networks to generalize makes it pretty stable with noise. The drawback is that the construction and training of such a network requires a lot of work.

5.1.2 Draw a baseline

To draw a baseline that describes everything you are not looking for is often more suitable if the features you are looking for are not so well defined, as with ST episodes. ST episodes come in all kinds of sizes and shapes and it might thus be easier to describe what is not an episode than what is. There are several ways to define the baseline in signals. Among the fastest and easiest is smoothing which can be done in at least three different ways [13]. In sliding regression, you fit a series of linear or polynomial functions to overlapping subsets of the data. In kernel methods the estimator \( \hat{f}(x) \) has the form of a weighted average of the data,

\[
\hat{f}(x) = \sum \frac{y_i K((x_i - x) / w)}{\sum K((x_i - x) / w)},
\]

where \( K \) is the smoothing kernel. This method will be used quite extensively later in this chapter to attenuate certain changes in the signal. A third method is the use of smoothing splines where a series of piecewise polynomials are fitted together so that the first and second derivatives are continuous. This method is part of our final solution to the problem of finding a baseline. In [14] self-organizing maps (a form of neural network) are used to find the baseline. Even if this method is quite efficient it is iterative and thus more time consuming than e.g. smoothing splines.

5.2 OUR ALGORITHM

The major problem with the methods mentioned above is that the episodes also contribute to the baseline. This makes them less distinguished and thus harder to identify. To solve this problem we would like to remove the episodes from the signal before we calculate our baseline. Herrmberger and
Zimmer [14] have proposed a method that combines the information you get from looking at the signal from a local prospective, that is the neighborhood around each point, with the global prospective. We have modified this approach so that it is better suited for our purpose of finding ST episodes. Instead of using a neural network for the global analysis we have used smoothing cubic splines since this method is faster and easier to work with. We start by describing three local features of the baseline:

1. The variations in the patient’s normal ST level are slow changes as opposed to the rapid changes that distinguish the episodes. Thus the slope of the curve gives us a clue to whether a data point belongs to the baseline or not.

2. Most patients have relatively few episodes implying that the ST level lies around the baseline most of the time. This means that the point density might be a good measure of whether a point belongs to the baseline or not. Since we also want the baseline to be reasonably stable we use an anisotropic Gaussian filter with the width larger than the height to measure the point density.

3. Since we want to minimize the impact of those peaks that are easily identified we perform what in image processing is referred to as a closing on our scatterplot. For a 2D curve that means we first use a running minimum to take away the peaks and then use a running maximum on the result to restore the displacement in time introduced by the running minimum. Using the distance to this very coarse baseline as a parameter gives us an opportunity to exclude all the obvious peaks from our baseline.

We now have to find a way to use this information in an optimal way in order to find the baseline.

### 5.2.1 Fuzzy logic and membership functions

Our first step is to introduce the concept of fuzzy set theory to describe the features of the baseline stated above. In conventional dual logic a statement can be true or false, and nothing in between. In terms of set theory this means that each single element can either belong to or not belong to a set, \( A \subseteq X \) or \( A \not\subset X \). Such a set can be described in different ways: one can either list the elements that belong to the set; describe the set analytically by stating the conditions for membership; or define the member elements by using the characteristic function, in which one indicates membership and zero nonmembership. For a fuzzy set, the characteristic function allows various degrees of membership for the elements of a given set. We define a fuzzy set [15]:

If \( X \) is a collection of objects denoted generically by \( x \), then a fuzzy set \( \tilde{A} \) in \( X \) is a set of ordered pairs:

\[
\tilde{A} = \{(x, \mu_x(x)) | x \in X\}
\]

\( \mu_x(x) \) is called the membership function or grade of membership of \( x \) in \( \tilde{A} \) that maps \( X \) to the membership space \( M \). (When \( M \) contains only the two points 0 and 1, \( \tilde{A} \) is nonfuzzy).

Let \( \tilde{B} \) equal “points that belong to the baseline” and we can define a membership function \( \mu_{f_x}(f_x(t)) \) for each of the features \( F \), mentioned above, that map some parameter \( f_x(t) \) describing this feature to the membership space \( M_F \).

#### 5.2.1.1 Fuzzification of the characteristic features of the baseline

Let us begin with the slope of the curve. We mentioned earlier that an episode is defined as a fast increase in the signal level followed by a decrease in the signal level. This means that to avoid including the episodes in our baseline we want the slope of the baseline to be low. The slope of the curve can be described by the parameter \( f_s(t_i) = (x_{i+1} - x_i)/(t_{i+1} - t_i) \) where \( x_i \) is the signal level at the time \( t_i \). Considering that this tends to amplify fast changes we use a smoothing filter to reduce its influence on the slope. Next we have to come up with a suitable membership function to map \( f_s \) to the membership space \( M_F \). We have decided to use exponential function,
Computerized analysis of ST recovery and ST episodes

\[ \mu_s^b(f_x) = k_s \cdot e^{-\frac{x^2}{2\sigma^2}}, \]

for two reasons: exponential functions have an attractive shape and are easy to calculate. The constant \( k_s \) is the maximum level of membership that any \( x \) can have. In this case it would make no sense to use \( k_s = 1 \) since the slope is always zero at the top of an episode. The second constant \( b_s \) decides how steep the membership function will be, i.e. how fast it will approach zero membership. We will show how these constants can be set later in this paper.

The point density of the curve is obtained by using an anisotropic Gaussian filter,

\[ f_s(t) = \sum_i e^{-\frac{(\mu_s - t)^2}{p}} \cdot e^{-\frac{(\mu_s - x)^2}{q}} \]

where \( p = 2 \cdot \text{std}_s^2 \) and \( q = 2 \cdot \text{std}_s^2 \). The filter is made anisotropic because we want to discriminate against changes in the signal level \( x \). The membership function for the point density looks distinctly different from the one for the slope but we keep the exponential function. The points on the baseline should have a high point density and we thus get,

\[ \mu_s^p(f_x) = k_p \cdot (1 - e^{-\frac{|\mu_s - x|^2}{2\sigma^2}}). \]

The constants \( k_p \) and \( b_p \) have the same interpretations as in \( \mu_s^b \); the new constant, \( a_p \), is the point density up to which we want the baseline membership to be zero.

Last we perform a closing on the curve and then use the distance to the resulting curve as a feature. The closing of a curve in two dimensions is simply a running minimum followed by a running maximum, that is \( g_t = \min\{|x_1, x_2, \ldots| \}, \) and \( h = \max\{|g_1, g_2, \ldots| \}. \) The distance thus becomes \( f_c(t) = \text{abs}\{x_i - h\}. \) Since the closing curve is supposed to resemble the baseline the points on the baseline should have a small distance \( f_c \) to this curve. We stick to the exponential functions and get,

\[ \mu_s^c(f_c) = k_c \cdot e^{-\frac{x^2}{2\sigma^2}}, \]

where the constants \( k_c \) and \( a_c \) have the same interpretation as earlier.

### 5.2.1.2 Optimization of the parameters \( k, b \) and \( a \)

Next we will try to find values for the constants \( k, b \) and \( a \) such that we gain maximum information about the baseline from the local analysis. We have decided to use the simpler and faster method of visually comparing the feature curves with the trend curves to see at what level the cut-off should be. The steepness of the membership function is then adjusted so that the function approximates this level. As for the maximum membership, \( k \), we have tried to make a guess as to how sure we can be that a point with the maximum value of the feature, e.g. no slope, is really part of the baseline. When we have decided on values for both constants we insert them in the membership function and plot the membership function and the trend together. The constants are then adjusted so that the membership function better distinguishes between those points that belong to the baseline and those that don’t. This procedure is then repeated for a number of trends to improve generality.

We start with the membership function \( \mu_s^b(f_x) \) that describes the baseline membership based on the slope of the curve. First we plot the feature function \( f_s(t) \) and the trend in the same picture trying to find values for the two constants, \( k_s \) and \( b_s \), such that the cut-off level take away all of the peaks and a minimum of the rest of the curve. These values are then used to calculate \( \mu_s^p(f_x) \) so that we can compare it to the trend. The constants can now be further adjusted for an even better resolution. As can
bee seen in figure 12 the baseline based only on the slope is not very good. The values on the constants used here and in the tests are $k_S=0.5$ and $b_S=2$.

**Figure 11** To avoid amplification of high frequency variations the trend is smoothed (a) before the slope (b) is calculated. The membership function (c) is applied to the slope resulting in the baseline membership (d) of each date point. The baseline, here showed together with the trend (e), is then calculated using the modified smoothing cubic splines algorithm.
Let us continue with the baseline membership based on the point density, $\mu_{D}^{b}(f_{x})$. Here we have three constants to determine: $k_{D}$, $a_{D}$, and $b_{D}$. The procedure is similar to that for the slope except we now have one more constant to determine. In this case we want the baseline membership to be high for points with high point density. Apart from deciding the maximum value, $k_{D}$, and how steep, $b_{D}$, we want the function to be, we also have to define a cut-off level, $a_{D}$, from which we want the baseline membership to be zero. In figure 13, graph (e) we can see that the baseline based only on the point density isn’t very good. It includes some peaks that might very well be episodes.

![Graph](image_url)

**Figure 12** The point density (b) for the trend (a) is calculated using the anisotropic Gaussian filter. The membership function (c) is then used to calculate the baseline membership (d) based on the point density. The baseline, here showed together with the trend (e), is then calculated using the modified smoothing cubic splines algorithm.

The membership function based on the distance to the curve resulting from the closing operation, $\mu_{C}^{b}(f_{x})$, can be seen in figure 14. We want the distance to the closing curve to be as short as possible and thus we only have to use two constants, $k_{C}$, and $b_{C}$ that determine the maximum and slope of the membership function. The procedure of finding suitable values for the constants is again similar to that used for the slope. As you can see in figure 14 this baseline looks a lot better than the two previous baselines. The reason for this is the choice of trend; the closing is very efficient on stable trends with few disturbances. If we had chosen a trend with more disturbances and rapid changes, the baselines based on the slope and the point density would have looked better than the baseline based on the closing.
Figure 13 In (a) we have plotted the median filtered trend together with the result of the closing procedure. The difference between the two curves seen in (a) is plotted in (b). The membership function (c) is applied to the difference and results in the baseline membership (d). The baseline, here showed together with the trend (e), is then calculated using the modified smoothing cubic splines algorithm.

5.2.1.3 Fusion of the membership functions

Now we have three membership functions that describe the baseline membership based on three different features. Our next step will be to fuse these into one membership function that describes the baseline membership based on all three of these features, that is, we want to find the best function \( \mu_\text{f} = \mu_\text{f}(\mu_\text{c}, \mu_\text{d}, \mu_\text{s}) \). We want this function to have the following properties: if one of the feature-bound membership functions is zero the resulting baseline membership should be zero; if all of the feature-bound membership functions are at their maximum the resulting baseline membership shouldn’t be higher than the largest feature-bound baseline membership. One candidate that has both these properties is \( \mu_\text{f} = r \cdot \mu_\text{c} \cdot \max(\mu_\text{d}, \mu_\text{s}) \)where

\[
\begin{align*}
r &= \frac{1}{k_c} \cdot \frac{1}{k_d} \cdot \frac{1}{k_s} \cdot \max(k_c, k_d, k_s),
\end{align*}
\]

so that the second property is valid.

If we fuse all our membership functions using the function proposed above and use it to calculate the baseline it looks like the one in figure 15.
Figure 14 The fusion of the baseline membership functions (a) result in a baseline membership function that well describe the baseline membership one would have guessed just looking at the trend. The result using our algorithm can be seen in (b).

The membership function is sparser since we have been able to extract more points that don’t belong to the baseline. Since we used $k=0.5$ in all three membership functions $r=4$ is used so that the total membership cannot be higher than the maximum individual membership.

### 5.2.2 Global approximation

Let us start with a brief review of some of the thoughts so far. Our goal is to define a baseline that describes everything in the ST-VM trend that is not an episode. To do this we want to use some smoothing method to filter away all fast changes. These methods involve one catch though, the episodes also contribute to the baseline making them less distinguishable and thus harder to identify.

We introduced the theory of fuzzy sets and used it to create a baseline membership function based on local features in the ST-VM trend. We will now use the baseline membership function to improve the performance of the smoothing cubic splines algorithm [16].

#### 5.2.2.1 Smoothing cubic splines

We will only give a very short description of smoothing cubic splines here (a more extensive description is provided in appendix A). A spline function is a curve constructed from polynomial segments that are subject to conditions of continuity at their joints. Imagine that we have a set of coordinates $(x_0, y_0), ..., (x_n, y_n)$ and we want to bridge the gap between adjacent points $(x_i, y_i), (x_{i+1}, y_{i+1})$ using the cubic functions $S_i$, $i=0,...,n-1$ to piece together a curve. Cubic splines use cubic polynomials and we require the first and second derivative to be continuous at their joints. The function $S_i$ can thus be written

$$S_i(x) = a_i(x-x_i)^3 + b_i(x-x_i)^2 + c_i(x-x_i) + d_i,$$

where $x$ ranges from $x_i$ to $x_{i+1}$. The first and second derivatives of this function are

$$S_i'(x) = 3a_i(x-x_i)^2 + 2b_i(x-x_i) + c_i,$$

$$S_i''(x) = 6a_i(x-x_i) + 2b_i,$$

If we wanted to interpolate between these data points it would be possible to find all the constants $a_i$, $b_i$, $c_i$, and $d_i$ with no additional conditions. In this case we want the spline function to be able to depart from the data points and we thus have to introduce some additional constraints in order to have a unique solution.

Smoothing can be seen as a compromise between keeping the error small and having a smooth function. The second derivative describes the smoothness of a function, the lower the second derivative the smoother the function. By adding up the errors and integrating the second derivative of the spline function we arrive at the following equation:
Computerized analysis of ST recovery and ST episodes

\( L = \lambda \sum_{i=1}^{n} (y_i - S_i)^2 + (1 - \lambda) \int_{t_0}^{t_f} (S'(x))^2 \, dx, \)

where \( S_i = S_i(x_i) \) and \( \lambda \in [0,1] \) is a parameter that controls the relation between the importance of being close to the data points and the smoothness of the curve. If \( \lambda = 0 \), smoothness is all that matters and the spline function \( S \) become a straight line; if \( \lambda = 1 \), being close to the coordinates is the only priority and we get an interpolating spline. Minimizing this function by setting its derivative to zero gives us another set of constraints needed to find a unique spline function. By introducing constraints on the first derivative at the beginning and end of the spline function we are again able to obtain the constants \( a, b, c, \) and \( d \).

5.2.2.2 Incorporating the baseline membership function

The next step is to use the baseline membership function to control the stiffness of the spline. If a data point doesn’t belong to the baseline (the baseline membership function equals zero) we don’t really care if the spline function departs from it introducing a large error. On the other hand, if a data point belong to the baseline for sure (baseline membership function equals one) we want the spline to interpolate that data point even if it has to make a sharp turn in order to reach it. In reality however, the baseline membership function isn’t accurate enough to be used this way. We thus let \( \lambda \) be a linear function of the baseline membership \( \mu \) so that \( \lambda(\mu_a) = a \cdot \mu + b \). The constant \( a \) is used to control how strong the dependence will be whereas \( b \) is used for computational reasons (if \( \lambda = 0 \) we will get a scaling problem).

5.2.2.3 Removal of square peaks

One apparent weakness in the algorithm described above is that none of the three features used in the local analysis is capable of removing square peaks, i.e. a very rapid rise followed by a period when the level is relatively stable and then a rapid decay. If they are short enough the closing will remove them but otherwise they will receive a quite high baseline membership at the top causing large errors over long sections of the trend due to the stiffness of the spline (see figure 16 plot c). These peaks are often associated with movement artifacts that arise as the patient shifts in bed or even sits up but can also be pathological changes. Whether this kind of peak should be included in the baseline could be discussed but in any case our algorithm doesn’t handle them very well. To solve this problem we have developed a supplementary filter that removes all peaks that fit the description above by setting the baseline membership function to zero under these peaks making the baseline ignore them completely. The other possible solution would be to define a new baseline at the top of the peak treating them as artifacts caused by movement. The problem with this approach is that the criterion for episodes is only defined for the patient in a supine position. Little information can thus be gained from this part of the trend anyway, so we might as well let the baseline ignore it.
Figure 15 This patient has a square peak at around 8:00 hours. If we disable the supplementary filter (c) that removes square peaks this peak will introduce errors before and after the peak due to the stiffness of the spline. In the baseline membership functions (b) and (d) you can see that the baseline membership is set to zero under the peak when the supplementary filter is active (b).

The filter uses the derivative to find areas that are relatively flat, are preceded by a steep increase, and terminated by a steep decline. Before the analysis a smoothing kernel is used to remove small but fast changes that tend to be amplified by the differentiation. When the trend is differentiated we use a counter to find all sequences where the absolute value of the derivative is lower than $1 \mu V/\text{min}$ for more than 20 minutes. In the next step we check if the derivative is higher than $2 \mu V/\text{min}$ during the 15 minutes preceding the beginning of the sequence and lower than $-2 \mu V/\text{min}$ during the 15 minutes after the sequence. If these conditions are fulfilled the baseline membership function is set to zero during the sequence, 15 minutes before and 15 minutes after.

5.2.3 Finding the episodes

Now that we have defined a baseline it is time to try and find the episodes. We thus define a threshold level (normally 50 $\mu V$ but it is a variable) above the baseline and every time the trend is above the threshold we have a candidate for an episode. The only further restrictions we apply on the episodes is that they should last more than one averaging period in order to sort out some disturbances, and that they shouldn’t last longer than 30 minutes. The reason for the upper limit is that we don’t want to put off a reocclusion as an episode. Besides if an ischemic episode lasts longer than 30 minutes something should probably be done about it. We also have to handle gaps in the trend. If an episode contains a gap of more than one averaging period the episode is divided in two. If there is less than one averaging period between two episodes, they are merged into one.

The algorithm uses three loops: first we identify every section where the trend is above the threshold taking into account the gaps. Then we check to see if the time between the sections is long enough, if not, the sections are merged into one. Next we make sure that the episodes are long enough but not too long. The sections that remain at the end of this procedure are returned as episodes. The first loop goes through the whole data vector looking for three specific situations. If the previous value is under the threshold and the current value is above we mark the beginning of an episode. Besides if an ischemic episode lasts longer than 30 minutes something should probably be done about it. We also have to handle gaps in the trend. If an episode contains a gap of more than one averaging period the episode is divided in two. If there is less than one averaging period between two episodes, they are merged into one.
new episode. Last we check for data points where the previous value is above the threshold but the current value is below. If this is the case we mark the end of the current episode. We have now gathered an array containing episodes but not all of these fulfill all the conditions for an episode. Next we go through all the episodes to see if the time distance between each of the episodes is long enough; if it isn’t the episodes are merged. The last loop makes sure that the episodes aren’t too long. If they are longer than 30 minutes they are excluded. The algorithm is described in pseudo code in appendix B.

5.3 Evaluating the results

The algorithm was developed in MATLAB and the MATLAB version was used during the testing. The total number of episodes was examined for 406 patients from the Accent II study. The records from the Accent II study list time at onset of the episode, duration of the episode, maximum ST-VM level during the episode, and area for the episodes in two time spans, 4-14 hours, and 14-24 hours. For those patients (114) where the study records contained episodes the number and localization of the episodes were examined. Time at onset, duration, and maximum ST-VM level were compared for each episode, both the ones in the Accent II study and those found by our algorithm. We also examined the baseline drawn by our algorithm for these 114 patients, trying to spot obvious errors.

5.3.1 Number and location of episodes

The comparison between the number of episodes found in the study and the number of episodes found by our algorithm shows a difference in most patients that have episodes. Because of this a direct comparison is hard to interpret. Instead we decided to divide the patients into groups with 0, 1-3, 4-6, and >7 episodes, which is a more clinically relevant way to view the results (the result from this comparison is presented in figure 17). As you can see most patients don’t exhibit episodes using either method. The rest of the patients show similar results using both methods which is good. There are also a large number (51) of patients that were not analyzed with regard to episodes in the Accent II study due to bad data quality. These patients are excluded in figure 17.

<table>
<thead>
<tr>
<th>C\P</th>
<th>0</th>
<th>1-3</th>
<th>4-6</th>
<th>&gt;7</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>210</td>
<td>18</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1-3</td>
<td>35</td>
<td>34</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>4-6</td>
<td>1</td>
<td>6</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>&gt;7</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>17</td>
</tr>
</tbody>
</table>

Figure 16 The diagram on the right shows how the number of episodes found by our algorithm (C) correlates with the number of episodes found in the Accent II study (P). The same results are summarized in the table on the right. The table only shows 355 patients. The reason is that we have excluded the 51 patients that were not analyzed with regard to episodes due to bad data quality.

For the 114 patients where the study records show episodes we compared time at onset, duration, and maximum of each episode. If the episodes overlapped the reason for the difference in time of onset and/or duration was noted. If they did not overlap the reason why the episode was excluded by the other method was noted. If an episode was split into several or several episodes merged into one the reason for that was noted. We also compared the maximum value of the episodes.
The result of this comparison was that there are differences in the number and timing of episodes in most patients. The maximum value of the episodes agrees in most cases but not all. There are several reasons for the differences in number and localization of the episodes. Even small differences in the baseline level often cause differences in the number of episodes found. Either the peak is over the threshold using one baseline and under using the other or one episode is split into several or several episodes merged into one depending on the level of the threshold. In our algorithm we introduced an upper limit for the duration of an episode at 30 minutes. In the Accent II study there is no such limit. In the study some peaks are disqualified as they are identified as disturbances. This is also the reason for the differences in the maximum value of the episode. The differences in the location of the episodes are either due to the two methods having in fact found different episodes or that the differences in the baseline level cause the time at onset and duration of the episode to vary.

The total number of episodes found by our algorithm in all 406 patients was calculated for different levels of the threshold in order to see how strong the dependence was. The interval 40 to 60 microvolts was used and the conclusion was that the dependence is quite strong (see figure 18). Thus, a closer examination of the threshold level against clinical end points in order to find the best possible threshold level might be worthwhile.

5.3.2 Evaluation of the baseline

In order to see how the algorithm performed in defining the baseline the baselines defined for the 114 patients with episodes were visually examined. The examination was done by the author, who is not an expert in the area, and thus only covers flaws that are clearly visible. In 98 patients (86%) the baseline showed no unwanted behavior. Of the remaining 16 (14%) patients 12 (11% of the total) patients had very fast changes in the baseline, or steps if you will (see figure 19). The algorithm is unable to handle these steps due to the fact that during fast changes the membership function is very low and the spline thus quite stiff. In three of the patients (3%) the reason was that the baseline membership function was equal to zero during the first hours of the trend. This leaves the smoothing splines algorithm with no guiding points. In the last patient the quality of the data was very low. The trend consisted of occasional data points, and the baseline membership is thus very low, and as a result the baseline was left drifting.

![Figure 18](image18.png)

Figure 18 The step in the patients ST-VM level at about 6:30 hours causes the baseline to deviate from the expected path before and after the step. The baseline membership (seen in the lower graph) is zero around the step making the spline stiff.
5.3.3 Comparison between baselines

In the records from the Accent II study the baseline levels used in the analysis are noted. A comparison between these and the continuous baseline defined by our algorithm are quite illustrative of the advantages of a continuous baseline for automated analysis. The number of faults introduced because of differences in the ST-VM level that are not likely to be episodes are clearly reduced. Some examples follow in figure 20.

![Figure 19](image)

**Figure 19** The first patient (top) has a ST-VM level that falls about 150 µV during twelve hours. Apart from this longtime trend the ST-VM level is quite stable and the baseline is easy to outline if you use a continuous baseline. If you try to fit a series of horizontal lines the task becomes a lot harder since there are no natural places to change the baseline level.

The second patient (middle) has a rising ST-VM level and several episodes. Despite this the continuous baseline is easy to outline and makes the episodes easy to spot. The horizontal baseline experiences the same problems as in the first patient.

The third patient (bottom) has a slightly elevated ST-VM level for about 14 hours. The horizontal baseline describes the elevated part pretty well but introduce large errors at both ends of the interval shown here.

5.4 Discussion

The method we have used to define the baseline has several good qualities. It is easy to understand since it works partly the same way one does when one is visually outlining a baseline. First we look for segments where the baseline is well defined; then we connect these segments, avoiding fast changes and steep slopes. As we mentioned earlier the algorithm is stable even if the data contains large gaps and a lot of disturbances.
The algorithm also has a few disadvantages; it is unable to handle steps in the trend especially in the beginning and at the end of the trend. This problem could perhaps be fixed with another supplementary filter that identifies the steps and handles them either by dividing the baseline into two separate baselines or by making it very soft at the gap, making it possible for the baseline to follow the step. The latter method can easily be implemented by setting $\lambda$ very high. Another drawback is that the algorithm is quite slow and needs a lot of data to perform well. These are qualities that make its performance worse in an online situation. We have done no tests on exactly how much data is needed for the algorithm to reach acceptable performance but it is in the range of hours. One option is to let the baseline be updated as more data becomes available.

We have chosen to use a continuous baseline for two reasons: it is a lot easier to define in computerized analysis, and the visual appearance is better. In reality it is difficult to know what the baseline should look like or even what it is. In the end it all comes down to the definition of an episode. We have chosen to define an episode as a fast rise in the ST-VM level of more than 50 $\mu$V that doesn’t last more than 30 minutes. The meaning of the word fast is defined by the parameters in the baseline membership functions. Using several horizontal baselines causes errors at the beginning and end of each baseline segment. If one uses only one horizontal line, should it describe the lowest ST-VM value that is not a disturbance or some other level? Unless the lowest value that is not a disturbance is used a horizontal baseline forces one to make compromises.

The method we have used to find values for the parameters $k$, $a$, and $b$ is by no means optimal. To obtain better values for these constants some kind of statistical optimization should be performed. Herrnberger and Zimmer [14] used a cost function,

$$C(M) = \frac{1}{2K} \sum_{i=1}^{K} (1-M_i)^2 b_i + M_i^2(1-b_i),$$

where $M_i = \mu'(f_i(t_i),a_i,b_i,k_i)$ that was minimized for a set of $K$ manually pre-classified data points from various plots. The data points were classified as baseline points ($b_i=1$) or non-baseline points ($b_i=0$).

In our visual representation of the analysis we have chosen to plot the actual ST-VM trend and the baseline in the same plot instead of simply subtracting the baseline from the trend, like we do in the analysis, and plot the result. There are several reasons for this, the strongest perhaps that we are not able to trust our algorithm 100%. The results still have to be double-checked by the doctor that is responsible and we don’t want to deprive him/her of any information about the trend.

The conclusion is that the algorithm does a great job in finding the episodes. If implemented in a commercial product it would be of big help in the prognosis of patients with ST episodes and hopefully lead to better care. Automated analysis also opens up new areas of research since it makes it easy to try different values for testing parameters. One example is the threshold level 50$\mu$V used to find episodes (see figure 18).
6 Final conclusions

Today coronary artery disease is one of the largest causes of death in the western world. Due to growing problems with overweight coronary artery disease will probably be an even larger cause of death in the future. This creates a great need for better and faster methods to diagnose and treat heart problems. As new techniques, new medications, and new tests work their way into the hospitals the amount of information presented to the doctor grows unmanageable. Processing all the information takes to much time and sometimes is not even possible. Much of the analysis that doctors and nurses perform in order to extract information from data can be automated rendering greater precision and saving time. This in turn leads to better and faster methods to diagnose diseases. Computerized analysis of medical data also opens up new windows in research. Many hypotheses that previously have been too farfetched or too costly to test can now easily be tested on large materials. The task isn’t quite as easy as it seems however, during their professional lives doctors and nurses collect a lot of subjective knowledge in their fields. This knowledge is often very hard to summarize in rules that can be implemented using conventional algorithms. This is one of the things that make this subject very interesting from an engineering point of view.

The algorithm used to find the ST recovery parameters, recovery at 90 minutes and time to 50% recovery, is by all means very conventional. The work has been done in the evaluation of the algorithm where a large number of patients have been tested both manually and with the algorithm. The results were compared and every single difference has then been studied and the reason noted. This has given us great insight in what have to be improved in order to have a useful tool. Of 155 patients tested 111 matched exactly. In 22 patients out of the remaining 44 we concluded that the reviewers had made mistakes in their analysis (not following their own criteria for the analysis). In 9 patients the reviewers concluded that the data was in such a condition that no analysis could be done. In 12 patients the reviewer identified the maximum found by the algorithm as a disturbance and thus chose another maximum. One patient had three maxims with exactly the same amplitude. The reviewers concluded that the first maxima was the only correct one (the others were disturbances). We consistently choose the last one to minimize the risk of giving false positive results. To improve the performance of the algorithm we have to minimize the number of disturbances in the data. We also need to find a way to check if the identified maxima is a disturbance e.g. by looking at the average beat used to derive the ST value.

The algorithm that is used to find ST episodes is a bit more complex. It combines the information in a number of local features, slope, point density, and the distance to the curve created by performing a closing on the trend, into one parameter that tells us how much each point belongs to the baseline. It then utilizes a common smoothing algorithm, smoothing cubic splines, which incorporates a more global perspective of the trend, to create the baseline. The result is a nice looking baseline that very often lays very close to the one the human eye sees when outlining an imaginary baseline. The algorithm is also quite stable. The results of the algorithm was compared to the results from the Accent II study, but the result was very poor. The main reason for this is that the methods to define the baseline are different. While we use a continuous baseline, a straight line is used in the Accent II study. Instead we did an extensive analysis of each episode found in both materials. The conclusion drawn from this analysis is that our method of using a continuous baseline gives a better result than using straight lines. Using a continuous baseline leads to fewer compromises and fewer errors. To really evaluate the algorithm a comparison of the results against hard endpoints like 30-day mortality has to be done.
Computerized analysis of ST recovery and ST episodes

References


Appendix A: Deriving the Equations for Smoothing Cubic Splines

A spline function is a curve constructed from polynomial segments that are subject to conditions of continuity at their joints. Imagine that we have a set of coordinates \((x_0, y_0), \ldots, (x_n, y_n)\) of the function \(y(x)\). We want to bridge the gap between adjacent points \((x_i, y_i), (x_{i+1}, y_{i+1})\) using the cubic functions \(S_i; i=0, \ldots, n-1\) to piece together a curve with continuous first and second derivatives.

The function \(S_i\) can be written

\[
S_i(x) = a_i (x-x_i)^3 + b_i (x-x_i)^2 + c_i (x-x_i) + d_i,
\]

where \(x\) ranges from \(x_i\) till \(x_{i+1}\). The first and second derivatives of this function are

\[
\begin{align*}
S'_i(x) &= 3a_i (x-x_i)^2 + 2b_i (x-x_i) + c_i, \\
S''_i(x) &= 6a_i (x-x_i) + 2b_i.
\end{align*}
\]

Note that \(d_i = y_i\) and the conditions of continuous first and second derivatives make it possible to extract the polynomial constants \(a_i, b_i, c_i\) for all \(i\) and thus the spline function \(S(x)\) is known.

Now imagine instead that our set of coordinates \((x_0, y_0), \ldots, (x_n, y_n)\) is from the function \(y_i=f(x_i)+e_i\) where \(e_i\) is some independent error. Since we want to approximate the function \(f(x)\) it would be unsuitable to interpolate these coordinates. Instead we want our function to be able to depart from the data points in order to obtain a smooth function. Since \(d_i = y_i\) is no longer true we need another criterion to be able to extract the polynomial constants. The solution is to find a spline function that minimizes

\[
L = \lambda \sum_{i=0}^{n} y_i^2 - S_i^2 + (1-\lambda) \int_{x_i}^{x_{i+1}} (S''(x))^2 \, dx,
\]

where \(S_i = S_i(x_i)\) and \(\lambda \in [0,1]\) is a parameter that controls the relation between the importance of being close to the data points and the smoothness of the curve. If \(\lambda=0\), smoothness is all that matters and the spline function \(S\) becomes a straight line; if \(\lambda=1\), being close to the coordinates is the only priority and we get an interpolating spline function.

Since \(S\) is a piecewise function the second part of \(L\) can be written

\[
\int_{x_i}^{x_{i+1}} (S''(x))^2 \, dx = \sum_{i=0}^{n} \int_{x_i}^{x_{i+1}} (S''(x))^2 \, dx,
\]

and as we are using cubic polynomial functions the integral over the second derivative in the interval \([x_i, x_{i+1}]\) is

\[
\int_{x_i}^{x_{i+1}} (S''(x))^2 \, dx = \frac{4h_i}{3} (b_i^2 + b_i b_{i+1} + b_{i+1}^2),
\]

where \(h_i = x_{i+1} - x_i\).

This means \(L\) can be written as

\[
L = \frac{3\lambda}{4(1-\lambda)} \sum_{i=0}^{n} (y_i - d_i)^2 + \sum_{i=0}^{n} h_i (b_i^2 + b_i b_{i+1} + b_{i+1}^2),
\]

where \(d_i = S_i(x_i)\).

In our case with the baseline the data points \((x_0, y_0), \ldots, (x_n, y_n)\) belong to a fuzzy set and we thus know the membership function \(\mu_y\). We want \(S(x)\) to interpolate the data points that belong to the baseline, that is, has \(\mu_y = 1\), and at the same time completely ignore the data points that don’t belong to the baseline, i.e. where \(\mu_y = 0\). Put in another way, in those sections of the curve where the data points for sure belongs to the baseline smoothness isn’t that important as long as it interpolates the data points. On the other hand, in those segments where the data points do not belong to the baseline it is important that the curve doesn’t make any sudden changes. This behavior can be obtained if we let the smoothness parameter \(\lambda\) depend on the membership function \(\mu_y\) so that \(\lambda = \lambda(\mu_y(x_i))\). \(L\) the becomes
If we use equation A.11 to express 

\[ L = \sum_{i=0}^{n} \frac{1}{v_i} (y_i - d_i) + \sum_{i=0}^{n} b_i (b_{i+1} - b_{i-1}) , \]

where \( v_i = t(1-x_i) / 3 \lambda_i \). Minimizing this function gives us \( n \) equations to obtain \( 2n \) parameters. By using the conditions for continuous first and second derivatives we can eliminate \( a_i \) and \( b_i \) and obtain the following equation

\[ b_{i+1} h_{i+1} + 2b_i h_{i+1} + b_i h_i = \frac{3}{h_i} (d_{i+1} - d_i) - \frac{3}{h_{i+1}} (d_i - d_{i-1}). \]

This gives another \( n-2 \) equations. By using the condition that the first derivative is zero at the beginning and end we get two more equations and are able to obtain \( b \) and \( d \). Equation A.8 and the conditions at the ends give us the following matrix equation

\[
\begin{bmatrix}
2h_0 & h_0 & 0 & \Lambda & 0 & 0 & 0 \\
h_0 & p_1 & h_1 & \Lambda & 0 & 0 & 0 \\
0 & h_1 & p_2 & \Lambda & 0 & 0 & 0 \\
M & M & M & O & M & M & M \\
0 & 0 & 0 & \Lambda & p_{n-2} & h_{n-2} & 0 \\
0 & 0 & 0 & \Lambda & h_{n-2} & p_{n-1} & h_{n-1} \\
0 & 0 & 0 & \Lambda & 0 & h_{n-1} & 2h_{n-1} \\
r_0 & r_0 & 0 & \Lambda & 0 & 0 & 0 \\
r_0 & r_1 & f_1 & r_1 & \Lambda & 0 & 0 \\
r_1 & r_2 & f_2 & r_2 & \Lambda & 0 & 0 \\
M & M & M & O & M & M & M \\
0 & 0 & 0 & \Lambda & f_{n-2} & r_{n-2} & 0 \\
0 & 0 & 0 & \Lambda & r_{n-2} & f_{n-1} & r_{n-1} \\
0 & 0 & 0 & \Lambda & 0 & r_{n-1} & -r_{n-1} \\
\end{bmatrix}
\begin{bmatrix}
b_0 \\
b_1 \\
b_2 \\
\vdots \\
b_{n-2} \\
b_{n-1} \\
b_n \\
d_0 \\
d_1 \\
d_2 \\
d_3 \\
d_{n-1} \\
d_n \\
\end{bmatrix}
\]

where

\[ p_i = 2(h_{i+1} + h_i), \]

\[ r_i = \frac{3}{h_i}, \]

\[ f_i = \frac{-3}{h_{i+1}} + \frac{3}{h_i} = -(r_{i+1} + r_i). \]

This can be written

\[ Mb = Q^d. \]

If we use equation A.11 to express \( L \) we get

\[ L = (y - d)^T A^{-1} (y - d) + b^T Mb , \]

where \( A=\text{diag}\{v_0, \ldots, v_n\} \). By using \( b=M^T Q^d \) we can write \( L \) as a function of \( d \) alone.

\[ L(d) = (y - d)^T A^{-1} (y - d) + d^T Q M^{-1} Q^d. \]

To minimize this function we differentiate it with respect to \( d \) and set the result to zero, which gives us

\[ - (y - d)^T A^{-1} + d^T Q M^{-1} Q = 0 , \]

\[ \text{(A.14)} \]
If we rewrite this as

\[(A.15) \quad \Lambda^{-1}(y - d) = QM^{-1}Q'd = Qb,\]

multiply with \(Q' \Lambda\) and use the fact that \(Mb = Q'd\) we get

\[(A.16) \quad (M + Q' \Lambda Q)b = Q'y.\]

Using equation A.16 to obtain \(b\) the second parameter \(d\) can be obtained from

\[(A.17) \quad d = y - \Lambda Qb.\]
Appendix B: Finding the episodes

The algorithm used to find the episodes after the baseline has been identified is described in pseudo code below.

```plaintext
function epifind { STVM, time, baseline }
    threshold ← baseline+50
    for i = 1:length (STVM)
        if STVM(i) >= threshold(i) & STVM(i-1) < threshold(i-1)
            ++n
            if time(i)-time(i-1) <= MAXGAP
                episode(n).start ← time(i-1)+threshold(i)*(time(i)-time(i-1))/(STVM(i)-STVM(i-1))
            else
                episode(n).start ← time(i)
            end
        elseif STVM(i) >= threshold & STVM(i-1) >= threshold & time(i)-time(i-1) > MAXGAP
            episode(n).end ← time(i-1)
            ++n
        elseif STVM(i) < threshold & STVM(i-1) >= threshold
            if time(i)-time(i-1) <= MAXGAP
                episode(n).start ← time(i-1)+threshold(i)*(time(i)-time(i-1))/(STVM(i)-STVM(i-1))
            else
                episode(n).start ← time(i)
            end
        end
    end
    if STVM(end) >= threshold & STVM(end-1) >= threshold
        --n
    end
    for j = 2:n
        k=j-a
        if episode(j).start-episode(k-1).end < MINBETWEEN
            episode(k-1).end ← episode(j).end
            ++a
        end
    end
    for l = 1:length (episode)
        if episode(l).end-episode(l).start>=MINTIME & episode(l).end-episode(l).start<=MAXTIME
            tempepisode(m).start ← episode(l).start
            tempepisode(m).end ← episode(l).end
            ++m
        end
    end
    episode ← tempepisode
    return episode
end
```