Cardiac Monitoring of Heart Failure Patients using Smartwatches

Master's Thesis in Medical Engineering

submitted
by

Luisa Susanne Graf

born 01.09.1998 in Neumarkt i.d.OPf.

Written at

Machine Learning and Data Analytics Lab
Department Artificial Intelligence in Biomedical Engineering
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

in Cooperation with

ProCarement GmbH - Forchheim, Germany

Advisors:
Katharina Jäger M.Sc., Madeleine Flaucher M.Sc., Dr.-Ing. Heike Leutheuser,
Prof. Dr. Björn Eskofier
(Machine Learning and Data Analytics Lab, FAU Erlangen-Nürnberg)
Dr. Sebastian Eckl, Patricia Trißler
(ProCarement GmbH)

Started: 01.06.2022

Finished: 01.12.2022

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.
Die Richtlinien des Lehrstuhls für Bachelor- und Masterarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Erlangen, den 30. November 2022

Übersicht

Die kontinuierliche Überwachung der Herzinsuffizienz (HF) ist von entscheidender Bedeutung, um eine Verschlechterung der Herzleistung frühzeitig zu erkennen. Smartwatches bieten eine einfache und benutzerfreundliche Möglichkeit zur Erfassung von Gesundheitsdaten, einschließlich eines Elektrokardiogramms (EKG).
Diese Arbeit untersucht den Einfluss von HF auf die Qualität eines Smartwatch-EKGs. Dabei wurden EKG Aufzeichnungen von 18 gesunden und 8 Personen mit HF analysiert. Die EKGs wurden mit der Apple Watch Serie 7 und der Withings Scanwatch aufgezeichnet. Merkmale wurden aus den EKGs mit Hilfe von zwei verschiedenen automatischen EKG-Segmentierungsalgorithmen (Neurokit und ECGdeli) extrahiert und mit Merkmalen eines Referenz-EKG-Signal verglichen. Die Spearman-Korrelation wurde verwendet, um die Beziehung zwischen den Parametern, der verschiedenen Geräte zu bestimmen. Herzfrequenz, QRS-Dauer und QT-Dauer zeigten eine hohe positive Korrelation ($r > 0,7$) zwischen Smartwatch und Referenzaufzeichnungen. Sie waren von der verwendeten Smartwatch und dem Segmentierungsalgorithmus abhängig. Eine insgesamt geringe bis mäßige Korrelation (Apple-Neurokit: $r = 0.41$, Apple-ECGdeli: $r = 0.5$, Withings-Neurokit: $r = 0.40$, Withings-ECGdeli: $r = 035$) konnte bei der Zusammenfassung aller berechneten Merkmale beobachtet werden. Ein Unterschied in der Gesamtkorrelation beim Vergleich von HF- und gesunden Gruppen wurde nicht erfasst, was darauf hindeutet, dass HF keinen Einfluss auf die Qualität der Smartwatch-EKGs hat. Insgesamt deuten die Ergebnisse darauf hin, dass die niedrigen Korrelationen auf die Ungenauigkeit der EKG-Segmentierungsalgorithmen zurückzuführen sind.
Außerdem wurden die aus den Smartwatches extrahierten Merkmale zwischen gesunden und HF-Teilnehmern verglichen. Einzelne Merkmale erreichten eine statistische Signifikanz, aber die Ergebnisse waren bei Verwendung unterschiedlicher Smartwatches oder Segmentierungsalgorithmen nicht beständig. Dies deutet darauf hin, dass die Aussagekraft der Ergebnisse gering ist.
Insgesamt hatten die Teilnehmer keine Probleme mit der Aufzeichnung von Smartwatch-EKGs und zeigten eine hohe Bereitschaft, ihre Herzaktivität regelmäßig mit Smartwatch-EKGs zu überwachen. Dies zeigt, dass Smartwatches ein großes Potenzial für den Einsatz in einem Telemonitoring-System für HF-Patienten haben, wenn die durch die Segmentierungsalgorithmen verursachten Einschränkungen gelöst werden können.

Abstract

Continuous monitoring of heart failure (HF) is crucial to detect cardiac deterioration early and provide good care for the patients. With the emerging trend of wearables, access to biomedical parameters is becoming more available. Smartwatches provide a simple and user friendly way for collecting health data, including an electrocardiogram (ECG).

This thesis aims at investigating the effect that HF has on the quality of a smartwatch ECG recorded by the Apple Watch Series 7 and the Withings Scanwatch. In the context of this work, smartwatch ECG recordings of 18 healthy people and 8 patients diagnosed with HF were analyzed. Features were extracted from the ECG recordings with the help of two different automatic ECG segmentation algorithms (Neurokit and ECGdeli). Features extracted from the smartwatch ECGs were compared to features extracted from a simultaneously recorded reference ECG signals. Spearman's correlation was used to determine the relationship between the features recorded by the different devices. Heart Rate, QRS duration and QT duration showed a high positive correlation ($r > 0.7$) between smartwatch and reference recordings. However, the strength of the correlation depended on the used smartwatch and segmentation algorithm. An overall low to moderate correlation (Apple-Neurokit: $r = 0.41$, Apple-ECGdeli: $r = 0.5$, Withings-Neurokit: $r = 0.40$, Withings-ECGdeli: $r = 035$) could be observed when combining all calculated features. Differences in overall correlation when comparing HF and healthy groups could not be observed, suggesting that HF has no impact on the quality of the smartwatch ECGs. Overall the results suggest that the low correlations are more likely caused by the inaccuracy of the ECG segmentation algorithms.

Furthermore, the features extracted from the smartwatches were compared between healthy and HF participants. Individual features achieved a statistical significance, but the results were not consistent when using different smartwatches or segmentation algorithms. This suggests a low power of the results.

Overall the participants had no problems recording smartwatch ECGs and showed high willingness to monitor their cardiac activity with smartwatch ECGs on a regular basis. This shows that smartwatches have a great potential in being used in a telemonitoring system for HF patients. In future work, limitations caused by the segmentation algorithms need to be solved.

Contents

1 Introduction **1**

2 Fundamentals **5**

2.1 The Electrocardiogram 5

2.1.1 Electrical Activity of the Heart 5

2.1.2 The ECG Signal 6

2.1.3 Leads 7

2.1.4 Electrocardiogram in Smartwatches 8

2.2 Heart Failure 9

2.2.1 Diagnosis 9

2.2.2 Types 10

2.2.3 Classification 11

3 Related Work **13**

3.1 ECG in identifying Heart Failure 13

3.2 Parameters indicating Decompensated Heart Failure 14

3.3 Validation of Smartwatch ECG Recordings 20

3.4 Research Goals 22

4 Methods **23**

4.1 Data Acquisition 23

4.1.1 Study Population 23

4.1.2 Study Components 24

4.1.3 Procedure 27

4.2 Data Preprocessing 29

4.3 Feature Calculation 30
4.4 Evaluation 34
4.4.1 Exclusion of Data 34
4.4.2 Correlation and Agreement 34
4.4.3 Differences 35

5 Results **37**
5.1 Feasibility to Record ECG and SpO2 37
5.2 Outlier Detection 38
5.3 Correlation and Agreement 38
5.4 Differences 47

6 Discussion **53**

7 Conclusion and Outlook **57**

A Patents **59**
A.1 Body-worn sensor for characterizing patients with heart failure 59
A.2 Wearable Device Electrocardiogram 60

B Additional Tables **61**

C Additional Figures **79**

D Questionnaire **85**

E Acronyms **89**

List of Figures **91**

List of Tables **95**

Bibliography **97**

Chapter 1

Introduction

Cardiovascular diseases were identified as the leading cause of death worldwide in 2019. In the United States, heart failure (HF) accounts for 9.9 % of deaths associated with cardiovascular disease [Tsa22]. HF is not a single pathological diagnosis, but a clinical syndrome. The heart is not able to provide the body with a sufficient amount of blood and oxygen, causing breathlessness, ankle swelling and fatigue [McD21]. While the estimated prevalence of HF from 1998 and onwards seemed relatively stable [Rie16], it is projected to rise by 46 % in adults from 2012 to 2030, estimating an increase from 2.4 % to 3.0 % in the total population [Tsa22]. With prevalence rising, the estimated overall costs of HF are also projected to grow, suggesting a total increase of 127 % from 30.7 billion to 69.8 billion US dollars in the US in the same period [Tsa22].

A worsening of cardiac output with symptoms severe enough for the patients to require immediate medical intervention is known as decompensated HF. Decompensated HF is associated with high hospitalization, rehospitalization and mortality rates[McD21]. To detect changes in the patient's physical as well as mental state, regular monitoring is recommended by national guidelines for current care of HF patients [Wis19]. If necessary, the patient's treatment can be adjusted accordingly. Additionally, more intensive monitoring can make it easier to follow the recommended lifestyle and medication. While advances in technology over the past decades make remote data collection for telemonitoring easy, the challenge lies in integrating that data into systems of care that increase the patients' experience of care. Telemontitoring can be considered a remote clinical service, for which data collection can

be done using simple telephone-calls, remote assessment by a nurse specialist, implanted devices, standalone home-based systems as well as wearable technology including smartwatches [Bra19].

With the emerging trend of wearables, especially smartwatches, access to biomedical parameters, such as heart rate and oxygen saturation, is becoming more available. Smartwatches can provide meaningful patient-reported parameters, and are therefore useful tools. However, they are not yet standardly integrated into telemonitoring systems for HF patients [Wer19]. Studies suggest that telemonitoring systems show beneficial outcomes when included in the care of HF patients regarding overall mortality and hospitalization rates [Cle05] [Koe18]. As hospital admissions are a big cost factor for HF treatment, a reduced hospitalization rate, brought forth by the use of telemonitoring systems, leads to the reduction of costs for overall HF treatment [Syd21]. Parameters such as heart rate, heart rate variability, oxygen saturation, and electrocardiograms (ECGs) are being investigated in different telemonitoring concepts [Sen21]. Additionally, they can already be captured by modern smartwatches such as the Apple Watch [App18]. The potential of self-recorded ECGs has previously been shown with the identification of atrial fibrillation [Per19]. Besides detecting abnormal heart rhythm, other valuable parameters, e.g. QRS duration, can be extracted from smartwatch-derived data. The importance of a prolonged QRS complex for the prognosis of HF patients has already been discussed extensively [Kas05]. Additionally, smartwatches are being analyzed for clinical accuracy and it has been established that they can record the baseline intervals accurately [Sag20].

However, the accuracy of smartwatches has not yet been established on HF patients, using an automatic ECG signal segmentation algorithm. Furthermore, relevant parameters that can be extracted from smartwatch data and are useful for detecting changes in the condition of HF patients need to be established. No conclusive evidence is present, for whether including smartwatches in telemonitoring systems for HF patients is justfied. Therefore, the goal of this master thesis is to extend the knowledge of how different parameters, extracted and calculated from commercial smartwatch data, can be used to monitor HF and detect decompensation early. The thesis is structured as follows: Chapter 2 presents the basic fundamentals of ECG recordings and the clinical syndrome HF. Related work in the field of diagnosing HF based on ECG recordings, parameters indicating worsening HF and validating smartwatch ECGs

is discussed in Chapter 3. In Chapter 4, the key methods and overall approach of this work, including details on the study, feature extraction and analysis process are described. Results of the data analysis are presented in Chapter 5. In Chapter 6 the overall results are discussed. Finally, Chapter 7 concludes the thesis and provides an outlook for possible future research.

Chapter 2

Fundamentals

The heart's purpose is to ensure the circulation of the blood in the human body. The heart contracts rhythmically as synchronized electrical currents spread through the heart muscle. According to our need for oxygen and nutrients, the heart adjusts its pumping rate leading to faster or lower heart rates. Cardiovascular diseases can limit the heart's functionality. The following sections provide basic information about the contraction mechanism of the heart and the electrical signal measured by the ECG associated with it. Additionally, information about the cardiovascular disease heart failure is given.

2.1 The Electrocardiogram

2.1.1 Electrical Activity of the Heart

The human heart is made up of four chambers, that can be segregated into the left and the right side. Each side has an atrium, one of the two upper chambers, and a ventricle, one of the two lower chambers. The atria are responsible for collecting blood from the body, while the ventricles are responsible for pumping blood into the body. The propagation of the signal for the heartbeat starts in the right atria with the firing of the sinoatrial (SA) node. The SA node functions as the heart's pacemaker. The firing of the SA node causes the atria to contract and pump blood into the ventricles. As the stimulus spreads through the atria, it reaches the atrioventricular (AV) junction, connecting atria and ventricles. With the firing of the AV node, the signal traverses the right and left bundle branches into the right and left ventricle, which are separated by the interventricular septum. This causes the ventricles to contract and

pump blood into the body. This automatically generated cardiac stimulus leads to a rhythmic contractile activity [Gol17].

The described cardiac activation process is also known as the depolarization of the heart. It is followed by the repolarization phase, representing the return of the heart muscle cells to the resting state. This continuous and repeated depolarization and repolarization is the mechanism that makes the human heart beat [Gol17]. Figure 2.1 shows a schematic representation of the heart as well as the signal propagation through it.

2.1.2 The ECG Signal

The electrical activity of the heart is measured by an ECG. A typical ECG recording, corresponding to one contraction of the heart, can be seen in Figure 2.1. There are six basic wave forms and segments recorded by the ECG. Each wave or deflection corresponds to a specific cardiac electrical activity [Gol17]:

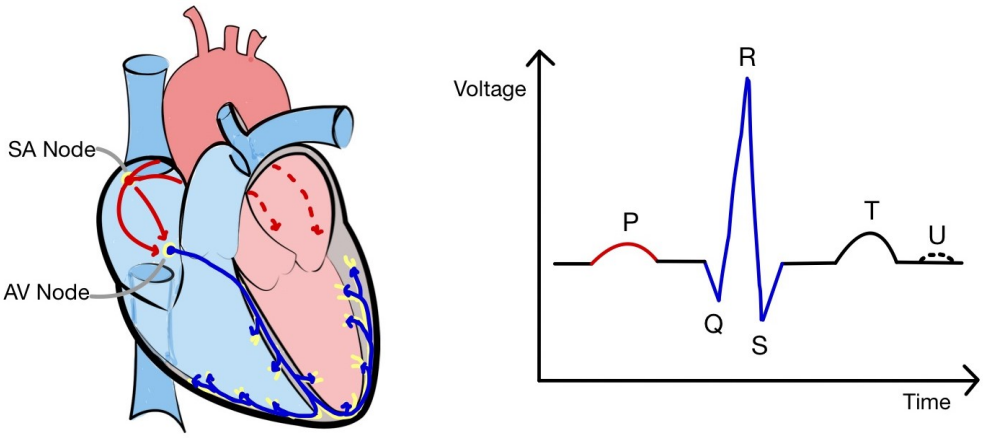

Figure 2.1: Signal propagation through the human heart and according ECG signal.

P wave: The P wave represents atrial depolarization. The SA node starts firing, indicating the start of the beat. The electrical signals spread through atria and causes it to contract.

P-Q segment: The P-Q segments represents the time the signal travels from the SA node to the AV node. This delay optimizes cardiac output, as it allows the ventricles to completely fill with blood before they contract.

QRS complex: The QRS complex shows the stimulation and depolarization of the ventricles as the AV node starts firing. The initial deflection of the QRS complex is negative, known as the Q wave. The Q wave corresponds to the depolarization of the interventricular septum. This is followed by a positive deflection, the R wave. The R wave represents the depolarization of the main mass of ventricles. The S wave is the final negative deflection of the complex and the last phase of ventricular depolarization at the base of the heart. However not every complex consists of all three waves or deflection may change and therefore different nomenclatures exist.

ST segment: The ST segment is the earliest phase of ventricular repolarization. It is isoelectric, meaning flat on the baseline, but can also be slightly elevated or depressed. It is difficult to determine precisely when the ST segment ends and the T wave starts. For clinical purposes an accuracy within 40 ms is usually acceptable.

T wave: The T wave is the mid-latter part of ventricular repolarization. The shape is slightly asymmetrical, with the peak closer to the end of the wave.

U wave: The last phase of ventricular repolarization is represented by the U wave, which is a small rounded deflection following the T wave. It usually has the deflection in the same direction as the T wave, but is not always present in a normal ECG recording.

Atrial repolarization is not recorded by the ECG due to the low amplitudes of the corresponding waves. The atrial repolarization occurs during the ventricular depolarization and the signal is masked by the large QRS complex.

2.1.3 Leads

The common way to measure the cardiac electrical activity is through the 12 standard ECG leads. They are comprised of connections and derivations. The leads record the differences in potential between the electrodes on the body. The individual leads display different views of the electrical activity, and therefore the form of the signal may differ for each lead. The leads can be divided into six limb (extremity) leads (I, II, III, aVR, aVL, aVF) and six chest (precordial) leads (V1 - V6). The limb leads can be further divided into three bipolar limb leads (I, II, III) and three augmented unipolar limb leads (aVR, aVL, aVF). Those divisions stem from their historic development by Einthoven and Goldberger [Gol17].

2.1.4 Electrocardiogram in Smartwatches

While the standard 12-lead ECG gives insight from different views and is considered the gold standard, it cannot be used for portable ECG devices. Many portable long-term ECGs consist of only two leads. In 2018, Apple introduced the first smartwatch (Apple Watch Series 4) with an integrated ECG measurement [App18]. The watch has two inbuild electrodes, one on the back of the watch and one on the digital crown. The ECG can be recorded by wearing the smartwatch on one arm while resting a finger from the other arm on the digital crown, thereby creating a closed circuit. The lead that is measured by the smartwatch corresponds to lead I of the standard 12-lead ECG. Lead I records the difference in voltage between the left arm and right arm electrodes as shown in Figure 2.2. Based on the ECG recording, the apple algorithm can detect heart rhythms such as the normal sinus rhythm as well as abnormal rhythms such as tachycardia, bradycardia and atrial fibrillation. Apple got the CE certification and FDA approval for the ECG App as well as for the irregular rhythm notification feature, making the software a medical product. The ECG App is cleared for users older than 22 years, while the irregular rhythm notification feature is cleared for users 22 years and older with no prior history of atrial fibrillation [App22].

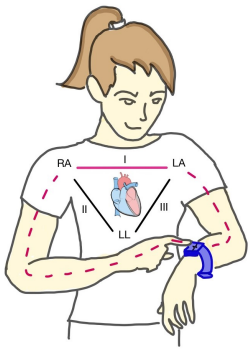

Figure 2.2: Smartwatch ECG. By placing a finger on the second electrode on the watch, lead I can be recorded.

Lead I can now also be recorded by other smartwatches, including among others, the smartwatches by Withings [Wit19]. The working principle is similar to the one of the Apple Watch with a need for a closed circuit between the arms due to the watch's integrated electrodes. The Withings Scanwatch has electrodes build into the main body of the watch, while another one is integrated in the the steel upper ring, the bezel, of the watch. When wearing the watch

(a) (b)

Figure 2.3: Smartwatch with ECG functions. (a) Apple Watch Series 4 [App18] (b) Withings Scanwatch [Wit22a]

and then touching both sides of the bezel, an ECG recording can be obtained. The Scanwatch has received CE medical certification in Europe and FDA clearance in the United States [Wit22b]. Other smartwatches, which are able to record an ECG and are certified for use in Europe include watches by Samsung [Sam21] and Fitbit [Fit20].

2.2 Heart Failure

2.2.1 Diagnosis

Heart failure is one of the leading causes of death worldwide [Tsa22]. It is caused by structural or functional abnormalities of the heart. The heart is therefore unable to pump a sufficient amount of blood through the body. Myocardial dysfunction is the most common cause, but other causes such as pathology of the valves or abnormalities of heart rhythm can be contributors. The first signs of HF are the onset of symptoms, such as shortness of breath, swelling of feet, ankles, and legs as well as overall tiredness [McD21].

Early detection is essential to treat HF and reduce mortality rates successfully. To confirm the presence of HF, investigative tests are performed. An ECG measurement is a simple non-invasive and non-expensive approach. It is a recommended diagnostic test in all patients with suspected chronic heart failure [McD21]. The ECG is used to identify rhtythm (to verify atrial fibrillation (AF)), left ventricular (LV) hypertrophy, P wave duration and morphology or fibrillatory waves, preexcitation, bundle-branch block, prior myocardial infarction, other

atrial arrhythmias, and to measure and follow the R-R, QRS, and QT intervals [Yan13]. A normal ECG recording makes the diagnosis of HF unlikely, and abnormal findings increase the likelihood of a HF diagnosis. However, the changes in an ECG are non-specific and insensitive for diagnosis of HF. To confirm the presence of HF, the concentration of natriuretic peptides is measured and echocardiography is carried out [McD21].

2.2.2 Types

As the heart can be segregated into the left and the right side, HF can present as left-sided failure or right-sided failure. Right-sided HF is less common and often occurs as a result of left-sided failure. Left-sided HF occurs when the left ventricle does not pump efficiently. It is typically categorized based on measurements of the left ventricular ejection fraction (LVEF). A LVEF smaller than 40% is called heart failure with reduced ejection fraction (HFrEF), also called systolic failure. This means, the left ventricle loses the ability to contract normally. A LVEF between 41 and 49% is known as heart failure with mildly reduced ejection fraction. If the LVEF is above 50%, it is classified as heart failure with preserved ejection fraction (HFpEF), also known as diastolic failure. This is caused by the left ventricle not being able to relax normally. Those distinctions are important as treatments differ [McD21] [Yan13].

Additionally, symptoms of HF can present gradually or suddenly. If the symptoms are severe enough to require immediate medical attention, HF is said to be acute. Acute heart failure (AHF) can be of acute nature, without previous HF diagnosis as well as an acute worsening of previously diagnosed HF. Differentiating those types is of importance as treatments and different mortality rates are associated with them [McD21]. Possible causes for AHF are cardiac causes including among other things ischemia, myocardial infarction, arrhythmia and myocarditis. Comorbidities, patient behavior and drug and therapy effects can also play a part in the development of AHF. Relevant comorbidities include infections, renal insufficiency, anemia, pulmonary embolism and thyroid dysfunction while patient behavior can be understood as compliance towards medical therapy and substance misuse. At times the cause of AHF cannot be identified [Wis19].

Acute decompensated heart failure (ADHF) occurs in patients preciously diagnosed with HF and history of cardiac dysfunction of LVEF and therefore often comes with a more gradual onset of symptoms. With 50 - 70% of all AHF presentations being classified as ADHF, it

is the most common type. The main cause of symptoms is progressive fluid accumulation, responsible for systemic congestion. Patients may also present symptoms and clinical signs of hypoperfusion. Other types of AHF inlcude acute pulmonary oedema, isolated right ventricular failure and cardiogenic shock [McD21].

2.2.3 Classification

Based on the severity of HF, the patients can be categorized by the New York Heart Association (NYHA) functional classification [Com94]. The NYHA classes are focused on severity of symptoms and limitations during physical activity. Patients are placed into one of four classes, according to their exercise capacity. It is a subjective assessment by the doctor and can also change over short periods of time.

Table 2.1: New York Heart Association (NYHA) Functional Classification [Com94]

Class	Symptoms
I	No limitation of physical activity. Ordinary physical activity does not cause symptoms.
II	Slight limitations of physical activity. Comfortable at rest. Ordinary physical activity results in symptoms.
III	Marked limitations of physical activity. Comfortable at rest. Less than ordinary activity causes symptoms.
IV	Unable to carry out any physical activity without discomfort. Symptoms of heart failure at rest.

Chapter 3

Related Work

3.1 ECG in identifying Heart Failure

ECGs are a recommended diagnostic test in patients with suspected HF [McD21]. However, studies have shown that ECGs have insufficient specificity in being an alone diagnosis, when evaluated by general practitioners or cardiologists [Dav06] [Gou07]. In recent years, different computer-aided diagnosis systems using ECG signals for classification of HF have therefore been proposed. The contributions are mostly based on time-domain, frequency-domain and non-linear features extracted from the signals and supported by machine learning algorithms.

Time domain methods and the nonlinear Pointcare plot method were used for heart rate variability (HRV) feature extraction by Khaled et al. [Kha06]. Additionally, they evaluated four different classification approaches for diagnosing congestive heart failure (CHF). Best results were obtained using voting k-Nearest Neighbor Classifier and Back propagation Neural Networks with normalized time domain HRV features as inputs. Sensitivity reached up to 97.90% and positive predictive accuracy was more than 98%.

Wang et al. [Wan18] compared different time-domain, frequency-domain and non-linear features of HRV for distinguishing the heart rhythm of HF patients with sinus rhythm of healthy subjects. They showed that all frequency-domain, nonlinear indices, and the time-domain index standard deviation of the NN interval (SDNN), had differentiating power for CHF patients and normal sinus rhythm subjects. Using a support-vector machine based classification algorithm, they achieved sensitivity of 91.31%, specificity of 90.04% and accuracy of 90.95%.

Bhurane et al. [Bhu19] used nonlinear features extracted from wavelet coefficients of frequency-localized filter bank for diagnosis of CHF. For automated classification of ECG signals from healthy people and HF patients, Support Vector Machine was employed. The approach was tested on four different datasets and obtained an accuracy $\geq 99.66\%$, sensitivity $\geq 99.82\%$, and specificity $\geq 99.28\%$ across all datasets.

The results of the studies show that ECG signals have great potential in differentiating healthy subjects from HF patients. Using such systems to facilitate the diagnosis of CHF in hospitals can reduce the time requirement and error rate associated with manual reading of large ECG signals [Bhu19].

The invention of wearable devices to measure ECGs plays an important role in extending the possibilities of HF diagnosis and monitoring through ECG signals. The patent US20160249858A1 (A.1) from 2016 describes a wearable sensor to characterize patients with HF. It used an ECG sensor to measure electrical cardiac activity and an impedance system to measure the stroke volume and cardiac output. Another example of an invention for wearable ECG systems is the patent US20160360986A1 (A.2). It also includes sensors to extract an ECG by wearing the device on a limb such as the wrist. The 'active two hand ECG' that can be measured works similarly to the working principle of the smartwatch ECGs.

3.2 Parameters indicating Decompensated Heart Failure

While research on diagnosing HF based on a ECG recording shows great potential, it is also being explored how different ECG parameters can give information about the state of a patient. Based on the information provided by duration, delay, presence or absence and morphology of the individual segments, abnormalities can be noticed. The worsening of a patient's state can then be identified and predicted according to those abnormalities and changes. The following sections give an overview of which parameters have prognostic value in foreseeing a decompensation in heart failure patients.

Table 3.1: Parameters, extractable from an ECG, identifying and predicting decompensation in heart failure patients.

Parameter	Description	Source
Heart Rate Recovery	Rate at which the heart recovers from exercise	[Are06] [Cah13]
Heart Rate Reserve	Difference between HR at peak exercise and at rest	[Ben13]
Heart Rate Turbulance	Behavior of HR after a premature ventricular contraction	[Moo06] [Cyg06] [Cyg08b] [Dis16]
Heart Rate Variability	Variation in the time interval between heart beats	[Nol98] [Gal00] [La 03] [Aro04]
(1) SDNN	(1) Standard deviation of all normal-to-normal RR intervals	
(2) SDANN	(2) Standard deviation of all 5-minute mean RR intervals	
(3) Total Power	(3) Sum of the energy in all frequency bands	
(4) LF Power	(4) Absolute power of the low-frequency band (0.04-0.15 Hz)	
(5) Ultra LF Power	(5) Absolute power of the ultra-low frequency band ($\leq$0.003 Hz)	
Heart Rate	Number of time the heart beats within a specified time period	[Ho10] [For15] [Zaf18]
T Wave Alternans	Beat-to-beat variation in the amplitude of the ST segment and T wave	[Ste08] [Mon12] [Yam18] [Ars18]
T Wave Morphology	Morphology of T Wave	[Lin09] [Hua09] [Ram17] [Isa21]
(1) Lead Dispersion	(1) Temporal variations in leads for a 3D vector T-loop	
(2) TCRT	(2) Total cosine between QRS and T wave	
(3) Morphology Restitution	(3) Morphological variation of the T wave per RR increment	
(4) Asymmetry	(4) Shift of peak of T wave to left or right	
(5) Notch	(5) Presence of notches, bulges or humps on the T wave	
(6) Flatness	(6) Sharpness or flatness of the T Wave	
QRS Duration	Time from beginning of the Q wave to end of the S wave	[Kal02] [Kas05] [Wan08] [Rav19]
QT Dynamicity	Variations in the QT segment	[Pat05] [Wat07] [Cyg08b] [Ram14] [Ars22]
(1) QT Duration	(1) Time from beginning of the Q wave to the end of the T wave	
(2) QTc Duration	(2) QT Duration corrected for HR	
(3) QT/RR slope	(3) Slope of QT/RR plots of the linear regression	

Heart Rate and Rhythm

Ho et al. [Ho10] showed that heart rate (HR) ≥ 70 beats per minute (bpm) is an independent predictor of all-cause mortality and HF hospitalization in a study with 9580 patients. Additionally, their findings showed that for every 10 bpm increase in HR at rest, the risk of major cardiovascular event increased by 8%. Ford et al. [For15] also used a threshold of 70 bpm. Like Ho et al., they associated every 10 bpm increase with a worsening in outcome. Every 10 bpm increase in resting HR was linked to a 31% increase in cardiovascular mortality or HF hospitalization, a 26% increase in all-cause mortality, a 27% increase in cardiovascular mortality, and a 32% increase in HF hospitalization. Not only tachycardia but also other heart rhythm abnormalities can be prognostic for cardiac decompensation. In a large study including 14946 patients, Zafrir et al. [Zaf18] independently associated AF with either HF hospitalization or hospitalization combined with mortality in HF patients.

Heart Rate Recovery and Heart Rate Reserve

HR Recovery describes the rate at which the heart recovers from exercise. Arena et al. [Are06] showed that HR Recovery < 6.5 bpm was predictive of death or hospitalization in a one year follow up. Their study included 87 patients with compensated HF. HR Recovery was calculated as the difference of the maximal HR during cardiopulmonary exercise testing and the heart rate at one minute after exercise. Similar findings were made by Cahalin et al. [Cah13], who showed that HR Recovery in the six-minute walking test is a strong predictor of major cardiac events, such as death or transplants in HFpEF and HFrEF patients. 258 patients were included in the study with a mean follow up of 22.8 months. HR Reserve was analyzed by Benes et al. [Ben13]. HR reserve was calculated as the difference between HR at peak exercise and at rest. Additionally, it was divided by the difference of age predicted maximal and resting HRs. In a group of 81 stable but advanced HF patients, they showed how impaired heart rate reserve (cut-off value of 0.38) was predictive of adverse outcomes.

Heart Rate Turbulence

Heart rate turbulence (HRT) describes the behavior of the HR after a premature ventricular contraction.

Moore et al. [Moo06] investigated the importance of HRT in CHF patients of NYHA classes II and III. The analysis included 358 patients with a follow up of five years. The data was recorded with a 24h Holter ECG. Moore et al. found that turbulence slope was as an independent predictor of death due to decompensated HF. The prognostic value of HRT was also analyzed by Cygankiewicz et al. [Cyg06]. 487 CHF patients of NYHA classes II and III were included in the study population. Abnormal turbulence onset and turbulence slope were evaluated and found to be independent predictor of NYHA class III or a LVEF < 40%. In an additional study, Cygankiewicz et al. [Cyg08b] found that abnormal turbulence slope (≤ 2.5ms/RR) and abnormal turbulence onset as well as abnormal turbulence slope were predictive for total mortality, sudden death and heart failure death. The study included 607 patients with a median follow up of 44 months. For both studies by Cygankiewicz et al. the data was recorded with a 24h, 3-lead Holter ECG. A more complete review and meta-analysis on the importance of HRT was performed by Disertori et al. [Dis16] in 2016. The authors showed that information gained by evaluating HRT are important in predicting total mortality, cardiac death and arrhythmic events in HF populations.

Heart Rate Variability

The prognostic value of HRV is well established. Nolan et al. [Nol98] associated SDNN, the standard deviation of all normal-to-normal RR intervals in the entire 24h recording, < 100 ms with higher all-cause mortality in HF patients. Additionally, SDNN was found to be an independent predictor of death due to progressive HF. The data was extracted from 24h ambulatory ECGs of 433 patients. Similar findings were made by Galinier et al. [Gal00]. They found a decreased SDNN to be an independent prognostic value for all-cause death and death due to progressive heart failure. Additionally, they showed the predictive value of lower daytime low frequency (LF) power for sudden death. Short term HRV (8-minutes) was analyzed by La Rovere et al. [La 03]. The authors showed that reduced low frequency power (< 13 ms^2) independently predicts sudden death in 444 CHF patients with a follow up of three years. Aronson et al. [Aro04] evaluated the 24h, 3-lead ECG data of 199 patients with decompensated HF and found higher mortality rates for patients with SDNN < 44 ms

and standard deviation of all 5-minute mean RR intervals (SDANN) < 37 ms. Patients with total power < 1.475 ms^2 and ultra-low frequency power < 1.100 ms^2 were at increased risk of death.

Oxygen Saturation

Weatherley et al. [Wea09] investigated the state of 337 patients, hospitalized for acute HF. They followed the patients for a mean of six months. Weatherley et al. associated early worsening of HF with lower oxygen saturation at admission. Masip et al. [Mas12] evaluated oxygen saturation for discriminating between patients with acute myocardial infarction with and without acute HF. 192 patients were included in the study with a mean follow up of 13 months. Masip et al. suggested that a baseline oxygen saturation lower than 93% may be considered a signal of acute HF. Gálvez-Barrón et al. [Gál19] explored the effect of oxygen saturation on HF patients in stable and exacerbation state. First measurements were taken at the hospital admission of the patients. The measurements in stable phase were measured a minimum of 30 days after discharge if the patient was in stable condition. The measurements were preformed in rest and during walking. The authors showed, that there was a significant difference between stable and exacerbation phase for oxygen saturation. Additionally, they concluded that effort situations may help improve the discriminatory power.

T Wave Alternans

T wave alternanss (TWAs) describe a beat-to-beat variation in the amplitude of the ST segment and T wave and therefore, reflect abnormalities in ventricular repolarization.

Based on the Holter ECG recordings of 493 hospitalized post-myocardial infarction patients with heart failure and/or diabetes with LV dysfunction, Stein et al. [Ste08] reported higher TWA to be a powerful predictor of sudden cardiac death with a stratification of $> 47\mu$V. Monasterio et al. [Mon12] showed, that TWA predicts total mortality, cardiac death and sudden cardiac death. In the study by Yamada et al. [Yam18] TWA were found to be predictive of cardiac events at one year, in patients with HF. The authors calculated the data based on a 24h Holter ECG. Arsensos et al. [Ars18] investigated the predictive value of TWA for shorter recordings of 30 minutes in 146 HF patients. He found TWA to be an independent total mortality predictor. While the studies link higher TWA values to higher cardiac events or mortality, the cut-off value for stratification ranges from 42μV [Ars18] to 58.5μV [Yam18].

T Wave Morphology

Lin et al. [Lin09] evaluated different T-wave morphology parameters in heart failure patients with and without life-threatening ventricular arrhythmias. By evaluating a standard 12-lead ECG, they found a significantly higher lead dispersion in the patients with ventricular arrhythmias than in the patients without. Huang et al. [Hua09] looked at a standard 12-lead ECG of 650 systolic heart failure patients with a follow up period of around 2.7 years. They found total cosine between QRS and T-wave (TCRT) as a cardiovascular mortality predictor for a cut-off point of -0.473. Ramirez et al. [Ram17] analyzed 24h, 3-lead ECG data of 651 HF patients recruited by the MUSIC trial. They found T-wave morphology restitution (TMR) > 0.040 to be predictive of sudden cardiac death. In the study by Isaksen et al. [Isa21] a morphology combination score (MCS) based on asymmetry, notch and flatness of the T-wave, was calculated from a 10 second 12-lead ECG in 270039 patients. The patient group in the study included HF patients but was not limited to that group. The authors found MSC, as well as the three morphology components, predictive of mortality independent of HR, QTc and baseline comorbidities.

QRS Duration

In a review by Kashani et al. [Kas05] it was assessed, that patients with a prolonged QRS complex are considered to have a poorer prognosis of outcome. They established that as the LV function worsens, QRS duration increases. Kalra et al. [Kal02] showed in a study that included 155 CHF patients, that HF patients with QRS duration < 120 ms have significantly better prognosis for medium and long-term even-free survival. Wang et al. [Wan08] showed that for patients with reduced LVEF , who were hospitalized for worsening HF and showed a prolonged QRS duration (> 120 ms) had a higher post discharge mortality and readmission rate. Their study included 2962 patients with reduced LVEF and hospitalized for worsening HF. Rav-Acha et al. [Rav19] found a QRS prolongation > 130 ms to be significantly and independently associated with mortality and hospitalization in HF patients with left ventricular dysfunction and thereby supporting the previous findings.

QT dynamicity

Pathak et al. [Pat05] analyzed the QT dynamicity in 175 CHF patients of NYHA classes II – III. The mean follow up was 29.9 months and data were recorded with a 24h Holter ECG. The authors calculated the interval of the onset of Q wave to the apex of the T wave (QTa) and the interval of the onset of Q wave to the end of the T wave (QTe). They found the 24h QTe/RR slope > 0.28 predictive for total mortality and sudden death. Watanabe et al. [Wat07] included 121 CHF patients in their study with a mean follow up of 34 months. Additionally, to a steeper QT/RR slope, they also found a longer QT interval to be predictive of cardiac and sudden death. The importance of QTe/RR slope was also suggested by results from Cygankiewicz et al. [Cyg08a] and Ramirez et al. [Ram14]. Evaluating the data of 542 patients over a median period of 44 months, Cygankiewicz et al. found daytime QTe/RR slope > 0.22 to be a predictor for increased total and cardiac mortality. Ramirez et al. results support the findings by again showing the predictive value of an increased QTe/RR slope for cardiac as well as sudden death. A recent study by Arsenos et al. [Ars22] showed the prognostic value of the corrected QT interval (QTc) when extracted from a 30 min ECG. In 145 patients with a mean follow up of 42.1 months, QTc interval was an independent predictor of total mortality in patients with heart failure.

3.3 Validation of Smartwatch ECG Recordings

With the emerging trend of wearables, especially smartwatches, access to biomedical parameters, such as heart rate and oxygen saturation, is becoming more available. Smartwatches can provide meaningful patient-reported parameters, and are therefore useful tools. The integration of ECG measurement systems in commercial smartwatches, plays a major factor in self-monitoring of the heart. The potential of self-recorded ECGs has previously been shown with the identification of atrial fibrillation [Per19]. However, before ECGs can be used for clinical purposes, their clinical accuracy has to be determined.

The Apple Watch Series 4 was evaluated for clinical accuracy for healthy people with no known cardiac history by Saghir et al. [Sag20]. The smartwatch and 12-lead reference ECG were recorded consecutively to avoid interferences. The measurements were typically recorded less than 60 seconds apart. HR, RR interval, PR interval, QRS interval, QT interval

and QTc interval were analyzed and strong (mean difference < 20 ms) and moderate (mean difference < 40 ms) agreement was found. Overall, the authors concluded that the Apple Watch produces an accurate 1-lead ECG in healthy adults.

While smartwatches typically record Einthoven's first lead, attempts of recording other limb leads as well as precordial leads have been made. Behzadi et al. [Beh20] recorded, Einthoven's first, second and third lead with the Apple Watch Series 4 by placing the smartwatch on additional places of the body. Interval durations, amplitudes and polarity of the waves were evaluated. The 12-lead reference ECG and Apple Watch ECG were recorded consecutively. Strong correlation between the ECG recordings of Apple Watch and reference system were obtained for all of the examined segments in all three leads. Sprenger et al. [Spr22] analyzed the amplitude and duration of QRS complex, T wave, P wave, PR interval and QT interval in precordial leads V1 through V6. Strong and significant correlations were observed for amplitudes, duration and polarities. The research group showed that the precordial leads could be obtained from smartwatches with similar reliability to a standard ECG in people with sinus rhythm.

The previously mentioned studies evaluated the feasibility and reliability of smartwatch ECGs on healthy individuals above the age of 18. However, studies that explore different populations than healthy adults, have been reported. Spaccarotella et al. [Spa21] analyzed the QT and QTc interval in healthy individuals as well as patients with the diagnosis of acute coronary syndrome and patients with ST-elevation myocardial infarction in three leads. All participants were in sinus rhythm and the ECG tracings were recorded simultaneously. The authors assert that the Apple Watch Series 4 can accurately measure the QT interval in both the healthy and patient group. Kobel et at. [Kob22] conducted a study for evaluating the accuracy of the Apple Watch in children with and without congenital heart disease. PR duration, QRS duration, QT duration, QRS amplitude, and T wave amplitude, extracted from the smartwatch ECG, showed excellent correlation with the reference system.

While the majority of studies investigate the reliability and accuracy of the the Apple Watch, the Withings Scanwatch has also been evaluated in two recent studies [Man22] [Pen22]. The Withings algorithm can not only detect AF but also gives information about the ECG intervals in recordings with good quality. However, Mannhart et al. [Man22] concluded that the automatic algorithm for identifying intervals needs further improvement to be useful in for clinical applications. Pengel et al. [Pen22] manually evaluated the ECG intervals and showed that the QTc interval was often underestimated and only acceptable in 51% of

the participants. While PR interval and QRS duration had a strong correlation ($r = 0.84$ and $r = 0.83$ respectively) for high quality ECG recordings, the QRS complex was still underestimated on average.

Besides the study by Mannhart et al. [Man22], all studies investigated the accuracy of smartwatch ECGs by manually analyzing the signals. Therefore, the idea of smartwatch ECGs being used for automatic analysis has not been discussed. Furthermore, to the best of our knowledge, no literature evaluating the accuracy of smartwatch ECG tracings recorded on HF patients exists.

3.4 Research Goals

The goal of this master thesis is to extend the knowledge on how different parameters, extracted and calculated from commercial smartwatch data, can be used to monitor HF and detect decompensation early. Therefore, it is necessary to identify ECG parameters, that can predict a cardiac decompensation and can be recorded with commercial smartwatches. Additionally, the validity of the specified parameters needs to be assessed on heart failure patients. For this purpose, a study was conducted with the aim to assess the feasibility and reliability of 1-lead ECG recordings from smartwatch devices. Furthermore, the significance of the extracted ECG parameters has to be established. Those tasks make up the research goals of this thesis.

Chapter 4

Methods

4.1 Data Acquisition

To assess the accuracy and reliability of smartwatch ECG recordings for HF patients a study was conducted at the Machine Learning and Data Analytics Lab from August to October 2022. Participants were asked to obtain ECG recordings with two different smartwatches in a resting phase as well as in a recovery phase. This study was approved by the ethics committee of the Friedrich-Alexander-Universität (application number: 22-237-S) and complies with the declaration of Helsinki.

4.1.1 Study Population

In total, 26 participants (18 healthy and 8 HF patients) were recruited for the study. The mean age of the healthy participants was 55.9 years. 11 healthy participants were female. The mean age of the HF group was 68.9 years and four HF participants were female. Demographic data for all participants can be found in Table 4.1. The most common comborbidity of the HF participants was AF. Additionally, seven HF participants were previously hospitalized for AHF. An overview of comorbidities of the HF patients is shown in Table 4.2.

Healthy participants were recruited using electronic flyers, which were distributed via an online platform promoting neighborhood assistance. Additionally, some participants were personally invited to participate in the study. HF patients were mostly recruited by ProCarement GmbH, a company which develops a telemonitoring system for HF patients and a cooperation partner of this project. To check the participants eligibility for the study, participants were

informed about the exclusion criteria in advance. Exclusion criteria included an age below 22, electrical implants, poor skin integrity, pregnant, and breastfeeding women. Furthermore, healthy participants should have no prior diagnosed heart diseases. HF patients should have a HF diagnosis. Additionally, all participants had to give signed consent and had to be able to participate in all phases of the study. A compensation for participation was not provided. Participants had the right to terminate their participation in the study at any given point of time, causing them no disadvantages.

Table 4.1: Demographics of healthy and heart failure (HF) participants.

Demographics of enrolled participants (n = 26)		
	Healthy (n = 18)	HF (n = 8)
Age	55.9 ±19.7	68.9 ±9.5
BMI	23.7 ±2.4	26.6 ±4.2
Female	11 (61.1%)	4 (50%)
Smoker	0 (0%)	0 (0%)
Smartwatch Users	3 (16.7%)	5 (62.5%)
NYHA I	-	2 (25%)
NYHA II		6 (75%)

Table 4.2: Comorbidities of heart failure (HF) patients.

Baseline Characteristics of HF Patients	
Acute Heart Failure	7 (87.5%)
Atrial Fibrillation	7 (87.5%)
Coronary Heart Disease	2 (25%)
Chronic Obstructive Pulmonary Disease	1 (12.5%)
Diabetes Mellitus	2 (25%)
Heart Valve Defect	5 (62.5%)
High Blood Pressure	4 (50%)
Kidney Failure	1 (12.5%)
Sleep Apnea	2 (25%)

4.1.2 Study Components

ECG Recordings

Smartwatch ECG measurements were conducted using the Apple Watch Series 7 and the Withings Scanwatch. The Apple Watch recorded the ECG with a sampling rate of 512 Hz. The Scanwatch recorded the ECG with a sampling rate of 300 Hz. The smartwatches recorded Einthoven's lead I of a 12-lead ECG for 30 seconds. Information about how the signals were filter by an internal algorithm is not provided by Apple or Withings. The recordings were saved in the Apple Health App and in the Withings Health Mate App, respectively. A reference ECG signal was recorded with the NeXus Mind Media and the according BioTrace+

Figure 4.1: ECG recording setup. A participant wearing the Apple Watch (left wrist) and Withings Scanwatch (right wrist). The electrodes of the reference ECG are connected just below the collarbone.

software [Min22]. The signal was sampled at 2048 Hz, the highest sampling rate allowed by the software. Five electrodes were attached to the participant's upper body, to record Einthoven's first and second lead. To achieve simultaneous recordings, the reference ECG was recorded throughout the entire study sessions. For synchronization of signals, markers were set at specific time steps (start smartwatch ECG1, end smartwatch ECG1, ...).

Oxygen Saturation

Oxygen saturation was measured using both smartwatches. The score was determined with the internal algorithm of the corresponding watch. The Apple Watch computes the oxygen saturation based on a 15 second recording, while the Withings Scanwatch measures for 30 seconds. The value was shown on the watch's display immediately after the measurement concluded and was noted by the study supervisor. If the measurement failed, it was repeated no more than two times. If the measurement still failed on the third try, no value was noted and the study would move on. A reference oxygen saturation measurement was not performed. While the Withings Scanwatch uses a medical-grade SpO2 sensor, the blood oxygen measurements of the Apple Watch are not intended for medical use.

Six-Minute Walking Test

The six-minute walking test (6MWT) is an exercise test, that was first used to evaluate patients with chronic obstructive pulmonary disease and respiratory failure. In 1985, Guyatt et al. [Guy85] published the first study on the use of the 6MWT for chronic heart failure patients. Since then, it has been widely investigated in populations with HF [Gia19]. A statement by the American Thoracic Society on the guidelines for the 6MWT was made in 2002 [Cra02]. Those guidelines support the use of the 6MWT for the comparison of pretreatment and posttreatment, evaluation of functional status, and prediction of morbidity and mortality in not only subjects with respiratory diseases but also those with HF. It is considered a sub-maximal exercise test, that reflects the functional exercise level for daily physical activities.

The standardized approach is based on the statement by the America Thoracis Society [Cra02]. The object of the test is to walk as far as possible for six minutes without running or jogging. The test should be performed on a corridor of at least 30 meters. The patients are permitted to determine their own pace and slow down or stop at any given time. They are asked to resume walking as soon as they are able again. During the breaks the patients may lean against a wall. The test conductor should give words of encouragement on a minutely basis. After six minutes the test ends and the distance covered is measured.

In this study, participants were asked to walk on a round course of around 25 meters, due to the lack of a long corridor. They were allowed to chance direction at any given time. While the distance was noted, the main reason for including the 6MWT in the study protocol, was to include an exercise phase that activates the heart and reflects daily physical activities. HR was monitored during the 6MWT with all three recording devices. To get a more continuous HR measurement with the smartwatches, a workout mode was activated and HR was monitored via photoplethysmography. During workout mode, the Apple Watch sampled the HR with approximately 12 values per minute, while the Scanwatch sampled the HR with around 14 values per minute. The reference system recorded an ECG throughout the 6MWT.

4.1.3 Procedure

To ensure repeatability, a study protocol was designed. The study included a resting, an exercise and a recovery phase in which smartwatch ECG recordings were obtained. A graphical representation of the study procedure is shown in Figure 4.2.

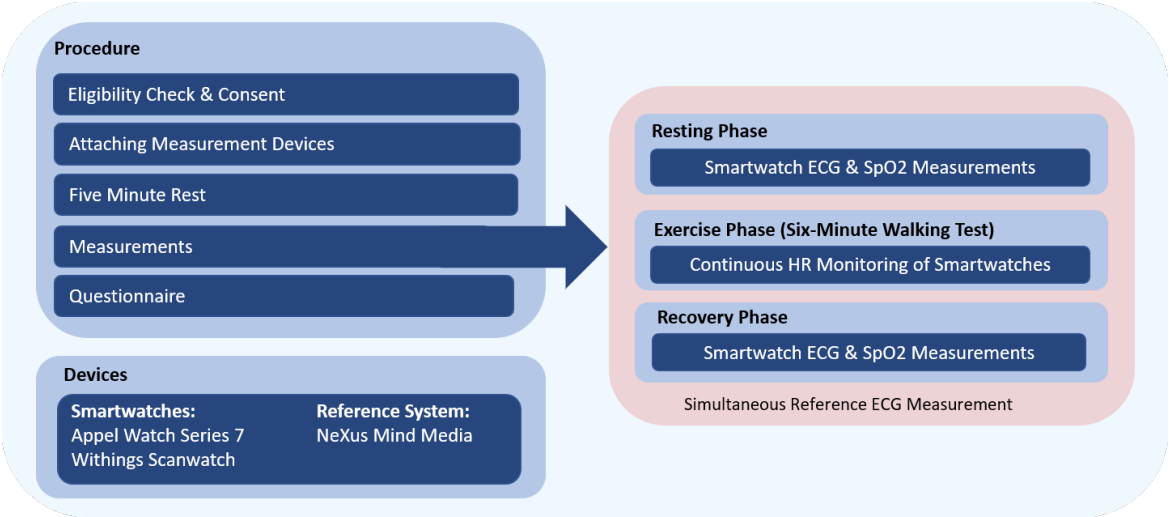

Figure 4.2: Schematic representation of the study procedure. The study included a resting, an exercise and a recovery phase.

After their arrival at the Machine Learning and Data Analytics Lab at the Friedrich-Alexander-University, the potential participants were welcomed, seated and asked to sign a declaration of consent. The study supervisor shortly explained the purpose and timeline of the study. The participants were introduced to the devices used for the measurements (Apple Watch Series 7, Withings Scanwatch and Nexus Mind Media). Once the participants were well informed, they were asked to attach both smartwatches to their wrists (left wrist: Apple Watch, right wrist: Withings Scanwatch). Specific skin regions were disinfected and prepared with special medical tape to ensure good electrode connection, before the electrodes for the reference system were attached to the participant's upper body. The reference ECG was recorded throughout the entire study session.

Participants were asked to relax for five minutes to ensure the body would be in a complete resting state. They were asked to not talk, not play on their phone or engage in any other distractions. Additionally, they were encouraged to close their eyes.

The first ECG was recorded with the Apple Watch. Immediately after, a second ECG tracing was obtained using the Scanwatch. Then, blood saturation was measured using Apple Watch and Scanwatch, respectively. Participants were then told to perform the 6MWT. The Nexus device for the reference ECG was fastened on the participant with a belt, so heart rate monitoring during the active session could be ensured. The workout 'Walking' was set on both watches. During the 6MWT, the study supervisor counted each lap and gave information about how much time was left and how much distance was covered. After six minutes, the study participants were asked to stop walking and stand where they had stopped. The distance was noted and workout session modes on the smartwatches were ended and saved. Participants were asked to remove the belt and sit down again. After 1.5 to 2 minutes, an ECG recording was obtained with the smartwatches. The order of the recordings obtained with Apple Watch and Withings Scanwatch alternated for each participant to ensure, that on average each ECG recording was acquired after the same amount of time had passed between recordings and walk. The same was done for the blood saturation measurements. All ECG and oxygen saturation measurements throughout the study followed the same recording protocol.

After the measurement the smartwatches were returned to the study supervisor and the electrodes were removed. The participants were asked to fill out a short online questionnaire, containing questions about demographics and smartwatch preferences. HF patients were additionally asked about their comorbidities and HF diagnosis. This included questions regarding time of first diagnosis, decompensations and symptoms. The study ended after completing the questionnaire. Participants were thanked for their participation.

4.2 Data Preprocessing

The smartwatch ECG signals were exported from the Apple Health App and the Withings Health Mate App and then manually saved in individual .csv files. The ECG signals recorded by all devices, including the reference system, were cleaned using a 0.5 Hz high-pass butterworth filter of order 5. This was followed by powerline filtering. The cleaned signal was used for the rest of the analysis.

The wave on-, offsets and peaks for the P, R and T waves were detected for each heartbeat of the ECG signal. In case the position of an on-, offsets, or peaks could not be determined, the value NaN was put down instead.

An outlier detection was performed for all wave onsets, peaks and offsets. This was done on three levels, as a statistical, a logical and a physiological outlier detection was conducted. The statistical outlier detection is based on the *z score*. The mean and standard deviation of the time difference from onsets, peaks and offsets to the R peak of the according heartbeat, was calculated. An outlier was detected if the *z score* was above a certain threshold. The threshold was set to 1.69, corresponding to the outer 5% of data assuming normal distribution. For the logical outlier detection, the onsets, peaks and offsets were compared to each other. This was done for P offsets and R onsets, and R offsets and T onsets. It was checked, whether the P offset occurred before the R onset and the R offset occurred before the T onset. Through visual evaluation it was determined, that the P offsets and R offsets were determined more accurately. Therefore the logical outlier detection aimed at catching wrongly determined R and T onsets. Additionally, it was checked, whether the onsets and offsets of the same wave occurred in the correct order. The final outlier detection was based on physiology. For this P onset, P peak, T peak and T offset outliers were detected. Various time differences were looked at and compared to normal ranges (PR interval: 0.12 - 0.20 seconds, QRS interval: 0.08 - 0.12 seconds and QT interval: 0.35 -0.43 seconds) [Hea]. If the calculated time difference exceeded 30% of the normal boarder ranges in either direction, that wave parameter was declared an outlier. For the wave to be included in the corresponding feature extraction and therefore in the analysis, neither onset, peak or offset were allowed to be classified as outliers.

The processing and segmentation of the signals were primarily based on *Neurokit2*, a Python toolbox for neurophysiological signal processing [Mak21] and *ECGdeli*, a Matlab toolbox for filtering and processing single or multilead ECGs [Pil20].

4.3 Feature Calculation

This section describes the feature calculation performed on the data recorded during the previously described study. The features used for this project are comprised of standard features extracted from ECG measurements and features with importance for HF patients. Standard features include, among others, wave duration and amplitudes. Features that have special importance for HF patients are based on literature findings. Those features were shown to have power in predicting HF decompensation and death. A review of such features has been outlined in Table 3.1. The described features are often calculated from a 24 hour ECG and multiple leads. As the smartwatch ECG is a 30 second, lead I ECG not all parameters were used for this analysis. Table 4.3 gives an overview which of the previously described features can be extracted from a smartwatch ECG.

Table 4.3: Parameters, extractable from a smartwatch ECG identifying and predicting decompensation in heart failure patients.

	Recording length (30 Second) is sufficient	**Lead 1 is sufficient**	**Included in Analysis**
HR Recovery	✓(max HR from PPG sensor)	✓(max HR from PPG sensor)	✓
HR Reserve	✓(max HR from PPG sensor)	✓(max HR from PPG sensor)	✓
HR Turbulence	premature ventricular contraction needed; unlikely in a 30 second recording	✓	✗
HR Variability	✓ SDNN, total power, LF power ✗ SDANN, Ultra LF Power	✓	✓ SDNN, total power, LF power
HR	✓	✓	✓
QRS Duration	✓	✓	✓
QT Dynamicity	✓ QT Duration, QTc Duration ✗ QT/RR slope (24 hours needed)	✓	✓ QT Duration, QTc Duration
T Wave Morphology	✓ Asymmetry, Notch, Flatness, Lead Dispersion, TCRT ✗ Morphology Resitution	✓ Asymmetry, Notch, Flatness ✗ Lead Dispersion, TCRT, Morphology Restitution	✓ Asymmetry and Flatness
TWA	in theory possible; effect is yet to be discussed	✓	✓

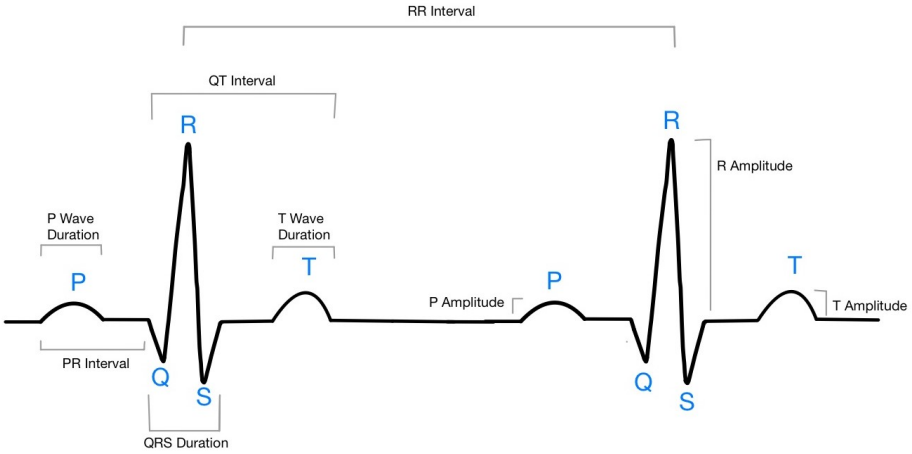

Figure 4.3: The ECG signal of a healthy human heart in sinus rhythm.

Heart Rate

The instantaneous HR or heart frequency was determined on the basis of the RR interval from one beat to the next. The mean, minimum and maximum heart rate were determined from the instantaneous HR. The mean HR was extracted for all 30 seconds recordings. The mean and maximum HR were computed for the 6 minute ECG recording of the reference system. The mean and maximum HR, recorded by the smartwatches during the 6MWT were manually derived from the downloaded health data.

Heart Rate Recovery and Reserve

HR Recovery was calculated as the difference between maximum HR during the 6MWT and the mean HR, obtained from the recordings in recovery phase. The recovery tracings were on average obtained at around 2 minutes after termination of the 6MWT. HR Reserve was calculated as the difference of maximal HR during the 6MWT and the resting HR, obtained as the mean HR from the ECG recording in the resting phase.

Heart Rate Variability

Extracted featured for the time domain include standard deviation of the NN interval (SDNN), square root of the mean of the squared successive differences between adjacent RR intervals (RMSSD), standard deviation of the successive differences between RR intervals (SDSD), proportion of RR intervals greater than 50 ms, out of the total number of RR intervals (pNN50) and HRV triangular index (HTI). Frequency domain parameters are LF, high frequncy (HFr),

low frequency normlalized (LFn), high frequncy normalized (HFn) and LF/HFr. Additionally the total power is calculated by addition of LF and HFr. The HRV parameters were caluclated using the *Neurokit2* Python toolbox [Mak21]. For the spectral density estimation the Lomb-Scargle method was used as this yields more reliable results for ECG recordings of 30 seconds [Weh21].

PR Interval

The PR interval is measured from the onset of the P wave to the onset of the Q wave. It is therefore also referred to as the PQ interval.

QRS Area

The QRS area is calculated as the integral from the onset of the Q wave to the offset of the S wave. It gives the net deflection of the QRS complex, which can be net positive or net negative.

QRS Duration

The QRS duration is measured from the beginning of the Q wave to the end of the S wave.

QT Interval

The QT interval is measured from the beginning of the QRS complex to the end of the T wave. Since the QT interval is a time duration interval, it is influenced by the R-R cycle length. To get meaningful values for the repolarization duration analysis, adjustments for heart rate have to be made. Various formulae have been developed to estimate the QT interval at a standard heart rate of 60 bpm, known as the corrected QT (QTc) interval. For this project the correction was done using the commonly used, exponential correction formula, known as Bazett formula [Baz97].

$$QTc = \frac{QT}{\sqrt{RR}} \tag{4.1}$$

T Wave Alternans

TWA were calculated using the modified moving average (MMA) method as described by Nearing and Richard [Nea02]. Alternatively, TWA can be calculated using a spectral method as done in by Sarzi Braga et al. [Sar04] and Monasterio et al. [Mon12]. For computing

TWA using the spectral method, the data is usually recorded during bicycle or treadmill exercises. The MMA method was designed to circumvent the requirements of the spectral method (stabilization of heart rate for several minutes by means of exercising). Using the MMA method, the ECG can also be computed at rest [Nea02].

Heartbeats were separated into odd and even groups. Next, a MMA computed beat is formed, using the current MMA beat and the following ECG beat in the series. According to their difference, the next MMA beat's value changes. MMA beats are continuously computed for each group. TWA were then calculated as the maximum absolute value of the difference between the odd and even MMA computed beats within the ST segment and T wave for each 15 second interval. All formulae used for the calculation are thoroughly described in [Nea02].

T Wave Morphology

Parameters reflecting the morphology of the T wave were based on the work of Isaksens et al. [Isa21].

Asymmetry refers to how much the peak of the T wave is shifted to the left or right. It is measured as the difference in the slopes of the ascending and descending part of the T wave according to Formula 4.2.

$$Asymmetry = \frac{\sum_{t_0}^{T}(asc'(t) - desc'(t))^2}{T - t_0} \tag{4.2}$$

$asc'(t)$ describes the slope of the rising part of the T wave, while $desc'(t)$ describes slope of the falling part. T indicates the maximal number of time steps, and t the discrete time steps. t_0 is first time step for which both $asc'(t)$ and $desc'(t)$ are defined.

The Flatness score [Isa21] describes how peaked or flattened the T wave is. High scores are found for flattened T waves while low scores describe peaked T waves. It is calculated using the second and fourth central moment as described in Formula 4.3.

$$Flatness = 1 - \frac{M_4}{M_2^2} \tag{4.3}$$

M_2 and M_4 correspond to the second and fourth central moment, respectively.

Wave Amplitude
Wave amplitudes are calculated as the difference from the baseline to the peak of the wave. P wave, R wave and T wave amplitude were evaluated.

Wave Duration
Wave durations were calculated as the time between their on- and offset. P wave and T wave durations were evaluated.

4.4 Evaluation

4.4.1 Exclusion of Data

ECG recordings, for which the smartwatch detected AF were not included in the analysis. This decision was based on the rhythm notification of the smartwatches [App22]. Therefore, one Apple Watch recording in recovery phase and one Withings Scanwatch recording in recovery phase, from two different healthy participants needed to be excluded. All data from one HF patient were excluded, as the patient displayed AF in all performed recordings.

Additionally, the quality of the Nexus recordings was evaluated. Recordings with insufficient quality were excluded. This decision was made based on visual inspection by the author.

4.4.2 Correlation and Agreement

A correlation analysis was performed to assess the relationship between reference ECG parameters and smartwatch ECG parameters. The Spearman correlation was used, as the data were non-normally distributed. A correlation coefficient r was computed for each parameter. Additionally, the average correlation coefficient was computed for all parameters. To account for underestimation of the correlation, caused by the skewness of the sampling distribution of the correlation coefficients, Fisher's z transform was used prior to averaging. The correlation coefficient takes values between $r = -1$ and $r = 1$. A negative value represents a inverse correlation, while a positive value represents a direct correlation. According to common guidelines [Muk12], $r \leq |0.3|$ was considered poor or negligible correlation, r in range

$|0.3 - 0.5|$ was considered a low correlation, r in range $|0.5 - 0.7|$ was considered moderate correlation, $r \geq |0.7 - 0.8|$ was considered high correlation, and $r \geq 0.9$ was considered very high correlation. For this comparison, a high to very high positive correlation is desirable.

As the correlation coefficient is a measure of association rather than agreement, it is not a sufficient statistic for comparing an established measurement method with a new technique in the clinical field [Bla86]. Therefore, Bland-Altman [Bla86] analysis was performed to identify the relative bias and the Limits of Agreement (LoA). Small bias and small ranges of LoA are desirable. The bias is calculated by taking the mean of the differences between the two measuring techniques. LoA represent the limits for which 95% of all differences will lie in. They are calculated as the mean of differences ± 1.96 standard deviation of differences. LoA can help decide, whether a new measuring technique is sufficient enough. If the differences within the LoA are not clinically important, the measuring techniques can be used interchangeably. Acceptable ranges have to be decided based on the measured values and the individual clinical application.

4.4.3 Differences

Comparison of Correlations

To establish the significance of the correlation of the extracted parameters, comparisons of correlation coefficients were made. The average correlation of all parameters was compared between the watches (Apple Watch and Withings Scanwatch), the participants (healhty and HF), the recording phases (resting and recovery) and the ECG segmentation algorithms (Neurokit and ECGdeli).

Independent and dependent two-tailed t-tests were used for evaluating mean accuracy differences in those groups. Independent t-tests were used to determine whether there is a statistically significant difference between the means in two unrelated groups. Dependent t-tests were used for related groups. The tests were performed on Fisher's z transformed correlation coefficients to account for the skewness of the sampling distribution and get normally distributed data. The results of the t-test are reported as follows: t(degree of freedom) = t-statistic, p = p-value, d = cohen-d effect size.

Fisher's r to z transform was also used to assess whether individual parameters differed in correlation between individual groups. The test statistic z [Hin03] is described in Formula 4.4.

$$z = \frac{z\_r1 - z\_r2}{\sqrt{\frac{1}{n1-3} + \frac{1}{n2-3}}} \tag{4.4}$$

z_r1 is the z transform of the first correlation coefficient and z_r2 is the z transform of the second correlation coefficient. $n1$ and $n2$ are the sample sizes used to compute the correlations. The test z statistic was converted to the p-value, which is the value reported for this test.

A statistically significant difference is observed for $p < 0.05$. The following notation was used to indicate statistical significance in Figures and Tables: ns: not significant, * for $p < 0.05$, *** for $p < 0.01$ and *** for $p < 0.001$. As the study is highly exploratory and it is more important to avoid a type II error than a type I error, the level of significance and p-values were not adjusted for multiple comparisons [Arm14].

Assessment of Differences between Healthy and HF Participants

To assess the meaningfulness of individual parameters, it was investigated whether there were any differences between the parameters for HF and healthy participants. As most of the data were non-normally distributed and the groups were independent, a two-tailed Mann-Whitney-U test was used. This analysis was performed on parameters extracted from the smartwatch recordings.

A statistically significant difference is again observed for $p < 0.05$. data are reported Due to the highly exploratory character of this study, a correction for multiple comparisons was not performed. All statistical analyses was done using $\mathrm{pingouin}$, an open-source statistical package in Python [Val18] and are reported as follows: U-value, p = p-value, CLES = common language effect size (CLES).

Chapter 5

Results

5.1 Feasibility to Record ECG and SpO2

After a short introduction on how to use the smartwatches, no participants had problems conducting recordings on their own. Only four participants had recorded an ECG with a smartwatch prior to this study. All four of them were HF patients.

Seven of the HF participants stated that they could imagine using the watches at least once a week to monitor their heart activity. Six of them could imagine using it daily. 16 healthy participants said they could see themselves using the smartwatch ECG regularly, if they were diagnosed with a heart condition, 12 of them on a daily basis. Only one HF participant and two healthy participants stated that they could not see themselves using it. For reasons they stated data privacy concerns and a feeling of constant worry. One participant did not give a reason.

While all participants were able to perform the SpO2 measurements, the devices often gave inconclusive results. On the first try, the SpO2 recording failed for 10 participants (HF = 7, healthy = 3) in resting phase and for seven participants (HF = 3, healthy = 4) in the recovery phase using the Withings Scanwatch. Detailed results on how often the Apple Watch measurement failed cannot be given, as inconclusive measurements are not saved by the Apple Health app. The recording was repeated three times at the most for each phase. For one participant (HF) it was not possible to obtain an SpO2 measurement in resting phase using the Apple Watch and for two participants (both HF) using the Withings Scanwatch. For

one participant (HF) it was not possible in the recovery phase when using the Apple Watch. When measuring with the Withings Scanwatch the recording in recovery phase failed for two participants (HF = 1, healthy = 1).

5.2 Outlier Detection

Table 5.1 shows the results of the outlier detection. Four 30 second recordings were recorded by the Nexus device, two by the Apple Watch and two by the Withings Scanwatch for each participant. The outlier detection has greatest impact on the exclusion of R waves.

Table 5.1: Table showing the results of the outlier detection. The numbers of detected waves before and after outlier detection are given for P, R and T wave. The results are shown for recordings obtained by the Nexus device, the Apple Watch and the Withings Scanwatch.

Recording Device	Algorithm	P Wave (before - after)	R Wave (before - after)	T Wave (before - after)
Nexus	Neurokit	3537 - 2949	3275 - 2087	3442 - 3036
	ECGdeli	3775 - 3165	3775 - 2184	3620 - 3176
Apple Watch	Neurokit	1469 - 1231	1421 - 960	1631 - 1442
	ECGdeli	1770 - 1479	1770 - 1122	1668 - 1458
Withings Scanwatch	Neurokit	1773 - 1501	1655 - 1147	1693 - 1459
	ECGdeli	1757 - 1410	1757 - 1318	1654 - 1368

5.3 Correlation and Agreement

ECG Data in Resting and Recovery Phase

Example ECG signals recorded with the smartwatches and reference system in the resting phase, are shown in Figure 5.1. The signal was segmented into single heartbeats and is portrayed as the mean with it's standard deviation. The signal of one healthy participant as well as of one HF participant is shown.

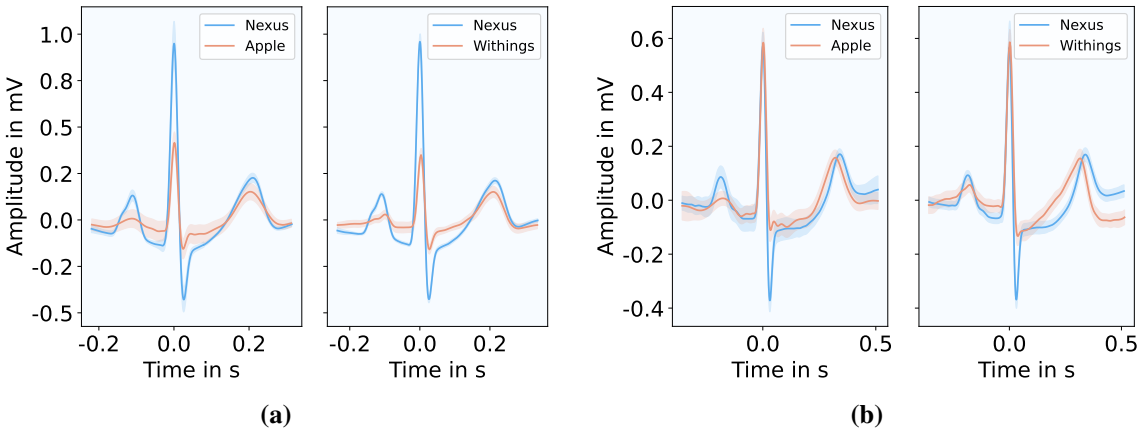

Figure 5.1: Mean heartbeats and their standard deviation of Nexus and smartwatch (Apple Watch (Apple), Withings Scanwatch (Withings) recordings. ECG tracings from two different participants (a) healthy participant (b) heart failure participant

The mean correlation coefficient betweeen Apple Watch and reference recording for all parameters over all phases and participants is $r = 0.51$ for ECG segmentation performed with the Neurokit algorithm, and $r = 0.49$ for ECG segmentation performed with the ECGdeli algorithm. The mean correlation coefficient between Withings Scanwatch and reference recording is $r = 0.40$ using the Neurokit algorithm, and $r = 0.36$ using the ECGdeli algorithm. This corresponds to a low or moderate positive overall correlation according to common guidelines [Muk12].

The mean correlation coefficient for all parameters, excluding amplitude and amplitude related parameters (P amplitude, R amplitude, T amplitude, T asymmetry, T flatness, TWA and QRS area), over all phases and participants is slightly higher but still corresponds to a low or moderate positive correlation (Apple-Neurokit: $r = 0.62$, Apple-ECGdeli: $r = 0.59$, Withings-Neurokit: $r = 0.50$ and Withings-ECGdeli: $r = 0.44$). The segmentation algorithms were compared on the whole dataset using the dependent t-test. Despite achieving a higher mean correlation using the Neurokit algorithm for both the Nexus-Apple Watch correlation and the Nexus-Withings Scanwatch correlation, there was no significant effect for both watches (Comparison of correlation computed on Nexus-Apple Watch recording: $t(24) = 0.29$, $p = 0.77$, $d = 0.05$, Comparison of correlation computed on Nexus-Withings Scanwatch: $t(24) = 0.89$, $p = 0.37$, $d = 0.14$). Therefore it is not clear which algorithm

is more accurate and the following analysis was performed for both. To make this section more readable, following abbreviations will be used to refer to the individual smartwatch recordings segmented with different algorithms: Apple Watch recordings segmented with the Neurokit or ECGdeli algorithm will be referred to as Apple-Neurokit and Apple-ECGdeli, respectively. Correspondingly, ECG recordings from the Withings Scanwatch will be referred to as Withings-Neurokit and Withings-ECGdeli.

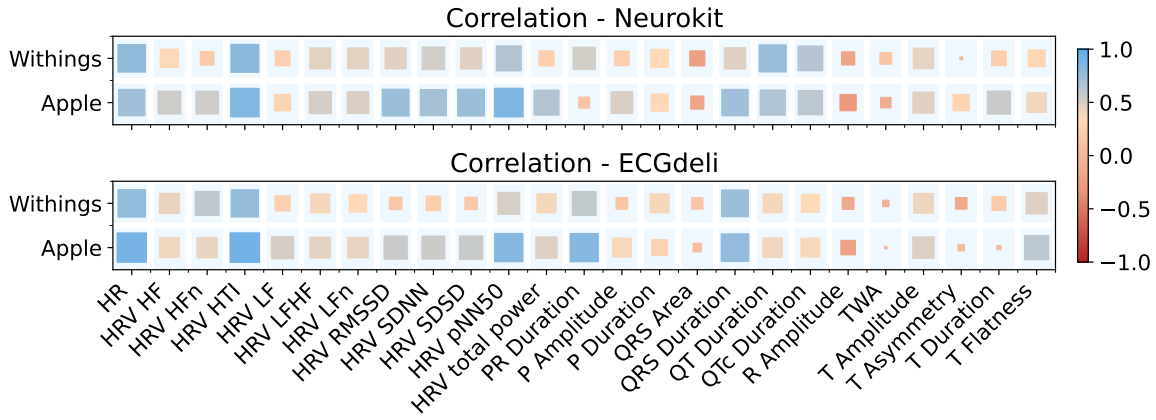

Figure 5.2: Spearman correlation coefficients of the individual parameters extracted from the ECG recordings. The coefficients were calculated for all recorded parameters using different ECG segmentation algorithms (Neurokit, ECGdeli) and watches (Apple Watch (Apple) and Withings Scanwatch (Withings)). Correlation coefficient is color and size coded. Blue squares indicate a moderate to very high positive correlation.

Correlation coefficients of the individual parameters can be seen in Figure 5.2. The correlation coefficients of smartwatch and reference parameters are shown over all samples. The size of the squares in the Figures 5.2 were mapped to represent the absolute correlation coefficient value, while the color represents the actual correlation coefficient value. Blue squares indicate a moderate to very high positive correlation between the parameters.

Features for which an overall high to very high positive correlation ($r \geq 0.7$) could be observed, include HR (for Apple-Neurokit, Apple-ECGdeli, Withings-Neurokit and Withings-ECGdeli), HRV HTI (for Apple-Neurokit, Apple-ECGdeli, Withings-Neurokit and Withings-ECGdeli), HRV RMSSD (only for Apple-Neurokit), HRV SDNN (only for Apple-Neurokit), HRV SDSD (only for Apple-Neurokit), HRV pNN50 (for Apple-Neurokit and

Apple-ECGdeli), PR Duration (only for Apple-ECGdeli), QRS Duration (for Appel-Neurokit, Apple-ECGdeli and Withings-ECGdeli) and QT Duration (only for Withings-Neurokit).

HR, QRS Duration and QT Duration, features with high to very high positive correlation are shown in scatter plots (Figure 5.3 and Figure 5.4) to visualize their correlation. Bland-Altman plots show their mean differences (bias) and limits of agreements (LoA). Figure 5.3 shows the scatter and Bland-Altman plots for the features computed from the Nexus and Apple Watch recordings, while Figure 5.4 shows the same for the features computed with the Nexus and Withings Scanwatch recordings. A full table of mean values, correlation coefficients, bias and LoA for all parameters extracted with different segmentation algorithms and recorded with different watches can be found in Appendix B.1.

Figure 5.2 additionally suggests, that some parameters extracted from the Apple Watch recordings have a higher correlation than the parameters extracted from the Scanwatch recordings. Individual parameters that achieved a statistically significant difference between the two watches were HRV pNN50 ($p = 0.024$), HRV RMSSD ($p = 0.038$), HRV SDSD ($p = 0.038$), HRV total power ($p = 0.019$) when using Neurokit and HRV HTI ($p = 0.029$), HRV RMSSD ($p = 0.034$), HRV SDSD ($p = 0.029$) and HRV pNN50 ($p = 0.005$) when using ECGdeli.

The dependent t-test was used to assess whether a difference in overall correlation between Apple Watch and Withings Scanwatch could be observed. An overall statistical significant difference ($t(17) = 2.58$, $p = 0.02$, $d = 0.57$) could be observed for the comparison of Apple Watch (mean correlation $r = 0.61$) and Withings Scanwatch (mean correlation $r = 0.50$) when when using the Neurokit segmentation algorithm. A significant difference ($t(17) = 3.25$, $p = 0.004$, $d = 0.56$) was also observed between Apple Watch (mean correlation $r = 0.59$) and Withings Scanwatch (mean correlation $r = 0.44$) when using the ECGdeli algorithm. Amplitude and amplitude related parameters as well as 6MWT data were excluded.

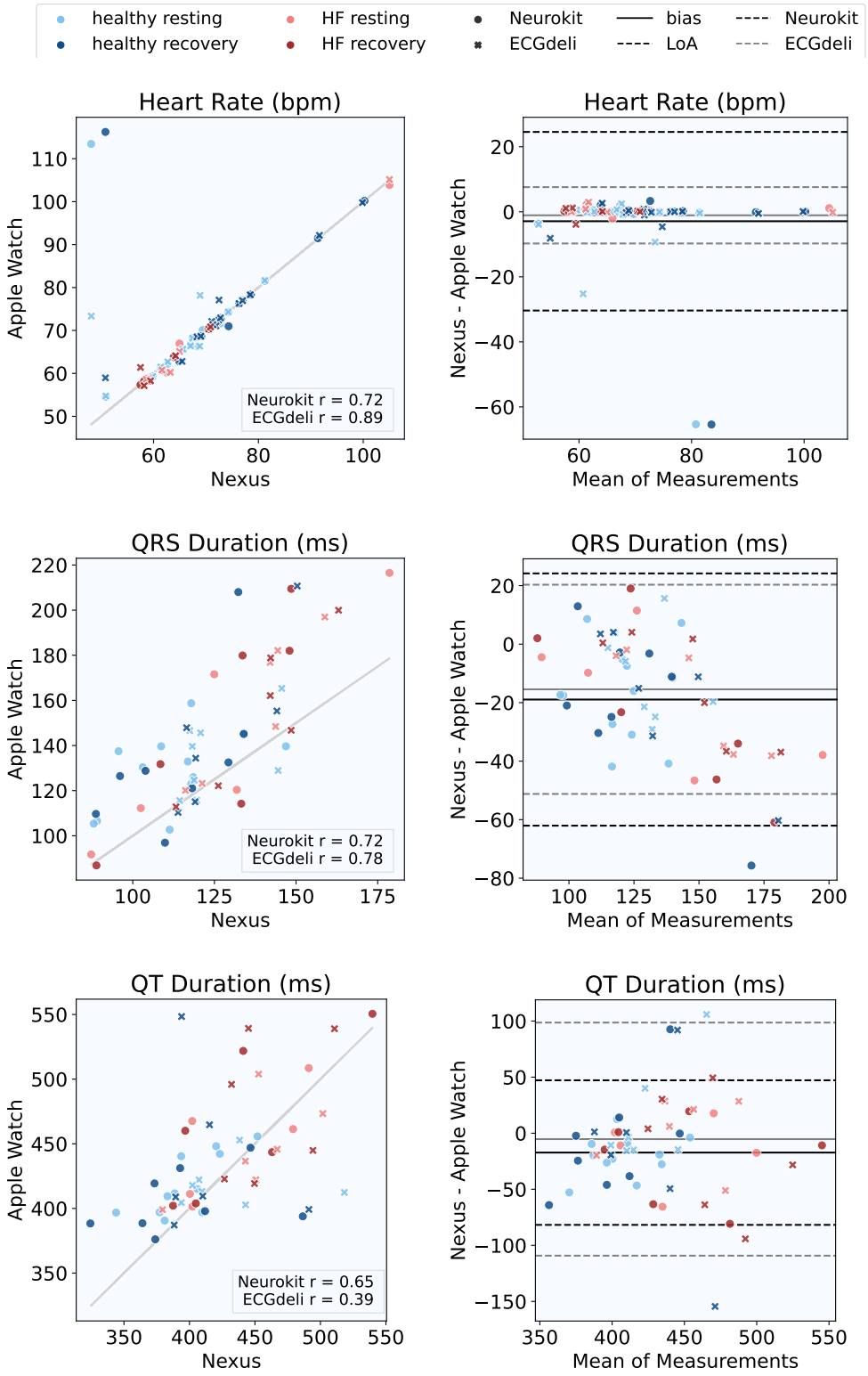

Figure 5.3: Scatter and Bland-Altman plots for comparison of Nexus and Apple Watch for heart rate and QRS and QT duration. Calculated from ECG recordings, using two different segmentation algorithms. Left: Scatter plot including identity line (grey) and Spearman correlation coefficient r. Right: Bland-Altman plots including bias (solid line) and Limits of Agreement (LoA) (dashed line).

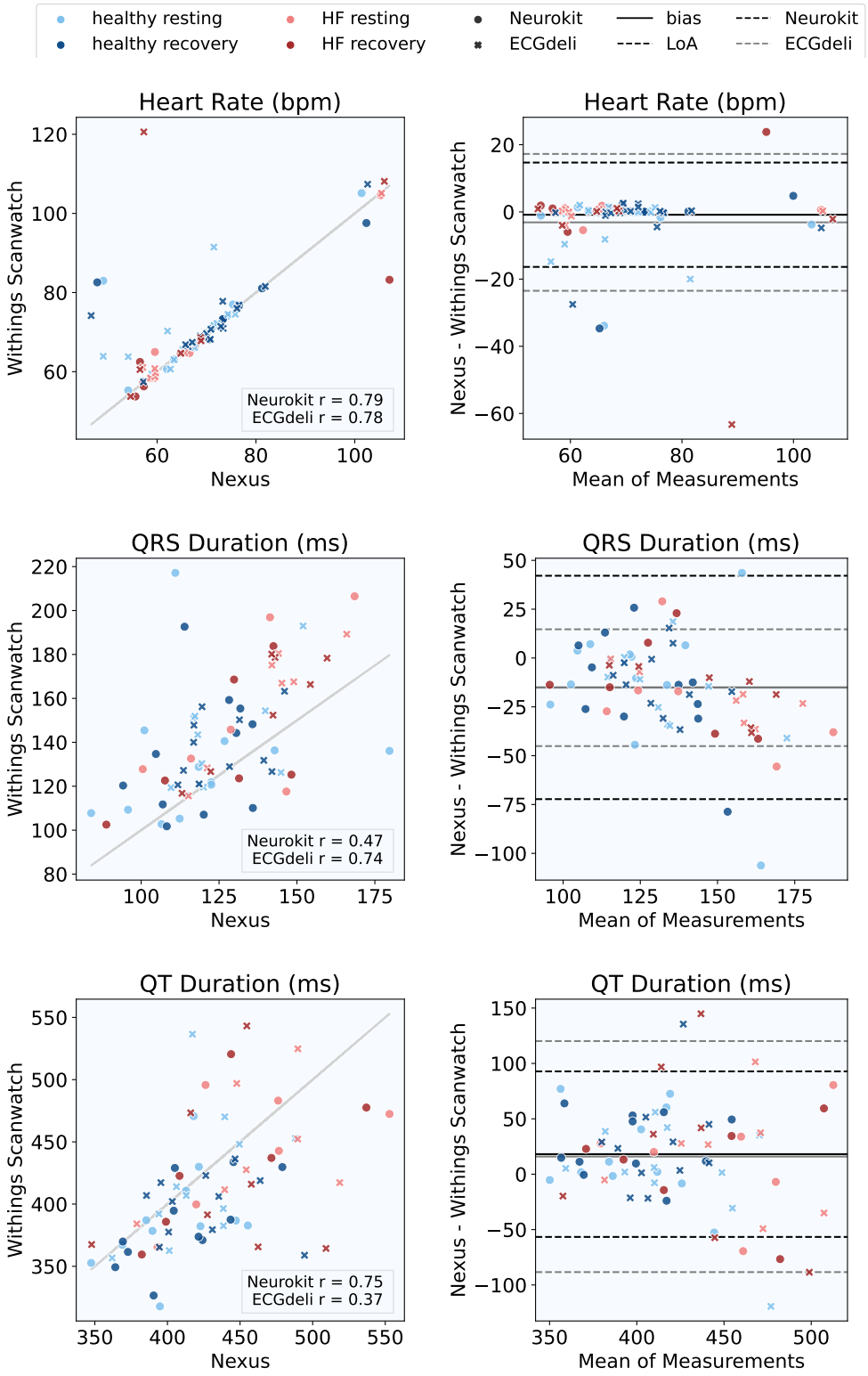

Figure 5.4: Scatter and Bland-Altman plots for comparison of Nexus and Withings Scanwatch for heart rate and QRS and QT Duration. Calculated from ECG recordings, using two different segmentation algorithms. Left: Scatter plot including identity line (grey) and Spearman correlation coefficient r. Right: Bland-Altman plots including bias (solid line) and Limits of Agreement (LoA) (dashed line)

Six Minute Walking Test Data

The mean HR and maximum HR were extracted during the 6MWT. Additionally, HR Reserve and HR Recovery were calculated. Those parameters were not included in the general analysis and are handled separately, as the HR measurements of the smartwatches do not stem from an ECG recording, but were measured through a photoplethysmogram (PPG) sensor during the walk.

The mean HR during the walk shows very high correlation between Apple Watch and Nexus recording for both algorithms (Neurokit: $r = 0.99$, $p = 2.36 \times 10^{-20}$, 95%CI $[0.97, 0.99]$, ECGdeli: $r = 0.99$, $p = 5.42 \times 10^{-21}$, 95%CI $[0.98, 1.0]$). For the Withings recording, the mean HR shows a low correlation for both algorithms (Neurokit: $r = 0.42$, $p = 0.04$, 95%CI $[0.02, 0.7]$, ECGdeli: $r = 0.40$, $p = 0.06$, 95%CI $[-0.01, 0.69]$). The maximum HR achieved a low correlation between Apple Watch and Nexus recording when segmented with the Neurokit algorithm ($r = 0.43$, $p = 0.03$, 95%CI $[0.04, 0.7]$) and a high correlation when segmented with the ECGdeli algorithms ($r = 0.76$, $p = 1.9 \times 10^{-5}$, 95%CI $[0.51, 0.89]$). However, the scatter plot in Figure 5.5 suggests that the agreement is also low when segmented with ECGdeli. A poor correlation between Withings Scanwatch and Nexus recording can be observed for the maximum HR for both algorithms (Neurokit: $r = 0.0$, $p = 4.93 \times 10^{-2}$, 95%CI $[-0.4, 0.4]$, ECGdeli: $r = 0.31$, $p = 0.14$, 95%CI $[-0.11, 0.63]$).

HR Reserve and HR Recovery also show poor to moderate correlations for all recordings. An exception is the correlation between Apple Watch and Nexus for HR Reserve when segmented with the ECGdeli algorithm ($r = 0.70$, $p = 1.97 \times 10^{-4}$, CI95% $[0.41, 0.86]$). However, based on the bias (83 bpm) and LoA (lower limit = 27 bpm, upper limit = 138 bpm) a high agreement is not suggested. A full table of correlation coefficients, mean differences and LoA can be found in Appendix B.2.

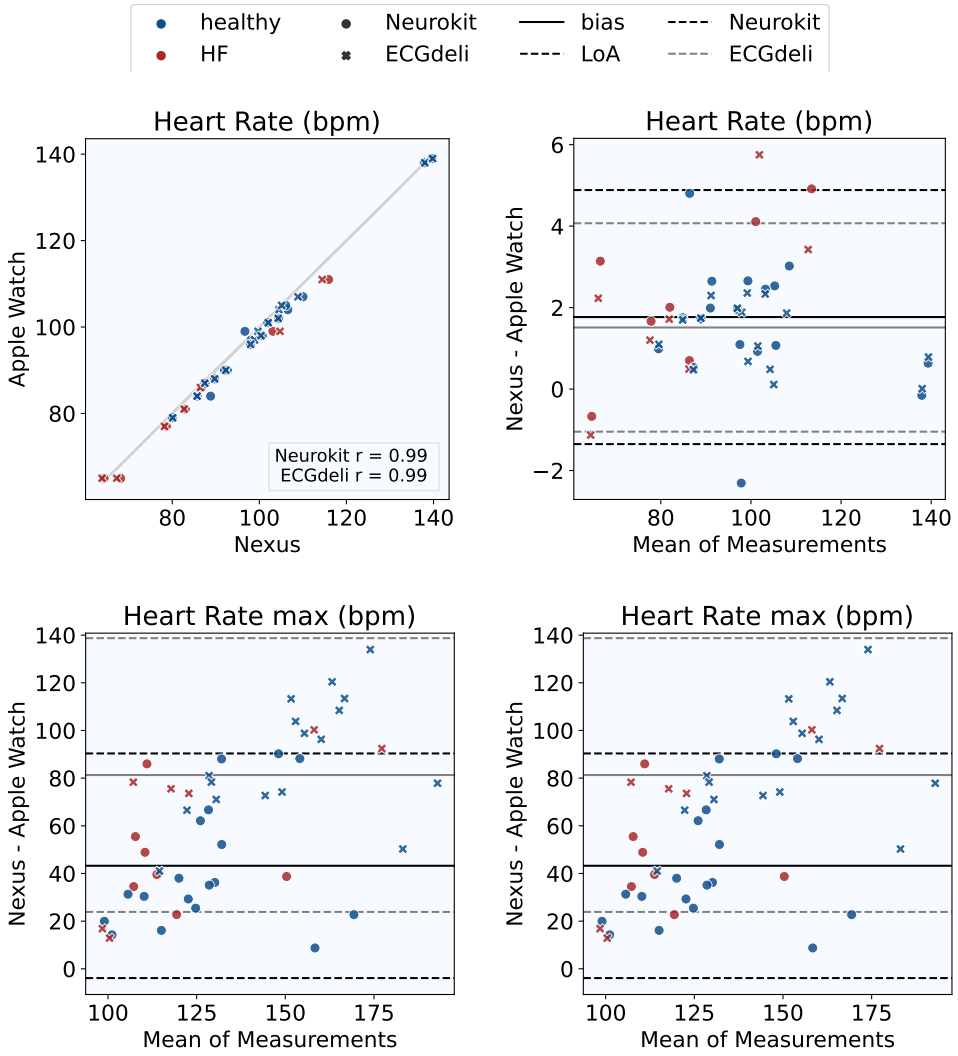

Figure 5.5: Scatter and Bland-Altman plots for comparison of Nexus and Apple Watch of mean and maximum heart rate recorded during the six-minute walk. Calculated using two different segmentation algorithms. Left: Scatter plot including identity line (grey) and Spearman correlation coefficient r. Right: Bland-Altman plots including bias (solid line) and Limits of Agreement (LoA) (dashed line).

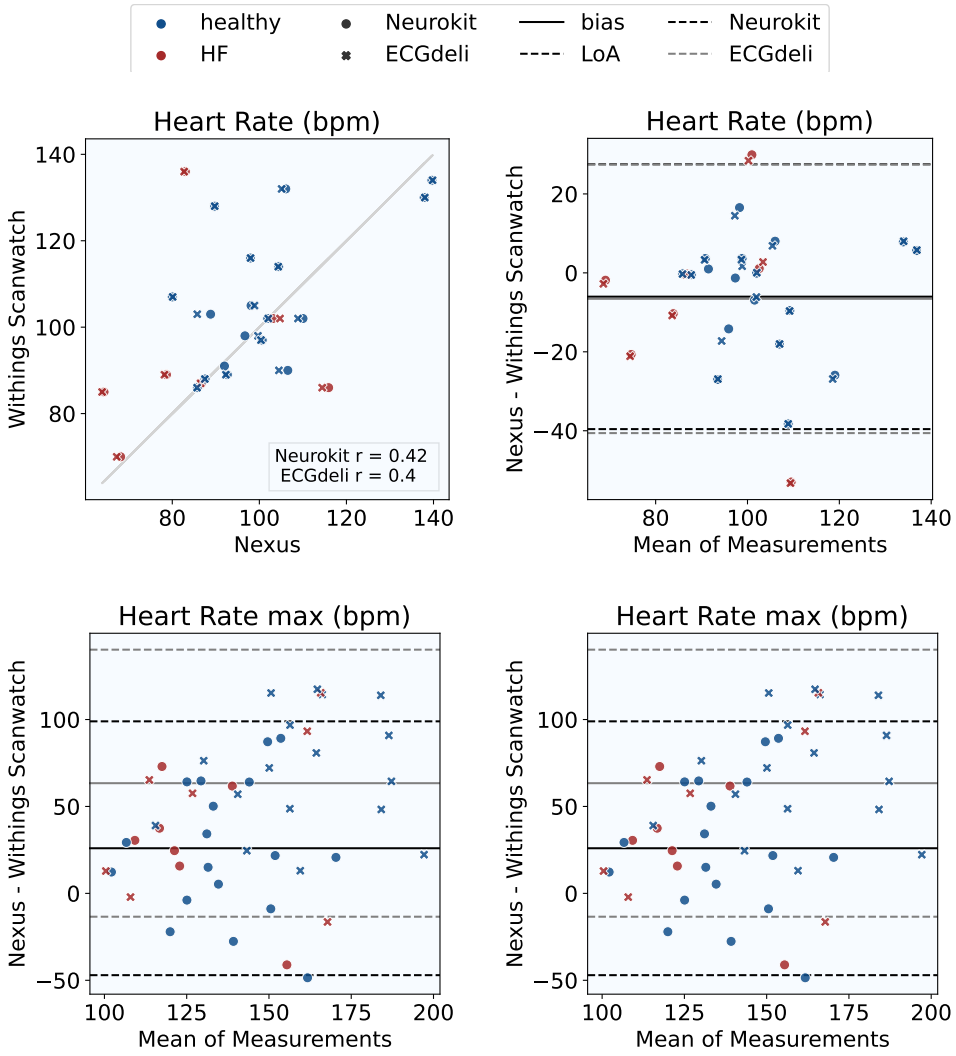

Figure 5.6: Scatter and Bland-Altman plots for comparison of Nexus and Withings Scanwatch of mean and maximum heart rate recorded during the six-minute walk. Calculated using two different segmentation algorithms. Left: Scatter plot including identity line (grey) and Spearman correlation coefficient r. Right: Bland-Altman plots including bias (solid line) and Limits of Agreement (LoA) (dashed line).

5.4 Differences

Comparison of Correlations

Breaking down the data set leads to the comparison of various groups. The individual correlation coefficients were calculated when differentiating between healthy and HF participants (see Figure 5.7). A clear trend for which group (healthy vs HF) achieved stronger correlation cannot be observed. However, individual parameters show significant differences in correlation. For Apple-Neurokit a significant difference in correlation between healthy and HF could be found for HR ($p = 0.0002$), P Duration ($p = 0.011$), QRS Duration ($p = 0.035$), and T Duration ($p = 0.004$). HRV pNN50 ($p = 0.013$), HRV HFr ($p = 0.002$), and HRV total power ($p = 0.002$) were significantly different for Apple-ECGdeli. P Duration ($p = 0.020$), R Amplitude ($p = 0.006$), and HRV SDNN ($p = 0.010$) for Withings-Neurokit and R Amplitude ($p = 0.002$) for Withings-ECGdeli. As can be seen in Figure 5.7, not all features have a higher correlation when extracted from the ECG signal of healthy participants. A full set of correlation coefficients, bias and LoA can be found in Appendix B.3 and B.4.

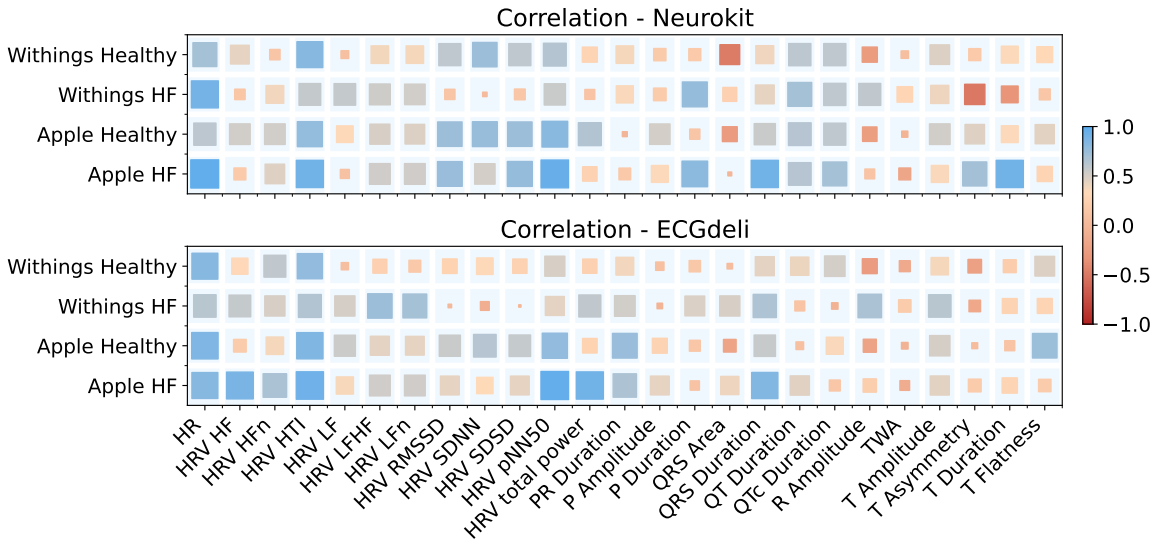

Figure 5.7: Spearman correlation coefficients of the individual ECG parameters for different participant groups (heart failure (HF) and healthy). The coefficients were calculated using different ECG segmentation algorithms (Neurokit and ECGdeli) and watches (Apple Watch (Apple) and Withings Scanwatch (Withings)). Correlation coefficient is color and size coded. Blue squares indicate a moderate to very high positive correlation.

The independent t-test was used, to evaluate the overall correlation difference between healthy and HF participants. Amplitude and amplitude related features and 6MWT data were excluded. While the HF group achieved an higher average correlation over the parameters ($r = 0.72$) compared to the healthy group ($r = 0.57$), a significant difference was not observed ($t(34) = -1.6$, $p = 0.12$, $d = 0.53$) for Apple-Neurokit. When comparing the correlations of the parameters for Apple-ECGdeli, a higher mean correlation was found for the HF group ($r = 0.66$) compared to the healthy group ($r = 0.52$) but showed no significant effect ($t(34) = -1.46$, $p = 0.15$, $d = 0.49$). For Withings-Neurokit, the healthy group showed a higher average correlation over the parameters ($r = 0.49$) compared to the HF group ($r = 0.44$). A significant effect was not observed ($t(34) = -0.48$, $p = 0.63$, $d = 0.16$). The average value of the correlation was not significantly different ($t(34) = -0.39$, $p = 0.70$, $d = 0.13$) between the healthy group ($r = 0.40$) and the HF group ($r = 0.43$) for Withings-ECGdeli.

Furthermore, the correlation coefficients were computed when differentiation between resting and recovery phase (see Figure 5.8). Slightly higher correlations for parameters extracted from an ECG recorded in resting phase, especially when segmenting with the ECGdeli algorithm, are suggested. However, individual features also achieved higher correlation coefficients in the recovery phase. For Apple-Neurokit a significant difference in correlation between resting and recovery phase could be found for HRV SDNN ($p = 0.015$). HR ($p = 0.002$) and HRV pNN50 ($p = 0.001$) were significantly different for Apple-ECGdeli, HRV SDNN ($p = 0.006$), HRV RMSSD ($p = 0.030$) and HRV SDSD ($p = 0.025$) for Withings-Neurokit and QT Duration ($p = 0.014$), HRV HFn ($p = 0.038$) and HRV LF ($p = 0.018$) for Withings-ECGdeli. The complete correlation coefficients for each parameter in the different recording phases can be found in Appendix B.5 and B.6.

A dependent t-test was performed to compare the correlation coefficients of the parameters between the resting group and the recovery group. Amplitude and amplitude related parameters and 6MWT data were excluded. For Apple-ECGdeli, the average correlation of the resting group ($r = 0.69$) was significantly higher ($t(17) = 2.31$, $p = 0.034$, $d = 0.51$) than the average correlation of the recovery group ($r = 0.52$). A significant difference ($t(17) = 3.18$, $p = 0.006$, $d = 0.79$) was also found between the resting group ($r = 0.54$) and recovery group ($r = 0.33$) for Withings-ECGdeli. The resting group showed a lower

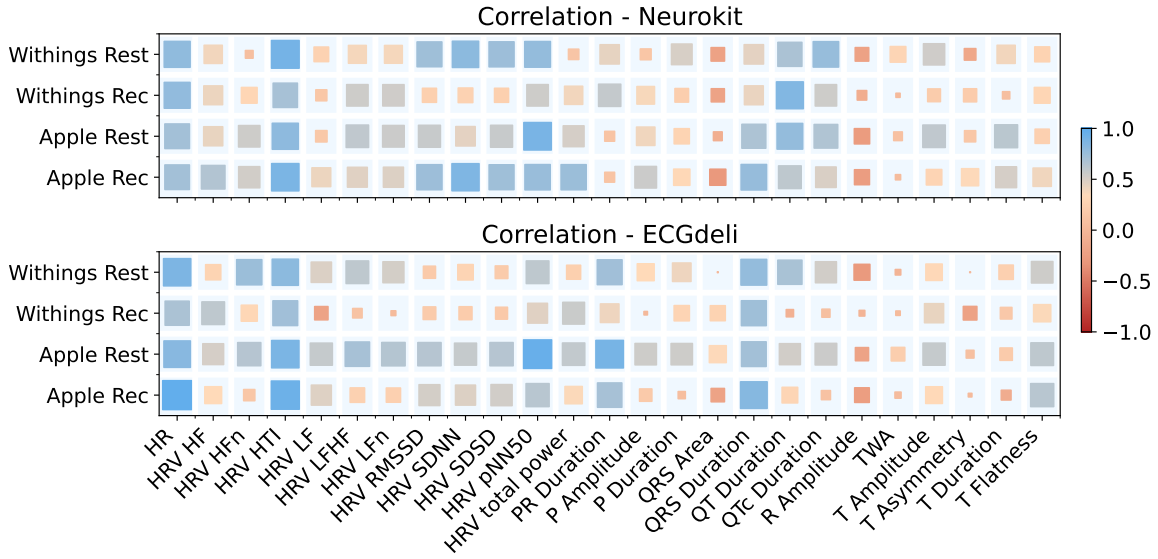

Figure 5.8: Spearman correlation coefficients of the individual ECG parameters for different recording phases (rest and recovery (rec)). The coefficients were calculated using different ECG segmentation algorithms (Neurokit and ECGdeli) an watches (Apple Watch (Apple) and Withings Scanwatch (Withings). Correlation coefficient is color and size coded. Blue squares indicate a moderate to very high positive correlation.

average correlation ($r = 0.58$) than the recovery group ($r = 0.63$) for Apple-Neurokit, but there was no significant effect ($t(17) = -1.12$, $p = 0.28$, $d = 0.25$). Despite finding a higher average correlation of the parameters for the resting group ($r = 0.58$) than in the recovery group ($r = 0.33$) for Withings-Neurokit a significant difference was not found ($t(17) = 1.77$, $p = 0.09$, $d = 0.46$).

Assessment of Differences between Healthy and HF Participants

A subset of different parameters showed a statistically significant difference in measurements between healthy and HF participants when analyzed with the Mann-Whitney-U test on the individual smartwatch recordings. Statistical significance depended on the segmentation algorithm, the phase and the watch. An overview can be found in Table 5.2. Boxplots of the statistically significant features can be found in Appendix C.

The data extracted from the 6MWT (HR mean, HR max, HR Reserve and HR Recovery) were also compared between the two different groups. Differences could be found for HR

mean when recorded with the Withings Scanwatch (Neurokit: $U - value = 96.0$, $p = 0.049$, CLES $= 0.76$ median healthy: 103 bpm and HF: 87 bpm, ECGdeli: $U - value = 91.0$, $p = 0.049$, CLES $= 0.76$, median healthy: 103 bpm and HF: 87 bpm). Additionally, a statistically significant difference was found for HR Recovery when recorded with the Apple Watch and segmented with ECGdeli ($U - value = 79.0$, $p = 0.021$ CLES $= 0.82$, median healthy: 33 bpm and HF: 21 bpm). Boxplots of the parameters can be found in Appendix C.

Oxygen saturation was compared between healthy and HF participants as shown in Figure 5.9. No significant difference could be seen in resting phase (Apple: $U - value = 65.0$, $p = 0.93$, CLES $= 0.51$ median healthy: 97% and HF: 97%, Withings: $U - value = 61.0$, $p = 0.66$, CLES $= 0.56$, median healthy: 98% and HF: 97.5%) or in recovery phase (Apple: $U - value = 71.5$, $p = 0.62$, CLES $= 0.57$, median healthy: 97.5% and HF: 97%, Withings: $U - value = 72.0$, $p = 0.14$, CLES $= 0.71$ median healthy: 98% and HF: 97%).

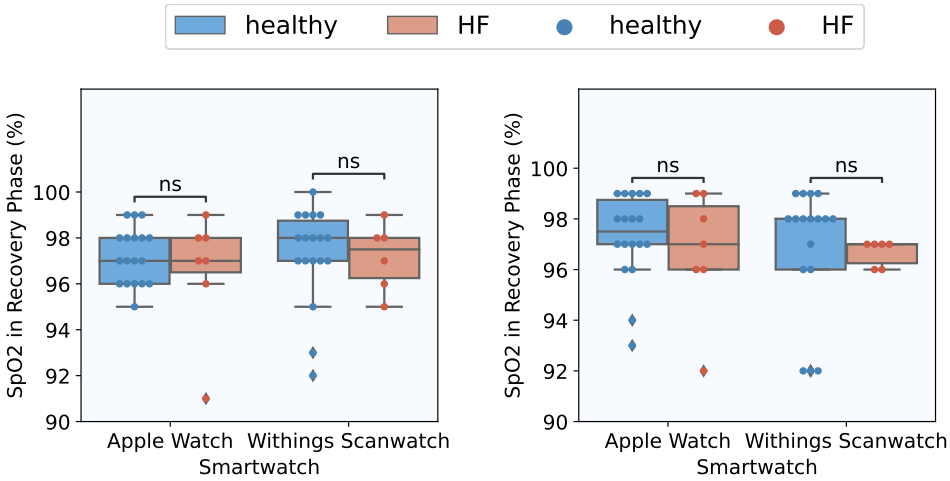

Figure 5.9: SpO2 in resting and recovery phase. Dots present the measurements of the individual participants. (ns: not significant)

Table 5.2: Features showing a statistically significant difference between HF and healthy participants. Evaluation was done using the Mann-Whitney-U test.

Resting Phase						Recovery Phase					
Feature	U -Value	p -Value	CLES	Median healthy	Median HF	Feature	U -Value	p -Value	CLES	Median healthy	Median HF
Apple Watch (Neurokit Segmentation)											
HR (bpm)	69.0	0.026	0.82	67	60	HR (bpm)	74.0	0.006	0.88	72	61
QRS Area (μVs)	8.0	0.002	0.89	3.64	13.36	QRS Area (μVs)	10.0	0.006	0.12	3.56	12.86
QT Duration (ms)	17.0	0.041	0.20	412	465	QT Duration (ms)	15.0	0.026	0.18	402	452
R Amplitude (μV)	15.0	0.018	0.67	416	618	R Amplitude (μV)	13.0	0.015	0.15	263	591
Apple Watch (ECGdeli Segmentation)											
HR (bpm)	47.0	0.027	0.85	66	60	HR (bpm)	63.0	0.009	0.88	72	63
QT Duration (ms)	20.0	0.046	0.22	418	446	HRV HFr (ms^2)	22.0	0.044	0.21	0.053	0.117
QTc Duration (ms)	19.0	0.037	0.21	442	469	HRV total power (ms^2)	21.0	0.036	0.59	0.059	0.125
						QRS Area (μVs)	11.0	0.006	0.12	2.34	12.19
						QT Duration (ms)	15.0	0.018	0.17	410	470
						R Amplitude (μV)	13.0	0.011	0.61	291	606
Withings Scanwatch (Neurokit Segmentation)											
HR (bpm)	90.0	0.005	0.88	70	59	QRS Area (μVs)	20.0	0.011	0.17	2.90	8.32
PR Duration (ms)	21.0	0.013	0.18	116	150	R Amplitude (μV)	27.0	0.040	0.23	215	481
QRS Area (μVs)	11.0	0.003	0.11	3.29	11.37						
QT Duration (ms)	21.0	0.036	0.21	387	458						
R Amplitude (μV))	14.0	0.008	0.14	257	449						
Withings Scanwatch (ECGdeli Segmentation)											
QRS Area (μVs)	29.0	0.041	0.23	1.21	11.20	QRS Area (μVs)	24.0	0.033	0.21	2.28	8.61
R Amplitude (μV)	29.0	0.041	0.23	263	438	R Amplitude (μV)	26.0	0.047	0.23	214	486

Chapter 6

Discussion

The aim of this study was to assess the feasibility and reliability of 1-lead ECG recordings from smartwatch devices on a healthy and a HF patient group. Additionally, the significance of the extracted features was investigated for differentiating healthy and HF participants.

The data indicates, that individual parameters can be captured accurately by smartwatch ECGs for the healthy and HF participants. HR, QRS duration and QT duration showed a high positive correlation ($r > 0.7$) between smartwatch and reference recordings. However, the strength of the correlation depended on the used smartwatch and segmentation algorithm. Furthermore, the analysis indicates, that HF has no influence on the accuracy of smartwatch ECGs. Instead, the results suggest, that differences between reference and smartwatch ECGs may be caused by the automatic ECG segmentation algorithms. Statistically significant differences in parameters between healthy and HF participants could be observed for individual parameters. However, the statistical significance of the parameters is not consistent throughout the devices or algorithms, when assessing the difference of HF and healthy. This suggest that the power of the results is low and statistical significance could have been achieved by chance.

Overall a low to moderate correlation could be observed for smartwatch and reference parameters. This can partly be explained, as parameters were extracted from recordings of two different leads. The feature extraction for the smartwatches was done using a lead I recording, whereas a lead II recording was used for the reference ECG. Lead I of the reference ECG was also recorded during the study, but was considered to be too poor quality to use for the analysis. While time intervals should not be affected by this, high correlations between

amplitudes and amplitude related features cannot be expected. Furthermore, the reliability of HRV parameters for ultra short recordings has not been established. The lack of correlation, especially for frequency domain features could be caused by the insufficient length of the ECG signal. While the reliability of ultra short term recordings used for HRV analysis is being investigated, there is a lack of consensus to their reliability [Pec18] [Sha20].

The accuracy of smartwatch ECGs has been explored on patients with acute coronary syndrome and ST-elevation myocardial infarction [Spa21], as well as on patients with congenital heart disease [Kob22]. Those studies suggest that the diseases have no effect on the accuracy and that basic intervals can be measured reliably by the Apple Watch. The effect of HF on the quality of smartwatch ECGs has, to the best of our knowledge, not yet been investigated prior to this project. The finding of this project which is that HF does not appear to affect the measurements and accuracy of the smartwatch devices complements the findings of Spaccaritella et al. [Spa21] and Kobel et al. [Kob22].

Multiple studies [Beh20] [Spa21] [Spr22] [Sag20] reported the bias and LoA for the examined ECG features additionally to the correlation coefficient to evaluate agreement. Overall the results of those studies suggest a better agreement for QRS, QT, PR and wave durations than the results of this project. The main difference to those studies is that the ECG signals were always analyzed manually by one or more experienced health care professionals as opposed to the automatic segmentation approach used for this project.

An inaccuracy in the segmentation algorithm is suggested by the mean values of typical time intervals. For QRS complexes, the normal duration ranges from 80 to 120 ms [Hea]. The mean value of QRS duration for this study lies between 118 and 147 ms, depending on recording device and segmentation algorithm. As the majority of the study population were healthy participants, a mean duration greater than a normal range seems highly unlikely. A similar phenomenon can be observed for the QT interval. Normal durations are between 350 and 430 ms [Hea], whereas the reported mean duration ranges from 404 to 443 ms. This presumable overestimation in duration was not only found in the smartwatches, but also in the Nexus recordings, suggesting that it is not caused by the signal quality of the smartwatches. A difference in overall accuracy could not be observed between the algorithms for either watch.

Mannhart et. al [Man22] concluded in the analysis of the automated QTc measurements

that the automated algorithm of the Withings Scanwatch needs improvement before being useful for clinical applications. This finding strengthens the theory that the automatic segmentation algorithms play an important role in the decline of the agreement between reference and smartwatch recordings.

The 6MWT was included as it reflects daily physical activity. As walking is a very natural activity in daily life it is important to assess the quality of measured data shortly after such activities. Thomson et al. [Tho19] investigated the HR measure accuracy of different smartwatches across different exercise intensities. They showed that the validity was high for very light intensities but decreased with the increase in intensity. Similar findings were made by Khushhal et al. [Khu17], who also stated that the Apple Watch heart rate sensor was accurate during walking but accuracy decreased with increasing exercise intensity. It can be stated, that the PPG sensors of the Apple Watch can very well identify the mean HR during the walk supporting the findings of Thomson et al. [Tho19] and Khushhal et al. [Khu17] that PPG sensors of the watches work well for low intensity workouts. Additionally it shows that HF has no effect on the PPG measurement of mean HR in a HF patient population. Poor results were found for the maximum HR as it was constantly underestimated by the smartwatches in all participants. HR Reserve and Recovery both depend on the maximum HR and therefore show poor results as well.

Due to a lot of independent variables such as different watches, segmentation algorithms and health conditions, it is difficult to state whether exercise influences the accuracy of the ECG in recovery phase. The presented results indicate that exercise influences the accuracy of individual parameters but that it cannot be generalized that recordings after an active session are of worse quality.

Although this project explored an automated analysis approach, the raw data and results were not not cleaned or validated by a clinical professional. Furthermore, the number of participants, especially HF patients, is low. This could affect the analysis of differentiating between healthy and HF. Including more patients would give more meaningful results. While the probable inaccuracy of the segmentation algorithm is a main weakness of the study, it shows that improvements of such algorithms are necessary to make them useful for telemonitoring systems.

Chapter 7

Conclusion and Outlook

One research goal was to validate ECG parameters extracted from smartwatch ECGs. A correlation and agreement analysis showed, that individual parameters achieved a high correlation. Those parameter include HR, QRS and QT duration. However, the strength of the correlation depends on the segmentation algorithm, recording phase and smartwatch. Those factors have not been addressed in detail and need to be further investigated. Furthermore, the results indicate that HF does not impact the accuracy of the recordings, but discrepancies between reference and smartwatch ECGs are caused by the segmentation algorithm. Finally, differences in the individual parameters between the healthy and HF group have been investigated. Results were again depended on the segmentation algorithm, recording phase and smartwatch and are of low power.

The results of this thesis show that there is a potential in using smartwatches to monitor cardiac activity in HF patients. There is no indication that HF has an effect on the quality of a smartwatch ECG signal and willingness of the participants to monitor their cardiac activity with smartwatch devices is high.

Different aspects leave room for future research. For subsequent projects the accuracy of automated segmentation algorithms needs to be evaluated more thoroughly. There is still potential for improvements in ECG segmentation algorithms. It is the key for an automated ECG analysis, which is required for telemonitoring systems. Parameters indicating worsening of HF have been identified in this project. Long term studies are needed to assess the meaningfulness of those parameters for risk stratification in HF patients, when computed

from a 30 seconds long recording. To make those features meaningful in the context of a HF telemonitoring system, the significance of the parameters would also need to be established in regards to worsening HF. In connection to this, a study with a larger number of HF patients and different NYHA classes could be used for the assessment. If these issues can be solved, smartwatches show great potential to be integrated into telemonitoring systems for HF patients.

Appendix A

Patents

A.1 Body-worn sensor for characterizing patients with heart failure

Publication Number	US20160249858A1
Date of Publication	September 1, 2016
Inventors	Matthew Banet, Susan Meeks Pede, Marshal Singh Dhillon, Kenneth Robert Hunt
Assignee	Baxter Healthcare SA Baxter International Inc
Abstract	The invention provides a sensor for measuring both impedance and ECG waveforms that is configured to be worn around a patient's neck. The sensor features 1) an ECG system that includes an analog ECG circuit, in electrical contact with at least two ECG electrodes, that generates an analog ECG waveform; and 2) an impedance system that includes an analog impedance circuit, in electrical contact with at least two (and typically four) impedance electrodes, that generates an analog impedance waveform.

Also included in the neck-worn system are a digital processing system featuring a microprocessor, and an analog-to-digital converter.During a measurement, the digital processing system receives and processes the analog ECG and impedance waveforms to measure physiological information from the patient. Finally, a cable that drapes around the patient's neck connects the ECG system, impedance system, and digital processing system.

A.2 Wearable Device Electrocardiogram

Publication Number	US20160360986A1
Date of Publication	December 15, 2016
Inventors	Daniel H. Lange
Assignee	Chronisense Medical Ltd
Abstract	Provided are a method and systems for measuring an electrocardiogram (ECG) using a wearable device. An example system includes the wearable device in a shape of a band worn on a limb (e.g., a wrist) of a patient. The wearable device includes an electrical sensor. The wearable device is operable to record, via the at least one electrical sensor, an electrical signal from the limb of the patient. The wearable device is operable to split the electrical signal into segments based on a reference signal. The reference signal includes an indication of onset times of heart beats. The segments are averaged to derive average ECG data of low signal-to-noise ratio. The wearable device includes an optical sensor operable to measure skin color beneath a pulsating artery of the limb. The reference signal includes a photoplethysmogram (PPG) signal recorded via the optical sensor simultaneously with the electrical signal.

Appendix B

Additional Tables

Table B.1: Overall electrocardiagram characteristics. Mean ±standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$.

Parameter (Unit)	Smartwatch	Algorithm	Nexus ECG	Smartwatch ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HR (bpm)	Apple	Neurokit	69 ±11	72 ±14	0.72*** [0.53, 0.84]	-3 (-30, 25)
		ECGdeli	69 ±11	70 ±11	0.89*** [0.8, 0.94]	-1 (-10, 8)
	Withings	Neurokit	69 ±13	70 ±11	0.79*** [0.65, 0.87]	-1 (-16, 15)
		ECGdeli	68 ±12	71 ±14	0.78*** [0.63, 0.87]	-3 (-23, 17)
HRV HF (ms^2)	Apple	Neurokit	0.06 ±0.03	0.06 ±0.03	0.53*** [0.27, 0.72]	-0.002 (-0.07, 0.07)
		ECGdeli	0.07 ±0.04	0.07 ±0.04	0.40** [0.11, 0.63]	0.003 (-0.09, 0.08)
	Withings	Neurokit	0.06 ±0.03	0.07 ±0.03	0.34* [0.07, 0.57]	-0.007 (-0.07, 0.06)
		ECGdeli	0.07 ±0.04	0.07 ±0.04	0.41** [0.15, 0.63]	-0.004 (-0.08, 0.07)
HRV HFn (n.u)	Apple	Neurokit	0.79 ±0.15	0.81 ±0.11	0.53*** [0.27, 0.72]	-0.02 (-0.26, 0.23)
		ECGdeli	0.79 ±0.13	0.82 ±0.10	0.42** [0.14, 0.64]	-0.03 (-0.31, 0.25)
	Withings	Neurokit	0.76 ±0.15	0.82 ±0.10	0.19 [-0.1, 0.45]	-0.05 (-0.34, 0.23)
		ECGdeli	0.75 ±0.14	0.78 ±0.13	0.59*** [0.36, 0.75]	-0.03 (-0.27, 0.20)

Table B.1 Continued

Parameter (Unit)	Smartwatch	Algorithm	Nexus ECG	Smartwatch ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HRV HTI (n.u)	Apple	Neurokit	6.06 ±2.42	5.91 ±2.58	0.84*** [0.72, 0.91]	0.15 (-3.92, 4.22)
		ECGdeli	6.23 ±2.61	5.95 ±2.35	0.90*** [0.82, 0.95]	0.28 (-1.86, 2.42)
	Withings	Neurokit	6.27 ±3.02	6.03 ±2.94	0.80*** [0.67, 0.89]	0.24 (-4.25, 4.73)
		ECGdeli	6.18 ±3.00	5.98 ±2.77	0.76*** [0.61, 0.86]	0.20 (-4.20, 4.61)
HRV LF/HF (n.u)	Apple	Neurokit	0.26 ±0.40	0.15 ±0.22	0.50** [0.22, 0.7]	0.11 (-0.64, 0.86)
		ECGdeli	0.22 ±0.38	0.15 ±0.22	0.44** [0.15, 0.67]	0.07 (-0.76, 0.90)
	Withings	Neurokit	0.29 ±0.49	0.12 ±0.19	0.45** [0.19, 0.66]	0.16 (-0.51, 0.83)
		ECGdeli	0.29 ±0.48	0.12 ±0.20	0.37* [0.09, 0.59]	0.17 (-0.62, 0.96)
HRV LF (ms^2)	Apple	Neurokit	0.007 ±0.006	0.006 ±0.004	0.27 [-0.05, 0.54]	0.002 (-0.01, 0.01)
		ECGdeli	0.007 ±0.006	0.006 ±0.004	0.51*** [0.23, 0.71]	0.001 (-0.01, 0.01)
	Withings	Neurokit	0.01 ±0.007	0.005 ±0.004	0.22 [-0.07, 0.48]	0.005 (-0.01, 0.02)
		ECGdeli	0.01 ±0.007	0.005 ±0.004	0.24 [-0.05, 0.49]	0.005 (-0.01, 0.02)
HRV LFn (n.u)	Apple	Neurokit	0.15 ±0.16	0.11 ±0.11	0.48** [0.2, 0.69]	0.04 (-0.23, 0.32)
		ECGdeli	0.13 ±0.15	0.10 ±0.11	0.43** [0.13, 0.66]	0.02 (-.0.29, 0.34)
	Withings	Neurokit	0.16 ±0.15	0.08 ±0.10	0.45** [0.18, 0.65]	0.07 (-0.15, 0.30)
		ECGdeli	0.16 ±0.15	0.08 ±0.10	0.31* [0.02, 0.55]	0.08 (-0.17, 0.33)
HRV RMSSD (ms)	Apple	Neurokit	43.27 ±33.86	64.06 ±106.75	0.74*** [0.56, 0.85]	-20.79 (-194.86, 153.28)
		ECGdeli	54.86 ±46.88	78.27 ±111.80	0.55*** [0.3, 0.73]	-23.40 (-226.03, 179.23)
	Withings	Neurokit	57.87 ±57.33	88.51 ±122.43	0.46** [0.2, 0.66]	-30.64 (-250.05, 188.77)
		ECGdeli	69.52 ±64.31	150.05 ±172.19	0.16 [-0.13, 0.42]	-80.53 (-431.11, 270.05)
HRV SDNN (ms)	Apple	Neurokit	45.47 ±32.33	57.70 ±64.27	0.70*** [0.51, 0.83]	-12.23 (-111.42, 86.97)
		ECGdeli	51.03 ±37.45	72.31 ±82.10	0.54*** [0.28, 0.72]	-21.29 (-169.62, 127.04)
	Withings	Neurokit	53.57 ±47.37	78.89 ±98.67	0.52*** [0.28, 0.7]	-25.32 (-197.49, 146.85)
		ECGdeli	57.20 ±43.20	120.79 ±139.03	0.21 [-0.08, 0.47]	-63.59 (-338.27, 211.09)
HRV SDSD (ms)	Appel	Neurokit	43.88 ±34.41	64.73 ±107.90	0.74*** [0.56, 0.85]	-20.85 (-196.55, 154.85)
		ECGdeli	55.67 ±47.63	79.48 ±113.94	0.56*** [0.31, 0.74]	-20.85 (-230.04, 182.43)
	Withings	Neurokit	58.63 ±58.11	89.68 ±124.34	0.46** [0.2, 0.66]	-31.05 (-253.83, 191.73)
		ECGdeli	70.53 ±65.26	151.89 ±174.38	0.15 [-0.14, 0.42]	-81.36 (-436.20, 273.47)
HRV pNN50 (%)	Apple	Neurokit	18.35 ±21.23	16.75 ±25.25	0.85*** [0.74, 0.92]	1.59 (-20.71, 23.90)
		ECGdeli	18.18 ±21.90	18.05 ±23.86	0.82*** [0.69, 0.9]	0.13 (-14.78, 15.05)
	Withings	Neurokit	19.89 ±24.22	21.68 ±27.48	0.64*** [0.44, 0.78]	-1.79 (-32.87, 29.29)
		ECGdeli	18.63 ±23.00	24.19 ±26.77	0.49*** [0.24, 0.68]	-5.56 (-41.65, 30.53)
HRV total power (ms^2)	Apple	Neurokit	0.06 ±0.03	0.07 ±0.03	0.64*** [0.41, 0.8]	-0.003 (-0.07, 0.07)
		ECGdeli	0.07 ±0.5	0.08 ±0.04	0.47** [0.18, 0.68]	-0.005 (-0.09, 0.08)
	Withings	Neurokit	0.07 ±0.03	0.07 ±0.03	0.23 [-0.06, 0.49]	-0.001 (-0.07, 0.07)
		ECGdeli	0.08 ±0.04	0.08 ±0.04	0.37** [0.1, 0.6]	-0.0001 (-0.08, 0.09)

Table B.1 Continued

Parameter (Unit)	Smartwatch	Algorithm	Nexus ECG	Smartwatch ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
PR Duration (ms)	Apple	Neurokit	157 ±37	175 ±30	0.12 [-0.25, 0.46]	-19 (-110, 73)
		ECGdeli	171 ±33	170 ±32	0.82*** [0.63, 0.92]	1 (-46, 47)
	Withings	Neurokit	162 ±41	141 ±33	0.51** [0.22, 0.72]	22 (-62, 105)
		ECGdeli	173 ±34	191 ±30	0.58*** [0.31, 0.77]	-18 (-82, 46)
P Amplitude (μV)	Apple	Neurokit	50 ±88	6 ±22	0.48** [0.21, 0.69]	44 (-111, 199)
		ECGdeli	34 ±98	3 ±29	0.36* [0.06, 0.6]	32 (-152, 216)
	Withings	Neurokit	51 ±82	36 ±17	0.22 [-0.07, 0.47]	15 (-140, 171)
		ECGdeli	45 ±91	22 ±22	0.14 [-0.15, 0.41]	23 (-156, 202)
P Duration (ms)	Apple	Neurokit	89 ±21	130 ±18	0.30* [-0.01, 0.56]	-40 (-83, 3)
		ECGdeli	139 ±16	151 ±13	0.25 [-0.05, 0.52]	-12 (-44, 20)
	Withings	Neurokit	93 ±20	75 ±14	0.33* [0.05, 0.56]	18 (-19, 55)
		ECGdeli	138 ±13	157 ±17	0.37** [0.1, 0.59]	-20 (-53, 13)
QRS Area (μVs)	Apple	Neurokit	11 ±19	8 ±10	-0.17 [-0.5, 0.19]	3 (-46, 51)
		ECGdeli	5 ±24	8 ±12	0.07 [-0.33, 0.44]	-3 (-53, 46)
	Withings	Neurokit	8 ±18	6 ±6	-0.23 [-0.52, 0.11]	2 (-39, 42)
		ECGdeli	2 ±25	4 ±7	0.13 [-0.21, 0.44]	-2 (-50, 46)
QRS Duration (ms)	Apple	Neurokit	118 ±21	137 ±33	0.72*** [0.49, 0.85]	-19 (-62, 24)
		ECGdeli	132 ±15	147 ±29	0.78*** [0.56, 0.9]	-15 (-51, 20)
	Withings	Neurokit	122 ±21	137 ±31	0.47** [0.16, 0.69]	-15 (-72, 42)
		ECGdeli	131 ±16	146 ±23	0.74*** [0.54, 0.86]	-15 (-45, 15)
QT Duration (ms)	Apple	Neurokit	412 ±46	429 ±41	0.65*** [0.38, 0.82]	-17 (-82, 47)
		ECGdeli	412 ±46	429 ±41	0.39 [-0.0, 0.68]	-5 (-109, 99)
	Withings	Neurokit	422 ±45	404 ±49	0.75*** [0.56, 0.87]	18 (-56, 93)
		ECGdeli	434 ±40	418 ±54	0.37* [0.04, 0.62]	16 (-88, 120)
QTc Duration (ms)	Apple	Neurokit	430 ±34	458 ±43	0.61** [0.32, 0.79]	-27 (-100, 45)
		ECGdeli	456 ±26	464 ±56	0.36 [-0.04, 0.66]	-8 (-112, 95)
	Withings	Neurokit	441 ±39	429 ±48	0.63*** [0.37, 0.8]	12 (-75, 99)
		ECGdeli	456 ±24	447 ±45	0.33* [0.0, 0.6]	9 (-77, 94)
R Amplitude (μV)	Apple	Neurokit	935 ±494	380 ±218	-0.27 [-0.53, 0.04]	555 (-616, 1726)
		ECGdeli	786 ±801	326 ±310	-0.21 [-0.48, 0.1]	461 (-1306, 2228)
	Withings	Neurokit	853 ±476	304 ±174	-0.17 [-0.44, 0.12]	549 (-496, 1595)
		ECGdeli	672 ±832	255 ±242	-0.14 [-0.41, 0.15]	417 (-1261, 2095)
TWA (μV)	Apple	Neurokit	71 ±77	317 ±424	-0.11 [-0.4, 0.21]	-246 (-1108, 617)
		ECGdeli	75 ±82	270 ±369	-0.01 [-0.34, 0.33]	-194 (-940, 552)
	Withings	Neurokit	83 ±109	420 ±659	0.14 [-0.16, 0.41]	-337 (-1611, 937)
		ECGdeli	75 ±102	197 ±264	-0.04 [-0.37, 0.3]	-122 (-688, 444)

Table B.1 Continued

Parameter (Unit)	Smartwatch	Algorithm	Nexus ECG	Smartwatch ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
T Amplitude (μV)	Apple	Neurokit	348 ±174	148 ±74	0.46** [0.17, 0.67]	201 (-95, 497)
		ECGdeli	353 ±175	137 ±84	0.47** [0.2, 0.68]	216 (-73, 504)
	Withings	Neurokit	328 ±186	113 ±56	0.43** [0.17, 0.65]	215 (-111, 541)
		ECGdeli	334 ±189	87 ±66	0.39** [0.12, 0.61]	246 (-93, 586)
T Asymmetry (n.u)	Apple	Neurokit	0.036 ±0.039	0.026 ±0.025	0.27 [-0.04, 0.53]	0.01 (-0.06, 0.08)
		ECGdeli	0.005 ±0.015	0.037 ±0.083	0.04 [-0.27, 0.34]	-0.031 (-0.18, 0.13)
	Withings	Neurokit	0.028 ±0.028	0.097 ±0.047	-0.006 [-0.3, 0.28]	-0.069 (-0.16, 0.03)
		ECGdeli	0.004 ±0.015	0.051 ±0.054	-0.13 [-0.3, 0.28]	-0.047 (-0.16, 0.07)
T Duration (ms)	Apple	Neurokit	135 ±22	165 ±23	0.55*** [0.29, 0.73]	-31 (-72, 11)
		ECGdeli	208 ±25	215 ±24	0.02 [-0.29, 0.32]	-7 (-70, 56)
	Withings	Neurokit	140 ±21	109 ±26	0.22 [-0.08, 0.48]	30 (-27, 87)
		ECGdeli	213 ±29	187 ±16	0.19 [-0.1, 0.45]	27 (-31, 85)
T Flatness (n.u)	Apple	Neurokit	-0.81 ±0.18	-0.94 ±0.34	0.39* [0.09, 0.62]	0.13 (-0.49, 0.76)
		ECGdeli	-0.72 ±0.15	-0.84 ±0.34	0.60*** [0.37, 0.77]	0.12 (-0.42, 0.66)
	Withings	Neurokit	-0.80 ±0.17	-1.17 ±0.18	0.29* [0.0, 0.54]	0.38 (-0.02, 0.78)
		ECGdeli	-0.71 ±0.15	-1.17 ±0.18	0.46* [0.2, 0.66]	0.32 (-0.31, 0.95)

Table B.2: Electrocardiagram characteristics during 6MWT. Mean ±standard deviation, spearman correlation coefficient and Bland-Altman analysis (bias, LoA) for 6MWT data. P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$

Parameter (Unit)	Smartwatch	Algorithm	Nexus ECG	Smartwatch ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HR mean (bpm)	Apple	Neurokit	97 ±17	95 ±17	0.99*** [0.97, 0.99]	2 (-1, 5)
		ECGdeli	97 ±18	95 ±18	0.99*** [0.98, 1.0]	2 (-1, 4)
	Withings	Neurokit	97 ±17	103 ±18	0.42* [0.02, 0.7]	-7 (-41, 27)
		ECGdeli	97 ±18	103 ±18	0.40 [0.02, 0.7]	-7 (-41, 27)
HR max (bpm)	Apple	Neurokit	147 ±24	103 ±20	0.43* [0.04, 0.7]	43 (-4, 90)
		ECGdeli	185 ±37	104 ±20	0.76*** [0.51, 0.89]	81 (24, 138)
	Withings	Neurokit	147 ±24	121 ±27	0.0 [-0.4, 0.4]	26 (-47, 99)
		ECGdeli	185 ±37	122 ±27	0.31 [-0.11, 0.63]	63 (-13, 140)
HR Recovery (bpm)	Apple	Neurokit	72 ±22	30 ±16	0.04 [-0.38, 0.45]	42 (-9, 93)
		ECGdeli	112 ±33	33 ±23	0.34 [-0.09, 0.67]	79 (11, 147)
	Withings	Neurokit	74 ±22	49 ±28	-0.28 [-0.61, 0.14]	25 (-56, 105)
		ECGdeli	110 ±40	46 ±31	0.4 [-0.02, 0.7]	64 (-14, 142)
HR Reserve (bpm)	Apple	Neurokit	78 ±20	35 ±19	0.34 [-0.08, 0.65]	44 (-1, 88)
		ECGdeli	119 ±33	36 ±16	0.70*** [0.41, 0.86]	83 (27, 138)
	Withings	Neurokit	79 ± 22	51 ± 29	0.08 [-0.33, 0.46]	27 (-45, 99)
		ECGdeli	117 ± 35	53 ± 28	0.31 [-0.1, 0.46]	63 (-13, 140)

Table B.3: Electrocardiagram characteristics of the Apple Watch and Nexus Recording for healthy and HF participants. Mean ±standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$.

Parameter (Unit)	Condition	Algorithm	Nexus ECG	Apple Watch ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HR (bpm)	healthy	Neurokit	70 ± 10	75 ± 14	0.6*** [0.31, 0.79]	-4 (-36, 28)
		ECGdeli	70 ± 10	72 ± 9	0.86*** [0.72, 0.93]	-2 (-11, 8)
	HF	Neurokit	66 ± 13	66 ± 12	0.97*** [0.9, 0.99]	0 (-2, 2)
		ECGdeli	66 ± 13	66 ± 13	0.83*** [0.48, 0.95]	0 (-3, 3)
HRV HF (ms^2)	healthy	Neurokit	0.06 ± 0.03	0.06 ± 0.03	0.52** [0.2, 0.74]	-0.0 (-0.06, 0.06)
		ECGdeli	0.06 ± 0.04	0.06 ± 0.04	0.2 [-0.18, 0.52]	-0.0 (-0.1, 0.09)
	HF	Neurokit	0.07 ± 0.03	0.08 ± 0.04	0.17 [-0.44, 0.68]	-0.0 (-0.09, 0.08)
		ECGdeli	0.08 ± 0.05	0.09 ± 0.05	0.88*** [0.62, 0.97]	-0.01 (-0.05, 0.04)
HRV HFn (n.u)	healthy	Neurokit	0.77 ± 0.17	0.81 ± 0.11	0.52** [0.2, 0.74]	-0.04 (-0.3, 0.23)
		ECGdeli	0.78 ± 0.15	0.82 ± 0.11	0.37* [0.02, 0.65]	-0.04 (-0.36, 0.27)
	HF	Neurokit	0.85 ± 0.07	0.82 ± 0.08	0.47 [-0.14, 0.82]	0.03 (-0.12, 0.18)
		ECGdeli	0.81 ± 0.07	0.81 ± 0.08	0.68* [0.17, 0.9]	-0.0 (-0.11, 0.11)
HRV HTI (n.u)	healthy	Neurokit	6.55 ± 2.39	6.09 ± 2.4	0.77*** [0.57, 0.88]	0.45 (-3.67, 4.58)
		ECGdeli	6.48 ± 2.46	6.19 ± 2.36	0.85*** [0.7, 0.92]	0.29 (-1.94, 2.52)
	HF	Neurokit	4.83 ± 2.03	5.44 ± 2.92	0.91*** [0.71, 0.98]	-0.61 (-4.13, 2.91)
		ECGdeli	5.6 ± 2.85	5.35 ± 2.21	0.92*** [0.74, 0.98]	0.26 (-1.64, 2.15)
HRv LF/HF (n.u)	healthy	Neurokit	0.32 ± 0.46	0.16 ± 0.24	0.49** [0.15, 0.73]	0.16 (-0.69, 1.01)
		ECGdeli	0.27 ± 0.44	0.16 ± 0.24	0.44* [0.08, 0.7]	0.11 (-0.84, 1.06)
	HF	Neurokit	0.12 ± 0.1	0.13 ± 0.16	0.53 [-0.11, 0.86]	-0.02 (-0.28, 0.25)
		ECGdeli	0.1 ± 0.09	0.12 ± 0.16	0.53 [-0.11, 0.86]	-0.03 (-0.3, 0.23)
HRV LF (ms^2)	healthy	Neurokit	0.01 ± 0.01	0.01 ± 0.0	0.34 [-0.04, 0.63]	0.0 (-0.01, 0.01)
		ECGdeli	0.01 ± 0.01	0.01 ± 0.0	0.54** [0.21, 0.76]	0.0 (-0.01, 0.01)
	HF	Neurokit	0.01 ± 0.0	0.01 ± 0.0	0.1 [-0.53, 0.66]	0.0 (-0.01, 0.01)
		ECGdeli	0.01 ± 0.0	0.01 ± 0.0	0.36 [-0.3, 0.79]	-0.0 (-0.01, 0.01)
HRV LFn (n.u)	healthy	Neurokit	0.17 ± 0.18	0.11 ± 0.12	0.48* [0.12, 0.72]	0.06 (-0.24, 0.37)
		ECGdeli	0.15 ± 0.17	0.11 ± 0.12	0.43* [0.07, 0.69]	0.04 (-0.31, 0.4)
	HF	Neurokit	0.1 ± 0.07	0.09 ± 0.1	0.53 [-0.11, 0.86]	-0.0 (-0.16, 0.16)
		ECGdeli	0.07 ± 0.06	0.09 ± 0.1	0.53 [-0.11, 0.86]	-0.02 (-0.18, 0.14)
HRV RMSSD (ms)	healthy	Neurokit	42.56 ± 32.59	65.45 ± 121.54	0.73*** [0.51, 0.87]	-22.89 (-219.74, 173.97)
		ECGdeli	53.77 ± 47.6	67.02 ± 107.54	0.55** [0.24, 0.76]	-13.25 (-204.53, 178.03)
	HF	Neurokit	45.03 ± 36.78	60.57 ± 54.17	0.75** [0.31, 0.92]	-15.54 (-110.51, 79.42)
		ECGdeli	57.59 ± 44.93	106.37 ± 117.18	0.43 [-0.19, 0.81]	-48.78 (-269.62, 172.06)
HRV SDNN (ms)	healthy	Neurokit	47.89 ± 33.14	58.49 ± 68.78	0.76*** [0.55, 0.88]	-10.59 (-109.56, 88.37)
		ECGdeli	53.23 ± 38.65	67.19 ± 81.62	0.63*** [0.34, 0.8]	-13.96 (-152.02, 124.11)
	HF	Neurokit	39.42 ± 29.36	55.73 ± 51.24	0.51 [-0.09, 0.84]	-16.31 (-115.62, 83.0)
		ECGdeli	45.52 ± 33.67	85.14 ± 81.9	0.33 [-0.3, 0.76]	-39.62 (-205.59, 126.36)

Table B.3 Continued

Parameter (Unit)	Condition	Algorithm	Nexus ECG	Apple Watch ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HRV SDSD (ms)	healthy	Neurokit	43.13 ± 33.09	66.13 ± 122.93	0.73*** [0.51, 0.87]	-23.0 (-221.9, 175.9)
		ECGdeli	54.54 ± 48.33	68.09 ± 109.69	0.57** [0.26, 0.77]	-13.54 (-208.48, 181.39)
	HF	Neurokit	45.76 ± 37.44	61.23 ± 54.34	0.76** [0.34, 0.93]	-15.47 (-110.3, 79.35)
		ECGdeli	58.5 ± 45.72	107.96 ± 119.26	0.43 [-0.19, 0.81]	-49.46 (-273.79, 174.87)
HRV pNN50 (%)	healthy	Neurokit	16.6 ± 17.86	14.55 ± 24.52	0.81*** [0.64, 0.91]	2.05 (-24.04, 28.14)
		ECGdeli	15.39 ± 17.89	14.95 ± 21.77	0.77*** [0.57, 0.89]	0.43 (-16.93, 17.8)
	HF	Neurokit	22.73 ± 27.44	22.25 ± 26.19	0.95*** [0.82, 0.99]	0.47 (-5.26, 6.2)
		ECGdeli	25.16 ± 28.43	25.79 ± 26.88	0.96*** [0.87, 0.99]	-0.63 (-5.32, 4.07)
HRV total power (ms^2)	healthy	Neurokit	0.06 ± 0.03	0.06 ± 0.03	0.64*** [0.36, 0.82]	-0.0 (-0.06, 0.05)
		ECGdeli	0.07 ± 0.04	0.07 ± 0.04	0.26 [-0.13, 0.58]	-0.0 (-0.09, 0.09)
	HF	Neurokit	0.08 ± 0.03	0.08 ± 0.04	0.25 [-0.41, 0.74]	-0.0 (-0.1, 0.1)
		ECGdeli	0.09 ± 0.05	0.1 ± 0.05	0.91*** [0.68, 0.98]	-0.01 (-0.06, 0.04)
PR Duration (ms)	healthy	Neurokit	144 ± 29	177 ± 29	-0.03 [-0.49, 0.44]	-34 (-114, 46)
		ECGdeli	155 ± 19	158 ± 31	0.75** [0.33, 0.92]	-3 (-55, 49)
	HF	Neurokit	176 ± 40	172 ± 31	0.17 [-0.45, 0.68]	4 (-85, 94)
		ECGdeli	188 ± 36	183 ± 28	0.67* [0.16, 0.9]	5 (-33, 43)
P Amplitude (μA)	healthy	Neurokit	59 ± 98	6 ± 23	0.51** [0.19, 0.74]	53 (-117, 223)
		ECGdeli	42 ± 111	11 ± 19	0.27 [-0.1, 0.58]	31 (-177, 238)
	HF	Neurokit	29 ± 47	7 ± 19	0.35 [-0.28, 0.77]	22 (-74, 117)
		ECGdeli	14 ± 46	-20 ± 37	0.43 [-0.19, 0.81]	34 (-71, 139)
P Duration (ms)	healthy	Neurokit	91 ± 18	129 ± 17	0.13 [-0.25, 0.47]	-37 (-81, 6)
		ECGdeli	135 ± 13	149 ± 14	0.15 [-0.23, 0.48]	-14 (-43, 15)
	HF	Neurokit	84 ± 26	132 ± 19	0.8** [0.43, 0.94]	-48 (-86, -9)
		ECGdeli	149 ± 18	156 ± 9	0.1 [-0.5, 0.64]	-6 (-42, 29)
QRS Area (μVS)	healthy	Neurokit	13.79 ± 14.59	3.25 ± 5.76	-0.27 [-0.64, 0.2]	10.54 (-22.74, 43.81)
		ECGdeli	10.19 ± 12.17	3.81 ± 6.77	-0.2 [-0.66, 0.37]	6.37 (-22.73, 35.48)
	HF	Neurokit	5.27 ± 24.58	17.27 ± 10.23	-0.02 [-0.61, 0.59]	-12.0 (-68.95, 44.94)
		ECGdeli	-1.96 ± 31.31	12.87 ± 13.58	0.41 [-0.22, 0.79]	-14.83 (-73.07, 43.41)
QRS Duration (ms))	healthy	Neurokit	113 ± 16	131 ± 24	0.55* [0.14, 0.8]	-18 (-57, 21)
		ECGdeli	126 ± 13	140 ± 25	0.56* [0.04, 0.84]	-14 (-51, 23)
	HF	Neurokit	126 ± 26	147 ± 44	0.91*** [0.68, 0.98]	-21 (-70, 28)
		ECGdeli	138 ± 15	156 ± 30	0.84*** [0.51, 0.95]	-17 (-51, 17)
QT Duration (ms)	healthy	Neurokit	398 ± 37	413 ± 24	0.63** [0.25, 0.84]	-15 (-80, 49)
		ECGdeli	423 ± 39	426 ± 41	0.07 [-0.5, 0.6]	-3 (-123, 117)
	HF	Neurokit	437 ± 47	457 ± 49	0.63* [0.04, 0.89]	-20 (-84, 44)
		ECGdeli	454 ± 35	462 ± 46	0.45 [-0.16, 0.82]	-7 (-90, 76)

Table B.3 Continued

Parameter (Unit)	Condition	Algorithm	Nexus ECG	Apple Watch ECG	Correlation Coefficient {[95%CI Interval]}	Bias (95% LoA)
QTc Duration (ms)	healthy	Neurokit	423 ± 29	455 ± 43	0.59** [0.19, 0.83]	-32 (-108, 44)
		ECGdeli	441 ± 20	450 ± 63	0.36 [-0.24, 0.76]	-9 (-129, 111)
	HF	Neurokit	444 ± 37	464 ± 43	0.7* [0.17, 0.92]	-20 (-84, 44)
		ECGdeli	472 ± 20	479 ± 41	0.15 [-0.47, 0.66]	-7 (-89, 75)
R Amplitude (μV)	healthy	Neurokit	1056 ± 469	302 ± 171	-0.26 [-0.56, 0.12]	754 (-301, 1810)
		ECGdeli	975 ± 692	243 ± 271	-0.21 [-0.53, 0.16]	732 (-842, 2307)
	HF	Neurokit	631 ± 418	574 ± 200	0.12 [-0.49, 0.65]	57 (-766, 879)
		ECGdeli	314 ± 858	532 ± 304	0.22 [-0.41, 0.7]	-219 (-1718, 1281)
TWA (μV)	healthy	Neurokit	74.24 ± 82.63	274.69 ± 404.35	-0.05 [-0.4, 0.32]	-200.45 (-1026.46, 625.56)
		ECGdeli	75.7 ± 90.23	209.39 ± 327.34	-0.05 [-0.44, 0.35]	-133.69 (-798.04, 530.65)
	HF	Neurokit	61.92 ± 55.71	430.85 ± 452.98	-0.17 [-0.7, 0.48]	-368.93 (-1280.19, 542.32)
		ECGdeli	74.73 ± 57.81	420.39 ± 421.14	-0.12 [-0.69, 0.55]	-345.66 (-1194.46, 503.14)
T Amplitude(μV)	healthy	Neurokit	399 ± 167	154 ± 71	0.52** [0.2, 0.74]	245 (-27, 518)
		ECGdeli	399 ± 167	148 ± 75	0.51** [0.18, 0.73]	251 (-21, 523)
	HF	Neurokit	210 ± 105	131 ± 79	0.35 [-0.31, 0.79]	79 (-136, 294)
		ECGdeli	236 ± 134	110 ± 97	0.45 [-0.17, 0.81]	127 (-126, 379)
T Asymmetry (n.u)	healthy	Neurokit	0.03 ± 0.03	0.03 ± 0.03	0.47** [0.13, 0.71]	-0.0 (-0.06, 0.05)
		ECGdeli	0.0 ± 0.01	0.04 ± 0.09	0.04 [-0.32, 0.4]	-0.04 (-0.23, 0.14)
	HF	Neurokit	0.07 ± 0.03	0.02 ± 0.02	0.71* [0.19, 0.92]	0.05 (-0.01, 0.11)
		ECGdeli	0.01 ± 0.02	0.02 ± 0.04	0.2 [-0.42, 0.69]	-0.01 (-0.09, 0.07)
T Duration (ms)	healthy	Neurokit	135 ± 20	163 ± 19	0.35 [-0.02, 0.63]	-28 (-71, 15)
		ECGdeli	201 ± 23	217 ± 25	0.12 [-0.25, 0.46]	-16 (-78, 45)
	HF	Neurokit	136 ± 25	173 ± 30	0.91*** [0.68, 0.98]	-37 (-72, -3)
		ECGdeli	227 ± 16	210 ± 19	0.29 [-0.34, 0.74]	17 (-20, 54)
T Flatness (n.u)	healthy	Neurokit	-0.77 ± 0.18	-0.91 ± 0.32	0.45* [0.11, 0.7]	0.14 (-0.42, 0.7)
		ECGdeli	-0.69 ± 0.12	-0.84 ± 0.37	0.74*** [0.52, 0.87]	0.15 (-0.39, 0.7)
	HF	Neurokit	-0.9 ± 0.13	-1.01 ± 0.38	0.28 [-0.38, 0.75]	0.11 (-0.66, 0.88)
		ECGdeli	-0.81 ± 0.19	-0.86 ± 0.23	0.19 [-0.43, 0.69]	0.05 (-0.45, 0.55)

Table B.4: Electrocardiagram characteristics of the Withings Scanwatch and Nexus Recording for healthy and HF participants. Mean ±standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$.

Parameter (Unit)	Condition	Algorithm	Nexus ECG	Withings ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HR (bpm)	healthy	Neurokit	70 ± 11	71 ± 10	0.7*** [0.47, 0.84]	-2 (-18, 14)
		ECGdeli	68 ± 10	71 ± 9	0.82*** [0.67, 0.91]	-2 (-15, 10)
	HF	Neurokit	67 ± 16	66 ± 13	0.89*** [0.69, 0.97]	1 (-11, 14)
		ECGdeli	67 ± 16	72 ± 21	0.62* [0.13, 0.86]	-5 (-37, 27)
HRV HF (ms^2)	healthy	Neurokit	0.06 ± 0.03	0.07 ± 0.04	0.43* [0.11, 0.67]	-0.01 (-0.08, 0.06)
		ECGdeli	0.06 ± 0.03	0.07 ± 0.03	0.32 [-0.02, 0.59]	-0.01 (-0.08, 0.07)
	HF	Neurokit	0.06 ± 0.03	0.06 ± 0.02	0.13 [-0.43, 0.62]	0.0 (-0.06, 0.06)
		ECGdeli	0.08 ± 0.04	0.08 ± 0.05	0.58* [0.07, 0.85]	-0.0 (-0.08, 0.07)
HRV HFn (n.u)	healthy	Neurokit	0.77 ± 0.15	0.81 ± 0.11	0.13 [-0.22, 0.45]	-0.04 (-0.33, 0.24)
		ECGdeli	0.75 ± 0.14	0.78 ± 0.13	0.59*** [0.31, 0.77]	-0.03 (-0.29, 0.22)
	HF	Neurokit	0.75 ± 0.15	0.83 ± 0.05	0.38 [-0.18, 0.76]	-0.08 (-0.37, 0.2)
		ECGdeli	0.75 ± 0.13	0.79 ± 0.12	0.5 [-0.04, 0.81]	-0.04 (-0.2, 0.13)
HRV HTI (n.u)	healthy	Neurokit	6.71 ± 2.7	6.32 ± 2.99	0.82*** [0.67, 0.91]	0.39 (-3.9, 4.68)
		ECGdeli	6.77 ± 2.92	6.24 ± 2.7	0.78*** [0.6, 0.88]	0.53 (-3.94, 4.99)
	HF	Neurokit	5.2 ± 3.46	5.34 ± 2.68	0.58* [0.07, 0.85]	-0.14 (-5.01, 4.74)
		ECGdeli	4.77 ± 2.7	5.35 ± 2.84	0.64* [0.17, 0.88]	-0.59 (-4.43, 3.26)
HRv LF/HF (n.u)	healthy	Neurokit	0.3 ± 0.55	0.12 ± 0.22	0.38* [0.04, 0.64]	0.15 (-0.53, 0.84)
		ECGdeli	0.3 ± 0.55	0.1 ± 0.13	0.24 [-0.11, 0.53]	0.2 (-0.69, 1.1)
	HF	Neurokit	0.27 ± 0.32	0.1 ± 0.08	0.53 [-0.03, 0.84]	0.17 (-0.45, 0.79)
		ECGdeli	0.25 ± 0.26	0.16 ± 0.3	0.73** [0.3, 0.91]	0.08 (-0.26, 0.43)
HRV LF (ms^2)	healthy	Neurokit	0.01 ± 0.01	0.01 ± 0.0	0.07 [-0.28, 0.41]	0.0 (-0.01, 0.02)
		ECGdeli	0.01 ± 0.01	0.0 ± 0.0	0.06 [-0.28, 0.39]	0.0 (-0.01, 0.02)
	HF	Neurokit	0.01 ± 0.01	0.01 ± 0.0	0.57* [0.03, 0.85]	0.01 (-0.01, 0.02)
		ECGdeli	0.01 ± 0.01	0.01 ± 0.0	0.5 [-0.07, 0.82]	0.01 (-0.0, 0.02)
HRV LFn (n.u)	healthy	Neurokit	0.16 ± 0.16	0.08 ± 0.11	0.37* [0.04, 0.64]	0.07 (-0.14, 0.28)
		ECGdeli	0.16 ± 0.17	0.07 ± 0.08	0.17 [-0.18, 0.48]	0.09 (-0.18, 0.36)
	HF	Neurokit	0.16 ± 0.14	0.08 ± 0.06	0.52 [-0.05, 0.83]	0.08 (-0.18, 0.33)
		ECGdeli	0.15 ± 0.12	0.1 ± 0.12	0.7** [0.24, 0.9]	0.06 (-0.11, 0.23)
HRV RMSSD (ms)	healthy	Neurokit	61.61 ± 61.78	84.08 ± 119.24	0.6*** [0.32, 0.78]	-22.48 (-209.56, 164.6)
		ECGdeli	66.74 ± 61.93	166.6 ± 189.79	0.26 [-0.09, 0.55]	-99.86 (-474.88, 275.17)
	HF	Neurokit	48.79 ± 43.36	99.25 ± 129.23	0.13 [-0.43, 0.62]	-50.46 (-329.62, 228.7)
		ECGdeli	76.27 ± 69.29	109.87 ± 109.05	0.02 [-0.52, 0.54]	-33.6 (-294.16, 226.96)
HRV SDNN (ms)	healthy	Neurokit	57.51 ± 50.26	71.43 ± 91.89	0.73*** [0.52, 0.86]	-13.92 (-142.0, 114.15)
		ECGdeli	56.31 ± 40.96	133.12 ± 154.68	0.34 [-0.0, 0.61]	-76.81 (-372.91, 219.29)
	HF	Neurokit	43.99 ± 37.81	97.0 ± 111.4	0.02 [-0.51, 0.55]	-53.0 (-293.09, 187.08)
		ECGdeli	59.36 ± 48.12	90.85 ± 83.05	-0.09 [-0.6, 0.46]	-31.48 (-231.92, 168.95)

Table B.4 Continued

Parameter (Unit)	Condition	Algorithm	Nexus ECG	Withings ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HRV SDSD (ms)	healthy	Neurokit	62.41 ± 62.61	85.24 ± 121.13	0.6*** [0.32, 0.78]	-22.82 (-212.9, 167.25)
		ECGdeli	67.71 ± 62.86	168.73 ± 192.24	0.25 [-0.1, 0.54]	-101.02 (-480.64, 278.6)
	HF	Neurokit	49.46 ± 43.99	100.48 ± 131.19	0.14 [-0.42, 0.63]	-51.02 (-334.32, 232.28)
		ECGdeli	77.37 ± 70.29	111.0 ± 110.22	-0.01 [-0.54, 0.53]	-33.62 (-296.95, 229.7)
HRV pNN50 (%)	healthy	Neurokit	19.08 ± 23.33	20.26 ± 27.55	0.64*** [0.38, 0.8]	-1.18 (-27.41, 25.06)
		ECGdeli	16.68 ± 21.58	22.92 ± 26.16	0.49** [0.19, 0.71]	-6.25 (-36.06, 23.57)
	HF	Neurokit	21.84 ± 26.14	25.12 ± 27.0	0.55* [0.03, 0.84]	-3.28 (-43.63, 37.07)
		ECGdeli	23.38 ± 25.51	27.27 ± 27.94	0.44 [-0.12, 0.79]	-3.88 (-51.74, 43.98)
HRV total power (ms^2)	healthy	Neurokit	0.07 ± 0.03	0.08 ± 0.04	0.28 [-0.07, 0.57]	-0.0 (-0.08, 0.07)
		ECGdeli	0.07 ± 0.03	0.07 ± 0.03	0.25 [-0.1, 0.54]	-0.0 (-0.08, 0.08)
	HF	Neurokit	0.07 ± 0.03	0.07 ± 0.02	0.12 [-0.46, 0.63]	0.01 (-0.06, 0.07)
		ECGdeli	0.09 ± 0.04	0.09 ± 0.05	0.59* [0.06, 0.86]	0.0 (-0.08, 0.08)
PR Duration (ms)	healthy	Neurokit	154 ± 40	129 ± 25	0.37 [-0.04, 0.68]	25 (-59, 108)
		ECGdeli	164 ± 23	183 ± 28	0.39 [-0.05, 0.7]	-20 (-74, 34)
	HF	Neurokit	177 ± 39	160 ± 35	0.36 [-0.21, 0.75]	16 (-65, 98)
		ECGdeli	188 ± 42	203 ± 28	0.52 [-0.01, 0.82]	-15 (-91, 60)
P Amplitude (μA)	healthy	Neurokit	60 ± 91	38 ± 18	0.18 [-0.16, 0.49]	22 (-152, 196)
		ECGdeli	57 ± 103	25 ± 20	0.08 [-0.26, 0.41]	32 (-170, 235)
	HF	Neurokit	30 ± 44	31 ± 14	0.2 [-0.37, 0.66]	-1 (-90, 88)
		ECGdeli	15 ± 42	15 ± 25	-0.04 [-0.56, 0.5]	0 (-90, 90)
P Duration (ms)	healthy	Neurokit	94 ± 18	73 ± 13	0.18 [-0.17, 0.48]	21 (-17, 58)
		ECGdeli	134 ± 10	155 ± 15	0.16 [-0.19, 0.47]	-21 (-53, 11)
	HF	Neurokit	90 ± 23	78 ± 14	0.76** [0.38, 0.92]	12 (-22, 46)
		ECGdeli	146 ± 15	164 ± 21	0.48 [-0.07, 0.81]	-18 (-53, 18)
QRS Area (μVS)	healthy	Neurokit	10.77 ± 14.46	3.64 ± 5.76	-0.48* [-0.75, -0.09]	7.13 (-25.93, 40.2)
		ECGdeli	8.92 ± 14.95	2.17 ± 5.4	0.04 [-0.39, 0.45]	6.75 (-24.65, 38.16)
	HF	Neurokit	2.47 ± 22.94	11.36 ± 4.39	0.24 [-0.39, 0.71]	-8.88 (-53.33, 35.57)
		ECGdeli	-8.46 ± 33.24	8.12 ± 9.04	0.5 [-0.04, 0.81]	-16.58 (-72.18, 39.01)
QRS Duration (ms))	healthy	Neurokit	119 ± 20	133 ± 28	0.39 [-0.02, 0.69]	-14 (-75, 47)
		ECGdeli	125 ± 13	138 ± 18	0.44* [0.03, 0.73]	-13 (-45, 19)
	HF	Neurokit	129 ± 22	146 ± 33	0.43 [-0.19, 0.81]	-17 (-66, 32)
		ECGdeli	140 ± 16	159 ± 25	0.66** [0.2, 0.88]	-19 (-44, 6)
QT Duration (ms)	healthy	Neurokit	408 ± 33	386 ± 36	0.6** [0.25, 0.82]	22 (-43, 88)
		ECGdeli	424 ± 32	410 ± 41	0.41 [-0.02, 0.71]	14 (-74, 103)
	HF	Neurokit	449 ± 53	439 ± 51	0.7* [0.21, 0.91]	10 (-77, 98)
		ECGdeli	450 ± 45	431 ± 57	0.12 [-0.44, 0.61]	18 (-107, 143)

Table B.4 Continued

Parameter (Unit)	Condition	Algorithm	Nexus ECG	Withings ECG	Correlation Coefficient {[95%CI Interval]}	Bias (95% LoA)
QTc Duration (ms)	healthy	Neurokit	436 ± 35	423 ± 50	0.6** [0.24, 0.82]	13 (-73, 98)
		ECGdeli	448 ± 21	440 ± 42	0.51* [0.12, 0.77]	9 (-65, 82)
	HF	Neurokit	450 ± 42	439 ± 41	0.6* [0.04, 0.87]	11 (-78, 101)
		ECGdeli	467 ± 25	458 ± 46	-0.05 [-0.57, 0.49]	9 (-92, 111)
R Amplitude (μV)	healthy	Neurokit	984 ± 456	260 ± 139	-0.28 [-0.57, 0.06]	724 (-287, 1735)
		ECGdeli	903 ± 680	206 ± 208	-0.27 [-0.56, 0.07]	697 (-786, 2179)
	HF	Neurokit	536 ± 359	410 ± 200	0.59* [0.09, 0.85]	125 (-413, 663)
		ECGdeli	111 ± 898	373 ± 275	0.68** [0.23, 0.89]	-262 (-1610, 1087)
TWA (μV)	healthy	Neurokit	97.41 ± 123.2	467.4 ± 740.88	0.07 [-0.28, 0.4]	-369.99 (-1809.38, 1069.4)
		ECGdeli	86.35 ± 118.51	144.47 ± 177.81	-0.15 [-0.52, 0.27]	-58.12 (-481.66, 365.42)
	HF	Neurokit	41.71 ± 13.14	286.6 ± 289.76	0.29 [-0.34, 0.74]	-244.89 (-800.18, 310.4)
		ECGdeli	49.44 ± 19.12	324.07 ± 372.7	0.19 [-0.5, 0.73]	-274.63 (-1004.29, 455.03)
T Amplitude(μV)	healthy	Neurokit	376 ± 188	118 ± 58	0.47** [0.16, 0.7]	258 (-69, 585)
		ECGdeli	369 ± 188	92 ± 67	0.35* [0.01, 0.62]	277 (-68, 622)
	HF	Neurokit	193 ± 91	101 ± 50	0.29 [-0.34, 0.74]	92 (-73, 257)
		ECGdeli	250 ± 165	76 ± 62	0.62* [0.13, 0.86]	174 (-103, 452)
T Asymmetry (n.u)	healthy	Neurokit	0.02 ± 0.02	0.1 ± 0.04	0.18 [-0.17, 0.49]	-0.08 (-0.16, 0.0)
		ECGdeli	0.0 ± 0.0	0.05 ± 0.06	-0.23 [-0.53, 0.13]	-0.05 (-0.16, 0.06)
	HF	Neurokit	0.05 ± 0.03	0.09 ± 0.04	-0.5 [-0.84, 0.1]	-0.04 (-0.15, 0.07)
		ECGdeli	0.01 ± 0.03	0.05 ± 0.05	-0.16 [-0.64, 0.4]	-0.04 (-0.16, 0.08)
T Duration (ms)	healthy	Neurokit	138 ± 22	109 ± 28	0.35* [0.01, 0.62]	29 (-27, 86)
		ECGdeli	207 ± 26	186 ± 16	0.2 [-0.14, 0.51]	21 (-33, 75)
	HF	Neurokit	143 ± 16	111 ± 18	-0.34 [-0.77, 0.29]	32 (-25, 90)
		ECGdeli	229 ± 30	188 ± 15	0.27 [-0.3, 0.7]	41 (-19, 101)
T Flatness (n.u)	healthy	Neurokit	-0.78 ± 0.17	-1.16 ± 0.19	0.31 [-0.03, 0.58]	0.38 (-0.02, 0.77)
		ECGdeli	-0.7 ± 0.15	-0.97 ± 0.29	0.48** [0.16, 0.7]	0.27 (-0.28, 0.82)
	HF	Neurokit	-0.84 ± 0.16	-1.22 ± 0.15	0.15 [-0.47, 0.66]	0.38 (-0.04, 0.8)
		ECGdeli	-0.73 ± 0.14	-1.16 ± 0.38	0.27 [-0.3, 0.7]	0.43 (-0.32, 1.17)

Table B.5: Electrocardiagram characteristics of the Apple Watch and Nexus Recording during resting and recovery phase. Mean ±standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$.

Parameter (Unit)	Recording Phase	Algorithm	Nexus ECG	Apple Watch ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HR (bpm)	Rest	Neurokit	67 ± 12	70 ± 14	0.71*** [0.4, 0.87]	-3 (-30, 24)
		ECGdeli	67 ± 11	69 ± 11	0.82*** [0.59, 0.92]	-2 (-13, 10)
	Recovery	Neurokit	71 ± 11	74 ± 14	0.7*** [0.39, 0.87]	-3 (-30, 25)
		ECGdeli	71 ± 11	71 ± 10	0.98*** [0.94, 0.99]	-1 (-5, 4)
HRV HF (ms^2)	Rest	Neurokit	0.06 ± 0.03	0.07 ± 0.04	0.42 [-0.01, 0.72]	-0.01 (-0.09, 0.08)
		ECGdeli	0.06 ± 0.04	0.06 ± 0.04	0.51* [0.1, 0.77]	-0.0 (-0.06, 0.06)
	Recovery	Neurokit	0.06 ± 0.03	0.06 ± 0.03	0.65** [0.3, 0.84]	0.0 (-0.05, 0.06)
		ECGdeli	0.07 ± 0.05	0.08 ± 0.05	0.33 [-0.12, 0.67]	-0.0 (-0.1, 0.1)
HRV HFn (n.u)	Rest	Neurokit	0.8 ± 0.13	0.81 ± 0.09	0.54* [0.14, 0.79]	-0.0 (-0.22, 0.21)
		ECGdeli	0.8 ± 0.13	0.82 ± 0.09	0.63** [0.27, 0.83]	-0.02 (-0.2, 0.15)
	Recovery	Neurokit	0.78 ± 0.17	0.81 ± 0.12	0.52* [0.12, 0.78]	-0.03 (-0.3, 0.24)
		ECGdeli	0.77 ± 0.14	0.81 ± 0.11	0.15 [-0.3, 0.54]	-0.04 (-0.38, 0.31)
HRV HTI (n.u)	Rest	Neurokit	6.34 ± 2.02	6.73 ± 2.58	0.79*** [0.55, 0.91]	-0.39 (-5.0, 4.22)
		ECGdeli	6.84 ± 2.39	6.47 ± 1.93	0.87*** [0.7, 0.95]	0.37 (-1.98, 2.71)
	Recovery	Neurokit	5.78 ± 2.73	5.09 ± 2.3	0.87*** [0.7, 0.95]	0.69 (-2.42, 3.8)
		ECGdeli	5.63 ± 2.67	5.43 ± 2.61	0.92*** [0.82, 0.97]	0.2 (-1.7, 2.09)
HRv LF/HF (n.u)	Rest	Neurokit	0.21 ± 0.28	0.14 ± 0.15	0.59** [0.18, 0.82]	0.07 (-0.35, 0.48)
		ECGdeli	0.2 ± 0.27	0.14 ± 0.14	0.69** [0.35, 0.87]	0.06 (-0.3, 0.42)
	Recovery	Neurokit	0.31 ± 0.49	0.16 ± 0.27	0.46* [0.02, 0.75]	0.15 (-0.81, 1.1)
		ECGdeli	0.24 ± 0.46	0.16 ± 0.28	0.25 [-0.22, 0.62]	0.08 (-1.02, 1.18)
HRV LF (ms^2)	Rest	Neurokit	0.01 ± 0.0	0.01 ± 0.0	0.15 [-0.32, 0.57]	-0.0 (-0.01, 0.01)
		ECGdeli	0.01 ± 0.01	0.01 ± 0.0	0.57* [0.16, 0.81]	0.0 (-0.01, 0.01)
	Recovery	Neurokit	0.01 ± 0.01	0.01 ± 0.0	0.41 [-0.04, 0.72]	0.0 (-0.01, 0.02)
		ECGdeli	0.01 ± 0.01	0.01 ± 0.0	0.47* [0.03, 0.75]	0.0 (-0.01, 0.01)
HRV LFn (n.u)	Rest	Neurokit	0.14 ± 0.15	0.1 ± 0.1	0.54* [0.12, 0.8]	0.03 (-0.2, 0.25)
		ECGdeli	0.13 ± 0.14	0.1 ± 0.09	0.64** [0.26, 0.85]	0.02 (-0.17, 0.22)
	Recovery	Neurokit	0.17 ± 0.17	0.11 ± 0.13	0.47* [0.04, 0.76]	0.06 (-0.26, 0.38)
		ECGdeli	0.13 ± 0.16	0.1 ± 0.13	0.24 [-0.23, 0.62]	0.03 (-0.38, 0.43)
HRV RMSSD (ms)	Rest	Neurokit	50.12 ± 30.31	81.46 ± 97.78	0.56** [0.17, 0.8]	-31.34 (-208.82, 146.14)
		ECGdeli	57.78 ± 40.04	92.56 ± 127.39	0.63** [0.27, 0.83]	-34.78 (-265.1, 195.55)
	Recovery	Neurokit	36.42 ± 35.79	46.65 ± 112.36	0.73*** [0.44, 0.89]	-10.24 (-178.3, 157.83)
		ECGdeli	51.95 ± 52.68	63.98 ± 91.45	0.51* [0.1, 0.77]	-12.03 (-179.58, 155.52)
HRV SDNN (ms)	Rest	Neurokit	48.1 ± 23.86	72.79 ± 61.93	0.44* [0.01, 0.73]	-24.69 (-143.36, 93.98)
		ECGdeli	51.53 ± 28.23	80.62 ± 88.87	0.57** [0.18, 0.8]	-29.09 (-194.38, 136.2)
	Recovery	Neurokit	42.84 ± 38.82	42.61 ± 63.01	0.86*** [0.68, 0.94]	0.24 (-66.11, 66.58)
		ECGdeli	50.52 ± 44.81	64.01 ± 73.8	0.47* [0.05, 0.75]	-13.49 (-140.84, 113.86)

Table B.5 Continued

Parameter (Unit)	Recording Phase	Algorithm	Nexus ECG	Apple Watch ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HRV SDSD (ms)	Rest	Neurokit	50.87 ± 30.86	82.25 ± 98.73	0.56** [0.17, 0.8]	-31.38 (-210.23, 147.47)
		ECGdeli	58.65 ± 40.71	94.08 ± 129.84	0.63** [0.27, 0.83]	-35.43 (-269.94, 199.08)
	Recovery	Neurokit	36.9 ± 36.31	47.22 ± 113.69	0.72*** [0.42, 0.88]	-10.32 (-180.33, 159.69)
		ECGdeli	52.7 ± 53.51	64.88 ± 93.17	0.52* [0.11, 0.78]	-12.18 (-182.57, 158.2)
HRV pNN50 (%)	Rest	Neurokit	26.03 ± 23.03	23.6 ± 25.94	0.88*** [0.72, 0.95]	2.43 (-19.39, 24.25)
		ECGdeli	25.4 ± 24.28	26.42 ± 26.79	0.95*** [0.87, 0.98]	-1.02 (-15.54, 13.49)
	Recovery	Neurokit	10.67 ± 15.9	9.91 ± 22.55	0.75*** [0.47, 0.89]	0.76 (-21.9, 23.42)
		ECGdeli	10.96 ± 16.26	9.67 ± 16.74	0.62** [0.26, 0.83]	1.29 (-13.68, 16.26)
HRV total power (ms^2)	Rest	Neurokit	0.06 ± 0.03	0.07 ± 0.04	0.51* [0.07, 0.78]	-0.01 (-0.09, 0.07)
		ECGdeli	0.07 ± 0.04	0.07 ± 0.04	0.58** [0.17, 0.82]	-0.01 (-0.06, 0.05)
	Recovery	Neurokit	0.07 ± 0.03	0.06 ± 0.03	0.74*** [0.44, 0.89]	0.0 (-0.04, 0.05)
		ECGdeli	0.08 ± 0.05	0.08 ± 0.05	0.35 [-0.11, 0.69]	-0.0 (-0.1, 0.1)
PR Duration (ms)	Rest	Neurokit	155 ± 37	174 ± 29	0.12 [-0.39, 0.56]	-19 (-108, 71)
		ECGdeli	165 ± 29	168 ± 30	0.88*** [0.65, 0.97]	-2 (-38, 33)
	Recovery	Neurokit	158 ± 38	177 ± 31	0.12 [-0.47, 0.63]	-18 (-113, 76)
		ECGdeli	177 ± 36	172 ± 34	0.69* [0.2, 0.91]	5 (-50, 59)
P Amplitude (μA)	Rest	Neurokit	43 ± 83	2 ± 24	0.4 [-0.04, 0.71]	41 (-109, 190)
		ECGdeli	34 ± 85	1 ± 30	0.54* [0.13, 0.79]	33 (-121, 187)
	Recovery	Neurokit	58 ± 92	10 ± 19	0.54* [0.15, 0.79]	48 (-113, 208)
		ECGdeli	34 ± 109	3 ± 28	0.18 [-0.27, 0.57]	30 (-180, 240)
P Duration (ms)	Rest	Neurokit	92 ± 17	130 ± 17	0.27 [-0.2, 0.64]	-38 (-78, 1)
		ECGdeli	136 ± 14	150 ± 11	0.53* [0.13, 0.78]	-13 (-38, 11)
	Recovery	Neurokit	87 ± 24	129 ± 18	0.31 [-0.14, 0.65]	-42 (-88, 4)
		ECGdeli	142 ± 17	152 ± 15	0.06 [-0.38, 0.48]	-10 (-48, 27)
QRS Area (μVS)	Rest	Neurokit	12.09 ± 17.26	6.84 ± 9.74	-0.09 [-0.55, 0.41]	5.24 (-36.51, 46.99)
		ECGdeli	5.65 ± 26.78	5.19 ± 10.28	0.34 [-0.24, 0.74]	0.46 (-43.94, 44.87)
	Recovery	Neurokit	9.16 ± 21.19	9.9 ± 10.43	-0.3 [-0.72, 0.27]	-0.74 (-54.85, 53.36)
		ECGdeli	3.33 ± 19.83	11.26 ± 11.8	-0.21 [-0.7, 0.41]	-7.93 (-61.45, 45.58)
QRS Duration (ms))	Rest	Neurokit	116 ± 23	133 ± 29	0.67** [0.28, 0.87]	-17 (-52, 18)
		ECGdeli	130 ± 15	145 ± 25	0.71** [0.29, 0.9]	-15 (-47, 18)
	Recovery	Neurokit	120 ± 20	141 ± 38	0.76** [0.38, 0.92]	-21 (-73, 30)
		ECGdeli	133 ± 16	150 ± 32	0.83*** [0.48, 0.95]	-16 (-56, 23)
QT Duration (ms)	Rest	Neurokit	410 ± 37	428 ± 32	0.77*** [0.44, 0.92]	-19 (-61, 24)
		ECGdeli	439 ± 39	431 ± 30	0.52 [-0.05, 0.83]	8 (-65, 81)
	Recovery	Neurokit	415 ± 54	430 ± 50	0.6* [0.1, 0.86]	-16 (-98, 67)
		ECGdeli	437 ± 40	457 ± 57	0.29 [-0.34, 0.74]	-19 (-143, 104)

Table B.5 Continued

Parameter (Unit)	Recording Phase	Algorithm	Nexus ECG	Apple Watch	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
QTc Duration (ms)	Rest	Neurokit	426 ± 32	446 ± 32	0.65** [0.23, 0.87]	-20 (-68, 28)
		ECGdeli	454 ± 30	448 ± 41	0.54 [-0.01, 0.84]	5 (-59, 70)
	Recovery	Neurokit	436 ± 35	472 ± 50	0.49 [-0.05, 0.81]	-36 (-126, 55)
		ECGdeli	458 ± 20	481 ± 64	0.1 [-0.5, 0.64]	-23 (-150, 104)
R Amplitude (μV)	Rest	Neurokit	946 ± 472	386 ± 225	-0.27 [-0.63, 0.18]	560 (-575, 1695)
		ECGdeli	801 ± 801	302 ± 340	-0.2 [-0.58, 0.25]	499 (-1231, 2229)
	Recovery	Neurokit	924 ± 515	374 ± 210	-0.26 [-0.62, 0.19]	550 (-655, 1755)
		ECGdeli	771 ± 800	349 ± 274	-0.25 [-0.62, 0.2]	422 (-1378, 2222)
TWA (μV)	Rest	Neurokit	47.94 ± 24.71	457.1 ± 517.07	0.09 [-0.37, 0.51]	-409.16 (-1421.6, 603.27)
		ECGdeli	52.87 ± 31.98	412.92 ± 473.4	0.22 [-0.3, 0.63]	-360.05 (-1271.81, 551.71)
	Recovery	Neurokit	92.83 ± 99.37	182.76 ± 243.02	0.03 [-0.4, 0.46]	-89.93 (-623.11, 443.25)
		ECGdeli	96.72 ± 106.13	134.39 ± 125.22	0.04 [-0.43, 0.5]	-37.67 (-358.14, 282.79)
T Amplitude (μV)	Rest	Neurokit	380 ± 193	170 ± 70	0.59** [0.2, 0.82]	210 (-103, 523)
		ECGdeli	387 ± 192	149 ± 93	0.57** [0.18, 0.8]	238 (-55, 532)
	Recovery	Neurokit	319 ± 148	127 ± 71	0.28 [-0.17, 0.64]	192 (-85, 469)
		ECGdeli	318 ± 147	126 ± 72	0.33 [-0.12, 0.67]	193 (-84, 469)
T Asymmetry (n.u)	Rest	Neurokit	0.03 ± 0.03	0.02 ± 0.02	0.15 [-0.32, 0.55]	0.0 (-0.06, 0.06)
		ECGdeli	0.0 ± 0.01	0.03 ± 0.07	0.07 [-0.37, 0.49]	-0.02 (-0.17, 0.12)
	Recovery	Neurokit	0.04 ± 0.04	0.03 ± 0.03	0.34 [-0.11, 0.67]	0.02 (-0.06, 0.09)
		ECGdeli	0.01 ± 0.02	0.04 ± 0.09	0.02 [-0.42, 0.44]	-0.04 (-0.22, 0.14)
T Duration (ms)	Rest	Neurokit	138 ± 26	163 ± 26	0.61** [0.23, 0.83]	-26 (-69, 17)
		ECGdeli	210 ± 23	214 ± 24	0.18 [-0.27, 0.57]	-4 (-58, 51)
	Recovery	Neurokit	132 ± 16	167 ± 20	0.51* [0.1, 0.77]	-35 (-73, 3)
		ECGdeli	207 ± 26	216 ± 23	-0.11 [-0.52, 0.34]	-10 (-80, 60)
T Flatness (n.u)	Rest	Neurokit	-0.79 ± 0.13	-0.93 ± 0.24	0.24 [-0.23, 0.62]	0.14 (-0.33, 0.61)
		ECGdeli	-0.73 ± 0.18	-0.82 ± 0.26	0.6** [0.23, 0.82]	0.1 (-0.24, 0.43)
	Recovery	Neurokit	-0.82 ± 0.21	-0.94 ± 0.4	0.39 [-0.05, 0.7]	0.13 (-0.62, 0.87)
		ECGdeli	-0.71 ± 0.12	-0.86 ± 0.4	0.62** [0.26, 0.83]	0.15 (-0.54, 0.84)

Table B.6: Electrocardiagram characteristics of the Withings Scanwatch and Nexus Recording during resting and recovery phase. Mean ±standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$.

Parameter (Unit)	Recording Phase	Algorithm	Nexus ECG	Withings ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HR (bpm)	Rest	Neurokit	68 ± 12	69 ± 12	0.78*** [0.56, 0.9]	-1 (-15, 12)
		ECGdeli	66 ± 10	68 ± 10	0.86*** [0.71, 0.94]	-2 (-12, 8)
	Recovery	Neurokit	70 ± 13	71 ± 10	0.78*** [0.54, 0.9]	0 (-18, 17)
		ECGdeli	70 ± 13	75 ± 16	0.68*** [0.37, 0.85]	-4 (-31, 23)
HRV HF (ms^2)	Rest	Neurokit	0.06 ± 0.02	0.06 ± 0.03	0.39 [-0.01, 0.68]	-0.0 (-0.06, 0.06)
		ECGdeli	0.07 ± 0.03	0.07 ± 0.04	0.27 [-0.14, 0.6]	-0.0 (-0.08, 0.08)
	Recovery	Neurokit	0.06 ± 0.04	0.07 ± 0.03	0.41* [0.0, 0.71]	-0.01 (-0.08, 0.06)
		ECGdeli	0.07 ± 0.04	0.08 ± 0.04	0.6** [0.25, 0.81]	-0.01 (-0.08, 0.06)
HRV HFn (n.u)	Rest	Neurokit	0.78 ± 0.13	0.83 ± 0.08	0.07 [-0.33, 0.45]	-0.05 (-0.34, 0.24)
		ECGdeli	0.77 ± 0.12	0.8 ± 0.15	0.74*** [0.49, 0.88]	-0.02 (-0.19, 0.15)
	Recovery	Neurokit	0.74 ± 0.17	0.8 ± 0.11	0.29 [-0.14, 0.63]	-0.06 (-0.34, 0.23)
		ECGdeli	0.73 ± 0.15	0.77 ± 0.11	0.3 [-0.13, 0.63]	-0.05 (-0.33, 0.24)
HRV HTI (n.u)	Rest	Neurokit	6.45 ± 3.19	6.05 ± 3.32	0.89*** [0.77, 0.95]	0.4 (-4.33, 5.13)
		ECGdeli	6.25 ± 2.93	6.06 ± 2.92	0.8*** [0.59, 0.91]	0.19 (-4.08, 4.47)
	Recovery	Neurokit	6.08 ± 2.8	6.02 ± 2.47	0.69*** [0.39, 0.86]	0.07 (-4.13, 4.26)
		ECGdeli	6.11 ± 3.08	5.9 ± 2.6	0.71*** [0.43, 0.87]	0.21 (-4.33, 4.76)
HRV LF/HF (n.u)	Rest	Neurokit	0.22 ± 0.25	0.09 ± 0.1	0.37 [-0.03, 0.67]	0.12 (-0.38, 0.62)
		ECGdeli	0.22 ± 0.25	0.12 ± 0.24	0.6** [0.26, 0.8]	0.1 (-0.25, 0.46)
	Recovery	Neurokit	0.38 ± 0.66	0.14 ± 0.26	0.54* [0.14, 0.79]	0.2 (-0.61, 1.02)
		ECGdeli	0.36 ± 0.65	0.12 ± 0.14	0.11 [-0.33, 0.5]	0.25 (-0.82, 1.31)
HRV LF (ms^2)	Rest	Neurokit	0.01 ± 0.01	0.01 ± 0.0	0.25 [-0.16, 0.59]	0.01 (-0.01, 0.02)
		ECGdeli	0.01 ± 0.01	0.0 ± 0.0	0.48* [0.11, 0.74]	0.01 (-0.0, 0.02)
	Recovery	Neurokit	0.01 ± 0.01	0.01 ± 0.0	0.14 [-0.31, 0.54]	0.0 (-0.01, 0.02)
		ECGdeli	0.01 ± 0.01	0.01 ± 0.0	-0.21 [-0.58, 0.23]	0.0 (-0.01, 0.02)
HRV LFn (n.u)	Rest	Neurokit	0.14 ± 0.12	0.07 ± 0.07	0.37 [-0.03, 0.67]	0.07 (-0.17, 0.3)
		ECGdeli	0.14 ± 0.13	0.07 ± 0.11	0.51** [0.15, 0.76]	0.07 (-0.12, 0.27)
	Recovery	Neurokit	0.18 ± 0.18	0.09 ± 0.12	0.53* [0.13, 0.78]	0.08 (-0.13, 0.29)
		ECGdeli	0.17 ± 0.18	0.08 ± 0.08	0.03 [-0.4, 0.44]	0.09 (-0.2, 0.39)
HRV RMSSD (ms)	Rest	Neurokit	66.76 ± 62.65	73.34 ± 100.8	0.72*** [0.45, 0.87]	-6.58 (-145.0, 131.83)
		ECGdeli	71.54 ± 60.21	115.86 ± 121.42	0.17 [-0.24, 0.53]	-44.32 (-313.01, 224.37)
	Recovery	Neurokit	48.2 ± 49.12	104.99 ± 140.42	0.23 [-0.2, 0.59]	-56.79 (-329.92, 216.34)
		ECGdeli	67.33 ± 68.42	187.22 ± 207.84	0.19 [-0.24, 0.56]	-119.9 (-528.07, 288.28)
HRV SDNN (ms)	Rest	Neurokit	57.98 ± 53.22	66.3 ± 83.73	0.8*** [0.6, 0.91]	-8.32 (-120.6, 103.96)
		ECGdeli	57.38 ± 42.54	96.62 ± 108.14	0.28 [-0.13, 0.61]	-39.24 (-265.28, 186.79)
	Recovery	Neurokit	48.77 ± 39.5	92.57 ± 111.08	0.25 [-0.18, 0.6]	-43.8 (-257.44, 169.83)
		ECGdeli	57.01 ± 43.9	147.07 ± 162.18	0.2 [-0.23, 0.56]	-90.06 (-401.12, 221.01)

Table B.6 Continued

Parameter (Unit)	Recording Phase	Algorithm	Nexus ECG	Withings ECG	Correlation Coefficient [95%CI Interval]	Bias (95% LoA)
HRV SDSD (ms)	Rest	Neurokit	67.68 ± 63.5	74.21 ± 102.24	0.73*** [0.47, 0.87]	-6.53 (-147.07, 134.0)
		ECGdeli	72.6 ± 61.13	117.29 ± 122.96	0.19 [-0.22, 0.54]	-44.69 (-316.63, 227.25)
	Recovery	Neurokit	48.8 ± 49.78	106.49 ± 142.69	0.23 [-0.2, 0.59]	-57.69 (-334.96, 219.57)
		ECGdeli	68.27 ± 69.41	189.5 ± 210.51	0.17 [-0.26, 0.54]	-121.23 (-534.34, 291.89)
HRV pNN50 (%)	Rest	Neurokit	23.78 ± 24.45	22.05 ± 27.9	0.77*** [0.54, 0.89]	1.73 (-23.65, 27.11)
		ECGdeli	20.76 ± 23.28	23.51 ± 26.33	0.6** [0.27, 0.8]	-2.75 (-32.96, 27.46)
	Recovery	Neurokit	15.66 ± 23.24	21.27 ± 27.0	0.54** [0.16, 0.78]	-5.61 (-40.37, 29.14)
		ECGdeli	16.32 ± 22.47	24.92 ± 27.21	0.46* [0.06, 0.73]	-8.6 (-49.32, 32.11)
HRV total power (ms^2)	Rest	Neurokit	0.07 ± 0.02	0.07 ± 0.03	0.13 [-0.28, 0.5]	0.0 (-0.06, 0.07)
		ECGdeli	0.08 ± 0.03	0.07 ± 0.04	0.23 [-0.18, 0.57]	0.01 (-0.08, 0.1)
	Recovery	Neurokit	0.07 ± 0.04	0.08 ± 0.03	0.39 [-0.06, 0.7]	-0.01 (-0.08, 0.07)
		ECGdeli	0.08 ± 0.04	0.08 ± 0.04	0.55** [0.17, 0.79]	-0.01 (-0.07, 0.06)
PR Duration (ms)	Rest	Neurokit	167 ± 45	138 ± 33	0.44 [-0.02, 0.74]	29 (-63, 121)
		ECGdeli	178 ± 34	195 ± 32	0.72** [0.36, 0.89]	-18 (-67, 31)
	Recovery	Neurokit	157 ± 36	143 ± 33	0.58* [0.15, 0.82]	14 (-56, 84)
		ECGdeli	169 ± 33	187 ± 27	0.4 [-0.08, 0.73]	-18 (-93, 57)
P Amplitude (μA)	Rest	Neurokit	42 ± 73	38 ± 18	0.14 [-0.27, 0.51]	4 (-138, 145)
		ECGdeli	40 ± 83	26 ± 20	0.33 [-0.07, 0.64]	14 (-141, 169)
	Recovery	Neurokit	61 ± 89	33 ± 16	0.36 [-0.06, 0.67]	28 (-139, 194)
		ECGdeli	50 ± 99	18 ± 23	0.01 [-0.4, 0.42]	33 (-168, 234)
P Duration (ms)	Rest	Neurokit	93 ± 20	76 ± 14	0.5* [0.13, 0.75]	18 (-17, 52)
		ECGdeli	138 ± 14	157 ± 18	0.4* [0.01, 0.69]	-20 (-54, 15)
	Recovery	Neurokit	92 ± 20	74 ± 13	0.22 [-0.21, 0.58]	19 (-21, 58)
		ECGdeli	138 ± 11	157 ± 17	0.27 [-0.16, 0.62]	-20 (-52, 12)
QRS Area (μVs)	Rest	Neurokit	8.19 ± 18.06	5.44 ± 7.34	-0.21 [-0.62, 0.28]	2.75 (-37.35, 42.85)
		ECGdeli	3.23 ± 24.76	5.19 ± 7.36	0.0 [-0.47, 0.47]	-1.96 (-49.93, 46.01)
	Recovery	Neurokit	7.65 ± 18.46	7.18 ± 5.24	-0.2 [-0.62, 0.31]	0.47 (-39.75, 40.69)
		ECGdeli	1.09 ± 25.71	3.77 ± 7.8	0.27 [-0.23, 0.65]	-2.68 (-50.66, 45.29)
QRS Duration (ms))	Rest	Neurokit	124 ± 24	139 ± 33	0.44 [-0.04, 0.75]	-15 (-79, 48)
		ECGdeli	131 ± 16	147 ± 26	0.77*** [0.48, 0.91]	-16 (-47, 14)
	Recovery	Neurokit	121 ± 17	136 ± 27	0.42 [-0.08, 0.75]	-15 (-65, 35)
		ECGdeli	131 ± 15	145 ± 21	0.73*** [0.39, 0.89]	-14 (-43, 15)
QT Duration (ms)	Rest	Neurokit	424 ± 47	407 ± 49	0.67** [0.29, 0.87]	17 (-63, 97)
		ECGdeli	437 ± 40	430 ± 50	0.68** [0.31, 0.87]	7 (-84, 98)
	Recovery	Neurokit	421 ± 44	402 ± 48	0.84*** [0.6, 0.94]	20 (-49, 88)
		ECGdeli	431 ± 39	406 ± 44	-0.06 [-0.51, 0.42]	24 (-89, 138)

Table B.6 Continued

Parameter (Unit)	Recording Phase	Algorithm	Nexus ECG	Withings ECG	Correlation Coefficient {[95%CI Interval]}	Bias (95% LoA)
QTc Duration (ms)	Rest	Neurokit	443 ± 44	430 ± 49	0.75*** [0.43, 0.91]	13 (-70, 96)
		ECGdeli	458 ± 28	453 ± 49	0.52* [0.07, 0.79]	5 (-84, 95)
	Recovery	Neurokit	439 ± 32	428 ± 46	0.53* [0.07, 0.81]	11 (-79, 102)
		ECGdeli	453 ± 19	441 ± 38	0.08 [-0.4, 0.53]	12 (-69, 93)
R Amplitude (μV)	Rest	Neurokit	861 ± 480	313 ± 173	-0.22 [-0.56, 0.2]	548 (-521, 1618)
		ECGdeli	686 ± 817	244 ± 258	-0.29 [-0.62, 0.11]	442 (-1260, 2145)
	Recovery	Neurokit	845 ± 471	294 ± 174	-0.11 [-0.5, 0.32]	550 (-469, 1569)
		ECGdeli	657 ± 847	267 ± 222	-0.04 [-0.44, 0.38]	390 (-1260, 2040)
TWA (μV)	Rest	Neurokit	83.49 ± 119.67	412.52 ± 745.45	0.29 [-0.13, 0.62]	-329.03 (-1786.76, 1128.71)
		ECGdeli	86.89 ± 135.7	167.64 ± 240.36	-0.04 [-0.49, 0.44]	-80.75 (-641.25, 479.75)
	Recovery	Neurokit	82.21 ± 95.81	428.65 ± 548.57	0.02 [-0.4, 0.44]	-346.44 (-1383.54, 690.66)
		ECGdeli	62.68 ± 29.5	230.66 ± 285.26	0.02 [-0.48, 0.51]	-167.98 (-726.89, 390.93)
T Amplitude(μV)	Rest	Neurokit	340 ± 194	119 ± 61	0.54** [0.18, 0.78]	221 (-108, 550)
		ECGdeli	338 ± 193	97 ± 67	0.31 [-0.11, 0.63]	241 (-110, 592)
	Recovery	Neurokit	315 ± 176	107 ± 50	0.21 [-0.24, 0.58]	208 (-115, 531)
		ECGdeli	329 ± 185	77 ± 63	0.43* [0.02, 0.71]	252 (-74, 578)
T Asymmetry (n.u)	Rest	Neurokit	0.02 ± 0.03	0.1 ± 0.03	-0.16 [-0.53, 0.26]	-0.07 (-0.17, 0.02)
		ECGdeli	0.0 ± 0.02	0.05 ± 0.06	0.0 [-0.4, 0.4]	-0.05 (-0.16, 0.07)
	Recovery	Neurokit	0.03 ± 0.03	0.1 ± 0.04	0.2 [-0.24, 0.57]	-0.06 (-0.16, 0.03)
		ECGdeli	0.0 ± 0.01	0.05 ± 0.05	-0.22 [-0.58, 0.21]	-0.05 (-0.16, 0.06)
T Duration (ms)	Rest	Neurokit	141 ± 22	111 ± 27	0.39 [-0.02, 0.68]	30 (-24, 84)
		ECGdeli	216 ± 30	191 ± 16	0.24 [-0.17, 0.58]	25 (-34, 85)
	Recovery	Neurokit	138 ± 19	108 ± 24	0.06 [-0.37, 0.47]	30 (-29, 89)
		ECGdeli	210 ± 27	182 ± 15	0.16 [-0.27, 0.54]	28 (-29, 85)
T Flatness (n.u)	Rest	Neurokit	-0.8 ± 0.15	-1.23 ± 0.17	0.26 [-0.16, 0.6]	0.43 (0.09, 0.76)
		ECGdeli	-0.73 ± 0.17	-1.06 ± 0.32	0.54** [0.17, 0.77]	0.33 (-0.29, 0.96)
	Recovery	Neurokit	-0.79 ± 0.19	-1.11 ± 0.17	0.29 [-0.15, 0.63]	0.32 (-0.11, 0.75)
		ECGdeli	-0.69 ± 0.12	-0.99 ± 0.34	0.34 [-0.08, 0.66]	0.3 (-0.33, 0.94)

Appendix C

Additional Figures

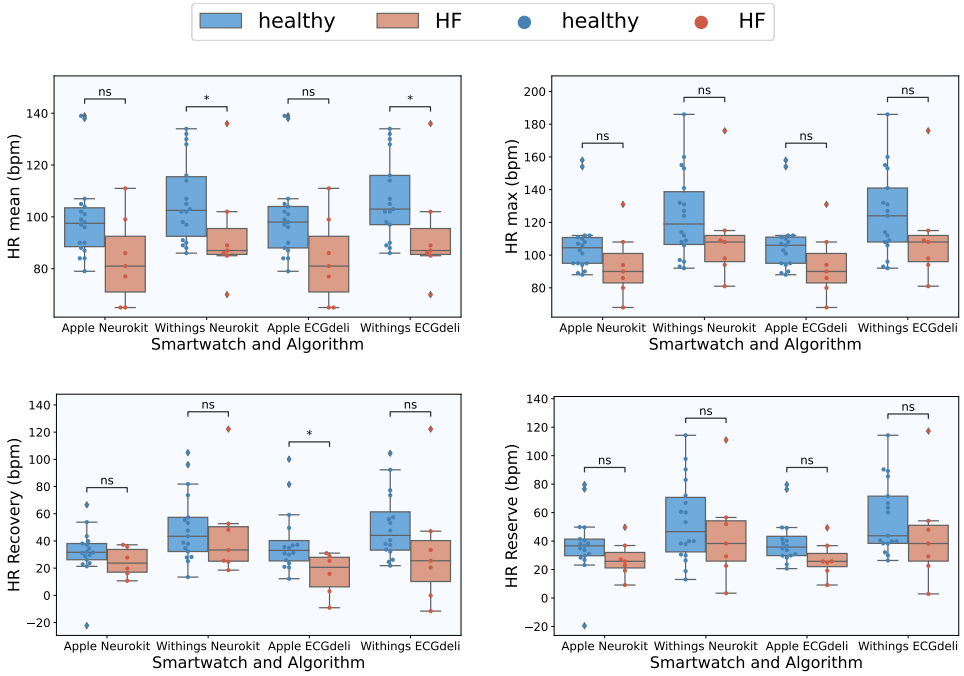

Figure C.1: 6MWT parameters from Apple Watch and Withings Scanwatch recordings when using Neurokit and ECGdeli segmentation algorithm comparing healthy and heart failure participants. Dots present the measurements of the individual participants. P-values for significance of correlation: ns: not significant, $*p < 0.05$, $***p < 0.01$, $***p < 0.001$

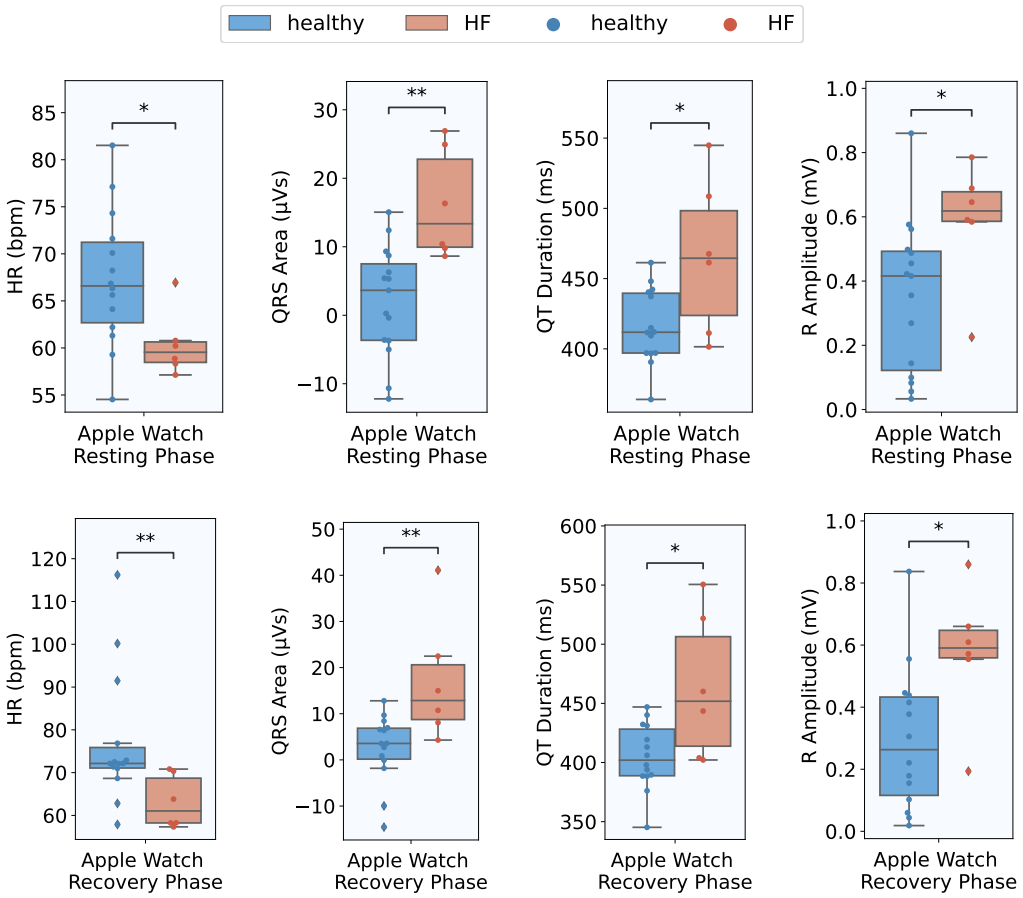

Figure C.2: ECG parameters in resting and recovery phase from the Apple Watch recordings showing statistical significant differences between healthy and heart failure participants when using the Neurokit segmentation algorithm. Dots present the measurements of the individual participants. P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$

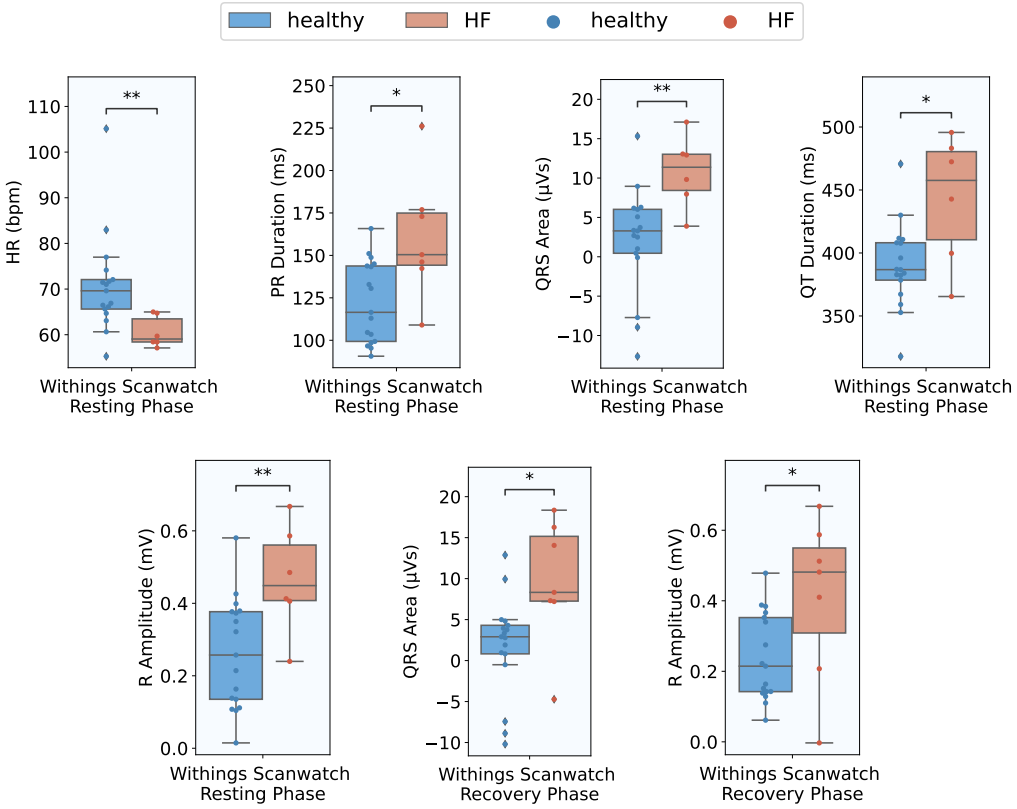

Figure C.3: ECG parameters in resting and recovery phase from the Withings Scanwatch recordings showing statistical significant differences between healthy and heart failure participants when using the Neurokit segmentation algorithm. Dots present the measurements of the individual participants. P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$

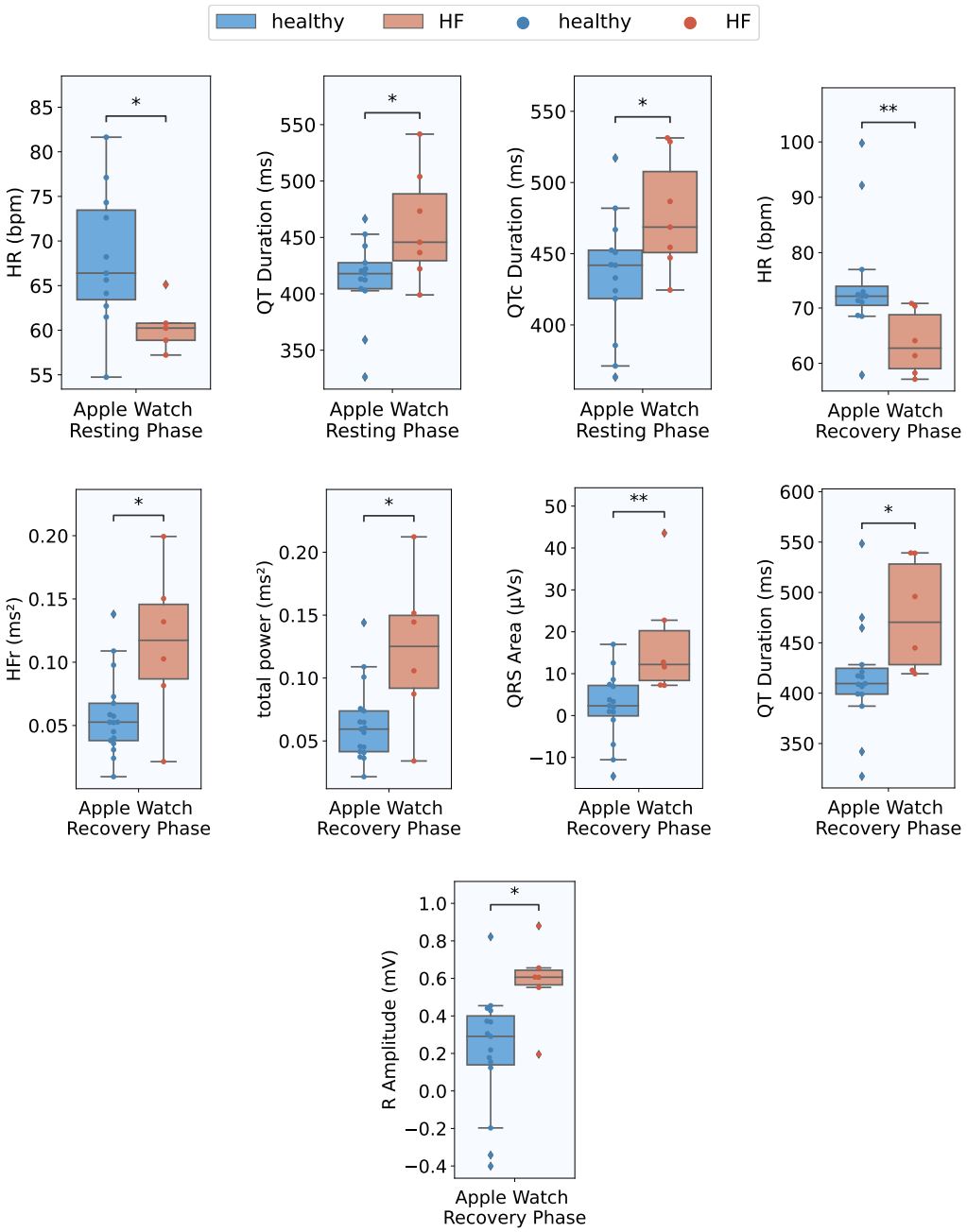

Figure C.4: ECG parameters in resting and recovery phase from the Apple Watch recordings showing statistical significant differences between healthy and heart failure participants when using the ECGdeli segmentation algorithm. Dots present the measurements of the individual participants. P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$

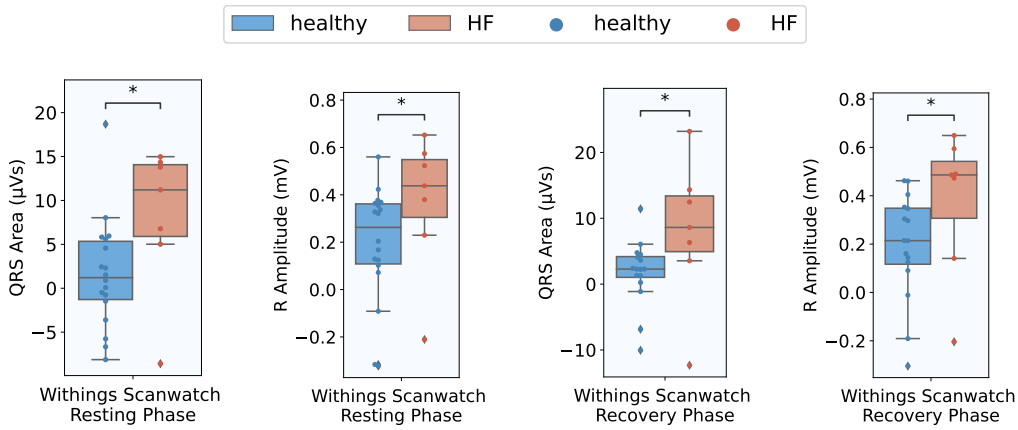

Figure C.5: ECG parameters in resting and recovery phase from the Withings Scanwatch recordings showing statistical significant differences between healthy and heart failure participants when using the ECGdeli segmentation algorithm. Dots present the measurements of the individual participants. P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$

Appendix D

Questionnaire

Fragebogen zur Studie „Überwachung der Herzaktivität von Herzinsuffizienz-Patienten mittels Smart Watches“

Alle Fragen die mit * gekennzeichnet sind, wurden nur von Herzinsuffizienzpatienten beantwortet.

Teil 1 - Demographische Daten

Geschlecht: Bitte geben Sie Ihr Geschlecht an.

- männlich
- weiblich
- divers

Alter: Bitte geben Sie Ihr Alter an.

Körpergröße: Bitte geben Ihre Körpergöße an (in cm).

Körpergewicht: Bitte geben Ihr Körpergewicht an (in kg).

Rauchverhalten: Rauchen Sie - wenn auch nur gelegentlich?

- ja
- nein

Körperliche Leistungsfähigkeit (NYHA)*:Bewerten Sie Ihre aktuelle körperliche Leisungsfähigkeit. Die zur Beurteilung der Stadien herangezogenen Symptome beinhalten Atemnot (Dyspnoe), häufiges nächtliches Wasserlassen (Nykturie), Zyanose (‚Blausucht'), allgemeine Schwäche und Müdigkeit, Brustenge (Angina pectoris) oder kalte Extremitäten, Ödeme.

- NYHA I: Keine Einschränkung der Belastbarkeit. Vollständiges Fehlen von Symptomen oder Beschwerden bei Belastung bei diagnostizierter Herzkrankheit.
- NYHA II: Leichte Einschränkung der Belastbarkeit. Beschwerdefreiheit in Ruhe und bei leichter Anstrengung, Auftreten von Symptomen bei stärkerer Belastung.
- NYHA III: Starke Einschränkung der Belastbarkeit. Beschwerdefreiheit in Ruhe, Auftreten von Symptomen bereits bei leichter Belastung.
- NYHA IV: Dauerhafte Symptomatik, auch in Ruhe.

Herzleistung*: Haben Sie eine diagnostizierte reduzierte Herzleistung (Herzinsuffizienz mit reduzierter linksventrikulärer Ejektionsfraktion, kleiner als 50%)?

- ja
- nein
- keine Angaben/nicht bekannt

Diagnosen*: Bitte geben Sie Ihre Diagnosen an.

- koronare Herzkrankheit
- Vorhofflimmern/-flattern (Herzrhythmusstörungen)
- Bluthochdruck
- Herzklappenfehler
- Periphere arterielle Verschlusskrankheit (pAVK)
- chronisch obstruktive Lungenerkrankung (COPD)
- Asthma
- schlafbezogene Atmungsstörungen (Schlafapnoe)
- Diabetes mellitus

- Depressionen
- chronische Nierenerkrankung
- weitere Diagnosen:

Akute Herzinsuffizienz*: Litten Sie schon mal an akuter Herzinsuffizienz (starke Verschlechterung des Zustandes, die zu einem Krankenhausaufenthalt führte)?

- ja, falls ja wie oft? und wann zuletzt?
- nein

Weiterführende Fragen

Smartwatch: Besitzen Sie eine Smartwatch oder einen Fitness-Tracker?

- ja, ich besitze Fitness-Tracker/Smartwatch und benutze sie auch regelmäßig
- ja, ich benutze sie aber nicht
- nein, ich besitze keinen Fitness-Tracker/Smartwatch

Funktionen: Falls Sie einen Fitness-Tracker/Smartwatch besitzen, welche Funktionen verwenden Sie oder haben Sie verwendet? (Mehrfachantworten sind möglich)

- Sportaktivitäten
- Schlafüberwachung
- Kalorienverbrauch
- Herzfrequenzmessung
- EKG-Messung
- Sauerstoffsättignungsmessung
- andere:

EKG: Haben Sie schon selbst ein EKG mithilfe einer Smartwatch gemessen (außerhalb dieser Studie)?

- ja, regelmäßig um meinen Zustand zu überwachen
- ja, aber nur als ich mich unwohl fühlte
- ja, nicht regelmäßig und einfach nur zum Spaß
- nein

Datenschutz: Haben Sie bedenken bzgl. des Datenschutzes bei der Verwendung von kommerziellen Smartwatches (z.B. Apple, Samsung oder Withings) für die Gesundheitsüberwachung?

- ja
- eher ja
- eher nein
- nein

Überwachung mittels Smartwatch: Könnten Sie sich (für Gesunde: im Falle einer Herzerkrankung) vorstellen regelmäßig Ihre Herzaktivität mit einer Smartwatch zu messen?

- ja
- mehrmals am Tag
- 1x täglich
- 1-4x in der Woche
- weniger oft
- nein, falls nein: warum nicht?

Telemonitoring*: Benutzen Sie bereits ein System um ihre Symptome/Messwerte zu überwachen?

- ja, falls ja: welches?
- nein

Arztbesuch: Würden Sie einen Arzt aufsuchen, wenn ihr Smartwatch EKG eine abnorme Messung anzeigt?

- ja, sofort
- ja, aber nur wenn es vermehrte abnorme Messungen gibt
- ja, aber nur wenn ich mich auch schlecht fühle
- nein

Appendix E

Acronyms

AF atrial fibrillation

AHF acute heart failure

AV atrioventricular

bpm beats per minute

CHF congestive heart failure

CLES common language effect size

ECG electrocardiogram

HF heart failure

HFr high frequncy

HFn high frequncy normalized

HFpEF heart failure with preserved ejection fraction

HFrEF heart failure with reduced ejection fraction

HR heart rate

HRT heart rate turbulence

HRV heart rate variability

HTI HRV triangular index

LF low frequency

LFn low frequency normlalized

LoA Limits of Agreement

LV left ventricular

LVEF left ventricular ejection fraction‘

SA sinoatrial

MMA modified moving average

NYHA New York Heart Association

pNN50 proportion of RR intervals greater than 50 ms, out of the total number of RR intervals

PPG photoplethysmogram

QTa the interval of the onset of Q wave to the apex of the T wave

QTc corrected QT interval

QTe the interval of the onset of Q wave to the end of the T wave

RMSSD square root of the mean of the squared successive differences between adjacent RR intervals

SDANN standard deviation of all 5-minute mean RR intervals

SDNN standard deviation of the NN interval

SDSD standard deviation of the successive differences between RR intervals

TWA T wave alternans

List of Figures

2.1 Signal propagation through the human heart and according ECG signal. . . 6

2.2 Smartwatch ECG. By placing a finger on the second electrode on the watch, lead I can be recorded. 8

2.3 Smartwatch with ECG functions. (a) Apple Watch Series 4 [App18] (b) Withings Scanwatch [Wit22a] 9

4.1 ECG recording setup. A participant wearing the Apple Watch (left wrist) and Withings Scanwatch (right wrist). The electrodes of the reference ECG are connected just below the collarbone. 25

4.2 Schematic representation of the study procedure. The study included a resting, an exercise and a recovery phase. 27

4.3 The ECG signal of a healthy human heart in sinus rhythm. 31

5.1 Mean heartbeats and their standard deviation of Nexus and smartwatch (Apple Watch (Apple), Withings Scanwatch (Withings) recordings. ECG tracings from two different participants (a) healthy participant (b) heart failure participant 39

5.2 Spearman correlation coefficients of the individual parameters extracted from the ECG recordings. The coefficients were calculated for all recorded parameters using different ECG segmentation algorithms (Neurokit, ECGdeli) and watches (Apple Watch (Apple) and Withings Scanwatch (Withings)). Correlation coefficient is color and size coded. Blue squares indicate a moderate to very high positive correlation. 40

5.3 Scatter and Bland-Altman plots for comparison of Nexus and Apple Watch for heart rate and QRS and QT duration. Calculated from ECG recordings, using two different segmentation algorithms. Left: Scatter plot including identity line (grey) and Spearman correlation coefficient r. Right: Bland-Altman plots including bias (solid line) and Limits of Agreement (LoA) (dashed line). 42

5.4 Scatter and Bland-Altman plots for comparison of Nexus and Withings Scanwatch for heart rate and QRS and QT Duration. Calculated from ECG recordings, using two different segmentation algorithms. Left: Scatter plot including identity line (grey) and Spearman correlation coefficient r. Right: Bland-Altman plots including bias (solid line) and Limits of Agreement (LoA) (dashed line) 43

5.5 Scatter and Bland-Altman plots for comparison of Nexus and Apple Watch of mean and maximum heart rate recorded during the six-minute walk. Calculated using two different segmentation algorithms. Left: Scatter plot including identity line (grey) and Spearman correlation coefficient r. Right: Bland-Altman plots including bias (solid line) and Limits of Agreement (LoA) (dashed line). 45

5.6 Scatter and Bland-Altman plots for comparison of Nexus and Withings Scanwatch of mean and maximum heart rate recorded during the six-minute walk. Calculated using two different segmentation algorithms. Left: Scatter plot including identity line (grey) and Spearman correlation coefficient r. Right: Bland-Altman plots including bias (solid line) and Limits of Agreement (LoA) (dashed line). 46

5.7 Spearman correlation coefficients of the individual ECG parameters for different participant groups (heart failure (HF) and healthy). The coefficients were calculated using different ECG segmentation algorithms (Neurokit and ECGdeli) and watches (Apple Watch (Apple) and Withings Scanwatch (Withings)). Correlation coefficient is color and size coded. Blue squares indicate a moderate to very high positive correlation. 47

5.8 Spearman correlation coefficients of the individual ECG parameters for different recording phases (rest and recovery (rec)). The coefficients were calculated using different ECG segmentation algorithms (Neurokit and ECGdeli) an watches (Apple Watch (Apple) and Withings Scanwatch (Withings). Correlation coefficient is color and size coded. Blue squares indicate a moderate to very high positive correlation. 49

5.9 SpO2 in resting and recovery phase. Dots present the measurements of the individual participants. (ns: not significant) 50

C.1 6MWT parameters from Apple Watch and Withings Scanwatch recordings when using Neurokit and ECGdeli segmentation algorithm comparing healthy and heart failure participants. Dots present the measurements of the individual participants. P-values for significance of correlation: ns: not significant, $*p < 0.05$, $***p < 0.01$, $***p < 0.001$ 79

C.2 ECG parameters in resting and recovery phase from the Apple Watch recordings showing statistical significant differences between healthy and heart failure participants when using the Neurokit segmentation algorithm. Dots present the measurements of the individual participants. P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$ 80

C.3 ECG parameters in resting and recovery phase from the Withings Scanwatch recordings showing statistical significant differences between healthy and heart failure participants when using the Neurokit segmentation algorithm. Dots present the measurements of the individual participants. P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$ 81

C.4 ECG parameters in resting and recovery phase from the Apple Watch recordings showing statistical significant differences between healthy and heart failure participants when using the ECGdeli segmentation algorithm. Dots present the measurements of the individual participants. P-values for significance of correlation: $*p < 0.05$, $***p < 0.01$, $***p < 0.001$ 82

C.5 ECG parameters in resting and recovery phase from the Withings Scanwatch recordings showing statistical significant differences between healthy and heart failure participants when using the ECGdeli segmentation algorithm. Dots present the measurements of the individual participants. P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$ 83

List of Tables

2.1 New York Heart Association (NYHA) Functional Classification [Com94] . 11

3.1 Parameters, extractable from an ECG, identifying and predicting decompensation in heart failure patients. 15

4.1 Demographics of healthy and heart failure (HF) participants. 24

4.2 Comorbidities of heart failure (HF) patients. 24

4.3 Parameters, extractable from a smartwatch ECG identifying and predicting decompensation in heart failure patients. 30

5.1 Table showing the results of the outlier detection. The numbers of detected waves before and after outlier detection are given for P, R and T wave. The results are shown for recordings obtained by the Nexus device, the Apple Watch and the Withings Scanwatch. 38

5.2 Features showing a statistically significant difference between HF and healthy participants. Evaluation was done using the Mann-Whitney-U test. 51

B.1 Overall electrocardiagram characteristics. Mean $\pm$standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$. . . . 61

B.2 Electrocardiagram characteristics during 6MWT. Mean $\pm$standard deviation, spearman correlation coefficient and Bland-Altman analysis (bias, LoA) for 6MWT data. P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$ 65

B.3 Electrocardiagram characteristics of the Apple Watch and Nexus Recording for healthy and HF participants. Mean ±standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$. 66

B.4 Electrocardiagram characteristics of the Withings Scanwatch and Nexus Recording for healthy and HF participants. Mean ±standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$. . . . 69

B.5 Electrocardiagram characteristics of the Apple Watch and Nexus Recording during resting and recovery phase. Mean ±standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$. 72

B.6 Electrocardiagram characteristics of the Withings Scanwatch and Nexus Recording during resting and recovery phase. Mean ±standard deviation, Spearman correlation coefficient and Bland-Altman analysis (bias, LoA). P-values for significance of correlation: *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$. 75

Bibliography

[App18] Apple. *Redesigned Apple Watch Series 4 revolutionizes communication, fitness and health*. en-US. 2018. URL: https://www.apple.com/newsroom/2018/09/redesigned-apple-watch-series-4-revolutionizes-communication-fitness-and-health/ (visited on 09/08/2022).

[App22] Apple. *Healthcare - Apple Watch*. en-GB. 2022. URL: https://www.apple.com/uk/healthcare/apple-watch/ (visited on 09/23/2022).

[Are06] Ross Arena, Marco Guazzi, Jonathan Myers, and Mary Ann Peberdy. "Prognostic value of heart rate recovery in patients with heart failure". In: *Am Heart J* 151 (Apr. 2006), 851.e7–13.

[Arm14] Richard A. Armstrong. "When to use the Bonferroni correction". In: *Ophthalmic and Physiological Optics* 34 (2014), pp. 502–508.

[Aro04] Doron Aronson, Murray A Mittleman, and Andrew J Burger. "Measures of heart period variability as predictors of mortality in hospitalized patients with decompensated congestive heart failure". In: *The American Journal of Cardiology* 93 (Jan. 2004), pp. 59–63.

[Ars18] Petros Arsenos, Konstantinos A. Gatzoulis, Polychronis Dilaveris, Skevos Sideris, and Dimitrios Tousoulis. "T wave alternans extracted from 30-minute short resting Holter ECG recordings predicts mortality in heart failure". In: *Journal of Electrocardiology* 51 (July 2018), pp. 588–591.

[Ars22] Petros Arsenos, Konstantinos A. Gatzoulis, Ageliki Laina, Ioannis Doundoulakis, Stergios Soulaidopoulos, Athanasios Kordalis, George Oikonomou, Konstantinos Triantafyllou, Nikolaos Fragakis, Vasillios Vasilikos, and Konstantinos Tsioufis. "QT interval extracted from 30-minute short resting Holter ECG recordings

predicts mortality in heart failure". In: *Journal of Electrocardiology* 72 (May 2022), pp. 109–114.

[Baz97] H. C. Bazett. "An Analysis of the Time-Relations of Electrocardiograms." In: *Annals of Noninvasive Electrocardiology* 2 (1997).

[Beh20] Amirali Behzadi, Alireza Sepehri Shamloo, Konstantinos Mouratis, Gerhard Hindricks, Arash Arya, and Andreas Bollmann. "Feasibility and Reliability of SmartWatch to Obtain 3-Lead Electrocardiogram Recordings". In: *Sensors (Basel)* 20 (Sept. 2020), p. 5074.

[Ben13] Jan Benes, Martin Kotrc, Barry A. Borlaug, Katerina Lefflerova, Petr Jarolim, Bela Bendlova, Antonin Jabor, Josef Kautzner, and Vojtech Melenovsky. "Resting Heart Rate and Heart Rate Reserve in Advanced Heart Failure Have Distinct Pathophysiologic Correlates and Prognostic Impact: A Prospective Pilot Study". In: *JACC: Heart Failure* 1 (June 2013), pp. 259–266.

[Bhu19] Ankit A. Bhurane, Manish Sharma, Ru San-Tan, and U. Rajendra Acharya. "An efficient detection of congestive heart failure using frequency localized filter banks for the diagnosis with ECG signals". In: *Cognitive Systems Research* 55 (June 2019), pp. 82–94.

[Bla86] J. M. Bland and D. G. Altman. "Statistical methods for assessing agreement between two methods of clinical measurement". In: *Lancet* 1 (Feb. 1986), pp. 307–310.

[Bra19] Darshan H Brahmbhatt and Martin R Cowie. "Remote Management of Heart Failure: An Overview of Telemonitoring Technologies". In: *Card Fail Rev* 5 (May 2019), pp. 86–92.

[Cah13] Lawrence P. Cahalin, Ross Arena, Valentina Labate, Francesco Bandera, Carl J. Lavie, and Marco Guazzi. "Heart rate recovery after the 6 min walk test rather than distance ambulated is a powerful prognostic indicator in heart failure with reduced and preserved ejection fraction: a comparison with cardiopulmonary exercise testing". In: *European Journal of Heart Failure* 15 (2013), pp. 519–527. (Visited on 09/12/2022).

[Cle05] John G. F. Cleland, Amala A. Louis, Alan S. Rigby, Uwe Janssens, and Aggie H. M. M. Balk. "Noninvasive Home Telemonitoring for Patients With Heart Failure at High Risk of Recurrent Admission and Death: The Trans-European Network-Home-Care Management System (TEN-HMS) study". In: *Journal of the American College of Cardiology* 45 (May 2005), pp. 1654–1664.

[Com94] New York Heart Association Criteria Committee. *Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels*. en. Little, Brown, 1994. ISBN: 978-0-316-60538-0.

[Cra02] Robert O. Crapo, Richard Casaburi, Allan L. Coates, Paul L. Enright, Neil R. MacIntyre, Roy T. McKay, Jack S. Wanger, R. Jorge Zeballos, Vera Bittner, and Carl Mottram. "ATS Statement". In: *Am J Respir Crit Care Med* 166 (July 2002). Publisher: American Thoracic Society - AJRCCM, pp. 111–117.

[Cyg06] Iwona Cygankiewicz, Wojciech Zareba, Rafael Vazquez, Montserrat Vallverdu, Juan Cino, Juan Cinca, Jesus Almendral, Jose R. Gonzalez Juanatey, Carlos Macaya, Mariano Valdes, Pere Caminal, and Antonio Bayes de Luna. "Relation of Heart Rate Turbulence to Severity of Heart Failure". In: *The American Journal of Cardiology* 98 (Dec. 2006), pp. 1635–1640.

[Cyg08a] Iwona Cygankiewicz, Wojciech Zareba, Rafael Vazquez, Jesus Almendral, Antoni Bayes-Genis, Miquel Fiol, Mariano Valdes, Carlos Macaya, Jose R. Gonzalez-Juanatey, Juan Cinca, Antoni Bayes De Luna, and on behalf of MUSIC Investigators. "Prognostic Value of QT/RR Slope in Predicting Mortality in Patients with Congestive Heart Failure". In: *Journal of Cardiovascular Electrophysiology* 19 (2008), pp. 1066–1072.

[Cyg08b] Iwona Cygankiewicz, Wojciech Zareba, Rafael Vazquez, Montserrat Vallverdu, Jose R. Gonzalez-Juanatey, Mariano Valdes, Jesus Almendral, Juan Cinca, Pere Caminal, and Antoni Bayes de Luna. "Heart rate turbulence predicts all-cause mortality and sudden death in congestive heart failure patients". In: *Heart Rhythm* 5 (Aug. 2008), pp. 1095–1102.

[Dav06] Clare Davenport, EYL Cheng, YTT Kwok, AHO Lai, T Wakabayashi, Christopher Hyde, and Martin Connock. "Assessing the diagnostic test accuracy of natriuretic peptides and ECG in the diagnosis of left ventricular systolic dys-

function: systematic review and meta-analysis". In: *British Journal of General Practice* 56 (Jan. 2006), pp. 48–56.

[Dis16] Marcello Disertori, Michela Masè, Marta Rigoni, Giandomenico Nollo, and Flavia Ravelli. "Heart Rate Turbulence Is a Powerful Predictor of Cardiac Death and Ventricular Arrhythmias in Postmyocardial Infarction and Heart Failure Patients". In: *Circulation: Arrhythmia and Electrophysiology* 9 (Dec. 2016). Publisher: American Heart Association, e004610.

[Fit20] Fitbit. *Identify AFib Early On with Fitbit's New On-Wrist ECG Technology*. en-US. Sept. 2020. URL: https://blog.fitbit.com/ecg-fitbit-sense/ (visited on 06/17/2022).

[For15] Ian Ford, Michele Robertson, Michel Komajda, Michael Böhm, Jeffrey S. Borer, Luigi Tavazzi, and Karl Swedberg. "Top ten risk factors for morbidity and mortality in patients with chronic systolic heart failure and elevated heart rate: The SHIFT Risk Model". In: *International Journal of Cardiology* 184 (Apr. 2015), pp. 163–169.

[Gal00] M Galinier, A Pathak, J Fourcade, C Androdias, D Curnier, S Varnous, S Boveda, P Massabuau, M Fauvel, J.M Senard, and J.P Bounhoure. "Depressed low frequency power of heart rate variability as an independent predictor of sudden death in chronic heart failure". In: *European Heart Journal* 21 (Mar. 2000), pp. 475–482.

[Gál19] César Gálvez-Barrón, Felipe Villar-Álvarez, Jesús Ribas, Francesc Formiga, David Chivite, Ramón Boixeda, Cristian Iborra, and Alejandro Rodríguez-Molinero. "Effort Oxygen Saturation and Effort Heart Rate to Detect Exacerbations of Chronic Obstructive Pulmonary Disease or Congestive Heart Failure". In: *Journal of Clinical Medicine* 8 (Jan. 2019). Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, p. 42.

[Gia19] Sophia Giannitsi, Mara Bougiakli, Aris Bechlioulis, Anna Kotsia, Lampros K. Michalis, and Katerina K. Naka. "6-minute walking test: a useful tool in the management of heart failure patients". In: *Ther Adv Cardiovasc Dis* 13 (Aug. 2019), p. 1753944719870084.

[Gol17] Ary L. Goldberger, Zachary D. Goldberger, and Alexei Shvilkin. *Clinical Electrocardiography: A Simplified Approach E-Book*. Elsevier Health Sciences, Mar. 2017. ISBN: 978-0-323-50877-3.

[Gou07] Barclay M Goudie, Rob I Jarvis, Peter T Donnan, Frank M Sullivan, Stuart D Pringle, Sanjay Jeyaseelan, and Allan D Struthers. "Screening for left ventricular systolic dysfunction using GP-reported ECGs". In: *Br J Gen Pract* 57 (Mar. 2007), pp. 191–195.

[Guy85] Gordon H. Guyatt, Michael J. Sullivan, Penelope J. Thompson, Ernest L. Fallen, Stewart O. Pugsley, D. Wayne Taylor, and Leslie B. Berman. "The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure". In: *Can Med Assoc J* 132 (Apr. 1985), pp. 919–923.

[Hea] School of Health Sciences. *Normal Duration Times - Normal Function of the Heart - Cardiology Teaching Package - Practice Learning - Division of Nursing - The University of Nottingham*. URL: https://www.nottingham.ac.uk/nursing/practice/resources/cardiology/function/normal_duration.php (visited on 11/06/2022).

[Hin03] Dennis E. Hinkle, William Wiersma, and Stephen G. Jurs. *Applied statistics for the behavioral sciences*. eng. 5th ed. Boston, Mass., [London]: Houghton Mifflin ; [Hi Marketing] (distributor), 2003.

[Ho10] Jennifer E. Ho, Vera Bittner, David A. DeMicco, Andrei Breazna, Prakash C. Deedwania, and David D. Waters. "Usefulness of Heart Rate at Rest as a Predictor of Mortality, Hospitalization for Heart Failure, Myocardial Infarction, and Stroke in Patients With Stable Coronary Heart Disease (Data from the Treating to New Targets [TNT] Trial)". In: *The American Journal of Cardiology* 105 (Apr. 2010), pp. 905–911.

[Hua09] Hui-Chun Huang, Lian-Yu Lin, Hsi-Yu Yu, and Yi-Lwun Ho. "Risk stratification by T-wave morphology for cardiovascular mortality in patients with systolic heart failure". In: *EP Europace* 11 (Nov. 2009), pp. 1522–1528.

[Isa21] Jonas L. Isaksen, Jonas Ghouse, Claus Graff, Morten S. Olesen, Anders G. Holst, Adrian Pietersen, Jonas B. Nielsen, Morten W. Skov, and Jørgen K. Kanters. "Elec-

trocardiographic T-wave morphology and risk of mortality". In: *International Journal of Cardiology* 328 (Apr. 2021), pp. 199–205.

[Kal02] Paul R. Kalra, Rakesh Sharma, Waqar Shamim, Wolfram Doehner, Roland Wensel, Aidan P. Bolger, Sabine Genth-Zotz, Mariantonietta Cicoira, Andrew J. S. Coats, and Stefan D. Anker. "Clinical characteristics and survival of patients with chronic heart failure and prolonged QRS duration". In: *International Journal of Cardiology* 86 (Dec. 2002), pp. 225–231.

[Kas05] Amir Kashani and S. Serge Barold. "Significance of QRS Complex Duration in Patients With Heart Failure". In: *Journal of the American College of Cardiology* 46 (Dec. 2005). Publisher: American College of Cardiology Foundation, pp. 2183–2192.

[Kha06] Alia Khaled, Mohamed Owis, and Abdalla Mohamed. "Employing Time-Domain Methods and Poincaré Plot of Heart Rate Variability Signals to Detect Congestive Heart Failure". In: (Jan. 2006).

[Khu17] Alaa Khushhal, Simon Nichols, Will Evans, Damien O. Gleadall-Siddall, Richard Page, Alasdair F. O'Doherty, Sean Carroll, Lee Ingle, and Grant Abt. "Validity and Reliability of the Apple Watch for Measuring Heart Rate During Exercise". In: *Sports Med Int Open* 1 (Oct. 2017), E206–E211.

[Kob22] M. Kobel, P. Kalden, A. Michaelis, F. Markel, S. Mensch, M. Weidenbach, F. T. Riede, F. Löffelbein, A. Bollmann, A. S. Shamloo, I. Dähnert, R. A. Gebauer, and C. Paech. "Accuracy of the Apple Watch iECG in Children With and Without Congenital Heart Disease". In: *Pediatr Cardiol* 43 (Jan. 2022), pp. 191–196.

[Koe18] Friedrich Koehler, Kerstin Koehler, Oliver Deckwart, Sandra Prescher, Karl Wegscheider, Bridget-Anne Kirwan, Sebastian Winkler, Eik Vettorazzi, Leonhard Bruch, Michael Oeff, Christian Zugck, Gesine Doerr, Herbert Naegele, Stefan Störk, Christian Butter, Udo Sechtem, Christiane Angermann, Guntram Gola, Roland Prondzinsky, Frank Edelmann, Sebastian Spethmann, Sebastian M. Schellong, P. Christian Schulze, Johann Bauersachs, Brunhilde Wellge, Christoph Schoebel, Milos Tajsic, Henryk Dreger, Stefan D. Anker, and Karl Stangl. "Efficacy of telemedical interventional management in patients with heart failure

(TIM-HF2): a randomised, controlled, parallel-group, unmasked trial". In: *Lancet* 392 (Sept. 2018), pp. 1047–1057.

[La 03] Maria Teresa La Rovere, Gian Domenico Pinna, Roberto Maestri, Andrea Mortara, Soccorso Capomolla, Oreste Febo, Roberto Ferrari, Mariella Franchini, Marco Gnemmi, Cristina Opasich, Pier Giorgio Riccardi, Egidio Traversi, and Franco Cobelli. "Short-Term Heart Rate Variability Strongly Predicts Sudden Cardiac Death in Chronic Heart Failure Patients". In: *Circulation* 107 (Feb. 2003). Publisher: American Heart Association, pp. 565–570.

[Lin09] Yen-Hung Lin, Lian-Yu Lin, Ying-Shren Chen, Hui-Chun Huang, Jen-Kuang Lee, Yi-Lwun Ho, Lin-Chu Liao, and Wen-Jone Chen. "The Association between T-Wave Morphology and Life-Threatening Ventricular Tachyarrhythmias in Patients with Congestive Heart Failure". In: *Pacing and Clinical Electrophysiology* 32 (2009), pp. 1173–1177.

[Mak21] Dominique Makowski, Tam Pham, Zen J. Lau, Jan C. Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and S. H. Annabel Chen. "NeuroKit2: A Python toolbox for neurophysiological signal processing". In: *Behav Res* 53 (Aug. 2021), pp. 1689–1696.

[Man22] D Mannhart, EH Hennings, ML Lischer, CV Vernier, JDF Du Fay De Lavallaz, SK Knecht, BS Schaer, SO Osswald, MK Kuehne, CS Sticherling, and PB Badertscher. "Clinical validation of automated QTc measurements from single lead ECG using a novel smartwatch". In: *EP Europace* 24 (May 2022), euac053.573.

[Mas12] Josep Masip, Maria Gayà, Joaquim Páez, Antoni Betbesé, Francisco Vecilla, Ruben Manresa, and Pilar Ruíz. "Pulse Oximetry in the Diagnosis of Acute Heart Failure". In: *Revista Española de Cardiología (English Edition)* 65 (Oct. 2012), pp. 879–884.

[McD21] Theresa A McDonagh, Marco Metra, Marianna Adamo, Roy S Gardner, Andreas Baumbach, Michael Böhm, Haran Burri, Javed Butler, Jelena Čelutkienė, Ovidiu Chioncel, John G F Cleland, Andrew J S Coats, Maria G Crespo-Leiro, Dimitrios Farmakis, Martine Gilard, Stephane Heymans, Arno W Hoes, Tiny Jaarsma, Ewa A Jankowska, Mitja Lainscak, Carolyn S P Lam, Alexander R Lyon, John J V McMurray, Alexandre Mebazaa, Richard Mindham, Claudio Muneretto, Massimo

Francesco Piepoli, Susanna Price, Giuseppe M C Rosano, Frank Ruschitzka, Anne Kathrine Skibelund, and ESC Scientific Document Group. "2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC". In: *European Heart Journal* 42 (Sept. 2021), pp. 3599–3726.

[Min22] MindMedia. *BioTrace+ Software | Biofeedback, neurofeedback und QEEG - Mind Media*. 2022. URL: https://www.mindmedia.com/de/produkte/biotrace-software/ (visited on 11/04/2022).

[Mon12] Violeta Monasterio, Pablo Laguna, Iwona Cygankiewicz, Rafael Vázquez, Antoni Bayés-Genís, Antoni Bayés de Luna, and Juan Pablo Martínez. "Average T-wave alternans activity in ambulatory ECG records predicts sudden cardiac death in patients with chronic heart failure". In: *Heart Rhythm* 9 (Mar. 2012), pp. 383–389.

[Moo06] Roger K.G. Moore, David G. Groves, Pauline E. Barlow, Keith A.A. Fox, Ajay Shah, James Nolan, and Mark T. Kearney. "Heart rate turbulence and death due to cardiac decompensation in patients with chronic heart failure". In: *European Journal of Heart Failure* 8 (2006), pp. 585–590.

[Muk12] MM Mukaka. "A guide to appropriate use of Correlation coefficient in medical research". In: *Malawi Med J* 24 (Sept. 2012), pp. 69–71.

[Nea02] Bruce D. Nearing and Richard L. Verrier. "Modified moving average analysis of T-wave alternans to predict ventricular fibrillation with high accuracy". In: *J Appl Physiol (1985)* 92 (Feb. 2002), pp. 541–549.

[Nol98] James Nolan, Phillip D. Batin, Richard Andrews, Steven J. Lindsay, Paul Brooksby, Michael Mullen, Wazir Baig, Andrew D. Flapan, Alan Cowley, Robin J. Prescott, James M.M. Neilson, and Keith A.A. Fox. "Prospective Study of Heart Rate Variability and Mortality in Chronic Heart Failure". In: *Circulation* 98 (Oct. 1998). Publisher: American Heart Association, pp. 1510–1516.

[Pat05] Atul Pathak, Daniel Curnier, Joëlle Fourcade, Jéme Roncalli, Phyllis K. Stein, Patricia Hermant, Marc Bousquet, Pierre Massabuau, Jean-Michel Sénard, Jean-Louis Montastruc, and Michel Galinier. "QT dynamicity: a prognostic factor for

sudden cardiac death in chronic heart failure". In: *European Journal of Heart Failure* 7 (2005), pp. 269–275.

[Pec18] Leandro Pecchia, Rossana Castaldo, Luis Montesinos, and Paolo Melillo. "Are ultra-short heart rate variability features good surrogates of short-term ones? State-of-the-art review and recommendations". In: *Healthcare Technology Letters* 5 (2018), pp. 94–100.

[Pen22] Lindsay K. D. Pengel, Daniëlle Robbers-Visser, Maarten Groenink, Michiel M. Winter, Mark J. Schuuring, Berto J. Bouma, and Jouke P. Bokma. "A comparison of ECG-based home monitoring devices in adults with CHD". In: *Cardiology in the Young* (July 2022). Publisher: Cambridge University Press, pp. 1–7.

[Per19] Marco V. Perez, Kenneth W. Mahaffey, Haley Hedlin, John S. Rumsfeld, Ariadna Garcia, Todd Ferris, Vidhya Balasubramanian, Andrea M. Russo, Amol Rajmane, Lauren Cheung, Grace Hung, Justin Lee, Peter Kowey, Nisha Talati, Divya Nag, Santosh E. Gummidipundi, Alexis Beatty, Mellanie True Hills, Sumbul Desai, Christopher B. Granger, Manisha Desai, and Mintu P. Turakhia. "Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation". In: *New England Journal of Medicine* 381 (Nov. 2019). Publisher: Massachusetts Medical Society, pp. 1909–1917.

[Pil20] A. Pilia, C Nagel, G. Lenis, S. Becker, O. Dössel, and A Loewe. *ECGdeli - An Open Source ECG Delineation Toolbox for MATLAB*. 2020. DOI: https://doi.org/10.5281/zenodo.3944621. URL: https://github.com/KIT-IBT/ECGdeli.

[Ram14] Julia Ramírez, Pablo Laguna, Antonio Bayés de Luna, Marek Malik, and Esther Pueyo. "QT/RR and T-peak-to-end/RR curvatures and slopes in chronic heart failure: Relation to sudden cardiac death". In: *Journal of Electrocardiology* 47 (Nov. 2014), pp. 842–848.

[Ram17] Julia Ramírez, Michele Orini, Ana Mincholé, Violeta Monasterio, Iwona Cygankiewicz, Antonio Bayés de Luna, Juan Pablo Martínez, Esther Pueyo, and Pablo Laguna. "T-Wave Morphology Restitution Predicts Sudden Cardiac Death in Patients With Chronic Heart Failure". In: *Journal of the American Heart Association* 6 (May 2017). Publisher: American Heart Association, e005310.

[Rav19] Moshe Rav-Acha, Ali Nujidat, Rivka Farkash, Aharon Medina, Michael Ilan, Marc Klutstein, Adi Butnaru, Tatyana Weitsman, Michael Glikson, and Tal Hasin. "Delayed prolongation of the QRS interval in patients with left ventricular dysfunction". In: *International Journal of Cardiology* 296 (Dec. 2019), pp. 71–75.

[Rie16] Evelien E. S. van Riet, Arno W. Hoes, Kim P. Wagenaar, Alexander Limburg, Marcel A. J. Landman, and Frans H. Rutten. "Epidemiology of heart failure: the prevalence of heart failure and ventricular dysfunction in older adults over time. A systematic review". In: *Eur J Heart Fail* 18 (Mar. 2016), pp. 242–252.

[Sag20] Nabeel Saghir, Arjun Aggarwal, Nisha Soneji, Victoria Valencia, George Rodgers, and Thomas Kurian. "A comparison of manual electrocardiographic interval and waveform analysis in lead 1 of 12-lead ECG and Apple Watch ECG: A validation study". In: *Cardiovascular Digital Health Journal* 1 (July 2020), pp. 30–36.

[Sam21] Samsung. *Samsung Expands Vital Blood Pressure and Electrocardiogram Tracking to Galaxy Watch3 and Galaxy Watch Active2 in 31 More Countries*. en. 2021. URL: https://news.samsung.com/global/samsung-expands-vital-blood-pressure-and-electrocardiogram-tracking-to-galaxy-watch3-and-galaxy-watch-active2-in-31-more-countries (visited on 06/17/2022).

[Sar04] Simona Sarzi Braga, Raffaella Vaninetti, Antonio Laporta, Anna Picozzi, and Roberto F. E Pedretti. "T wave alternans is a predictor of death in patients with congestive heart failure". In: *International Journal of Cardiology* 93 (Jan. 2004), pp. 31–38.

[Sen21] S. Senarath, G. Fernie, and A. Roshan Fekr. "Influential factors in remote monitoring of heart failure patients: A review of the literature and direction for future research". In: *Sensors* 21 (2021).

[Sha20] Fred Shaffer, Zachary M. Meehan, and Christopher L. Zerr. "A Critical Review of Ultra-Short-Term Heart Rate Variability Norms Research". In: *Frontiers in Neuroscience* 14 (2020).

[Spa21] Carmen Anna Maria Spaccarotella, Serena Migliarino, Annalisa Mongiardo, Jolanda Sabatino, Giuseppe Santarpia, Salvatore De Rosa, Antonio Curcio, and

Ciro Indolfi. “Measurement of the QT interval using the Apple Watch”. In: *Sci Rep* 11 (May 2021). Number: 1 Publisher: Nature Publishing Group, p. 10817.

[Spr22] Nora Sprenger, Alireza Sepehri Shamloo, Jonathan Schäfer, Sarah Burkhardt, Konstantinos Mouratis, Gerhard Hindricks, Andreas Bollmann, and Arash Arya. “Feasibility and Reliability of Smartwatch to Obtain Precordial Lead Electrocardiogram Recordings”. In: *Sensors (Basel)* 22.3 (Feb. 2022), p. 1217.

[Ste08] Phyllis K Stein, Devang Sanghavi, Peter P. Domitrovich, Robert A. Mackey, and Prakash Deedwania. “Ambulatory ECG-Based T-Wave Alternans Predicts Sudden Cardiac Death in High-Risk Post-MI Patients with Left Ventricular Dysfunction in the EPHESUS Study”. In: *Journal of Cardiovascular Electrophysiology* 19 (2008), pp. 1037–1042.

[Syd21] Hanna Sydow, Sandra Prescher, Friedrich Koehler, Kerstin Koehler, Marc Dorenkamp, Sebastian Spethmann, Benjamin Westerhoff, Christoph J. Wagner, Sebastian Liersch, Herbert Rebscher, Stefanie Wobbe-Ribinski, Heike Rindfleisch, Falk Müller-Riemenschneider, Stefan N. Willich, and Thomas Reinhold. “Cost-effectiveness of noninvasive telemedical interventional management in patients with heart failure: health economic analysis of the TIM-HF2 trial”. In: *Clin Res Cardiol* (Dec. 2021).

[Tho19] Elizabeth A. Thomson, Kayla Nuss, Ashley Comstock, Steven Reinwald, Sophie Blake, Richard E. Pimentel, Brian L. Tracy, and Kaigang Li. “Heart rate measures from the Apple Watch, Fitbit Charge HR 2, and electrocardiogram across different exercise intensities”. In: *J Sports Sci* 37 (June 2019), pp. 1411–1419.

[Tsa22] Connie W. Tsao, Aaron W. Aday, Zaid I. Almarzooq, Andea Z. Beaton, Marcio S. Bittencourt, ..., and Seth S. Martin. “Heart Disease and Stroke Statistics—2022 Update: A Report From the American Heart Association”. In: 145 (Feb. 2022). Publisher: American Heart Association, e153–e639.

[Val18] Raphael Vallat. “Pingouin: statistics in Python”. In: *The Journal of Open Source Software* 3.31 (Nov. 2018), p. 1026.

[Wan08] Norman C. Wang, Aldo P. Maggioni, Marvin A. Konstam, Faiez Zannad, Holly B. Krasa, John C. Burnett, Liliana Grinfeld, Karl Swedberg, James E. Udelson, Thomas Cook, Brian Traver, Christopher Zimmer, Cesare Orlandi, Mihai Gheo-

rghiade, and for the Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study With Tolvaptan (EVEREST) Investigators. “Clinical Implications of QRS Duration in Patients Hospitalized With Worsening Heart Failure and Reduced Left Ventricular Ejection Fraction”. In: *JAMA* 299 (June 2008), pp. 2656–2666.

[Wan18] Yan Wang, Shoushui Wei, Shuai Zhang, Yatao Zhang, Lina Zhao, Chengyu Liu, and Alan Murray. “Comparison of time-domain, frequency-domain and non-linear analysis for distinguishing congestive heart failure patients from normal sinus rhythm subjects”. In: *Biomedical Signal Processing and Control* 42 (Apr. 2018), pp. 30–36.

[Wat07] Eiichi Watanabe, Tomoharu Arakawa, Tatsushi Uchiyama, MaoQing Tong, Kenji Yasui, Hiroshi Takeuchi, Toshiaki Terasawa, Itsuo Kodama, and Hitoshi Hishida. “Prognostic significance of circadian variability of RR and QT intervals and QT dynamicity in patients with chronic heart failure”. In: *Heart Rhythm* 4 (Aug. 2007), pp. 999–1005.

[Wea09] Beth Davison Weatherley, Olga Milo-Cotter, G. Michael Felker, Nir Uriel, Edo Kaluski, Zvi Vered, Christopher M. O’Connor, Kirkwood F. Adams, and Gad Cotter. “Early worsening heart failure in patients admitted with acute heart failure – a new outcome measure associated with long-term prognosis?” In: *Fundamental & Clinical Pharmacology* 23 (2009), pp. 633–639.

[Weh21] D. Wehler, H. F. Jelinek, A. Gronau, N. Wessel, J. F. Kraemer, R. Krones, and T. Penzel. “Reliability of heart-rate-variability features derived from ultra-short ECG recordings and their validity in the assessment of cardiac autonomic neuropathy”. In: *Biomedical Signal Processing and Control* 68 (July 2021), p. 102651.

[Wer19] Stefanie Maria Werhahn, Henning Dathe, Thorsten Rottmann, Thomas Franke, Dan Vahdat, Gerd Hasenfuß, and Tim Seidler. “Designing meaningful outcome parameters using mobile technology: a new mobile application for telemonitoring of patients with heart failure”. In: *ESC Heart Failure* 6 (2019), pp. 516–525.

[Wis19] Bundesärztekammer (BÄK); Kassenärztliche Bundesvereinigung (KBV); Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). *NVL Chronische Herzinsuffizienz – Langfassung, 3. Auflage*. 2019.

[Wit19] Withings. *Move ECG Is Now Available in Europe*. en-US. Sept. 2019. URL: https://blog.withings.com/en/2019/09/06/move-ecg-now-available-in-europe/ (visited on 06/17/2022).

[Wit22a] Withings. *Fitness trackers and hybrid smartwatches by Withings*. en. 2022. URL: https://www.withings.com/de/en/watches (visited on 11/04/2022).

[Wit22b] Withings. *ScanWatch - What is Electrocardiogram (ECG)?* en-US. 2022. URL: https://support.withings.com/hc/en-us/articles/360004559098-ScanWatch-What-is-Electrocardiogram-ECG- (visited on 09/23/2022).

[Yam18] Shinya Yamada, Akiomi Yoshihisa, Yu Sato, Takamasa Sato, Masashi Kamioka, Takashi Kaneshiro, Masayoshi Oikawa, Atsushi Kobayashi, Hitoshi Suzuki, Takafumi Ishida, and Yasuchika Takeishi. "Utility of heart rate turbulence and T-wave alternans to assess risk for readmission and cardiac death in hospitalized heart failure patients". In: *Journal of Cardiovascular Electrophysiology* 29 (2018), pp. 1257–1264.

[Yan13] Clyde W. Yancy, Mariell Jessup, Biykem Bozkurt, Javed Butler, Donald E. Casey, Mark H. Drazner, Gregg C. Fonarow, Stephen A. Geraci, Tamara Horwich, James L. Januzzi, Maryl R. Johnson, Edward K. Kasper, Wayne C. Levy, Frederick A. Masoudi, Patrick E. McBride, John J. V. McMurray, Judith E. Mitchell, Pamela N. Peterson, Barbara Riegel, Flora Sam, Lynne W. Stevenson, W. H. Wilson Tang, Emily J. Tsai, and Bruce L. Wilkoff. "2013 ACCF/AHA Guideline for the Management of Heart Failure". In: *Journal of the American College of Cardiology* 62 (Oct. 2013). Publisher: American College of Cardiology Foundation, e147–e239.

[Zaf18] Barak Zafrir, Lars H. Lund, Cecile Laroche, Frank Ruschitzka, Maria G. Crespo-Leiro, Andrew J. S. Coats, Stefan D. Anker, Gerasimos Filippatos, Petar M. Seferovic, Aldo P. Maggioni, Manuel De Mora Martin, Lech Polonski, José Silva-Cardoso, Offer Amir, and ESC-HFA HF Long-Term Registry Investigators. "Prognostic implications of atrial fibrillation in heart failure with reduced, mid-range, and preserved ejection fraction: a report from 14 964 patients in the European Society of Cardiology Heart Failure Long-Term Registry". In: *Eur Heart J* 39 (Dec. 2018), pp. 4277–4284.