

Comparison of manually and automatically derived ECG-based left ventricular hypertrophy parameters and their predictive value in survival analysis

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Biomedical Engineering

by

Corina Contawe

Registration Number 1129874

to the Institute of Electrodynamics, Microwave and Circuit Engineering

at the TU Wien

Advisor: Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Eugenijus Kaniusas

Co-advisor: Univ.Lektor Dipl.-Ing. Dr.techn. Christopher Mayer

Vienna, 8th June, 2017

Corina Contawe

Declaration of Authorship

Corina Contawe
Kanitzgasse 13-19/1/2 1230 Vienna, Austria

I hereby declare that I have written this Master Thesis independently, that I have completely specified the utilized sources and resources and that I have definitely marked all parts of the work - including tables, maps and figures - which belong to other works or to the internet, literally or extracted, by referencing the source as borrowed.

Vienna, 8th June, 2017

Corina Contawe

Erklärung zur Verfassung der Arbeit

Corina Contawe
Kanitzgasse 13-19/1/2 1230 Vienna, Austria

Hiermit erkläre ich, dass die vorliegende Arbeit gemäß dem Code of Conduct - Regeln zur Sicherung guter wissenschaftlicher Praxis, insbesondere ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel, angefertigt wurde. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder in ähnlicher Form in anderen Prüfungsverfahren vorgelegt.

Wien, 8. Juni 2017

Corina Contawe

Acknowledgements

To my mom and dad: thanks to all your eternal love and encouragement, I am where I am today, and I am forever grateful. I love you both, always.

I would like to express my gratitude to the members of Cardiovascular Diagnostics team at the Austrian Institute of Technology for giving me the opportunity to carry out this thesis. Special thanks to Christopher Mayer and Stefan Hagmair for all their guidance and input during all stages of the thesis process. I would also like to thank Professor Eugenijus Kaniusas at the TU Vienna for his inspiring lectures and for overseeing this work.

A very special thank you to my cousin Roland and my good friend Huem for your friendship and guidance during my entire stay in Vienna. I would also like to thank my siblings, Carol, Carl and Vicki, for being there for me. Thanks also to all my friends who have been with me through everything, you know who you are.

I owe my deepest gratitude to Marcus, my fiancé, my partner, and my best friend, who helped me through the hardest times life has shown me so far and supported me endlessly. For your selfless support, I am forever indebted. I love you.

Many endless thanks,

Corina

Kurzfassung

Linksventrikuläre Hypertrophie (LVH) ist insbesondere unter Dialysepatienten mit einem hohen Risiko an kardiovaskulärer Sterblichkeit verbunden. Da LVH reversibel ist, wäre eine akkurate Entdeckungsmethode für diese Kohorte von großer Bedeutung. Die Elektrokardiographie (EKG) ist eine universelle und kostengünstige Methode, welche in vielen medizinischen Einrichtungen zum Einsatz kommt. Die Weiterentwicklung und Verbesserung EKG-basierter LVH-Parameter zur Analyse der LVH-Inzidenz ist daher von großem Interesse. Das Standardverfahren für die Berechnung der EKG-Parameter besteht in der manuellen Analyse und Auslesung der EKG-Messungen. Eine automatische Methode zur Berechnung der EKG-basierten LVH-Parameter ist vorteilhaft gegenüber der manuellen Auslesung, da letztere anfällig für menschliche Fehler und darüber hinaus zeitaufwändig ist.

Ziel dieser Masterarbeit ist es, eine automatische Berechnungsmethode für die Parameter zu entwickeln und die Ergebnisse mit den manuellen Auslesungen eines erfahrenen Mediziners zu vergleichen. Des Weiteren werden die Ergebnisse, in Kombination mit Follow-Up Daten, in Hinblick auf ihren Vorhersagewert mittels Ereigniszeitanalyse beurteilt. Folgende vier Parameter wurden für den Vergleich ausgewählt: Sokolow-Lyon-Voltage (SLV), Cornell Voltage (CV), Cornell-Voltage-Produkt (CVP) und Novacode-Estimate. Für die Detektion der EKG Merkmale wurden der AIT ECGsolver Algorithmus, sowie QRS-Detektoren von PhysioNet verwendet. Für die Berechnung der LVH-Parameter wurde ein Algorithmus entwickelt, welcher automatisiert Amplituden und Intervalle von 335 Patienten der ISAR Studie berechnet. Mit Hilfe üblicher Grenzwerte wurden die Parameter dichotomisiert. Sowohl die Ergebnisse der Parameter, als auch die dichotomen Merkmale wurden mit den manuellen Kriterien des Mediziners anhand von 335 Patienten der ISAR Studie verglichen.

Die Kriterien SLV und CV, sowie der weibliche Novacode estimate für Frauen stimmten in beiden Methoden gut überein (Parameter und LVH-Ergebnis). Der McNemar-Test zeigte keinen statistisch signifikanten Unterschied zwischen den Messmethoden, welche auf den dichotomisierten Werten beruhen. Als Problem stellte sich die konsistente Überschätzung der QRS-Dauer durch automatisierte Methoden im Vergleich zur manuellen Auslesung dar. Die Parameter CVP sowie der Novacode estimate für Männer hängen beide von der QRS-Dauer ab und zeigten weniger Übereinstimmung zwischen den Messmethoden. Der

McNemar-Test lieferte statistisch signifikante Unterschiede für den Novacode estimate für Männer, jedoch nicht für den Parameter CVP. Die Ereigniszeitanalyse ergab CV, CVP und den Novacode estimate als signifikante Prädiktoren der Gesamtmortalität.

Zusammenfassend zeigte sich, dass automatische Methoden der LVH-Parameter-Berechnung mit Einschränkungen sinnvoll sind. Eine Optimierung der QRS-Dauer Detektion würde diese Methode erheblich verbessern. Für eine erweiterte Version dieser Arbeit würden die folgenden Untersuchungen in Frage kommen: Vergleich der dichotomen Ergebnisse mit präzisen LVH-Parametern, LVH-Parameter während und außerhalb der Dialysezeit und ein erneuter Vergleich der Ergebnisse mit manuellen Werten, welche mittels Standardmethoden bestimmt wurden.

Abstract

The presence of left ventricular hypertrophy (LVH) is linked to high risk of cardiovascular mortality, especially in dialysis patients. With the knowledge that LVH is reversible, an accurate method of detection in this cohort would be invaluable. Electrocardiography (ECG) is an universal, inexpensive procedure done at most healthcare settings. It is of great interest to further develop and improve ECG-based LVH parameters for assessing LVH incidence. The standard procedure for determining the parameters is to derive them from manual measurements of the electrocardiogram. Manual measurements are subject to human error and time consuming, thus it would advantageous to develop automatic methods of computing ECG-based LVH parameters.

The aim of this thesis is to develop an automatic method of calculating the parameters and compare the results with manual measurements done by an experienced physician. Furthermore, the results, in combination with follow-up data, will be assessed for their predictive value in survival analysis. Four parameters were chosen to be automatically compared, the Sokolow-Lyon Voltage (SLV), the Cornell Voltage (CV), the Cornell Voltage Product (CVP), and the Novacode estimate. The AIT ECGsolver and QRS detectors taken from PhysioNet were used for detection of ECG waveforms and features. An algorithm was developed to compute the criteria by automatically evaluating amplitudes and intervals on a subset of 335 patients of the ISAR study. Recommended thresholds were applied to convert the calculated variables into dichotomized measurements. Both the criteria results and the threshold outcomes were compared with the manually derived matched beat measurements annotated by the physician.

The voltage dependent criteria, SLV, CV, and the female Novacode estimate, demonstrated very good agreement between the two methods, both in measurements and LVH outcome. McNemar's test for these criteria showed no statistically significant difference between the measurement methods based on the dichotomized outcomes. Consistent overestimation of the QRS interval by automatic methods in comparison to the manual method proved to be problematic. The CVP and male Novacode are both dependent on QRS duration and demonstrated less agreement between the method measurements. McNemar's test showed for the male Novacode that there is a statistically significant difference between the methods, but not for the CVP. Interestingly, the survival analysis showed that CV, CVP, and the automatically derived Novacode were statistically signifi-

cant risk predictors for all-cause mortality. This is in contrast to published studies which report no significant LVH criteria for all-cause mortality.

Automatic methods of parameter calculation proved to be useful, but with limitations. Improvements in QRS interval detection would greatly refine the method. Future opportunities to expand the work done in this thesis would be a comparison of the binary results with accurate left ventricular mass (LVM) measurements, a repeat of the work during known dialysis treatment times (dialysis versus non-dialysis), and a recomparison using a standard method of manual QRS interval measurement.

Contents

Kurzfassung	ix
Abstract	xi
Contents	xiii
1 Introduction	1
1.1 Motivation	1
1.2 Aim of the Thesis	2
1.3 Methodological Approach	2
1.4 Structure of the Thesis	3
2 Background	5
2.1 Physiological Background	5
2.1.1 Anatomy and Physiology of the Left Ventricle	5
2.1.2 Left Ventricular Hypertrophy	6
2.2 Methods of LVH Detection	7
2.2.1 Cardiac Imaging	7
2.2.2 Electrocardiogram	8
2.2.3 ECG Interpretation Challenges for Dialysis Patients	9
2.2.4 ECG-based Parameters for LVH	10
Sokolow-Lyon Voltage	10
Cornell Voltage	11
Cornell Voltage Product	11
Novacode Estimate	11
Thresholds for presence of LVH	12
3 Methods	13
3.1 Clinical Data Collection	13
3.2 Manual Annotation	14
3.3 Automatic Annotation	14
3.3.1 ECG Extraction using the AIT ECGsolver	14
3.3.2 Signal Pre-Processing	15
3.3.3 Automatic Amplitude Evaluation	15

3.3.4	Automatic QRS Duration Evaluation	18
3.4	Parameters	20
3.5	Statistical Methods	20
4	Results	23
4.1	Baseline Data	23
4.2	Amplitude Comparison	23
4.2.1	RaVL Amplitude Comparison	25
4.2.2	SV3 Amplitude Comparison	26
4.3	QRS Duration Comparison	27
4.3.1	Manual and AIT ECGsolver Comparison	27
4.3.2	MIT GQRS and SQRS Annotators Comparison	28
4.3.3	Combined MIT Annotators and Manual QRS Duration Comparison	30
4.4	Criteria Statistics	31
4.4.1	Sokolow-Lyon Voltage Criteria	31
4.4.2	Cornell Voltage Criteria	33
4.4.3	Cornell Voltage Product Criteria	35
4.4.4	Novacode Criteria	36
4.4.5	Novacode Criteria Men	37
4.4.6	Novacode Criteria Women	38
4.5	Survival Analysis	40
5	Discussion	43
5.1	Amplitude Comparison	43
5.1.1	rSR' Complex	44
5.2	QRS Duration Comparison	45
5.3	Parameters	46
5.3.1	Sokolow-Lyon Voltage	46
5.3.2	Cornell Voltage	47
5.3.3	Cornell Voltage Product	48
5.3.4	Novacode	48
	Novacode Male	48
	Novacode Female	49
5.4	Survival Analysis	50
6	Conclusion	51
A	QRS dependent criteria with ECGsolver durations.	53
	List of Figures	55
	List of Tables	56
	Bibliography	57

Introduction

1.1 Motivation

The incidence of cardiovascular diseases in patients suffering from chronic kidney disease (CKD), and in particular those who have progressed to end-stage renal disease (ESRD), is prevalent [11]. Patients suffering from ESRD are required to undergo either dialysis treatment or a kidney transplant. In those who receive dialysis, cardiovascular diseases are the leading cause of mortality [11]. Left ventricular hypertrophy (LVH) is a condition that is frequently seen in dialysis patients and is indicative of present or future cardiovascular problems [29]. The characteristic increased thickness of the left ventricular mass (LVM) in sufferers of LVH is an adaptive compensation to the increased afterload due to pathology such as hypertension in ESRD patients [23]. With the knowledge that LVH may be reversible [32], early detection and assessment of LVH give clinicians the ability to provide better, specific care for dialysis patients.

Though modern imaging techniques such as echocardiography and cardiac computer tomography (CT) have proven to be the most accurate methods for measuring LVM, the value of electrocardiography in the detection of LVH must not be taken for granted. Unlike echocardiography and cardiac CT, which directly measure cardiac muscle size, ECG takes into account the increased electrical activity of the heart due to the enlarged myocardium mass. The status of ECG as a worldwide standard in basic patient care at relatively low costs makes it the most convenient and valuable method for diagnosing LVH. Continuous research on the use of ECG in detecting LVH has resulted in the development of numerous ECG-based LVH parameters or criteria. The diagnostic value of these criteria are constantly being revised and developed. The most trusted method of evaluating LVH parameters is by manual measurement of specific deflections in a standard 12-lead ECG. Although ECG-based LVH parameters have been shown to have a low sensitivity (many false negatives), they also display a very high specificity (few

false positives). This high specificity combined with the ubiquity of ECG in health settings makes these criteria an extremely useful method for determining the likelihood of LVH and furthermore, the possibility of subsequent cardiovascular problems. Thus, it is advantageous for the medical community to investigate LVH prevalence through the use of ECG-derived parameters and to further improve and evaluate the prognostic value of these criteria.

As a result, the use of ECG-based LVH parameters is important for improved patient care, especially for patients of dialysis. The development of an automatic method for LVH parameter calculation would add to the value of these criteria as the method can be used to measure various specific time points such as dialysis versus non-dialysis periods. As cardiovascular disease continues to be a widespread epidemic, the detection of LVH as an indicator of cardiovascular pathology is imperative.

1.2 Aim of the Thesis

The data acquired for this thesis are taken from a study conducted by the Klinikum rechts der Isar, Technische Universität München and in partnership with the AIT Austrian Institute of Technology GmbH [36]. The objective of this study is to observe a cohort of dialysis patients that is representative of the current epidemiological population of dialysis centers in western countries. The purpose of this thesis is to use the ECG data collected from the subjects and derive automatic LVH parameters and compare the results with provided manual measurements. With provided manually derived ECG-based LVH parameters, statistical tests will be carried out to evaluate how well the automatic measurements work out. Additionally, survival analysis using follow-up data taken two years after the initial recordings will be performed. Manual and automatically derived ECG-based LVH parameters will be used to evaluate their predictive value for dialysis patients.

1.3 Methodological Approach

The ECG data for this study have been recorded by a 24-hour 12-lead Holter ECG recorder. The implementation of the thesis is performed in the MATLAB[®] 2015b environment to analyze the provided data and derive LVH parameters. The ECG-based parameters used to characterize LVH, and suggested by the study by Covic et al. [11], are the Sokolow-Lyon Voltage, Cornell Voltage, Cornell Voltage Product, and Novacode estimate. An experienced physician, blinded to the patient's clinical status, has manually measured the same ECG-based LVH parameters for single beats. The automatically derived parameters of matched beats are then computed using the developed algorithm. Comparison of the parameter derivatives (manual vs. automatic) will be carried out using Bland-Altman plots. Finally, Cox regression models and the Kaplan-Meier method will be performed for survival analyses of the cohort data.

1.4 Structure of the Thesis

Chapter 1: Introduction

The first chapter of this thesis introduces the topic of LVH and the use of ECG as a valuable method of LVH detection. The motivation for the work and the methodology are briefly explained.

Chapter 2: Background

The background chapter presents the literature research conducted during the course of this thesis. The main sections of the background chapter describe the physiological principles that lead to LVH and the methods with which LVH can be identified. The method of ECG is explained in detail, including several parameters developed for LVH detection and challenges with ECG interpretation for dialysis patients.

Chapter 3: Methods

The methods chapter describes in detail the methodology applied to evaluate automatically derived ECG-based LVH parameters, compare the automatic results with the manually derived parameters, and perform survival analyses on the cohort data using follow-up data.

Chapter 4: Results

The results of the data analysis are present in chapter four. The comparison of amplitude, duration and parameter measurements between manual and automatic methods are displayed. In addition, the results from the survival analysis using follow-up data are shown.

Chapter 5: Discussion

In the discussion chapter, the findings in chapter four are analyzed and interpreted. The limitations, in addition to the significance of the results are discussed.

Chapter 6: Conclusion

The conclusion chapter provides a review of the approach and findings of this work. An examination of the achievements and challenges encountered during the process of this thesis, in addition to suggestions for further investigation, are given.

Background

The purpose of this chapter is to give some background context to the problem that is investigated in this thesis. The chapter is divided into two main sections: physiological background and methods of LVH detection. The former delves into the physiological background of the heart and LVH as a cardiovascular condition. The latter section describes the methods used for LVH detection, including ECG-based LVH parameters.

2.1 Physiological Background

2.1.1 Anatomy and Physiology of the Left Ventricle

The human heart is one of the body's most important organs and is the driving force for the supply of nutrients and removal of waste throughout the entire body. It is anatomically divided into four chambers, as shown in figure 2.1. Blood enters the heart through the atria, the two upper chambers, and is expelled through larger lower chambers, called ventricles.

The human heart functions as a complementary double pump for the pulmonary and systemic circulation systems. These two types of circulation are associated with the two sides of the heart, the right and the left. In a process known as pulmonary circulation, the right atrium receives deoxygenated blood from both the inferior and superior vena cava and then transfers the blood to the right ventricle to be pumped out of the heart to the lungs [19]. In systemic circulation, oxygenated blood from the lungs re-enters to the left heart through the pulmonary arteries and is received in the left atrium [19]. After being moved to the left ventricle, blood is then pumped through the aorta to be distributed to the whole body, with the exception of the lungs [19]. Due to the extensive scope in which blood from the systemic system must travel in comparison to blood expelled from the pulmonary system, the left ventricle consists of a much thicker wall than the right ventricle [19], as illustrated in figure 2.1.

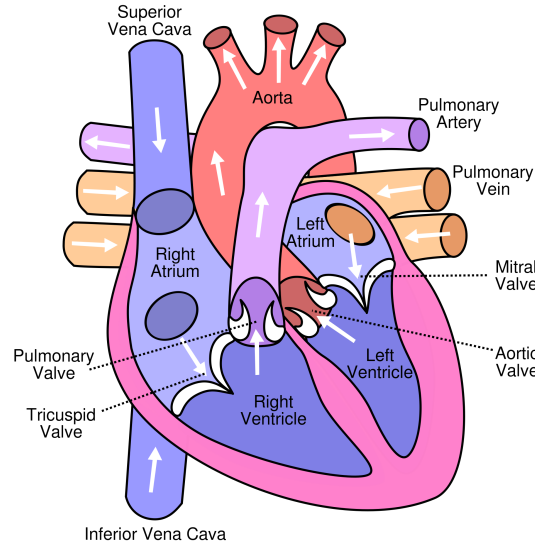


Figure 2.1: The four chambers of the human heart with the major vessels [43].

The heart wall is composed of three major layers, the endocardium, the myocardium, and the epicardium [19], as shown in figure 2.1. The middle layer, the myocardium or cardiac muscle tissue, comprises the majority of the heart wall. The cardiac myocytes are the primary cells of the myocardium and are responsible for the mechanical contraction of the heart [25]. At the cellular level, the cardiac myocytes consists of myofibrils that are composed of myofilaments. The myofibrils are ordered in a parallel arrangement of contractile units called sarcomeres [25]. Sarcomeres are distinct elements containing the proteins actin and myosin that respond in a mechanochemical fashion. When a chemical signal is released, a conformational change occurs between the proteins causing the sarcomere unit to shorten. The simultaneous shortening of all the sarcomeres in the cardiac muscle causes a contraction. The contraction of the ventricular myocardium causes the pumping of blood out of the ventricle.

2.1.2 Left Ventricular Hypertrophy

When the circulatory system is faced with an increased afterload over time, the heart must find a way to compensate. One way to offset the increase in pressure and load is to increase the pumping power of the left ventricle. This is achieved by increasing the myocardium in the wall of the left ventricle, and thus, increasing LVM [25].

Hypertrophy describes the enlargement of an organ or tissue due to an increase in the volume of the cells forming it. As an adaptive compensation, the left ventricle increases in mass by hypertrophy and the condition is known as left ventricular hypertrophy (LVH) [25]. By creating bigger cardiac myocytes, the myocardium in turn multiplies its number of sarcomere units and therefore, the contractile power of the ventricle. Though

this increase in size may also be due to consistent extreme physical exertion, as happens in the case of the so-called athlete's heart [35], in dialysis patients it is usually due to existing pathology.

LVH itself is not a disease, but a condition that signifies serious warning signs about a patient's cardiovascular health. In dialysis patients, 75% are found to have LVH [8]. LVH is a common feature in several cardiovascular complications. One of the main causes of LVH is high blood pressure (BP), also known as hypertension. Some studies have shown that up to 86% of hemodialysis patients have been diagnosed with hypertension [1], most likely causing their ESRD condition. Increased BP causes LVH by creating greater amounts of work and stress on the left ventricular muscle to pump the same volume of blood. Other cardiovascular pathologies that increase the amount of work for the muscle are those that increase the amount of afterload in the left ventricle such as aortic stenosis and aortic insufficiency [25].

As the myocardium layer grows as a result of pathology, the left ventricle becomes weaker, stiffer, and less elastic [26], therefore leading to worse cardiovascular conditions. In the case of physical exertion, these changes do not occur in the muscle and the hypertrophy is easily reversible. Pathological LVH has also been shown to not be permanent [20], but if left untreated, may lead to congestive heart failure, stroke, and cardiovascular mortality. Studies have shown that by controlling blood pressure, a regression in LVH can be achieved [31, 27].

2.2 Methods of LVH Detection

Over the years various methods have been used for detecting LVH in vivo. The validation of the different methods have been carried out in experiments by comparing in vivo results to cadaver measurements of LVM.

2.2.1 Cardiac Imaging

Medical images of the heart are the most accurate methods for detecting increases in heart wall thickness [11]. Cardiac imaging encompasses a variety of techniques that include, but are not limited to, echocardiography, cardiovascular magnetic resonance imaging (MRI), and cardiac computer tomography (CT). Though the previously mentioned methods are not typically used for initial clinical diagnosis of LVH, studies which compare LVH incidence to postmortem data show that they display the most accurate results [20]. In addition to their accuracy, the methods are non-invasive and provide reliable results. The downside of these techniques tend to be cost and availability. Cardiac MRI and CT imaging techniques are very expensive methods and therefore are not commonly used to initially diagnose LVH. Also, to use cardiac CT as a method of LVH detection would expose patients to unneeded levels of radiation that would be undesirable.

2.2.2 Electrocardiogram

Electrocardiography is a common technique used by medical facilities to monitor the electrical activity of the heart [10]. Therefore, in contrast to medical imaging, rather than directly measuring heart wall thickness, LVH detection by ECG relates a combination of features from the electrocardiogram to LVM. These features include specific voltages and durations. Various studies have developed equations that use a combination of features to yield parameters for LVH. These parameter equations are discussed further in the next subsection.

Electrocardiography is a fairly easy method that has become a standard in modern hospitals since its invention in the early 20th century [15]. The electrocardiogram is the recording of the potential difference measurements between electrodes placed on the body over a certain period of time. In a standard 12-lead ECG, ten electrodes are placed on the body of the patients and twelve leads are recorded [10]. Figure 2.2 shows the electrode placement for the 12-lead ECG.

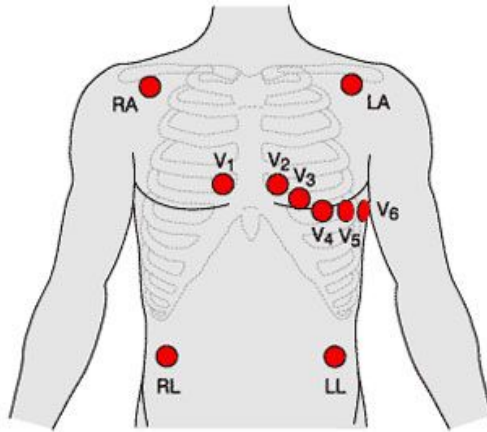


Figure 2.2: The placement of the ten electrodes on the body for a 12-lead ECG [37].

Three electrodes located on the limbs of the body are used to derive the three limb leads by recording the potential difference between the electrodes [10]. The upper segment of figure 2.3 shows how the limbs leads, labeled I, II, and III, are derived using Einthoven's triangle [10]. These same three limb electrodes are used to calculate the augmented limb leads, as shown in the lower segment of figure 2.3. These leads, called aVF, aVL, and aVR, are derived by taking the potential difference between the electrode and the center of Einthoven's triangle [10]. The remaining six electrodes, called the precordial electrodes, are placed on the chest and the precordial leads V1 - V6 are derived.

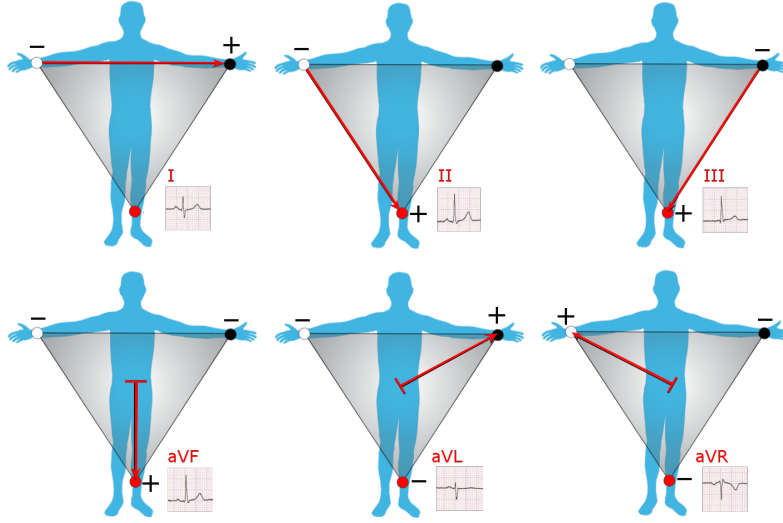


Figure 2.3: The derivation of the limb and augmented limb leads using Einthoven's triangle [46].

The familiar shape of the ECG is divided into segments as shown in figure 2.4. The most distinguishable characteristic of an ECG are the features known as the P wave, the QRS complex, and the T wave.

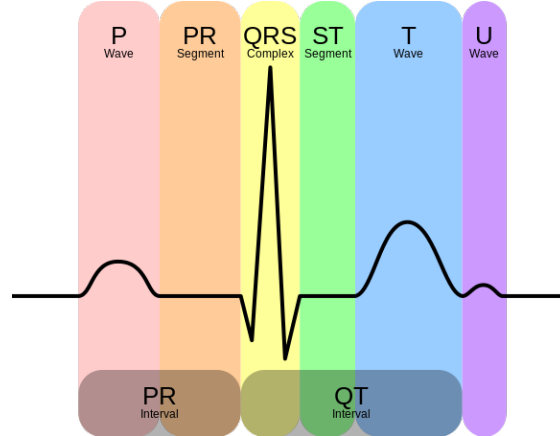


Figure 2.4: ECG schematic with labeled complexes [45].

2.2.3 ECG Interpretation Challenges for Dialysis Patients

ECG interpretation has been shown to be especially challenging for ESRD patients on dialysis [33]. For patient receiving hemodialysis, a treatment is typically done three times a week [13]. During hemodialysis, the patient's blood is removed, cleaned through a

dialyser, and returned to the body. During the cleaning of the blood, a dialysate, or dialyser solution, flows through the dialyser in the opposite direction from the blood [13]. The difference in electrolyte concentration is the driving force for removal of waste in the blood [13]. Consequently, the amount of fluids and electrolytes in the blood for hemodialysis patients is highly dependent on treatment times. Poulikakos et al. have shown that these fluctuations in fluid and electrolyte amounts affect ECG amplitudes and intervals and therefore pose a problem in ECG interpretation for hemodialysis patients [33]. Nevertheless, Poulikakos et al. revealed that by considering dialysis times, ECG readings for dialysis patients are still meaningful and useful for identifying cardiovascular conditions and future risk stratification [33].

2.2.4 ECG-based Parameters for LVH

Over the years, several parameters based on ECG signals have been developed for detecting LVH. Many ECG-based LVH parameters were derived by investigating changes in wave forms in the 12-lead ECG due to an increased LVM. These changes normally include a larger and longer QRS complex and abnormalities in the T and P waveforms [20]. Many of the early parameters are based on simple calculations between QRS voltages of different leads. Over the years, different methodologies have been developed to further improve the sensitivity and specificity of ECG-based LVH parameters. Types of parameters include methods based on voltages, voltage-product, and regression formula. The parameters that are investigated in this work were chosen based on the results from Covic et al. [11] who have applied the following ECG-based LVH parameters for risk prediction in ESRD patients. The following notation will be used in the subsequent parameter equations: $R(V_5)$, which corresponds to the R peak amplitude of the precordial lead V5.

Sokolow-Lyon Voltage

First published in the American Heart Journal in 1949 by physicians Maurice Sokolow and Thomas Lyon, the Sokolow-Lyon Voltage (SLV) [39] has become one of the most common criteria for quantitatively evaluating the mass of the left ventricle. The parameter is measured by adding the amplitudes of the S wave in the lead V1 and the maximal R wave in either the V5 or V6 leads. Presence of LVH is determined when the addition of these voltages is above a defined threshold. In their original publication, Sokolow and Lyon reported the SLV criteria had a sensitivity of 32% and a specificity of 100% [39] for their study population. The SLV parameter is characterized as a voltage criteria since it depends only on ECG voltages. The voltages are measured as absolute values and in millivolts according to the following equation:

$$SLV \text{ (mV)} = S(V_1) + \max(R(V_5), R(V_6)). \quad (2.1)$$

Cornell Voltage

Casale et al. [9] sought to develop an improved ECG-based criteria to evaluate LVH. In their 1985 publication they proposed the Cornell Voltage (CV) based on the amplitudes of the S wave in lead V3 and the R wave in lead aVL. As opposed to the SLV criteria, the CV defines different thresholds based on the sex of the subject, as shown in table 2.1, therefore increasing sensitivity and specificity in comparison to the SLV [9]. The results of Casale et al. demonstrated the performance of both criteria on the same patient test series. The SLV criteria gave a sensitivity of 20% and a specificity of 93%, while the CV performed with a sensitivity of 41% and a specificity of 98% [9]. Like the SLV, the Cornell Voltage is a voltage-based ECG parameter and the amplitudes are absolute and measured in millivolts. CV is defined as follows:

$$CV \text{ (mV)} = R(aVL) + S(V_3). \quad (2.2)$$

Cornell Voltage Product

According to the 1992 publication by Molloy et al. [29], a combination of the Cornell Voltage with QRS duration proved to be a more accurate assessment of LVH in comparison to the CV parameter alone. By measuring post-mortem LVM and comparing the results of the Cornell Voltage parameter and the Cornell Voltage Product (CVP) parameter, Molloy et al. found that the parameter-duration product displayed a significantly improved sensitivity in comparison to the CV criteria [29]. In their publication, Molloy et al. reported that the CV resulted in a sensitivity of 36%, while the CVP performed with a sensitivity of 51% [29]. Both criteria performed at a matched specificity of 95% [29]. As a result, the CVP is also a commonly used ECG-based parameter for LVH detection. While the voltage calculation remains the same as the CV, the CVP equation includes a factor for QRS duration which is measured in milliseconds. The CVP is defined for men and women as:

$$CVP \text{ (mV} \cdot \text{ms)} = \begin{cases} (R(aVL) + S(V_3)) \times \text{QRS duration, men} \\ (R(aVL) + S(V_3) + 0.8 \text{ mV}) \times \text{QRS duration, women.} \end{cases} \quad (2.3)$$

Novacode Estimate

The Novacode estimate developed by Rautaharju et al. in their 1998 publication [34] uses scores derived from regression equations. Unlike the relatively simple previously mentioned parameters, the Novacode estimate for LVM uses a combination of electrocardiographic factors of numerous amplitudes of different leads and QRS duration. The Novacode also takes into account the positive or negative deflection properties of the T wave, where abnormalities in the T wave form deflection in certain leads may correspond to LVH [34]. The calculation of the Novacode estimate for LVH considers race and gender of the

patient, therefore increasing the sensitivity of the parameter [3]. Covic et al. published that the Novacode estimate performed with a sensitivity for LVH of 61% [11]. Although, this increase in sensitivity was at a cost of specificity which was reported as 48.9% [11]. QRS duration is measured in milliseconds and amplitudes are measured in microvolts and are absolute. The following equations are used to calculate the Novacode estimate for LVM index (LVMI). The (*) symbol denotes that the amplitude may be the S or Q or QS waveform, depending on whichever has the maximum value.

Men:

$$LVMI (g/m^2) = -36.4 + 0.01 \times R(V_5) + 0.02 \times S(V_1) + 0.028 \times S(III)^* \\ + 0.182 \times Tneg(V_6) - 0.148 \times Tpos(aVR) + 1.049 \times QRS \text{ duration}.$$

White women:

$$LVMI (g/m^2) = 88.5 + 0.018 \times R(V_5) + 0.112 \times Q(I)^* + 0.053 \times S(V_5)^* \\ + 0.108 \times Tpos(V_1) - 0.094 \times Tpos(V_6) + 0.170 \times Tneg(aVF).$$

Black women:

$$LVMI (g/m^2) = 88.5 + 0.018 \times R(V_5) + 0.112 \times Q(I)^* + 0.053 \times S(V_5)^* \\ + 0.108 \times Tpos(V_1) - 0.094 \times Tpos(V_6) + 0.170 \times Tneg(aVF). \quad (2.4)$$

Thresholds for presence of LVH

Table 2.1 displays the threshold values used for determining presence of LVH in ECG measurements of patients. These thresholds convert the continuous criteria variables to dichotomized measurements for clinical practice.

Criteria	Threshold for LVH presence
Sokolow-Lyon Voltage [39]	$SLV \geq 3.5 \text{ mV}$
Cornell Voltage [9]	$CV \geq \begin{cases} 2.8 \text{ mV, men} \\ 2.0 \text{ mV, women} \end{cases}$
Cornell Voltage Product [29]	$CVP \geq 244.0 \text{ mV} \cdot \text{ms}$
Novacode LVMI [34]	$LVMI > \begin{cases} 131 \text{ gm/m}^2, \text{ men} \\ 115 \text{ gm/m}^2, \text{ white women} \end{cases}$

Table 2.1: Thresholds for LVH detection from ECG-based Criteria

Methods

This chapter provides an overview of the methodology used during the scope of this thesis. The procedure for collecting data, extracting the desired time frame of the ECG signal, filtering the raw signal, evaluating the required features of the ECG, using these features to calculate the pre-chosen parameters, and finally applying statistical methods are described. Figure 3.1 illustrates a flowchart of the data analysis process.

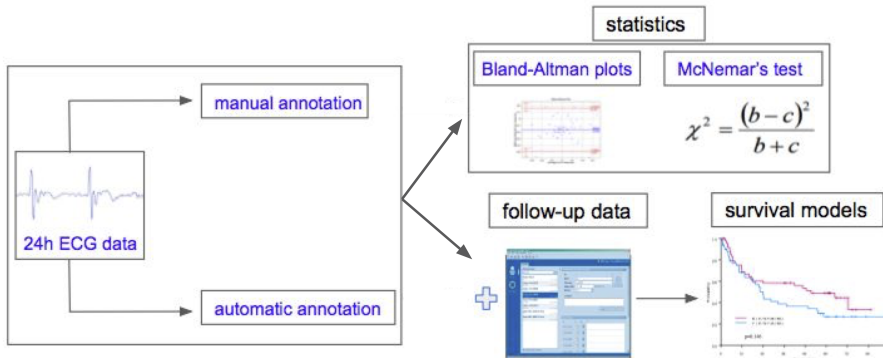


Figure 3.1: Data analysis flowchart.

3.1 Clinical Data Collection

The ECG data used in this thesis was obtained from a study in which patients were monitored using a Lifecard CF digital Holter recorder (Delmar Reynolds/Spacelabs, Nuremberg, Germany) for 24 hours [36]. The ECG was recorded with a paper speed of 25 mm per second and a voltage of 10 mm per mV. Along with each patient's ECG record, patient specifics such as an identification number and sex and age were also noted.

3.2 Manual Annotation

The manual annotations used for comparison in this thesis were taken by an experienced physician, blinded to each patient's condition. The manual measurements consisted of amplitudes and intervals measured in millimeters for a single ECG heartbeat from each patient. Using the manually measured amplitudes and durations, the four described ECG-based LVH criteria in section 2.2.4 were calculated for each patient. The identification of LVH for each criteria was also recorded using the thresholds described in table 2.1. A reference time and the specific heartbeat number relative to the reference time was given. In order to use the manually derived data for comparison, a short script was written to import the data and convert the measurements into standard units for potentials (mV) and intervals (ms).

3.3 Automatic Annotation

3.3.1 ECG Extraction using the AIT ECGsolver

Using the reference time and the number of the relative heartbeat for each patient ID noted by the physician, an extraction algorithm was used to retrieve a fragment of the ECG signals and annotations from the AIT ECGsolver [5, 18]. An ECG signal frame was extracted in order to use neighboring beats for baseline interpolation. The length of the extraction was based on the the relative heartbeat. If the examined heartbeat was four or less from the reference time, then a six second frame was taken. Otherwise, a ten second frame was extracted. This threshold was chosen based on an average adult resting heart rate of 60 beats per minute [4], ensuring that the beat in question was included in the extraction and minimizing processing time.

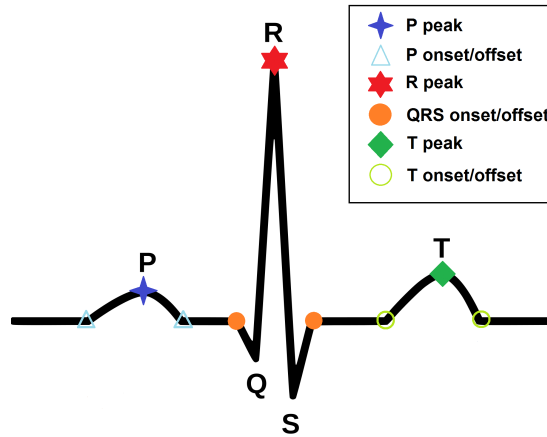


Figure 3.2: A heartbeat complex with the points given in the annotation file.

The resulting extraction included a frame of the raw 12-lead ECG, an annotation file, and a description file for each patient. The annotation file included the indexes of detected peaks and points of the ECG for each heartbeat complex that has been extracted. Figure 3.2 displays one heartbeat complex and the labeled index points that are given in the annotations files. The description file included the sample frequency, the patient's specifics, and the lead names.

3.3.2 Signal Pre-Processing

To further analyze the raw ECG data, unwanted artifacts and noise must be removed. These typically includes muscle and respiratory activity, movement, and powerline interference [40] which interfere between a minimum of 0.12 Hz (respiratory) [2] and a maximum of 50/60 Hz (electrical frequency) [24]. Muscle noise can cause severe problems as low-amplitude waveforms can be obstructed. This is not associated with narrow band filtering, but more complex since the spectral content of noise overlaps with the PQRST complex [40]. But due to its repetitive nature, the ECG signal can be filtered using ensemble averaging [21]. To filter low and high frequency signals together, a bandpass filter is applied between 0.05 and 40 Hz to minimize distortion to the ST-segment.

Other interferences include baseline wandering interferences typically caused by impedance mismatch, motion artifacts or respiratory breathing during the ECG recording [38]. To apply polynomial fitting, the QRS complexes and PQ intervals need to be localized. To eliminate the nonlinear trend, a polynomial of order six is fit to the signal. The resulting interpolation is then subtracted from the original signal to detrend baseline shift.

3.3.3 Automatic Amplitude Evaluation

Using the filtered ECG data and extracted annotation file from the AIT ECGsolver for each patient, an algorithm was developed to automatically determine the required amplitudes specific to the following ECG-based parameters: Sokolow-Lyon Voltage, Cornell Voltage, Cornell Voltage Product, and Novacode estimate for LVMI (see section 2.2.4).

The peak amplitude was defined as the vertical difference between the peak ECG value and the isoelectric line, or baseline, of the ECG signal. Figure 3.3 exemplifies the method of determining the R amplitude.

Isoelectric Line Evaluation

Figure 3.4 illustrates an example of a QRS complex that the physician has chosen to investigate manually. The isoelectric baseline of the complex is not at the zero line, so any amplitude automatically determined by taking the displacement of the peak of the R complex from the zero line would be inaccurate. In order to correctly determine an automatic amplitude of the R deflection, a linear baseline was fitted between the QRS onset of the investigated complex and the QRS onset of the subsequent complex. The interpolated value of this fitted line at the index at which the R peak occurs gave the

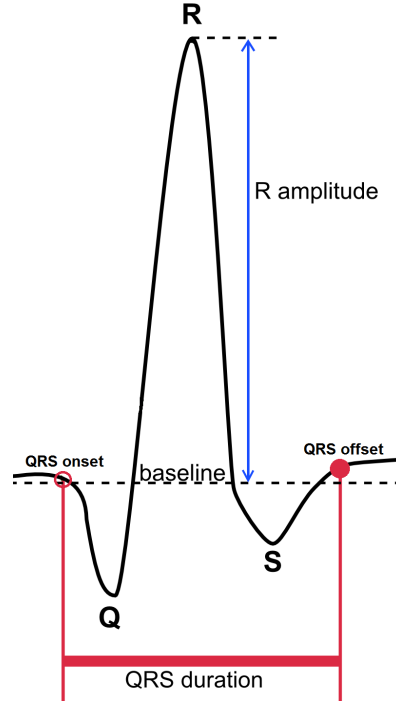


Figure 3.3: A heartbeat complex with the baseline displayed and the method of measuring R amplitude and QRS interval (modified from [22]).

correct reference value to determine the displacement of the R deflection and therefore the R amplitude.

The interpolation of the baseline value at the index of the deflection is calculated by:

$$y^* = y_0 + (x^* - x_0) \frac{y_1 - y_0}{x_1 - x_0}, \quad (3.1)$$

where y^* is equal to the interpolated value at the index, x^* , of the desired peak, (x_0, y_0) correspond to the QRS onset index of the beat in question and the corresponding ECG value at that index, and finally (x_1, y_1) are the index and corresponding ECG value of the QRS onset of the subsequent beat.

In the event that the subsequent complex was unavailable, because the extraction algorithm failed to find a value, an algorithm was used which successively iterates away from the investigated QRS onset point until the closest detected QRS onset for linear fitting was found. If no other QRS onset was available in the extracted frame, then the T wave offset of the examined heartbeat was used. In the case that the T wave was also not detected, the QRS offset of the same beat was used. For any investigated beat, a detected QRS onset is always accompanied by its complementary offset, thus the algorithm ensures that there would always be a point for baseline interpolation.

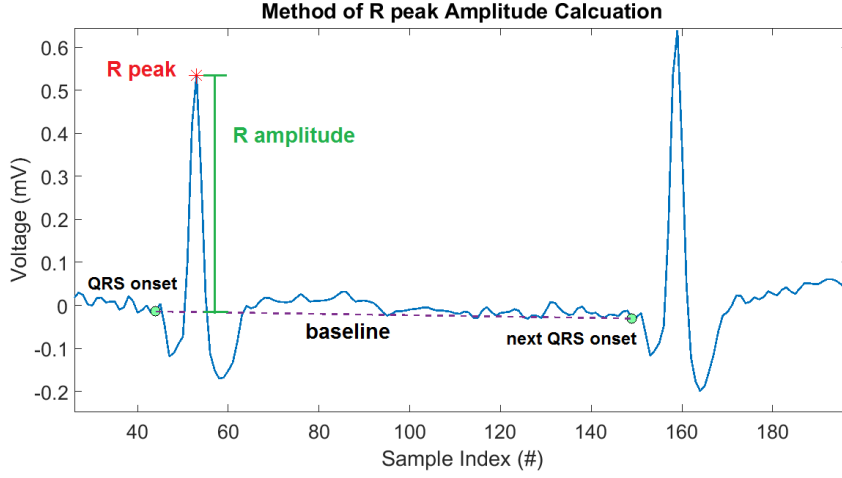


Figure 3.4: An example ECG heartbeat and the method of amplitude calculation.

Once a second point was found for linear fitting with the QRS onset of the examined beat, the position, or index, of second value determined if an interpolation or an extrapolation was carried out. In the case that the index of the peak amplitude to be determined, x^* , should lay between the examined QRS onset, x_0 , and the second point, x_1 such that $x_0 < x^* < x_1$, an interpolation and equation 3.1 was used.

If the second value used for linear fitting had a smaller index than the examined beat, thus was from a complex that occurs before the one analyzed, an extrapolation of the baseline value at the desired peak was performed instead. In this case, the extrapolated baseline value for peak amplitude calculation was given by:

$$y^* = y_{-1} + (y_0 - y_{-1}) \frac{x^* - x_{-1}}{x_0 - x_{-1}}, \quad (3.2)$$

where y^* is the extrapolated value which is used as a reference for the baseline, x^* is the index at which the peak deflection occurs, (x_0, y_0) are the index and ECG value at the QRS onset of the analyzed beat, and (x_{-1}, y_{-1}) are the index and ECG value of the second point used for baseline fitting that occurs before the analyzed beat. In this case, $x_{-1} < x_0 < x^*$.

Tukey's Test for Baseline Value Plausibility

Before the interpolated or extrapolated baseline value was used for amplitude calculation, a plausibility check was applied in order to determine if the value was reasonable. This was done by an algorithm that uses Tukey's test for outliers [41]. The input of the test was the vector of all ECG values within the extracted frame and the baseline value that had been determined by the method described above. Tukey's test was carried out by first determining the lower Q_1 and upper Q_3 quantiles of all ECG values within the

extracted frame. Then the interquantile range (IQR) was evaluated by, $IQR = Q_3 - Q_1$. An interval with an upper and lower threshold was defined by [41]:

$$[Q_1 - k \times IQR, Q_3 + k \times IQR]. \quad (3.3)$$

Where the constant k is given the value $k = 1.5$, as proposed by Tukey himself [41]. If the proposed baseline value was not within this range, it was determined to be an outlier. In this case, the mean of all the ECG data within the extracted frame was taken as the reference value for amplitude calculation instead.

Calculation of ECG Amplitudes

Using the reference value, y^* , that was determined by the method described above and the ECG value, y , corresponding to the index of the peak, x^* , the amplitude of the desired ECG deflection was calculated as:

$$\text{Amplitude (mV)} = \text{abs}(y - y^*), \quad (3.4)$$

where $\text{abs}()$ denotes the absolute value.

3.3.4 Automatic QRS Duration Evaluation

As shown in figure 3.3, the QRS onset and QRS offset were required for computing a QRS duration. Multiple detectors, including the AIT ECGsolver, were used in this thesis for detecting QRS offset and onset indices for QRS duration calculation.

AIT ECGsolver QRS detection

Although the AIT ECGsolver was able to detect a multitude of QRS onsets and offsets, there was a slight inaccuracy of this detection so far. During the course of this thesis an adaptation to the AIT ECGsolver was implemented by adding one index unit to each QRS onset value and subtracting two index units from the QRS offset. Nevertheless, the updated AIT ECGsolver extracted annotations for the QRS onset and offset indices were found to still be overestimating the QRS interval in comparison to the manually measured intervals (see Results 4.3.1). Due to this overestimation, additional QRS detectors were used for computing QRS duration. Two open-source QRS detectors were chosen, the **sqrs** and **gqrs** detectors [17].

WaveForm DataBase QRS detectors

As a part of the open-source PhysioNet project available online, the WaveForm DataBase (WFDB) software package [17] offers a MATLAB toolbox that includes various algorithms for ECG signal analysis.

The **sqrs** detector aims to locate QRS complexes in ECG signals. For every detected

slope, a threshold crossing of opposite sign in a specified time window is looked up [16]. If a sufficient interval has elapsed after the last threshold crossing, a QRS complex has been found and the annotation is recorded [16]. For single slope detections, a baseline shift/noise is assumed and the output is discarded.

The **gqrs** detector is a highly sensitive QRS detector, which is most effective for regular rhythms, but fails to predict patterns that are highly irregular, e.g. atrial fibrillation [30]. The detection works on the basis of R-peak intervals [30].

Calculation of QRS Duration

The QRS duration of a QRS complex was defined as the time between the QRS onset and QRS offset, as depicted in figure 3.3. Since the sample frequency of the signal, f_s (Hz), and the indices of the QRS onset and QRS offset were known, duration time, given in the unit of ms, was calculated by:

$$\text{QRS duration (ms)} = \frac{\text{idx}_{\text{QRS offset}} - \text{idx}_{\text{QRS onset}}}{f_s} \times 1000. \quad (3.5)$$

When calculating the criteria, the amplitudes used in the equations refer to specific leads. For example, in the case of the CV criteria in equation (2.2), $R(aVL)$ refers to the R amplitude in the aVL lead in the examined beat of the patient. In the case of the QRS duration, the equation does not specify a lead. The method used to calculate a single QRS duration in a 12-lead ECG was a variation of the one proposed by Bauer et al. [6].

According to Bauer et al., one duration is calculated for the limb and augmented limb leads (LALL) and a second for the precordial leads (PL). The LALL refer to the first through sixth leads (I - aVF), and the PL refer to the seventh through twelfth leads (V1 - V6).

Instead of taking the maximum of the two calculated durations, as in Bauer et al., the median was used to derive the final QRS duration for the single beat in order to avoid the influence of outliers. This is especially important for the automatic algorithm, since one large outlier would falsify the interval duration. By manually measuring QRS intervals, as in [6], an experienced cardiologist would disregard these outliers. The set of leads contained within LAL and PL were defined as:

$$\begin{aligned} \text{LAL} &= \{I, II, III, aVR, aVL, aVF\} \\ \text{PL} &= \{V_i\}_{i=1-6}. \end{aligned} \quad (3.6)$$

The median of the QRS offset indices and the median of the QRS indices were taken for each set. The difference between the median QRS offset and the median QRS onset was calculated as QRS interval of the set.

$$\begin{aligned} \text{QRS interval}_{LALL} &= \text{median}(\text{idx}_{\text{QRS offset}, i})_{i \in LALL} - \text{median}(\text{idx}_{\text{QRS onset}, i})_{i \in LALL}. \\ \text{QRS interval}_{PL} &= \text{median}(\text{idx}_{\text{QRS offset}, i})_{i \in PL} - \text{median}(\text{idx}_{\text{QRS onset}, i})_{i \in PL}. \end{aligned} \quad (3.7)$$

The median of the two was taken as the QRS duration of the heartbeat.

$$\text{QRS duration (ms)} = \frac{\text{median}(\text{QRS interval}_{LALL}, \text{QRS interval}_{PL})}{f_s} \times 1000. \quad (3.8)$$

3.4 Parameters

Using the automatically determined amplitudes and durations for each patient, four different ECG-based parameters were calculated. These same four parameters were calculated by the physician using the same exact equations described in section 2.2.4. Also included in the aforementioned section is a table of thresholds for each criteria that determine the incidence of LVH based on the criteria's equation. The automatically calculated parameter results, based on automatically determined voltages and durations, were compared to the manually determined parameters. Also the outcome of LVH incidence for each parameter was compared to the manually determined LVH incidence. The results were analyzed using the statistical methods described in the following section.

3.5 Statistical Methods

Bland-Altman Plots

A Bland-Altman plot, or difference plot, is common way to present the agreement between two measurement methods. To visualize possible trends in differences and show agreement between manually derived parameters M and the corresponding automatically derived parameters A , the differences between the two measurements were plotted against their respective average. The resulting graph, G , as shown in figure 3.5, features three distinct lines on the plot, the arithmetic mean μ , and the two limits of agreement $\mu \pm 1.96 \sigma$ [7].

$$G(x, y) = \left(\frac{M + A}{2}, M - A \right). \quad (3.9)$$

McNemar's Test

Possible matches between the paired set of binary labels for automatic and manual LVH incidence was calculated by evaluating the contingency table of the two sets. Given a 2×2 contingency table for N subjects as it can be seen in table 3.1, a chi-squared distribution was obtained by [28]:

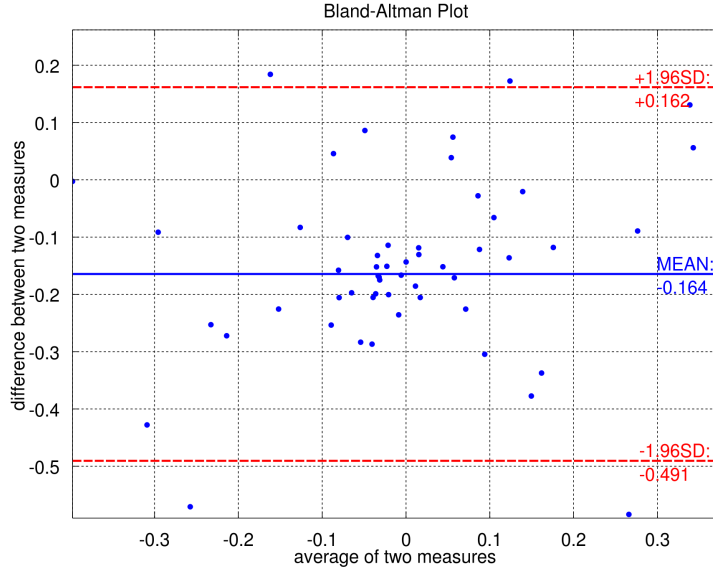


Figure 3.5: A Bland-Altman plot with sample data (dots), mean (solid line), and limits of agreement (dashed lines) [44].

$$\chi^2 = \frac{(b - c)^2}{b + c}. \quad (3.10)$$

The null hypothesis was rejected for a significance of 0.05 of the χ^2 result, meaning that the marginal proportions are significantly differing.

Table 3.1: An example McNemar's 2×2 contingency table.

	Test 2 positive	Test 2 negative	Row total
Test 1 positive	a	b	a + b
Test 1 negative	c	d	c + d
Column total	a + c	b + d	N

For small values in the discordant cells, b and c , such that $(b + c) < 25$, χ^2 is not well approximated by the chi-squared distribution [47]. In these cases a continuity corrected version of McNemar's test, also known as McNemar's test with Yates correction, is used instead [14].

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c}. \quad (3.11)$$

Survival Models:

All-cause mortality was used as endpoint for survival analysis. Kaplan-Meier curves were used for graphically representing estimated cumulative survival. Figure 3.6 displays a sample Kaplan-Meier survival plot for patients with two different conditions. Survival time started at baseline measurement and patients were censored at time of transplantation or at time of loss to follow-up. Kaplan-Meier curves (absent vs. present) were compared using a logrank test. Dichotomization for LVH identification (absent/present) was done using thresholds from literature (see section 2.2.4). Cox proportional hazard models were calculated univariate and adjusted consequently using the following covariates: sex, age, diabetes mellitus and arterial hypertension (model 1), and sex, age, adapted version of the Charlson Comorbidity Index (CCI) for ESRD [1], Albumin, Ca-Pho-Product and dialysis vintage (model 2).

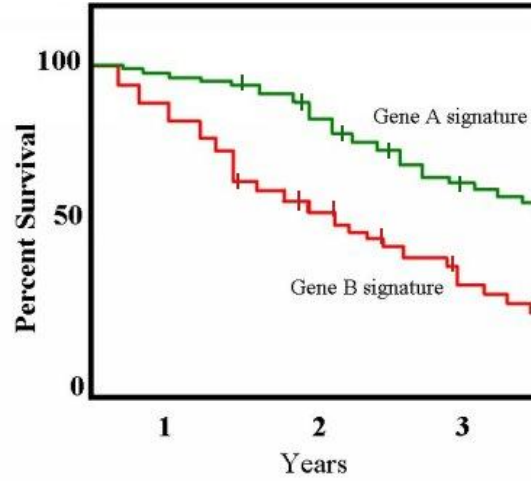
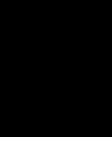


Figure 3.6: An example Kaplan-Meier plot for patients with two different conditions [42].

A comparison of predictive power between manually annotated data, automatically classified data, and ground truth follow-up data can be obtained by building a proportional hazards model. The underlying concept of the survival model is the relation between the time that passed before an event occurs and one or more covariates that are associated with the given time quantity. Given the observed time Y_i for a subject i , the realized values of covariates $X_i = X_{i1}, \dots, X_{ip}$, and the β parameters of the model, the hazard function can be written as [12]:

$$\lambda(t|X_i) = \lambda_0(t) \exp(\beta_1 X_{i1} + \dots + \beta_p X_{ip}) = \lambda_0(t) \exp(X_i \beta). \quad (3.12)$$

Covariates are multiplicatively related to the hazard, e.g in a simple case a treatment with certain drugs might halve a subject's hazard at any time t , but not vary the baseline hazard. This does not double the subject's life time.



Results

The chapter presents the results collected from the comparison of the automatically and manually derived ECG-based LVH parameters. The results include a table of baseline characteristics of the investigated cohort, a comparison of the manually and automatically derived amplitudes, a comparison of the different methods of QRS duration calculation, and the statistical results of the manual and automatic method comparison for each ECG-based LVH parameter. Each comparison is further investigated through the use of Bland-Altman plots and McNemar 2×2 contingency tables. Finally, the results of the survival analysis through Cox regression models and the Kaplan-Meier method are displayed.

4.1 Baseline Data

In the ISAR study, a total of 540 prevalent dialysis patients were included between 2010 and 2013 [36]. Thereof, 381 patients had a 24h ECG [18] and a subset of 335 ECGs were manually annotated by an experienced physician blinded to the patient's clinical status. These were included for the comparison of manual and automatic annotations. The baseline characteristics of this population are shown in table 4.1.

4.2 Amplitude Comparison

The amplitudes measured by the physician to calculate the manual parameters are included in the annotations. To exemplify the agreement with the automatically evaluated amplitudes, those used to calculate the CV Criteria are compared with the manually derived ones. These are the R wave in lead aVL and S wave in lead V3. The results of the comparisons are shown in the following tables and Bland-Altman plots.

Table 4.1: Baseline characteristics of the investigated cohort of 335 subjects.

Cohort number	335
Age (years)	68.2 (54 - 77)
Sex - male (% male)	226 (67%)
Body weight (kg)	74 (65.3 - 85.5)
Height (m)	1.71 \pm 0.086
Body mass index (kg/m^2)	25 (22.5 - 28.5)
Follow-up (days)	1291 (727 - 1702)
All-cause mortality, n (%)	122 (36%)
adapted Charlson Cormorbidity Index ()	3 (1 - 6)
Dialysis vintage (mo)	42.3 (21.8 - 75.3)
UFV (ml)	1.69 \pm 1.15
Kt/V ()	1.45 (1.22 - 1.64)
Ca-Pho-Product ($mmol^2/l^2$)	3.77 (3.07 - 4.6)
Presence of diabetes, n (%)	127 (38%)
Presence of hypertension, n (%)	313 (93%)
Hemoglobin (g/dl)	11.8 \pm 1.16
Total protein (g/dl)	6.65 \pm 0.538
Serum albumin (g/dl)	4 \pm 0.413
Total cholesterol (mg/dl)	180 \pm 43.9
HDL cholesterol (mg/dl)	43 (37 - 53)
LDL cholesterol (mg/dl)	113 \pm 35.8
Triglycerides (mg/dl)	157 (112 - 218)
Use of antihypertensive drugs, n (%)	301 (90%)
ACE inhibitors, n (%)	123 (37%)
ARBs, n (%)	74 (22%)
CCBs, n (%)	124 (37%)
Vasodilators, n (%)	47 (14%)
Beta-blockers, n (%)	217 (65%)
Diuretics, n (%)	203 (61%)
Other hypertensive drugs, n (%)	45 (13%)
Statins, n (%)	115 (34%)

UFV = Ultrafiltration Volume; Kt/V = K - dialyzer clearance of urea, t - dialysis time, V - volume of distribution of urea;

Ca-Pho-Product = Calcium Phosphate Product; HDL = High-density lipoprotein; LDL = Low-density lipoprotein; ACE = Angiotensin converting enzyme; ARBs = Angiotensin II receptor blockers; CCBs = Calcium channel blockers.

4.2.1 RaVL Amplitude Comparison

A total of 333 RaVL amplitudes were automatically evaluated by the method of amplitude calculation described in section 3.3.3. A comparison between the detectable RaVL amplitudes with the manual annotations is displayed in table 4.2. The statistical comparison between each method gave a mean and median that show good agreement. The Bland-Altman plot in figure 4.1 displays an average difference of -0.0036 mV between manual and automatic measurements.

Statistics	RaVL amplitude	
	Manual	Automatic
Amplitude measurements, n	333	333
Mean \pm SD, mV	0.44 ± 0.42	0.45 ± 0.43
Median (IQR), mV	0.35 (0.11 - 0.61)	0.35 (0.11 - 0.60)
Bland-Altman Plot Statistics	Manual - Automatic	
Difference mean (\bar{d}), mV	-0.0036	
Difference SD (s), mV	0.0835	
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	[-0.17, 0.16]	
Regression line	$y = -0.024x + 0.0071$	

Table 4.2: Statistical comparison of the manually measured amplitudes versus the automatically derived amplitudes for the R wave in Lead aVL.

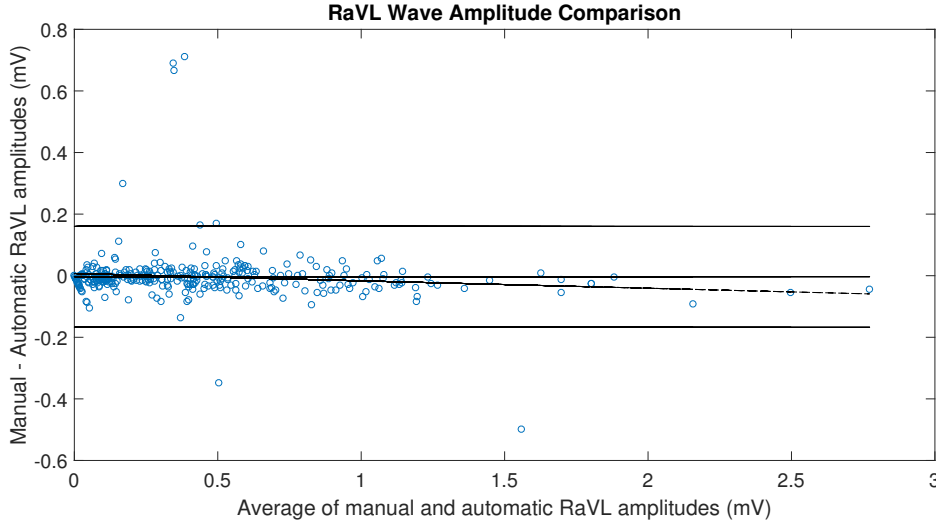


Figure 4.1: Comparison of the manually and automatically derived amplitudes for the RaVL wave.

4.2.2 SV3 Amplitude Comparison

A total of 334 SV3 amplitudes were automatically evaluated for comparison. Thus the S wave in lead V3 was undetected for one patient in the cohort and a parameter could be not automatically evaluated. The statistical comparison in table 4.3 shows that the mean and the median in agreement. The Bland-Altman plot for the SV3 comparison in figure 4.2 shows that the mean difference is -0.0098 mV and the linear fit has a slope of -0.018.

Statistics	SV3 amplitude	
	Manual	Automatic
Amplitude measurements, n	334	334
Mean \pm SD, mV	0.77 ± 0.63	0.78 ± 0.64
Median (IQR), mV	0.66 (0.33 - 1.08)	0.66 (0.34 - 1.09)
Bland-Altman Plot Statistics	Manual - Automatic	
Difference mean (\bar{d}), mV	-0.0098	
Difference SD (s), mV	0.066	
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	[-0.14, 0.12]	
Regression line	$y = -0.018x + 0.0044$	

Table 4.3: Statistical comparison of the manually measured amplitudes versus the automatically derived amplitudes for the S wave in Lead V3.

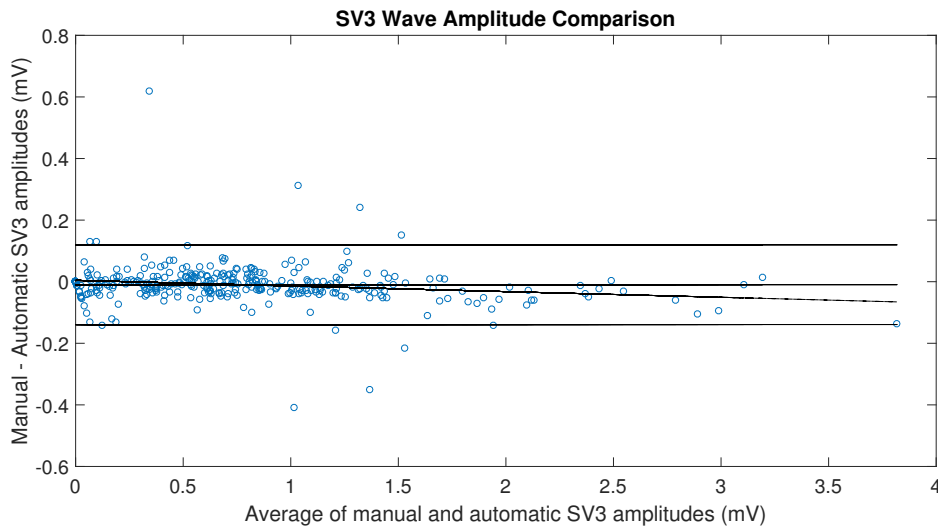


Figure 4.2: Comparison of the manually and automatically derived amplitudes for the SV3 wave.

4.3 QRS Duration Comparison

Multiple methods were used to determine the most accurate method for automatic QRS duration calculation. The methods investigated were the AIT ECGsolver [5], the MIT GQRS annotator [30], and the MIT SQRS annotator [16]. Each method was compared with the manually derived durations through statistical analysis and Bland-Altman plots. The results were used to determine the best method for further calculation of the parameters in which QRS duration is a variable.

4.3.1 Manual and AIT ECGsolver Comparison

For the cohort data, an automatically evaluated duration could be determined using the AIT ECGsolver for all subjects. The statistical comparison in table 4.4 shows that there are large differences between durations derived from the AIT ECGsolver and the manual annotations done by the physician. The mean of the manual durations is 97 ± 24 ms while the automatically computed durations yield a mean of 118 ± 30 ms. The median values with interquartile range for the manual durations is 92 (84 - 104) and 113 (96 - 135) for the automatic durations.

The Bland-Altman plot in figure 4.3 displays the agreement between the measurements and the values within the figure are found in table 4.4. On average, the manual durations are 20.44 ms (± 21.04) less than those computed automatically. The lower and upper limits of agreement are found at -61.68 and 20.81, respectively. The linear fit in figure 4.3 has a negative slope of -0.26 and a y-intercept of 7.49.

Statistics	<u>Manual vs AIT ECGsolver</u>	
	Manual	AIT ECGsolver
Total durations, n	335	335
Mean \pm SD, ms	97 ± 24	118 ± 30
Median (IQR), ms	92 (84 - 104)	113 (96 - 135)
Bland-Altman Plot Statistics	Manual - Automatic	
Difference mean (\bar{d}), ms	-20.44	
Difference SD (s), ms	21.04	
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	[-61.68, 20.81]	
Regression line	$y = -0.26x + 7.49$	

Table 4.4: Statistical comparison of the manually measured QRS durations versus the automatically derived durations using the AIT ECGsolver.

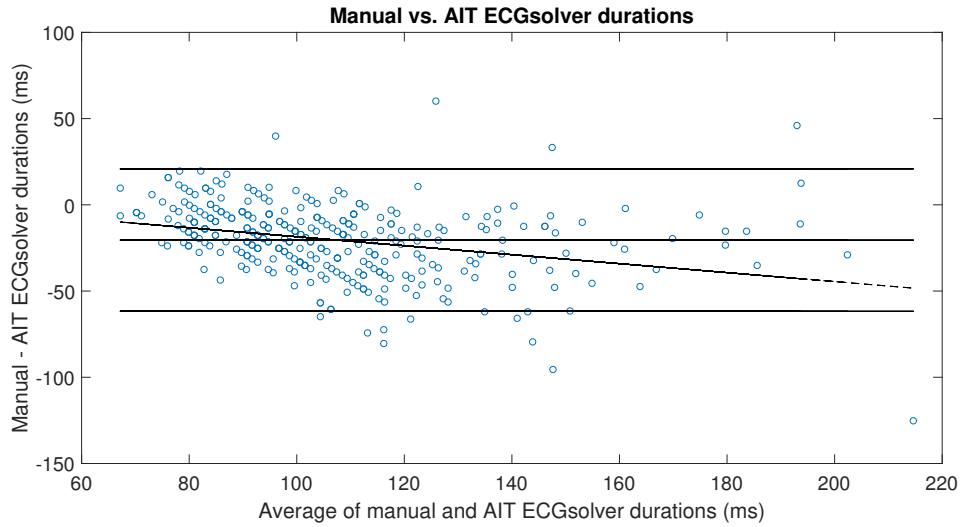


Figure 4.3: Bland-Altman plot comparing the manually and automatically derived QRS durations.

4.3.2 MIT GQRS and SQRS Annotators Comparison

A statistical comparison of each detector method with the manual durations is found in table 4.5. The limitation of the MIT detectors was the number of duration samples that can be derived from the data. While the AIT ECGsolver was able to detect QRS features for all subjects of the cohort, the SQRS and GQRS detectors were limited to 256 and 213 samples, respectively.

Comparing table 4.5 with table 4.4, it can be seen that both MIT detectors give rise to durations that show better agreement with the manual durations than the those derived from the AIT ECGsolver. The SQRS detector gives an average duration of 105 ± 20.7 ms in comparison to 94.5 ± 21.4 ms for the manual durations. The average duration for the GQRS detector is 108 ± 22.3 ms and the average manual duration for the same sample group is 99.5 ± 24.7 ms. Medians with interquartile range are also reported in table 4.5. A comparison of the durations derived for each detector for the sample group of 153

Statistics	Manual vs SQRS Detector		Manual vs GQRS Detector	
	Manual	SQRS	Manual	GQRS
Total durations, n	256	256	213	213
Mean \pm SD, ms	94.5 ± 21.4	105 ± 20.7	99.5 ± 24.7	108 ± 22.3
Median (IQR), ms	88 (80 - 102)	102 (91.8 - 112)	92 (84 - 108)	102 (91.8 - 117)

Table 4.5: Statistical comparison of the manually measured durations versus the MIT SQRS detector durations and MIT GQRS detector durations respectively.

subjects in which they intersect is shown in table 4.6. The medians with interquartile range are exactly the same for both the SQRS and GQRS at 102 (93.8 - 113) ms. The means vary slightly with the SQRS and GQRS detectors yielding means of 105 ms and 106 ms, respectively. The histogram in figure 4.4 displays that most durations derived from each method show no difference. Figure 4.4 shows a regression line with a minimal negative slope of -0.0017 and a mean difference of -0.077 ms between SQRS and GQRS.

Statistics	SQRS vs GQRS MIT Detectors	
	SQRS	GQRS
Total durations, n	153	153
Mean \pm SD, ms	105 \pm 20	106 \pm 20
Median (IQR), ms	102 (93.8 - 113)	102 (93.8 - 113)
Bland-Altman Plot Statistics	SQRS - GQRS	
Difference mean (\bar{d}), ms	-0.077	
Difference SD (s), ms	1.25	
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	[-2.52, 2.36]	
Regression line	$y = -0.0017x + 0.11$	

Table 4.6: Statistical comparison of SQRS and GQRS MIT detectors.

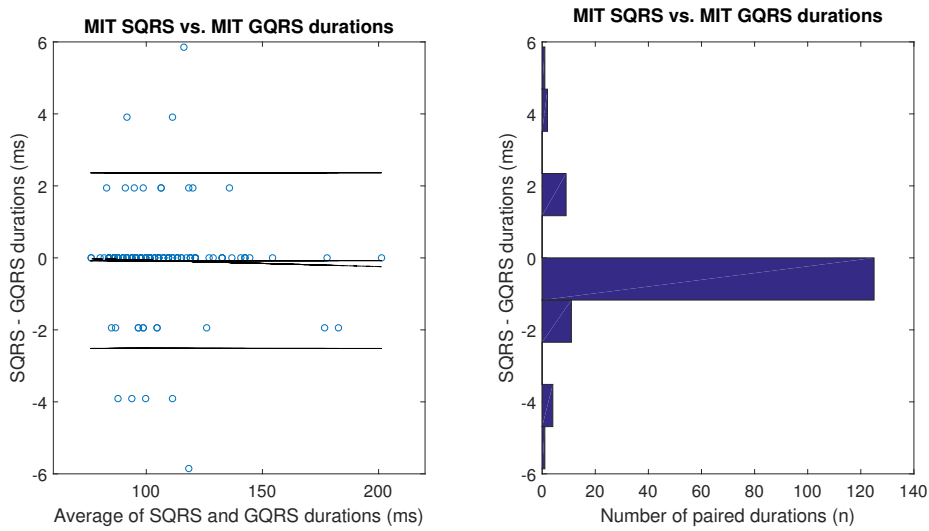


Figure 4.4: Bland-Altman plot and histogram comparing the SQRS and GQRS detectors.

4.3.3 Combined MIT Annotators and Manual QRS Duration Comparison

Since the results in section 4.3.2 show durations derived from both MIT detectors are nearly identical, it is determined that the results between the detectors can be used interchangeably. In order to get the largest sample number of durations from the cohort, a combination of detector durations is used. This yields a sample number of 316 durations

Statistics	<u>Manual vs comb. MIT Detectors</u>	
	Manual	comb. MIT Detectors
Total durations, n	316	316
Mean \pm SD, ms	96.8 ± 23.5	107 ± 22.1
Median (IQR), ms	92 (84 - 104)	102 (91.8 - 113)
Bland-Altman Plot Statistics	Manual - Automatic	
Difference mean (\bar{d}), ms	-9.78	
Difference SD (s), ms	14.94	
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	[-39.07, 19.51]	
Regression line	$y = 0.067x - 16.63$	

Table 4.7: Statistical comparison of the manually derived QRS durations with the combined MIT detectors.

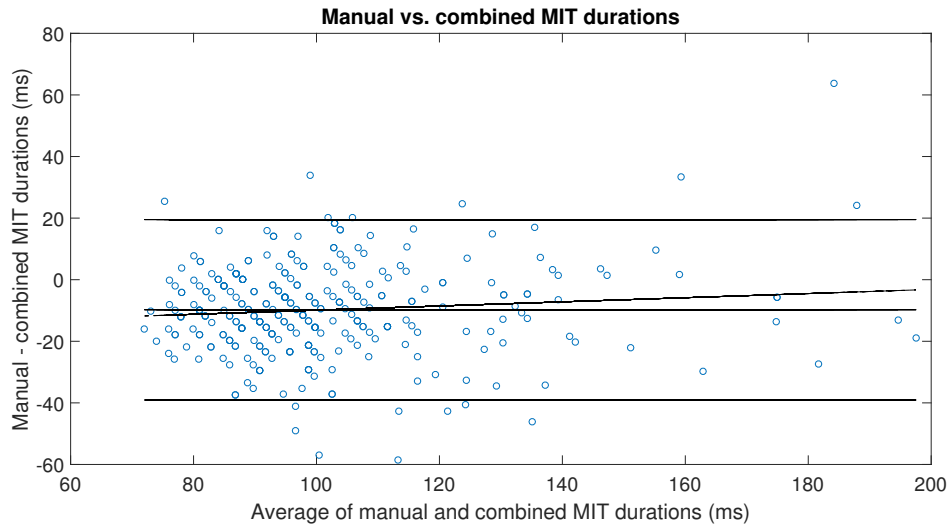


Figure 4.5: Bland-Altman plot comparing the manually derived durations with the combined MIT detectors result.

that is compared in table 4.7. The difference mean and standard deviation between the manual measurements and the combined MIT detector measurements is much lower than the one reported in table 4.4 at -9.78 ± 14.94 ms. The Bland-Altman plot in figure 4.5 shows a slightly positive regression line with a slope of -0.067 and a y-intercept at -16.63 and with limits of agreement from -39.07 to 19.51 ms.

4.4 Criteria Statistics

After careful verification of automatically derived amplitudes and QRS durations, the computed LVH parameters, as described in section 2.2.4, were analyzed. Table 4.9 compares the statistical results for each method of parameter calculation. The total number of patients signifies the amount of patients from the total cohort that were available for calculation of the parameter. Since several ECG features are needed to calculate each of the LVH parameters, the total number of patients varies for each criteria. The mean with standard deviation, median with interquartile range, and the number of positive patients for LVH with percentage is listed for each as well. The Novacode parameter is further displayed separately between men and women due to the different equations used to derive the parameter.

4.4.1 Sokolow-Lyon Voltage Criteria

Of the available cohort, it is possible to automatically determine the Sokolow-Lyon parameter for 321 patients. The manually derived mean voltage is 1.62 mV with a standard deviation of 0.91 mV and 1.64 mV (± 0.92) for automatic computation as shown in table 4.9.

Bland-Altman Plot & McNemar's Test Statistics	Manual - automatic SLV
Difference mean (\bar{d}), mV	-0.013
Difference SD (s), mV	0.11
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	$[-0.22, 0.19]$
Regression line	$y = -0.018x + 0.016$
McNemar's test p value	undefined

Table 4.8: Bland-Altman statistics for the manually and automatically derived SLV.

The median voltage is 1.42 mV for manual measurement and 1.45 mV for automatic computation. LVH prevalence is 13% for both methods. The Bland-Altman plot in figure 4.6 displays the agreement between the two methods and shows the fitted regression line between the boundaries of 2σ with a negative slope of -0.018 . From the plot it can be seen that the difference between the manual and automatic SLV value is larger than 0.7 mV for one subject. The majority of outliers are in the range of $[0.2, 0.6]$ mV.

Statistics	<u>Sokolow-Lyon Voltage</u>	
	Manual	Automatic
Total patients	321	321
Mean \pm SD, mV	1.62 \pm 0.91	1.64 \pm 0.92
Median (IQR), mV	1.42 (0.96 - 2.11)	1.45 (0.96 - 2.14)
LVH prevalence (%)	13 (4%)	13 (4%)
	<u>Cornell Voltage</u>	
	Manual	Automatic
Total patients	332	332
Mean \pm SD, mV	1.22 \pm 0.81	1.23 \pm 0.82
Median (IQR), mV	1.08 (0.66 - 1.58)	1.08 (0.68 - 1.62)
LVH prevalence (%)	26 (7.8%)	25 (7.5%)
	<u>Cornell Voltage Product</u>	
	Manual	Automatic
Total patients	313	313
Mean \pm SD, mV\cdotms	122 \pm 94.7	132 \pm 102
Median (IQR), mV\cdotms	95 (60.2 - 148)	105 (67.9 - 160)
LVH prevalence (%)	34 (10.9%)	41 (13.1%)
	<u>Novacode</u>	
	Manual	Automatic
Total patients	226	226
Mean \pm SD, gm/m²	101 \pm 42.8	106 \pm 43.6
Median (IQR), gm/m²	96.7 (74.1 - 117)	100 (82.1 - 123)
LVH prevalence (%)	41 (18.1%)	51 (22.6%)
	<u>Novacode - Men</u>	
	Manual	Automatic
Total patients	160	160
Mean \pm SD, gm/m²	106 \pm 42.9	115 \pm 42
Median (IQR), gm/m²	97.8 (76.8 - 124)	105 (90.7 - 130)
LVH prevalence (%)	30 (18.8%)	40 (25%)
	<u>Novacode - Women</u>	
	Manual	Automatic
Total patients	66	66
Mean \pm SD, gm/m²	88.7 \pm 39.9	84.8 \pm 39.9
Median (IQR), gm/m²	89.7 (66.7 - 108)	84.2 (62 - 104)
LVH prevalence (%)	11 (16.7%)	11 (16.7%)

Table 4.9: Statistics for all criteria calculated.

The Bland-Altman statistical results in table 4.8 show a negative difference mean of -0.013 ms and a difference standard deviation of 0.11 ms. The 2σ boundaries are in the range of $[-0.22, 0.19]$. The resulting p value from the McNemar test is undefined as there are no discordant values in table 4.10. The table also shows that the automatic and manual method each detect 13 positive and 308 negative subjects for LVH.

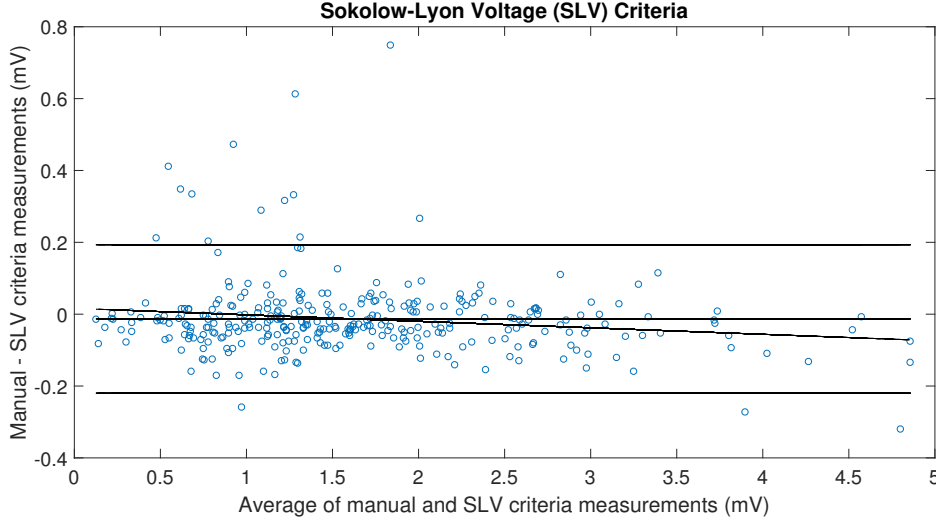


Figure 4.6: Bland-Altman plot comparing the manually and automatically derived SLV.

Table 4.10: McNemar's 2×2 contingency table for SLV results.

	Manual positive	Manual negative	Row totals
Automatic positive	13	0	13
Automatic negative	0	308	308
Column totals	13	308	321

4.4.2 Cornell Voltage Criteria

For the Cornell Voltage criteria, 332 parameters are able to be automatically determined and compared to the manual measurements. The manually derived mean voltage is 1.22 mV with a standard deviation of 0.805 mV and 1.23 mV (± 0.815) for automatic computation as it can be seen in table 4.9. The median voltage is 1.08 mV for manual measurement and automatic computation. The LVH prevalence is 7.8% for the manual method and 7.5% for automatic computation.

Figure 4.7 shows the relation between the two methods and the fitted regression line between the boundaries of 2σ with a negative slope of -0.013 . Most of the outliers are in the range of $[0.3, 0.6]$ mV with a maximum outlier of 0.75 mV. The Bland-Altman statistical results in table 4.11 show a negative difference mean of -0.013 ms and a differ-

4. RESULTS

ence standard deviation of 0.11 ms. The 2σ boundaries are in the range of $[-0.23, 0.20]$. The p value from McNemar's test is ~ 1.00 . In table 4.12, McNemar's 2×2 contingency table shows 26 positive LVH subjects for manual measurement and 25 for automatic computation.

Bland-Altman Plot & McNemar's Test Statistics	Manual - automatic CV
Difference mean (\bar{d}), mV	-0.013
Difference SD (s), mV	0.11
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	$[-0.23, 0.20]$
Regression line	$y = -0.013x + 0.002$
McNemar's test p value	~ 1.00

Table 4.11: Bland-Altman statistics for the manually and automatically derived CV values.

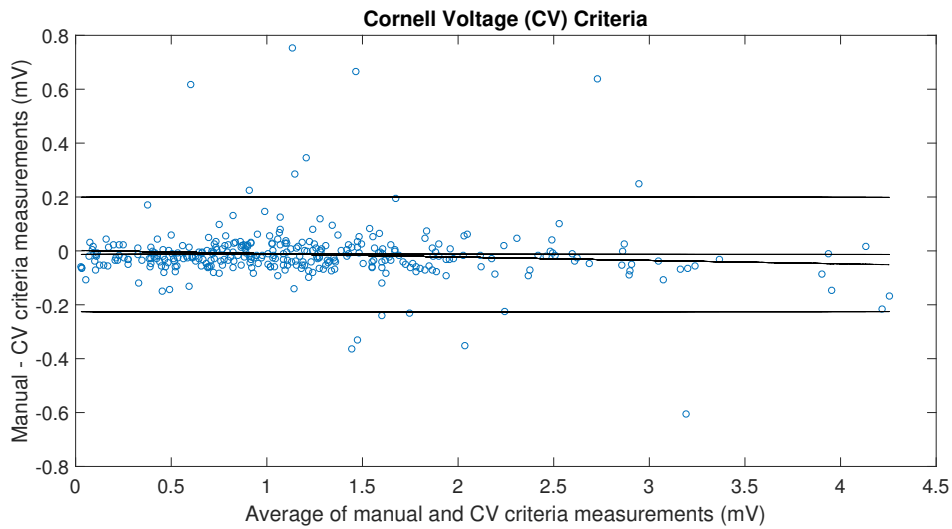


Figure 4.7: Bland-Altman plot comparing the manually and automatically derived CV.

Table 4.12: McNemar's 2×2 contingency table for CV.

	Manual positive	Manual negative	Row totals
Automatic positive	25	0	25
Automatic negative	1	306	307
Column totals	26	306	332

4.4.3 Cornell Voltage Product Criteria

The Cornell Voltage Product criteria is an adaptation of the CV Criteria. The criteria are essentially the same except for a QRS duration factor that is multiplied to obtain the CVP. However, the lack of detectable QRS onset and QRS offset features for several ECG signals severely limits the number of total patients that are comparable. Therefore, of the 324 total patients used to compare the CV criteria, only 313 are used for the comparison of the CVP criteria. Table 4.9 shows that the QRS duration factor greatly changed the agreement between the two measurement methods. LVH prevalence for the sample size is detected as 34 patients (10.9%) for the manual measurements and 41 patients (13.1%) for automatic measurements.

Bland-Altman Plot & McNemar's Test Statistics	Manual - automatic CVP
Difference mean (\bar{d}), mV·ms	-10.76
Difference SD (s), mV·ms	23.38
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	[-56.58, 35.06]
Regression line	$y = -0.075x - 1.24$
McNemar's test p value	0.07

Table 4.13: Bland-Altman statistics for the manually and automatically derived CVP.

The manually derived mean voltage product is 122 mV ms with a standard deviation of 94.7 mV·ms and 132 mV·ms (± 102) for automatic computation as it can be seen in table 4.9. The table also shows that the median with interquartile range is 95 (60.2 - 148) and 105 (67.9 - 160) for manual and automatic measurements, respectively. The regression fit has a negative slope of -0.075 with an intercept of -1.24 as shown in table 4.13. The statistical results of the Bland-Altman plot are -10.76 ms for the difference mean and 23.38 ms for the difference standard deviation. Table 4.14 shows 34 positive LVH subjects for the manual method and 41 for the automatic method. The result of the Yates corrected McNemar's test yields a p value of 0.07.

Table 4.14: McNemar's 2×2 contingency table for CVP.

	Manual positive	Manual negative	Row totals
Automatic positive	32	9	41
Automatic negative	2	270	272
Column totals	34	279	313

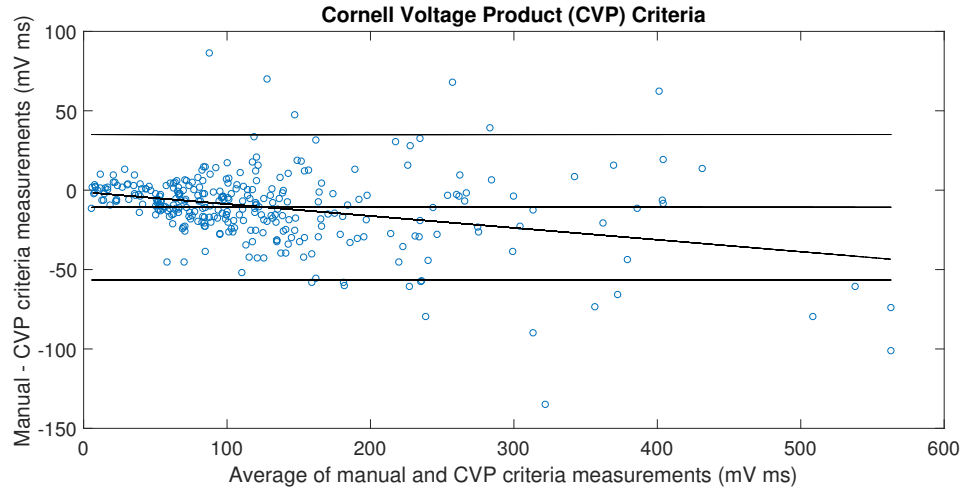


Figure 4.8: Bland-Altman plot comparing the manually and automatically derived CVP.

4.4.4 Novacode Criteria

Table 4.15 gives the values found in figure figure 4.9. The regression line shows a negative slope of -0.019 and an intercept of -3.03 . The 2σ boundaries are in the range of $[-35.97, 25.89]$. In total, the automatic method detects 51 positive LVH subjects and the manual method 41 subjects as shown in table 4.16. The result of McNemar's test gives a p value of 0.034.

The equation of the Novacode Criteria greatly differs depending on the sex of the patient. This results in two trends in the measurement comparison, as seen in Bland-Altman plot in figure 4.9. In order to correctly analyze the agreement between the two methods, two separate plots were created for each sex.

Bland-Altman Plot & McNemar's Test Statistics	Manual - automatic NC
Difference mean (\bar{d}), gm/m^2	-5.04
Difference SD (s), gm/m^2	15.78
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	$[-35.97, 25.89]$
Regression line	$y = -0.019x - 3.03$
McNemar's test p value	0.034

Table 4.15: Bland-Altman statistics for the manually and automatically derived Novacode values.

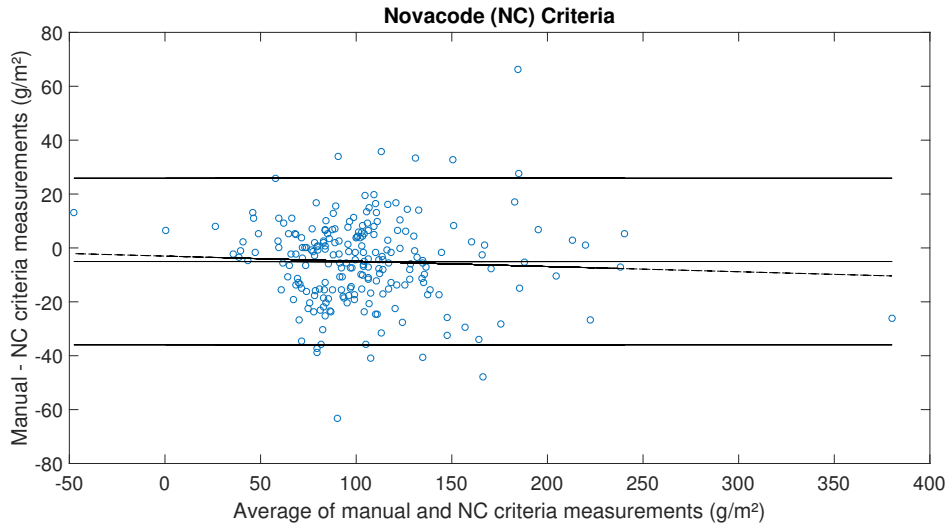


Figure 4.9: Bland-Altman plot comparing the manually and automatically derived Novacode Criteria.

Table 4.16: McNemar's 2×2 contingency table for Novacode.

	Manual positive	Manual negative	Row totals
Automatic positive	37	14	51
Automatic negative	4	171	175
Column totals	41	185	226

4.4.5 Novacode Criteria Men

The equation for the male Novacode criteria is dependent on six features of the ECG, including QRS duration. This large number of required amplitudes for parameter calculation, in comparison to the other calculated criteria, results in 160 male patients that are manually and automatically determined. The incidence of LVH greatly differs between the two methods with manual measurements displaying a prevalence of 30 subjects and automatic measurements displaying a prevalence of 40 subjects, as shown in table 4.18.

The Bland-Altman statistics in table 4.17 that describe the plot in figure 4.10, show a difference mean of -8.72 and a difference standard deviation of 16.88 . The regression fit has an intercept of -11.33 and a slope of 0.02 . McNemar's contingency table shows a total of 30 male subjects for the manual method and 40 positive subjects for the automatic method. Using the corrected McNemar's test gives a p value of 0.024 .

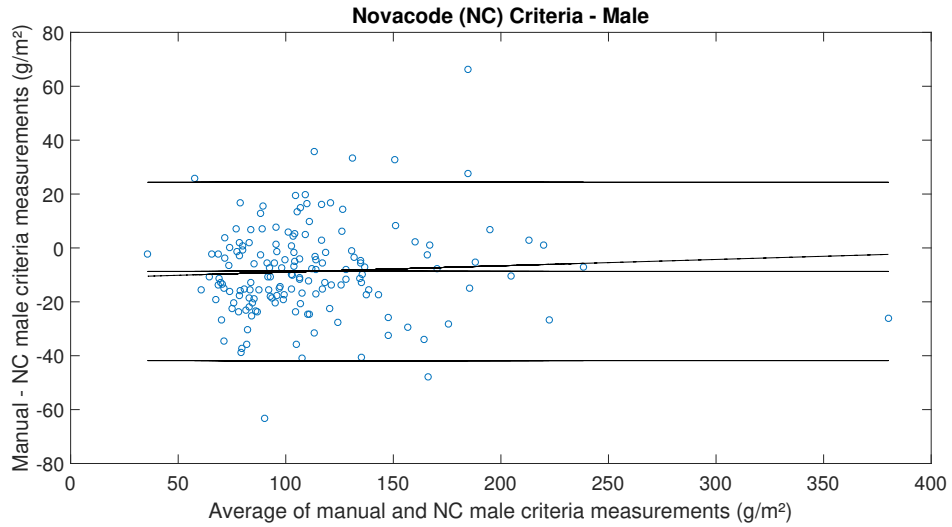


Figure 4.10: Bland-Altman plot comparing the manually and automatically derived Novacode Criteria for male patients.

Bland-Altman Plot & McNemar's Test Statistics	Manual - automatic male NC
Difference mean (\bar{d}), gm/m^2	-8.72
Difference SD (s), gm/m^2	16.88
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	$[-41.81, 24.37]$
Regression line	$y = 0.02x - 11.33$
McNemar's test p value	0.024

Table 4.17: Bland-Altman statistics for the manually and automatically derived male Novacode values.

Table 4.18: McNemar's 2×2 contingency table for Male Novacode.

	Manual positive	Manual negative	Row totals
Automatic positive	27	13	40
Automatic negative	3	117	120
Column totals	30	130	160

4.4.6 Novacode Criteria Women

The female Novacode Criteria is calculated for 66 female patients. Like the male Novacode parameter, the female equation is also dependent on many variables. The female Novacode equation includes eight features, but does not include a QRS duration. As a result, the

Bland-Altman plot in figure 4.11 shows better agreement than the male counterpart. Also it can be seen in table 4.9, both measurement methods determined the same LVH prevalence at 16.7%.

The Bland-Altman results in table 4.19 show a difference mean of 3.89 and a difference standard deviation of 7.11. The regression line has a negative slope of -0.00074 . McNemar's contingency table in table 4.20 shows eleven positive subjects for both automatic and manual measurements. The results of the corrected McNemar test is a p value of 0.48.

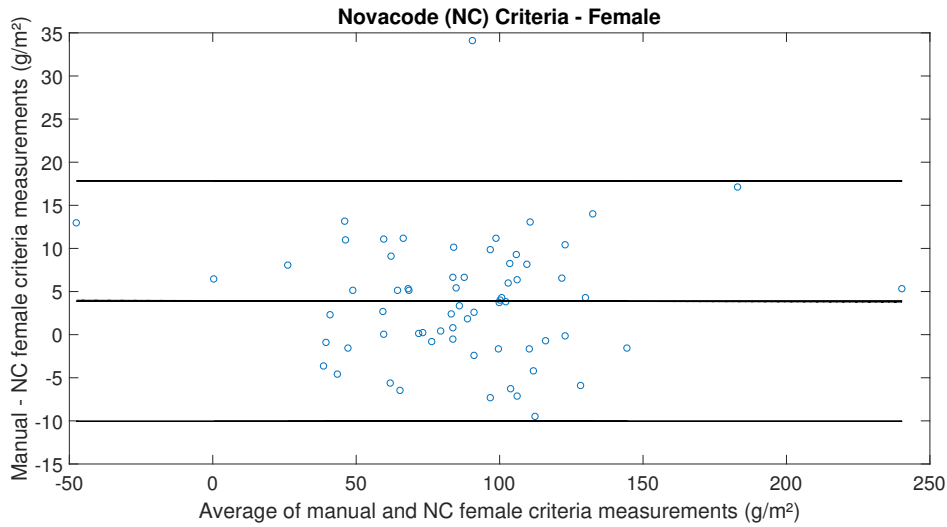


Figure 4.11: Bland-Altman plot comparing the manually and automatically derived Novacode Criteria for female patients.

Bland-Altman Plot & McNemar's Test Statistics	Manual - automatic female NC
Difference mean (\bar{d}), gm/m^2	3.89
Difference SD (s), gm/m^2	7.11
$[\bar{d} - 1.96s, \bar{d} + 1.96s]$	$[-10.04, 17.82]$
Regression line	$y = -0.00074x + 3.96$
McNemar's test p value	0.48

Table 4.19: Bland-Altman statistics for the manually and automatically derived female Novacode values.

Table 4.20: McNemar's 2×2 contingency table for female Novacode.

	Manual positive	Manual negative	Row totals
Automatic positive	10	1	11
Automatic negative	1	54	55
Column totals	11	55	66

4.5 Survival Analysis

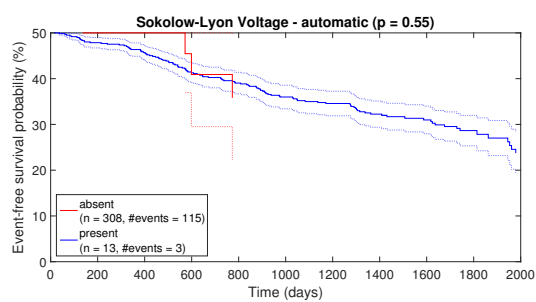
In total 122 events occurred during a median follow-up time of 1291 (727 - 1702) days. In figure 4.12 (a)-(d), Kaplan-Meier curves for the dichotomized LVH criteria (absent vs. present) based on the thresholds from literature (see section 2.2.4) are shown. Log-rank test revealed significant differences for automatically determined CV ($p = 0.003$), CVP ($p < 0.001$), and Novacode ($p = 0.008$), but not for SLV ($p = 0.55$). Results for Novacode showed large discrepancies for men ($p = 0.003$) and women ($p = 0.81$). Data for manually measured LVH criteria are similar (data not shown). Results for SLV need to be considered with caution, since number of events in group with LVH is rather small.

The results of the Cox proportional hazards models are shown in table 4.21, with hazard ratios (HR) and confidence intervals (CI). In univariate Cox regression, automatically calculated and dichotomized CV (HR = 2.23, 95% CI = (1.30; 3.85), $p = 0.004$), CVP (HR = 2.39, 95% CI = (1.52; 3.76), $p < 0.001$) and Novacode (HR = 1.91, 95% CI = (1.18; 3.11), $p = 0.009$) were significantly related to all-cause mortality. CV, CVP and Novacode remained statistically significant predictors after adjustment (model 1 and 2). Results for manually measured LVH criteria were similar with higher hazard ratios for CVP (univariate HR = 2.94, 95% CI = (1.84; 4.71), $p < 0.001$; after adjustment (model 2) HR = 2.31, 95% CI = (1.41; 3.77), $p < 0.001$). For detailed results see table 4.21.

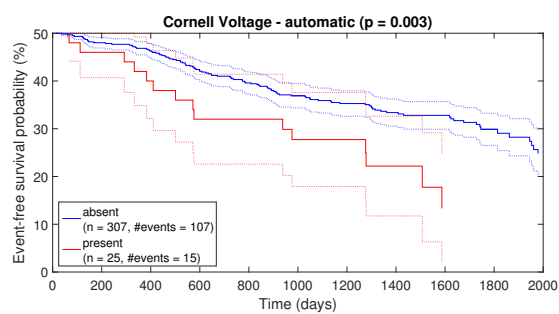
Table 4.21: Cox proportional hazards models reported as median with 95% confidence interval (median (95% CI)).

	HR univariate	p univariate	HR Model 1	p Model 1	HR Model 2	p Model 2
SLV - automatic	0.70 (0.22 - 2.22)	0.55				
CV - automatic	2.23 (1.30 - 3.85)	0.004	2.27 (1.31 - 3.94)	0.004	1.81 (1.03 - 3.19)	0.04
CVP - automatic	2.39 (1.52 - 3.76)	<0.001	2.24 (1.42 - 3.54)	<0.001	2.35 (1.46 - 3.80)	<0.001
NC - automatic	1.91 (1.18 - 3.11)	0.009	1.88 (1.14 - 3.07)	0.01	1.73 (1.04 - 2.86)	0.03
SLV - manual	0.70 (0.22 - 2.22)	0.55				
CV - manual	2.07 (1.20 - 3.57)	0.009	2.02 (1.16 - 3.50)	0.01	1.69 (0.96 - 2.97)	0.07
CVP - manual	2.94 (1.84 - 4.71)	>0.001	2.48 (1.54 - 3.99)	>0.001	2.31 (1.41 - 3.77)	<0.001
NC - manual	1.54 (0.91 - 2.62)	0.11	1.45 (0.85 - 2.46)	0.17	1.32 (0.76 - 2.28)	0.32

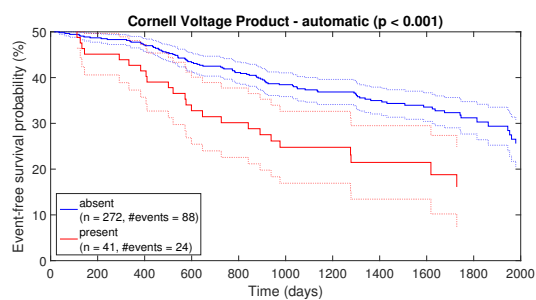
HR = hazard ratio; SLV = Sokolow-Lyon Voltage; CV = Cornell Voltage; CVP = Cornell Voltage Product; NC = Novacode



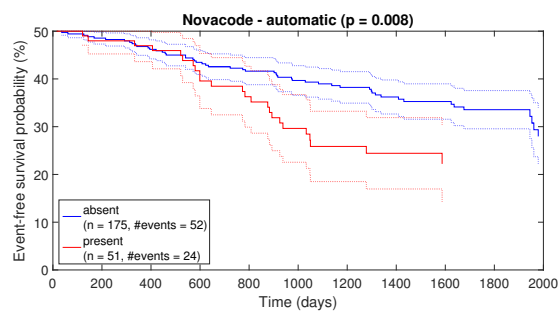
(a) Sokolow-Lyon Voltage



(b) Cornell Voltage



(c) Cornell Voltage Product



(d) Novacode Estimate

Figure 4.12: The Kaplan-Meier plots for each parameter.

Discussion

This chapter interprets and analyzes the results given in the previous chapter. The discussion chapter considers the drawbacks of using an automatic method for calculating ECG-based parameters for LVH detection, but also discusses the limitations of the manual method.

5.1 Amplitude Comparison

The number of subjects in which manually derived and automatically derived criteria could be compared is limited to the AIT ECGsolver detection of peaks. In single beat comparison, the case of an undetectable peak in the automatic method leads to the inability to evaluate the corresponding amplitude, resulting in indeterminable parameters in which the amplitude is a variable. For criteria which are dependent on many amplitudes for evaluation, such as the Novacode estimate, the likelihood that an automatic parameter cannot be derived increases. The consequence of this limitation can be seen in table 4.9. Here it is shown that the number of total patients that could be compared between the manual and automatic methods for each parameter differs from each other and from the original 335 available ECG manual annotations. Also, as the complexity of the parameter equation increases, there is a reduction in total patients that are computed for the parameter. As a result, the total number of subjects that could be compared for the Novacode estimate is severely reduced from the 335 available ECG manual annotations to only 226 subjects.

Considering the comparison of amplitude evaluation in section 4.2, both the R(aVL) and the S(V3) show a very small but negative difference mean, signifying, on average, a slight overestimation of amplitudes on the magnitude of microvolts. One reason for this, especially for small values, could be manual overestimation due to difficulty of measurement. Though, both Bland-Altman plots in figure 4.1 and in figure 4.2 show outliers that occur outside of the limits of agreement. Investigation of these outliers

show that there are challenges that must be refined in the automatic method of feature detection and amplitude calculation.

The main problem is the computation of the baseline to compute amplitudes. While the physician considers the baseline of the whole ECG frame, the automatic method of baseline calculation is restricted to the examined beat and following QRS complex. Therefore, the automatic baseline evaluation is vulnerable to mis-detection of QRS onset values. Slightly higher or lower detected points have large effects on the alignment of the baseline and thus the calculated amplitude amount. To overcome this challenge, it would be advantageous to develop an automatic baseline calculation method that considers the whole frame of the ECG and not just two neighboring complexes. Also, better detection of onset points would be valuable.

Another limitation of the automatic calculation of amplitudes is the mis-detection of peaks. Investigation of the outliers that occur in figure 4.1 discovered that mis-detection of the R peak value gives rise to inaccurate R amplitudes. This mis-detection is attributed to normal variants in certain leads of the ECG. Normal variants are ECG complexes that differ from the normal ECG waveform described in section 2.2.2 and displayed in figure 2.4. They are termed normal because they are known variations in particular leads of the 12-lead ECG that may either occur naturally or due to certain pathologies. The rsR' normal variant was able to be detected and correctly annotated manually by an experienced physician, but was falsely detected by the automatic method.

5.1.1 rsR' Complex

The rsR' complex is a known variant of the V1 lead in the ECG. As of yet, the AIT ECGsolver is unable to detect variations of the QRS forms. The difference between the waveforms of a normal QRS complex and the rsR' variant is shown in figure 5.1.

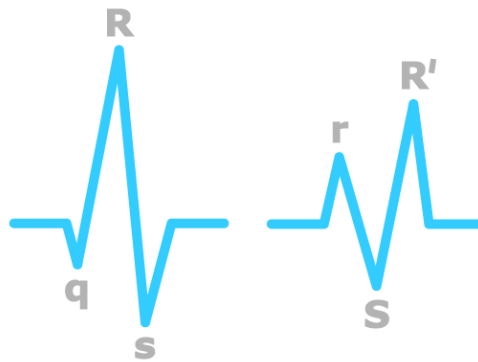


Figure 5.1: The normal QRS complex waveform (left) and the rsR' variation (right).

The correct detection of the rsR' complex form is essential in the calculation of the

Sokolow-Lyon criteria as the S amplitude in the V1 lead is needed. As seen in figure 5.1, the rsR' complex form contains two R peaks and the S deflection can be found between them. This is in contrast to the normal QRS form where the S deflection is found as the minimum between the R peak and the QRS offset. After comparison between the automatic detection and the manual annotations, it was found that the AIT ECGsolver would detect the R' of the rsR' complex as the R peak. This would lead to a mis-detection of the S deflection and an incorrect value for the automatic amplitude of the SV1 waveform.

Therefore, with improvements that consider normal variants of certain waveforms and whole-frame baseline evaluation, the automatic method of amplitude calculation would be an invaluable tool.

5.2 QRS Duration Comparison

When initially calculating the CVP and Novacode parameters, large discrepancies between the automatically and manually determined criteria were found. During inspection of the individual variables used to calculate these parameters, a large difference in the manually annotated QRS duration and the automatically determined QRS duration using AIT ECGsolver was found for single beats. This is shown by the Bland-Altman plot in figure 4.3, which compares the manual QRS durations and the AIT ECGsolver QRS durations. The plot shows a difference mean between the manual and automatic durations of -20.44 ms. Thus, the automatic method was on average evaluating a QRS duration that was around 20 ms more than the corresponding manually derived duration. Therefore, the AIT ECGsolver was consistently overestimating the QRS duration in comparison to the manual annotations. As a result, it was decided that in order to get an accurate automatically determined parameter, another method of QRS onset and offset detection was needed.

The open-source algorithms GQRS and SQRS from the PhysioNet Web page proved to be better for QRS onset and offset detection. The statistical comparison of each QRS detector with the manually determined durations is shown in table 4.5. Though these algorithms solved the issue of large overestimation of the QRS duration, it was found that the algorithms could not detect QRS onset and offset points for many of the patients in comparison to the AIT ECGsolver. This greatly reduced the number of patients in which a valid parameter could be calculated. It was shown that the SQRS detector cannot detect QRS duration features for 79 subjects and the GQRS lacks information for 122 subjects. In order to remedy this problem, a combination of results from the two algorithms, SQRS and GQRS, is used in order to maximize the number of valid samples.

To be able to use a combination of detectors, it was necessary to determine that duration results for each detector were essentially the same. The Bland-Altman plot in figure 4.4 displays a comparison between the SQRS and GQRS results for durations that were detected by both algorithms. The majority of paired samples showed no difference between the durations computed from each detector and therefore it determined that the

results from each detector could be used interchangeably. Since the SQRS detectors was able to detect more samples than the GQRS, the combined results are mostly given by the SQRS detector. For the patients in which the SQRS detector failed, but the GQRS detector was able to obtain a duration, the GQRS duration was included in the combined results. The Bland-Altman plot for the comparison of the combined results with the manual parameter is shown in figure 4.5. With the combined MIT detector durations, the estimation of the manual durations greatly improved, but was still overestimated.

Another reason for the duration discrepancies between the manual and automatic measurements is the difference in the method of evaluation. For the manual method used by the physician, one arbitrary interval within the frame was used as the final duration. The inconsistency of manual QRS duration evaluation poses challenges in the development of an automatic method. The automatic method uses a variation on the QRS interval evaluation technique according to Bauer et al. [6] (see section 3.3.4). Therefore, it would be useful to reevaluate the manual parameters with a similar method of manual interval measurement and repeat the comparison.

5.3 Parameters

The resulting parameters are calculated using durations from a combination of the MIT detectors. The parameters were also calculated with the durations derived from the AIT ECGsolver. The duration dependent criteria statistical comparison can be found in appendix A.

In general, the results show that there is good agreement between the methods when just voltages are taken into account. This is demonstrated by the very good agreement of the SLV, SV, and female Novacode results where all equations are just dependent on voltage amplitudes and not duration. On the other hand, the limitation of the QRS duration measurements do not show good agreement but have the potential for improvement.

5.3.1 Sokolow-Lyon Voltage

A comparison of the manually and automatically derived SLV parameters in table 4.9 show that the mean and median of both methods display very good agreement. The slightly negative difference mean line in figure 4.6 indicates that the automatic method is on average overestimating the SLV parameter in comparison to the manual method by 13 μV . This corresponds to a manual measurement difference of 0.13 mm. One reason for this could be the slight overestimation of the amplitudes themselves in comparison to the manual measurements taken as described in due to reasons discussed in section 5.1.

The matched pairs outside of the 95% limits of agreement in figure 4.6 are mostly positive (above the 2σ boundaries), indicating that the manual SLV parameter is for the majority of outliers greater than the automatically derived parameter, especially for small values. This difference can be attributed to the misinterpretation of the QRS complex as described in section 5.1.1. Due to the mis-detection of the R' peak as an R peak by the

AIT ECGsolver, the incorrect S wave that is calculated would be much smaller, and close to null, in comparison to the true value. Thus, since the SLV parameter is calculated as a sum of the max(RV5, RV6) and S(V1) wave, the incorrect, smaller value of S(V1) would give automatic parameters that are much smaller than the correctly detected manual ones. Thus, the majority of outliers, in which this problem occurs, are found above the confidence interval in the Bland-Altman plot.

The McNemar 2×2 contingency table for SLV shows that there are no discordant values. The values of b and c in table 4.10 are both zero, meaning the methods show a concordance of 100%. Thus, the result of McNemar's test is undefined and otherwise unnecessary. Though the table shows that both methods detected the same exact amount of LVH incidence at 13 (4%), it is also worth noting that the number of positive cases with respect to the sample size is very small. From literature research it has been found that the SLV displays the lowest sensitivity of the four tested parameters (see section 2.2.4), but consistently performs with a high specificity. This would explain the rather low number of patients found positive for LVH in comparison to the amount detected using the other parameter methods. The high tolerance for false negatives would also explain why the automatically computed SLV parameter detected the same amount of LVH incidence, though there were errors in detecting the S(V1) amplitude. Consequently, it would be expected that should the results of the SLV be compared to an accurate LVM measurement, they would display few false positives, but many false negatives.

5.3.2 Cornell Voltage

Similarly to the SLV results, the mean and median values for each method in the CV parameter comparison show very good agreement. The Bland-Altman plot for the CV comparison in figure 4.7 also displays a difference mean line at -0.013 mV, meaning that the average difference between the measurement methods is in the range of 13 microvolts or 0.13 mm. The sample pairs that lie outside of the limits of agreement are not skewed to one side. Therefore, it is not obvious whether one method is overestimating or underestimating the other.

The calculation of the CV parameter is only dependent on voltages and not durations. Upon examining the CV amplitude comparison results for R(aVL) and S(V3) in section 4.2, there is also very good agreement between the manual and automatic measurements for amplitudes. Few outliers exist outside of the limits of agreement for both amplitude comparisons. Possible reasons for discrepancies in the amplitude measurement between the methods are discussed in section 5.1.

Table 4.12 shows that manual method detected one more incidence of LVH than the automatic method. The result of McNemar's test with Yates correction is a p-value of ~ 1.00 , implying that there is no statistically significant difference between the measurement methods in calculating CV. Though, also similar to the SLV, this result may be due to the low sensitivity of the CV parameter. The LVH prevalence rate, 7.8% and 7.9% for

manual and automatic respectively, is also relatively low, making it likely that there are incidences of LVH not detected by the CV parameter.

5.3.3 Cornell Voltage Product

The introduction of the QRS duration as a variable in the parameter equation limited the number of sample pairs that could be compared. Even so, the LVH incidence rose for both the manual and automatic methods at 10.9% and 13.1%, respectively. This is most probably attributed to the increased sensitivity of the CVP parameter in comparison to the CV parameter. Table 4.14 shows that automatic method detects seven more incidences of LVH than the manual method. This is in contrast to the CV parameter results in which the manual method detected one more LVH case than the automatic method. The cause of this reversal in the automatic method detecting more LVH incidence can only be attributed to QRS duration.

As discussed in section 5.2, the automatic method consistently overestimated the QRS duration in comparison to the manual method. As a result, the difference mean line for the CVP in figure 4.8 shows that the automatic measurements are on average about 11 mV·ms more than the manual measurements. This difference is also observed in table 4.9, where the mean of the manual measurements is 10 mV·ms less than the automatic measurements, at 122 and 132 mV·ms respectively. The possible reasons for overestimation of the QRS interval by the detectors is explained in section 5.2.

The results of the McNemar's test with a Yates correction of 1.0 yields a p-value of 0.07, suggesting that there is no statistically significant difference between the methods of CVP evaluation. Though, this would be improved with better QRS duration calculation. The higher sensitivity of the CVP parameter, in comparison to the SLV and CV, and the high specificity make it a useful criteria for LVH detection. Since the agreement of the amplitude measurements is strong and McNemar's test shows that there is no statistical significant difference, the automatic computation of the CVP parameter would be a valuable method for LVH detection.

5.3.4 Novacode

Due to the individual Novacode equations used for male and female subjects, it is necessary to examine the criteria results separately based on sex.

Novacode Male

The number of sample pairs that could be evaluated for the Novacode estimate for males, 160, was much smaller than the those considered for the other criteria. This can be attributed to the numerous amount of variables, including QRS duration, needed to calculate the parameter. The average difference between the measurements is -8.72 gm/m^2 , indicating that again the automatic method is overestimating the parameter in comparison to the manual method. Since the female Novacode shows good agreement,

though it is also dependent on a large amount of variables, but not including the QRS duration, it can be said with confidence that the most probable reason for the overestimation is the influence of the QRS duration.

For the sample size, the male Novacode criteria displayed a larger rate of LVH prevalence in comparison to the other parameters, at 18.8% and 25% for manual and automatic methods, respectively. This high detection of LVH for such a relatively small sample group is most probably due to the higher sensitivity of the Novacode estimate. Although it must be taken with caution as the Novacode estimate also exhibits a much lower specificity than the other criteria. The McNemar test applied on the discordants in cells b and c of table 4.18, yields a p-value of 0.024, signifying that there is a statistically significant difference in the measurement methods. The relatively large number of LVH subjects found to be positive in the automatic method, but not in the manual method, 13 of 160, is most likely due to the larger QRS duration measured for the automatic method. It is expected that this number would reduce with improvement of QRS interval detection.

As a result, the Novacode male criteria would very much benefit from a refinement of the QRS duration evaluation. Especially considering that the Novacode estimate has shown to yield the highest sensitivities of the four parameters tested, the automatic evaluation of the Novcode criteria is of considerable value.

Novacode Female

In contrast to the male Novacode, the female Novacode results showed good agreement when comparing the mean and median values of the manual and automatic method. The average difference between the measurement values is positive at 3.89 gm/m^2 . Thus, as opposed to the other criteria, the female Novacode was on average underestimated by the automatic method in comparison to the manual method. As discussed, the female Novacode only contains variables dependent on voltages and no durations. Therefore this underestimation must be due to challenges in measuring amplitudes in various leads. One cause could be the underestimation of the T wave amplitude as a variable, which is not included in the SLV, CV, or CVP calculation.

One reason for the extremely small number of female Novacode compared pairs is due to the sample size being a subset of the total Novacode sample size, which itself is dependent on the available amount of QRS durations. It would be of interest to repeat the comparison not considering the availability of the variables used for the male Novacode. This would likely increase the number of the parameters compared.

The contingency table in table 4.20 shows that one discordant occurs in each b and c cells and gives a McNemar's test p-value of 0.49. Thus the comparison indicates that the difference in the measurement methods is not statistically significant. Therefore, with the results that there is good agreement between the measurement methods and the strong performance of the Novacode estimate as a LVH predictor, the automatic method

of female Novacode evaluation has extremely potential value. Further testing with larger sample groups would be needed for verification.

5.4 Survival Analysis

Preliminary results of survival analysis using Kaplan-Meier curves and Cox regression models revealed unforeseen results. CV, CVP and automatically, but not manually, determined Novacode were statistically significant risk predictors for all-cause mortality (univariate and after adjustment). These findings are contradictory to results presented by Covic et al., who reported no significant LVH criterium for all-cause mortality and just Novacode for cardiovascular mortality [11]. These discrepancies need further detailed investigation.

Reasons for these discrepancies could be: different strategies for adjustment, mixture of dialysis methods in Covic et al. (i.e., hemodialysis and continuous ambulatory peritoneal dialysis) vs. just hemodialysis in the current study, differences in baseline data (e.g., age - Covic et al.: $51.1 \pm 15.7SD$ years vs. ISAR: 65.2 ± 15.3), inclusion of pacemaker patients in the ISAR study, different times of extracted ECG segments in the current study and thus influences due to circadian patterns and especially due to hemodialysis and its fluid management need to be expected, and finally differences in number of events (Covic et al.: 76 (18%) deaths vs. ISAR: 122 (36%)). Results for SLV need to be considered with caution, since number of events in group with LVH is rather small.

Conclusion

In conclusion, the development of an automatic method of ECG-based LVH parameters proved to be useful, but with limitations. Improvements in the detection of features and the method of manual interval measurement would greatly refine the comparison of manually and automatically derived parameters. These improvements would include implementing normal variant detection in the AIT ECGsolver, designing a baseline evaluation method that considers a larger ECG time frame, and using a standard method for evaluating manual intervals. These changes would be essential to avoiding ECG mis-interpretation and possible mis-diagnosis and provide a better comparison between the methods.

Generally, the automatic method of ECG-based LVH parameters was shown to be congruent with the manual method for parameters dependent on only voltages. This is demonstrated by the comparisons of the SLV, CV, and female Novacode parameters and the amplitude comparisons of the R(aVL) and S(V3) peaks. With the corrections mentioned above, the automatic method of voltage criteria calculation would prove to be as valuable as the manual method.

The evaluation of accurate QRS intervals would greatly improve the manual and automatic derivation of duration dependent parameters such as the CVP and male Novacode. Both the manual and automatic methods of duration calculation are in need of revision. It would be especially important since literature research has shown that these parameters demonstrate the highest sensitivities for LVH detection. By comparing the agreement of the CV and female Novacode with their duration dependent counterparts, the CVP and male Novacode respectively, it is evident that there is potential for very good agreement between the methods, on the condition that QRS interval evaluation is improved.

Contrary to other literature, the unexpected survival analysis results that CV, CVP, and automatically derived Novacode were significant risk predictors for all-cause mortality is an interesting find. This would signify that CV and CVP, in addition to the Novacode

estimate, are all valuable ECG-based parameters for LVH-detection. Further investigation with refined methods of automatic CV and CVP evaluation and comparison with manual parameters would be a logical next step to this thesis.

It would also be of interest to conduct a further study to compare the manually and automatically derived parameters with an accurate method of LVM determination, such as cardiac imaging. This would lead to the calculation of the specificity and sensitivity of each automatic parameter which can be compared with published literature.

As the studied cohort consists of dialysis patients, it must also be noted that the ECG interpretation and LVH outcomes must be taken with caution due to the dependency of ECG waveforms on dialysis treatment times. Due to the knowledge that ECG interpretation varies with dialysis times, it would be valuable to extend this research to computing the same parameters at controlled times rather than arbitrary chosen times, as was done in this thesis. For example, further research can be done on this cohort to compare the automatically and manually derived ECG-based LVH parameter values at dialysis and non-dialysis treatment times. Therefore, the effects of treatment times can be investigated and used to make specific improvements on parameter evaluation for dialysis patients.

ECG-based LVH parameters are an invaluable method of LVH detection. Automatic methods to derived these parameters show potential to increase their value as manual mistakes such as mis-recording can be avoided. Though, there are many adaptations that must be made to the automatic method. These parameters will continue to be relevant as long as cardiovascular diseases and associated pathologies such as CKD continue to be prevalent. The convenience and accessibility of ECG add to the vital importance of continuing to improve and refine ECG-based LVH parameters for the benefit of society.

QRS dependent criteria with ECGsolver durations.

	Cornell Voltage Product	
	Manual	Automatic
Total patients	332	332
Mean \pm SD, mV\cdotms	123 \pm 96.2	150 \pm 123
Median (IQR), mV\cdotms	95.7 (59.8 - 153)	111 (68.5 - 187)
LVH prevalence (%)	38 (11.4%)	58 (17.5%)
	Novacode	
	Manual	Automatic
Total patients	232	232
Mean \pm SD, gm/m²	101 \pm 42.6	115 \pm 49.6
Median (IQR), gm/m²	96.7 (74 - 117)	110 (84.3 - 133)
LVH prevalence (%)	42 (18.1%)	72 (31%)
	Novacode - Men	
	Manual	Automatic
Total patients	166	166
Mean \pm SD, gm/m²	106 \pm 42.7	128 \pm 47.9
Median (IQR), gm/m²	97.8 (76.7 - 124)	122 (94.2 - 151)
LVH prevalence (%)	31 (18.7%)	61 (36.7%)

Table A.1: Statistics for duration dependent criteria calculated with AIT ECGsolver derived durations.

List of Figures

2.1	Four chambers of the human heart with the major vessels.	6
2.2	Placement of the ten electrodes on the body for a 12-lead ECG.	8
2.3	Derivation of the limb and augmented limb leads using Einthoven's triangle. .	9
2.4	ECG schematic with labeled complexes.	9
3.1	Data analysis flowchart.	13
3.2	Heartbeat complex with the points given in the annotation file.	14
3.3	Heartbeat complex with baseline.	16
3.4	ECG heartbeat and the method of amplitude calculation.	17
3.5	Example Bland-Altman plot.	21
3.6	Example Kaplan-Meier Plot.	22
4.1	Manually and automatically derived amplitudes for the RaVL wave.	25
4.2	Manually and automatically derived amplitudes for the SV3 wave.	26
4.3	Bland-Altman plot: Manually and automatically derived QRS durations. . .	28
4.4	Bland-Altman plot: SQRS and GQRS detectors	29
4.5	Bland-Altman plot: Manual durations and combined MIT detectors.	30
4.6	Bland-Altman plot: Manual and automatic SLV.	33
4.7	Bland-Altman plot: Manual and automatic CV.	34
4.8	Bland-Altman plot: Manual and automatic CVP.	36
4.9	Bland-Altman plot: Manual and automatic Novacode.	37
4.10	Bland-Altman plot: Manual and automatic male Novacode.	38
4.11	Bland-Altman plot: Manual and automatic female Novacode.	39
4.12	Kaplan-Meier plots for each parameter.	41
5.1	Normal QRS complex waveform and the rsR' variation.	44

List of Tables

2.1	Thresholds for LVH detection from ECG-based Criteria	12
3.1	Example McNemar's 2×2 contingency table.	21
4.1	Baseline characteristics of the investigated cohort.	24
4.2	Statistical comparison for the R wave in Lead aVL.	25
4.3	Statistical comparison for the S wave in Lead V3.	26
4.4	Statistical comparison of durations using the AIT ECGsolver.	27
4.5	Statistical comparison of the manually measured durations versus the MIT SQRS detector durations and MIT GQRS detector durations respectively. . .	28
4.6	Statistical comparison of SQRS and GQRS MIT detectors.	29
4.7	Statistical comparison of the manually derived QRS durations with the com- bined MIT detectors.	30
4.8	Bland-Altman statistics for SLV.	31
4.9	Statistics for all criteria calculated.	32
4.10	McNemar's 2×2 contingency table for SLV results.	33
4.11	Bland-Altman statistics for CV.	34
4.12	McNemar's 2×2 contingency table for CV.	34
4.13	Bland-Altman statistics for CVP.	35
4.14	McNemar's 2×2 contingency table for CVP.	35
4.15	Bland-Altman statistics for Novacode.	36
4.16	McNemar's 2×2 contingency table for Novacode.	37
4.17	Bland-Altman statistics for male Novacode.	38
4.18	McNemar's 2×2 contingency table for Male Novacode.	38
4.19	Bland-Altman statistics for female Novacode.	39
4.20	McNemar's 2×2 contingency table for female Novacode.	40
4.21	Cox proportional hazards models.	40
A.1	Statistics for duration dependent criteria with AIT ECGsolver durations. . .	53

Bibliography

- [1] R. Agarwal, J. Flynn, V. Pogue, M. Rahman, E. Reisin, and M. R. Weir. Assessment and management of hypertension in patients on dialysis. *Journal of the American Society of Nephrology*, pages ASN-2013060601, 2014.
- [2] A. Aguirre, G. Wodicka, C. Maayan, and D. Shannon. Interaction between respiratory and RR interval oscillations at low frequencies. *Journal of the autonomic nervous system*, 29(3):241–246, 1990.
- [3] D. K. Arnett, P. Rautaharju, R. Crow, A. R. Folsom, L. G. Ekelund, R. Hutchinson, H. A. Tyroler, G. Heiss, A. Investigators, et al. Black-white differences in electrocardiographic left ventricular mass and its association with blood pressure (the ARIC study). *The American journal of cardiology*, 74(3):247–252, 1994.
- [4] A. H. Association. Target heart rates. <http://www.heart.org/HEARTORG/HealthyLiving/PhysicalActivity/Target-Heart-Rates>, 2015.
- [5] M. Bachler, C. Mayer, B. Hametner, S. Wassertheurer, and A. Holzinger. Online and offline determination of QT and PR interval and QRS duration in electrocardiography. In *Joint International Conference on Pervasive Computing and the Networked World*, pages 1–15. Springer, 2012.
- [6] A. Bauer, M. A. Watanabe, P. Barthel, R. Schneider, K. Ulm, and G. Schmidt. QRS duration and late mortality in unselected post-infarction patients of the revascularization era. *European heart journal*, 27(4):427–433, 2006.
- [7] J. M. Bland and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet (London, England)*, 1:307–310, 1986.
- [8] R. Bregman, C. Lemos, R. Pecoits Filho, H. Abensur, S. Draibe, M. G. Bastos, and M. E. Canziani. Left ventricular hypertrophy in patients with chronic kidney disease under conservative treatment. *Jornal Brasileiro de Nefrologia*, 32(1):85–90, 2010.
- [9] P. N. Casale, R. B. Devereux, P. Kligfield, R. R. Eisenberg, D. H. Miller, B. S. Chaudhary, and M. C. Phillips. Electrocardiographic detection of left ventricular hypertrophy: development and prospective validation of improved criteria. *Journal of the American College of Cardiology*, 6:572–580, 1985.

- [10] M. B. Conover. *Understanding Electrocardiography*. Mosby, 8th edition, 2002.
- [11] A. C. Covic, L. D. Buimistriuc, D. Green, A. Stefan, S. Badarau, and P. A. Kalra. The prognostic value of electrocardiographic estimation of left ventricular hypertrophy in dialysis patients. *Ann Noninvasive Electrocardiol*, 18(2):188–198, 2013.
- [12] D. R. Cox. Regression models and life-tables. In *Breakthroughs in statistics*, pages 527–541. Springer, 1992.
- [13] J. T. Daugirdas, P. G. Blake, and T. S. Ing. *Handbook of dialysis*, volume 236. Lippincott Williams & Wilkins, 2007.
- [14] A. L. Edwards. Note on the “correction for continuity” in testing the significance of the difference between correlated proportions. *Psychometrika*, 13(3):185–187, 1948.
- [15] W. Einthoven. The string galvanometer and the human electrocardiogram. *KNAW Proceedings*, 6:107–115, 1903.
- [16] W. Engelse and C. Zeelenberg. A single scan algorithm for qrs-detection and feature extraction. *Computers in cardiology*, 6(1979):37–42, 1979.
- [17] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), 2000 (June 13).
- [18] S. Hagmair, M. C. Braunisch, M. Bachler, C. Schmaderer, A.-L. Hasenau, A. Bauer, K. D. Rizas, S. Wassertheurer, and C. C. Mayer. Implementation and verification of an enhanced algorithm for the automatic computation of RR-interval series derived from 24 h 12-lead ECGs. *Physiological Measurement*, 38(1):1, 2016.
- [19] J. E. Hall and A. C. Guyton. *Textbook of medical physiology*. Saunders/Elsevier, Philadelphia, Pa., 10th edition, 2000.
- [20] E. W. Hancock, B. J. Deal, D. M. Mirvis, P. Okin, P. Kligfield, and L. S. Gettes. AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: Part v: electrocardiogram changes associated with cardiac chamber hypertrophy a scientific statement from the american heart association electrocardiography and arrhythmias committee, council on clinical cardiology; the american college of cardiology foundation; and the heart rhythm society endorsed by the international society for computerized electrocardiology. *Journal of the American College of Cardiology*, 53(11):992–1002, 2009.
- [21] M. R. Homaeinezhad, M. Erfanian Moshiri-Nejad, and H. Naseri. A correlation analysis-based detection and delineation of ECG characteristic events using template waveforms extracted by ensemble averaging of clustered heart cycles. *Computers in biology and medicine*, 44:66–75, 2014.

- [22] M. Häggström. Medical gallery of mikael häggström 2014. https://en.wikipedia.org/wiki/QRS_complex#/media/File:QRS_complex.png, 2014. [Online; accessed March 3, 2017].
- [23] R. E. Katholi and D. M. Couri. Left ventricular hypertrophy: Major risk factor in patients with hypertension: Update and practical clinical applications. *Int J Hypertens*, 2011, 2011.
- [24] C. Levkov, G. Mihov, R. Ivanov, I. Daskalov, I. Christov, and I. Dotsinsky. Removal of power-line interference from the ECG: a review of the subtraction procedure. *BioMedical Engineering OnLine*, 4(1):50, 2005.
- [25] B. H. Lorell and B. A. Carabello. Left ventricular hypertrophy: Pathogenesis, detection, and prognosis. *Circulation*, 102:470–479, July 2000.
- [26] T. F. Lüscher. Of stiff and weak ventricles. *European heart journal*, 36:2545–2547, 2015.
- [27] J. Mathew, P. Sleight, E. Lonn, D. Johnstone, J. Pogue, Q. Yi, J. Bosch, B. Sussex, J. Probstfield, S. Yusuf, and H. O. P. E. H. Investigators. Reduction of cardiovascular risk by regression of electrocardiographic markers of left ventricular hypertrophy by the angiotensin-converting enzyme inhibitor ramipril. *Circulation*, 104:1615–1621, 2001.
- [28] Q. McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157, 1947.
- [29] T. J. Molloy, P. M. Okin, R. B. Devereux, and P. Kligfield. Electrocardiographic detection of left ventricular hypertrophy by the simple QRS voltage-duration product. *J. Am. Coll. Cardiol.*, 20(5):1180–1186, Nov 1992.
- [30] G. B. Moody. gqrs, gqpost - QRS detector and post-processor. <https://www.physionet.org/physiotools/wag/gqrs-1.htm>, 2015. [Online; accessed 30-January-2017].
- [31] P. M. Okin, R. B. Devereux, S. Jern, S. E. Kjeldsen, S. Julius, M. S. Nieminen, S. Snapinn, K. E. Harris, P. Aurup, J. M. Edelman, H. Wedel, L. H. Lindholm, B. Dahlöf, and L. S. Investigators. Regression of electrocardiographic left ventricular hypertrophy during antihypertensive treatment and the prediction of major cardiovascular events. *JAMA*, 2004.
- [32] I. P. Panidis, M. N. Kotler, J. F. Ren, G. S. Mintz, J. Ross, and P. Kalman. Development and regression of left ventricular hypertrophy. *J. Am. Coll. Cardiol.*, 3(5):1309–1320, 1984.
- [33] D. Poulikakos and M. Malik. Challenges of ECG monitoring and ECG interpretation in dialysis units. *Journal of Electrocardiology*, 49(6):855–859, 2016.

- [34] P. M. Rautaharju, A. Z. LaCroix, D. D. Savage, S. G. Haynes, J. H. Madans, H. K. Wolf, W. Hadden, J. Keller, and J. Cornoni-Huntley. Electrocardiographic estimate of left ventricular mass versus radiographic cardiac size and the risk of cardiovascular disease mortality in the epidemiologic follow-up study of the first national health and nutrition examination survey. *The American journal of cardiology*, 62:59–66, 1988.
- [35] J. Rawlins, A. Bhan, and S. Sharma. Left ventricular hypertrophy in athletes. *European Heart Journal-Cardiovascular Imaging*, 10(3):350–356, 2009.
- [36] C. Schmaderer, S. Tholen, A.-L. Hasenau, C. Hauser, Y. Suttman, S. Wassertheurer, C. C. Mayer, A. Bauer, K. D. Rizas, S. Kemmner, K. Kotliar, B. Haller, J. Mann, L. Renders, U. Heemann, and M. Baumann. Rationale and study design of the prospective, longitudinal, observational cohort study "rISk strAtification in end-stage Renal disease" (ISAR) study. *BMC nephrology*, 17:161, Oct. 2016.
- [37] N. Scuffy. 12 lead ECG placement. <http://nurse-scuffy.tumblr.com/post/106888865453/celestialmedicine-12-lead-ecg-placement>, 2015. [Online; accessed March 3, 2017].
- [38] S.-W. Shin, K.-S. Kim, C.-G. Song, J.-W. Lee, J.-H. Kim, and G.-W. Jeung. Removal of baseline wandering in ecg signal by improved detrending method. *Bio-Medical Materials and Engineering*, 26(s1):S1087–S1093, 2015.
- [39] M. Sokolow and T. P. Lyon. The ventricular complex in left ventricular hypertrophy as obtained by unipolar precordial and limb leads. *American heart journal*, 37:161–186, 1949.
- [40] L. Sörnmo and P. Laguna. *Bioelectrical signal processing in cardiac and neurological applications*, volume 8. Academic Press, 2005.
- [41] J. W. Tukey. *Exploratory data analysis*. Addison-Wesley, Reading, Massachusetts, 1977.
- [42] Wikipedia, the free encyclopedia. A representative kaplan-meier survival plot. https://commons.wikimedia.org/wiki/File:Km_plot.jpg, 2005. [Online; accessed March 3, 2017].
- [43] Wikipedia, the free encyclopedia. Diagram of the human heart. <https://en.wikipedia.org/wiki/File:Diagramofthehumanheart.svg>, 2006. [Online; accessed March 3, 2017].
- [44] Wikipedia, the free encyclopedia. Bland-alman plot with ci's on loa. https://commons.wikimedia.org/wiki/File:Bland-Alman_Plot_with_CI%27s_on_LOA.png, 2014. [Online; accessed March 3, 2017].

- [45] Wikipedia, the free encyclopedia. Normal ecg/ekg complex with labels. https://en.wikipedia.org/wiki/File:EKG_Complex_en.svg, 2014. [Online; accessed March 3, 2017].
- [46] Wikipedia, the free encyclopedia. Derivation of the limb leads. <https://en.wikipedia.org/wiki/File:LimbleadsofEKG.svg>, 2015. [Online; accessed March 3, 2017].
- [47] F. Yates. Contingency tables involving small numbers and the χ^2 test. *Supplement to the Journal of the Royal Statistical Society*, 1(2):217–235, 1934.